

# Performance limitation analysis in visual servo systems: bounding the location error introduced by image points matching

Graziano Chesi and Ho Lam Yung

**Abstract**— Visual servoing consists of positioning a robot end-effector based on the matching of some object features in the image. However, due to the presence of image noise, this matching can never be ensured, hence introducing an error on the final location of the robot. This paper addresses the problem of estimating the worst-case location error introduced by image points matching. In particular, we propose some strategies for computing upper bounds and lower bounds of such an error according to several possible measures for certain image noise intensity and camera-object configuration. These bounds provide an admissible region of the sought worst-case location error, and hence allow one to establish performance limitation of visual servo systems. Some examples are reported to illustrate the proposed strategies and their results.

**Index Terms**— Visual servoing, Image noise, Positioning accuracy, Convex optimization.

## I. INTRODUCTION

Eye-in-hand visual servoing consists of positioning a robot end-effector in a desired location by exploiting the visual information of some reference objects provided by a camera mounted on the end-effector itself. In particular, in the teaching-by-showing approach the camera (and, hence, the robot) is located in a location of interest, called desired location, from which some reference objects are visible. The view of the camera in this location, called desired view, is hence stored. Then, the camera is moved to another location of the scene from which the same reference objects are visible. The target is hence to steer the camera from its current location to the desired location by exploiting as feedback information the view of the camera in the current location, called current view, and the desired view previously stored.

In order to allow the camera to reach the desired location in the teaching-by-showing approach, several methods have been proposed, such as position-based visual servoing (see e.g. [1]), image-based visual servoing (see e.g. [2]), and 2 1/2 D visual servoing [3]. Other methods have proposed partition of the degrees of freedoms (see e.g. [4], [5]), global motion plan via navigation functions (see e.g. [6]), control invariant with respect to intrinsic parameters (see e.g. [7], [8]), use of complex image features via image moments (see e.g. [9]), switching strategies for ensuring the visibility constraint (see e.g. [10]), generation of circular-like trajectories for minimizing the trajectory length (see e.g. [11]), and path-planning methods for taking into account constraints (see e.g. [12]–[16]). See also [17]–[20].

G. Chesi and H. L. Yung are with the Department of Electrical and Electronic Engineering, University of Hong Kong. Contact information: please see <http://www.eee.hku.hk/~chesi>

In visual servoing the goal condition is defined as the matching of some object features in the image. However, this matching can never be ensured due to the presence of image noise, and this introduces an error on the final location of the robot. See also [21] which investigates the effect of image noise on the control law of position-based visual servoing and 2 1/2 D visual servoing.

This paper addresses the estimation of the worst-case location error introduced by image points matching. Specifically, we consider the computation of upper bounds and lower bounds of such an error, and propose to this end a strategy based on linear matrix inequality (LMI) for the former, and a strategy based on barrier functions for the latter. In these strategies the location error can be defined according to several possible measures, and it is a function of the image noise intensity and camera-object configuration. The derived upper bounds and lower bounds delimit an admissible region of the sought worst-case location error, and hence allow one to establish performance limitation of visual servo systems. Some examples are reported to illustrate the application of the proposed approach. This paper extends our previous results in [22] by introducing the derivation of lower bounds of the sought worst-case location error, allowing for the use of several different measures used to quantify such an error, and exploiting a complete parametrization for the rotational part which does not present singularities.

The organization of the paper is as follows. Section II defines the problem and notation. Section III describes the computation of the proposed upper and lower bounds for different measures. Section IV presents some illustrative examples. Lastly, Section V reports some concluding comments and possible directions for future research.

## II. FRAMEWORK DEFINITION

### A. Notation and problem statement

The notation exploited in this paper is as follows:

- $\mathbb{R}$ : real number set;
- $SO(3)$ : set of all rotation matrices of size  $3 \times 3$ ;
- $\mathbf{0}_n$ : null vector of size  $n \times 1$ ;
- $\mathbf{0}_{m \times n}$ : null matrix of size  $m \times n$ ;
- $\mathbf{I}_n$ : identity matrix of size  $n \times n$ ;
- $\mathbf{e}_i$ : null vector of size  $3 \times 1$  with the  $i$ th component set to 1;
- $\|\mathbf{X}\|$ : euclidean norm of vector/matrix  $\mathbf{X}$ ;
- $\|\mathbf{X}\|_\infty$ : infinity norm of vector/matrix  $\mathbf{X}$ ;
- $\mathbf{X}^T$ : transpose of a vector/matrix  $\mathbf{X}$ ;
- s.t.: subject to.

Let us consider an eye-in-hand visual servo system, such as a robotic manipulator or a mobile platform with a camera mounted on an end-effector. The task consists of controlling the robot so that the end-effector reaches a desired location by exploiting in closed-loop the camera view of some object features. In particular, the motion is terminated when the object features in the current camera view satisfactorily match those in the desired camera view, which have been previously recorded.

Let us observe that image noise unavoidably affects the image projections of the object features, and in particular it does each time in a different way being random image noise. Hence, even supposing that the camera has perfectly reached the desired location, the object features in the current camera view can be (and in generally are) different from those ones previously recorded from the same location. Vice versa, even supposing that the object features in the current camera view perfectly match those in the desired camera view, the current location can be (and in generally is) different from the desired one. This means that image noise unavoidably introduces a final location error in visual servoing as the goal condition is defined via visual matching.

In this paper we address the problem of bounding this error, which depends on the level of image noise, on the camera parameters, and on the object features.

### B. Mathematical formulation

Let  $F^{abs}$  be an absolute frame in the 3D space. We denote with  $F = (\mathbf{O}, \mathbf{c})$  the frame of the current camera of the visual servo system expressed with respect to the absolute frame  $F^{abs}$ , where  $\mathbf{O} \in SO(3)$  is a rotation matrix which defines the orientation, and  $\mathbf{c} \in \mathbb{R}^3$  is a vector which defines the translation. Similarly we denote with  $F^* = (\mathbf{O}^*, \mathbf{c}^*)$  the frame of the desired camera of the visual servo system.

Let  $\mathbf{q}_1, \dots, \mathbf{q}_N \in \mathbb{R}^3$  be a set of 3D points expressed with respect to the absolute frame  $F^{abs}$ . The  $i$ th 3D point  $\mathbf{q}_i$  projects onto the camera frame  $F$  at the point  $\mathbf{p}_i = (p_{i,1}, p_{i,2}, 1)^T \in \mathbb{R}^3$  expressed in homogeneous coordinates and given by

$$d_i \mathbf{p}_i = \mathbf{A} \mathbf{O}^T (\mathbf{q}_i - \mathbf{c}) \quad (1)$$

where  $d_i$  is the depth of the point with respect to  $F$ , and  $\mathbf{A} \in \mathbb{R}^{3 \times 3}$  is the upper-triangular matrix containing the camera intrinsic parameters:

$$\mathbf{A} = \begin{pmatrix} f_x & s & u_x \\ 0 & f_y & u_y \\ 0 & 0 & 1 \end{pmatrix}, \quad (2)$$

being  $f_x, f_y \in \mathbb{R}$  the focal lengths,  $u_x, u_y \in \mathbb{R}$  the coordinates of the principal point, and  $s \in \mathbb{R}$  the aspect ratio. Similarly, the  $i$ th 3D point  $\mathbf{q}_i$  projects onto the camera frame  $F^*$  at the point  $\mathbf{p}_i^* = (p_{i,1}^*, p_{i,2}^*, 1)^T \in \mathbb{R}^3$  expressed in homogeneous coordinates and given by

$$d_i^* \mathbf{p}_i^* = \mathbf{A} \mathbf{O}^{*T} (\mathbf{q}_i - \mathbf{c}^*) \quad (3)$$

where  $d_i^*$  is the depth of the point with respect to  $F^*$ .

The motion (or camera pose) between  $F$  and  $F^*$  can be described by the pair

$$(\mathbf{R}, \mathbf{t}) \in SO(3) \times \mathbb{R}^3 \quad (4)$$

where  $\mathbf{R}$  and  $\mathbf{t}$  are the rotational and translational components respectively given by

$$\mathbf{R} = \mathbf{O}^{*T} \mathbf{O} \quad (5)$$

while  $\mathbf{t}$  is given by

$$\mathbf{t} = \frac{\mathbf{O}^{*T} (\mathbf{c} - \mathbf{c}^*)}{\|\mathbf{O}^{*T} (\mathbf{c} - \mathbf{c}^*)\|} \quad (6)$$

(these components are expressed with respect to the desired camera frame  $F^*$ ).

In this paper we consider the situation where the object features used in the teaching-by-showing approach described in Section II-A are points, which is the typical case. Hence, let  $\mathbf{p}, \mathbf{p}^* \in \mathbb{R}^{2N}$  be the vectors containing the image projections of the 3D points  $\mathbf{q}_1, \dots, \mathbf{q}_N$  according to

$$\begin{aligned} \mathbf{p} &= (p_{1,1}, p_{1,2}, \dots, p_{N,1}, p_{N,2})' \\ \mathbf{p}^* &= (p_{1,1}^*, p_{1,2}^*, \dots, p_{N,1}^*, p_{N,2}^*)' \end{aligned} \quad (7)$$

The goal condition of the visual servoing in Step 3 can be expressed as

$$\|\mathbf{p} - \mathbf{p}^*\|_\infty \leq \varepsilon \quad (8)$$

where  $\varepsilon \in \mathbb{R}$  is a threshold chosen to limit the distance between  $\mathbf{p}$  and  $\mathbf{p}^*$  (for example, via the infinity norm).

The problem addressed in this paper can be formulated as computing upper and lower bounds of the worst-case location error introduced by image points matching through the goal condition (8). In particular we consider the worst-case rotational error

$$\begin{aligned} s_r(\varepsilon) &= \sup_{(\mathbf{R}, \mathbf{t}) \in SO(3) \times \mathbb{R}^3} \mu(\mathbf{R}) \\ \text{s.t. } &\|\mathbf{p} - \mathbf{p}^*\|_\infty \leq \varepsilon \end{aligned} \quad (9)$$

and the worst-case translational error

$$\begin{aligned} s_t(\varepsilon) &= \sup_{(\mathbf{R}, \mathbf{t}) \in SO(3) \times \mathbb{R}^3} \nu(\mathbf{t}) \\ \text{s.t. } &\|\mathbf{p} - \mathbf{p}^*\|_\infty \leq \varepsilon \end{aligned} \quad (10)$$

where  $\mu : SO(3) \rightarrow \mathbb{R}$  and  $\nu : \mathbb{R}^3 \rightarrow \mathbb{R}$  are some norms of interest (we will consider several cases in the next sections).

In the sequel we will consider without loss of generality that the desired camera frame  $F^*$  coincides with the absolute frame  $F^{abs}$ .

### III. BOUNDS COMPUTATION

First of all, let us observe how image noise affects the goal condition (8). Let us denote with  $\hat{\mathbf{p}}^*$  and  $\hat{\mathbf{p}}$  the estimates of  $\mathbf{p}^*$  and  $\mathbf{p}$  corrupted by image noise, in particular according to

$$\begin{cases} \hat{\mathbf{p}} &= \mathbf{p} + \mathbf{n} \\ \hat{\mathbf{p}}^* &= \mathbf{p}^* + \mathbf{n}^* \end{cases} \quad (11)$$

where  $\mathbf{n}, \mathbf{n}^* \in \mathbb{R}^{2N}$  are vectors containing image noise (for instance due to image quantization, lighting, features extraction, etc...). Suppose that  $\mathbf{n}, \mathbf{n}^*$  are bounded by

$$\begin{aligned} \|\mathbf{n}\|_\infty &\leq \zeta \\ \|\mathbf{n}^*\|_\infty &\leq \zeta \end{aligned} \quad (12)$$

where  $\zeta \in \mathbb{R}$  is a bound of the image noise intensity in both current and desired views. Then, it clearly follows that

$$\|\hat{\mathbf{p}} - \hat{\mathbf{p}}^*\|_\infty \leq \varepsilon \text{ guarantees only } \|\mathbf{p} - \mathbf{p}^*\|_\infty \leq \varepsilon + 2\zeta \quad (13)$$

i.e., one cannot guarantee that the real image error  $\|\mathbf{p}_{true} - \mathbf{p}_{true}^*\|_\infty$  converges to a value smaller than  $2\zeta$ . This clearly motivates the investigation of the location error introduced by image points matching.

Therefore, from now on we will consider the computation of the errors

$$\begin{cases} s_r(\delta) \text{ and } s_t(\delta) \\ \delta = \varepsilon + 2\zeta \end{cases} \quad (14)$$

where  $\delta \in \mathbb{R}$  represents the total image error.

Before proceeding, let us parameterize the rotation matrix through the Euler parameter as follows:

$$\begin{cases} \mathbf{R} = \Omega(\mathbf{a}) \\ \|\mathbf{a}\| = 1 \end{cases} \quad (15)$$

where  $\mathbf{a} = (a_1, \dots, a_4)^T \in \mathbb{R}^4$  is a unit-norm vector which represents the Euler parameter of  $\mathbf{R}$ , and  $\Omega(\mathbf{a})$  is the matrix function

$$\Omega(\mathbf{a}) = \begin{pmatrix} a_1^2 - a_2^2 - a_3^2 + a_4^2 & 2(a_1a_2 - a_3a_4) \\ 2(a_1a_2 + a_3a_4) & -a_1^2 + a_2^2 - a_3^2 + a_4^2 \\ 2(a_1a_3 - a_2a_4) & 2(a_2a_3 + a_1a_4) \\ 2(a_1a_3 + a_2a_4) & 2(a_2a_3 - a_1a_4) \\ -a_1^2 - a_2^2 + a_3^2 + a_4^2 \end{pmatrix}. \quad (16)$$

Let us observe that this parametrization is complete, in the sense that:

- 1) for any unit-norm vector  $\mathbf{a} \in \mathbb{R}^4$  it follows that  $\Omega(\mathbf{a})$  is a rotation matrix in  $SO(3)$ ;
- 2) for any rotation matrix  $\mathbf{R} \in SO(3)$  there exists a vector  $\mathbf{a} \in \mathbb{R}^4$  with unit-norm such that  $\mathbf{R} = \Omega(\mathbf{a})$ .

#### A. Upper bounds

Let us consider first the computation of upper bounds of  $s_r(\delta)$  and  $s_t(\delta)$  in (9)–(10). We will show that this step can be solved by exploiting convex optimizations. Indeed, consider the constraint  $\|\mathbf{p} - \mathbf{p}^*\|_\infty < \delta$  in the computation of  $s_r(\delta)$  and  $s_t(\delta)$ . From (1), (3), (5) and (6) it follows that for the  $i$ th point we can write

$$\mathbf{p}_i - \mathbf{p}_i^* = \mathbf{A} \frac{\Omega(\mathbf{a})\mathbf{q}_i + \mathbf{t}}{\mathbf{e}_3^T (\Omega(\mathbf{a})\mathbf{q}_i + \mathbf{t})} - \mathbf{A} \frac{\mathbf{q}_i}{\mathbf{e}_3^T \mathbf{q}_i} \quad (17)$$

where it has been taken into account that  $\mathbf{q}_i$  is expressed with respect to the desired camera frame  $F^*$ , which coincides with the absolute frame  $F^{abs}$ . Hence, we have that

$$\|\mathbf{p}_i - \mathbf{p}_i^*\|_\infty \leq \delta \quad (18)$$

if and only if

$$\begin{cases} |f_{i,3}g_{i,1} - f_{i,1}g_{i,3}| \leq \delta f_{i,3}g_{i,3} \\ |f_{i,3}g_{i,2} - f_{i,2}g_{i,3}| \leq \delta f_{i,3}g_{i,3} \\ g_{i,3} > 0 \end{cases} \quad (19)$$

where  $f_{i,j} \in \mathbb{R}$  is a constant and  $g_{i,j}$  is a polynomial (in  $\mathbf{a}$  and  $\mathbf{t}$ ) given by

$$\begin{aligned} f_{i,j} &= \mathbf{e}_j^T \mathbf{A} \mathbf{q}_i \\ g_{i,j} &= \mathbf{e}_j^T \mathbf{A} (\Omega(\mathbf{a})\mathbf{q}_i + \mathbf{t}). \end{aligned} \quad (20)$$

This means that the constraint  $\|\mathbf{p} - \mathbf{p}^*\|_\infty < \delta$  can be compactly expressed by defining the polynomials

$$h_{i,j,k} = (-1)^k (f_{i,3}g_{i,j} - f_{i,j}g_{i,3}) + \delta f_{i,3}g_{i,3}. \quad (21)$$

Indeed:

$$\|\mathbf{p} - \mathbf{p}^*\|_\infty \leq \delta \quad (22)$$

if and only if

$$h_{i,j,k} \geq 0 \quad \forall (i, j, k) \in \mathcal{I}. \quad (23)$$

where  $\mathcal{I}$  is the set

$$\begin{aligned} \mathcal{I} &= \{(i, j, k) : i = 1, \dots, N, j = 1, 2, k = 1, 2\} \\ &\cup \{(i, j, k) : i = 1, \dots, N, j = 3, k = 1\}. \end{aligned} \quad (24)$$

In fact, let us observe that the first two constraints in (20) are recovered by the first set on the right hand side of (24), while the third constraint is recovered by the second set on the right hand side of (24) since  $h_{i,3,1} = \delta f_{i,3}g_{i,3}$ . Therefore,  $s_r(\delta)$  and  $s_t(\delta)$  can be rewritten as

$$s_r(\delta) = \sup_{\substack{\mathbf{a} \in \mathbb{R}^4, \|\mathbf{a}\|=1, \mathbf{t} \in \mathbb{R}^3 \\ \text{s.t. (23)}}} \mu(\Omega(\mathbf{a})) \quad (25)$$

and

$$s_t(\delta) = \sup_{\substack{\mathbf{a} \in \mathbb{R}^4, \|\mathbf{a}\|=1, \mathbf{t} \in \mathbb{R}^3 \\ \text{s.t. (23)}}} \nu(\mathbf{t}) \quad (26)$$

The next step consists of solving (25)–(26) via convex optimizations. To this end, let us introduce the polynomials

$$\begin{aligned} b_r &= \gamma - \mu(\Omega(\mathbf{a})) - \sum_{(i,j,k) \in \mathcal{I}} u_{i,j,k} h_{i,j,k} \\ &\quad - u(1 - \|\mathbf{a}\|^2) \\ b_t &= \gamma - \nu(\mathbf{t}) - \sum_{(i,j,k) \in \mathcal{I}} u_{i,j,k} h_{i,j,k} \\ &\quad - u(1 - \|\mathbf{a}\|^2) \end{aligned} \quad (27)$$

where  $u, u_{i,j,k} \in \mathbb{R}$  are auxiliary polynomials and  $\gamma \in \mathbb{R}$  is an auxiliary scalar to be determined. Let  $\mathbf{v}_b$  be a vector containing any base for the polynomials  $b_r, b_t$ , and let  $\mathbf{v}_u$  be a similar vector for the polynomials  $u_{i,j,k}$ . Then, the polynomials  $b_r, b_t, u_{i,j,k}$  can be expressed as

$$\begin{cases} b_r = \mathbf{v}_b^T \mathbf{B}_r \mathbf{v}_b \\ b_t = \mathbf{v}_b^T \mathbf{B}_t \mathbf{v}_b \\ u_{i,j,k} = \mathbf{v}_u^T \mathbf{U}_{i,j,k} \mathbf{v}_u \end{cases} \quad (28)$$

where  $\mathbf{B}_r, \mathbf{B}_t$ , and  $\mathbf{U}_{i,j,k}$  are any symmetric matrices of suitable dimensions satisfying (28). Lastly, let  $\mathbf{L}(\alpha)$  be any linear parametrization of the linear set

$$\mathcal{L} = \{\mathbf{L} = \mathbf{L}^T : \mathbf{v}_b^T \mathbf{L}(\alpha) \mathbf{v}_b = 0\} \quad (29)$$

where  $\alpha$  is a free vector. The representation (28) is known as square matricial representation (SMR) and has been introduced in [23]. In [24], simple algorithms for the computation of SMR matrices as  $\mathbf{B}_r$ ,  $\mathbf{B}_t$ , and  $\mathbf{U}_{i,j,k}$  as well as the function  $\mathbf{L}(\alpha)$  are provided. In practice, one builds the vectors  $\mathbf{v}_b$  and  $\mathbf{v}_u$  by freely choosing any possible permutation of the admissible monomials in  $b_r, b_t, u_{i,j,k}$ . Then, one introduces the free matrix variables  $\mathbf{U}_{i,j,k}$ , hence defining the polynomials  $u_{i,j,k}, b_r, b_t$ . One hence builds possible matrices  $\mathbf{B}_r$  and  $\mathbf{B}_t$  (both depending affine linearly on  $\mathbf{U}_{i,j,k}$ ) through simple coefficients equalization of the equations in (28). Lastly, one constructs the matrix  $\mathbf{L}(\alpha)$  by selecting any possible parametrization of the linear set  $\mathcal{L}$ .

Hence, we finally have that upper bounds of  $s_r(\delta)$  and  $s_t(\delta)$  can be obtained as

$$\begin{aligned} s_r^+(\delta) &= \min_{\gamma, \alpha, \mathbf{U}_{i,j,k}} \gamma \\ \text{s.t. } \left\{ \begin{array}{l} \mathbf{B}_r + \mathbf{L}(\alpha) \geq 0 \\ \mathbf{U}_{i,j,k} \geq 0 \quad \forall (i,j,k) \in \mathcal{I} \end{array} \right. \end{aligned} \quad (30)$$

and

$$\begin{aligned} s_t^+(\delta) &= \min_{\gamma, \alpha, \mathbf{U}_{i,j,k}} \gamma \\ \text{s.t. } \left\{ \begin{array}{l} \mathbf{B}_t + \mathbf{L}(\alpha) \geq 0 \\ \mathbf{U}_{i,j,k} \geq 0 \quad \forall (i,j,k) \in \mathcal{I} \end{array} \right. \end{aligned} \quad (31)$$

In fact, the constraints in (30) and (31) ensure that the polynomials  $b_r, b_t, u_{i,j,k}$  are non-negative, and hence that  $\gamma > \mu(\Omega(\mathbf{a}))$  in (30) and  $\gamma > \nu(\mathbf{t})$  in (31) for all values of  $\mathbf{a}, \mathbf{t}$  such that  $h_{i,j,k} \geq 0$  and  $\|\mathbf{a}\| = 1$ .

Let us observe that (30) and (31) are minimizations with linear costs and linear matrix inequality (LMI) constraints, and hence are convex optimizations [25]. For more details about the SMR the reader is referred to [26]. Other applications of the SMR in computer vision can be found in [27], [28].

### B. Lower bounds

In Section III-A we have derived upper bounds of the worst-case errors  $s_r(\delta)$  and  $s_t(\delta)$ . In this section we consider the computation of lower bounds of these errors. The idea is to generate a sequence of camera poses  $(\mathbf{R}, \mathbf{t})$  such that:

- 1) the condition  $\|\mathbf{p} - \mathbf{p}^*\|_\infty \leq \delta$  holds true for all values of the sequence, i.e. the camera pose  $(\mathbf{R}, \mathbf{t})$  yields to an admissible image error;
- 2) the sequence approaches the sought worst-case error (either  $s_r(\delta)$  or  $s_t(\delta)$ ).

In fact, the first condition ensures that each camera pose of the sequence provides a lower bound of the sought worst-case error, while the second condition forces the sequence to provide, possibly, tight lower bounds.

To this end, let us define the functions

$$\psi_r = \begin{cases} \mu(\Omega(\mathbf{a})) & \text{if } w \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (32)$$

and

$$\psi_t = \begin{cases} \nu(\mathbf{t}) & \text{if } w \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (33)$$

where  $w$  and  $w_1$  are given by

$$\begin{cases} w = \min\{w_1, w_2\} \\ w_1 = 1 - a_1^2 - a_2^2 - a_3^2 \\ w_2 = \delta - \|\mathbf{p} - \mathbf{p}^*\|_\infty \text{ evaluated with} \\ \mathbf{a} = (a_1, a_2, a_3, \sqrt{w_1})^T \end{cases} \quad (34)$$

Let us observe that  $w$  is a barrier-like function since it becomes negative whenever any of the following conditions holds:

- 1) the parameter  $\mathbf{a}$  is not admissible as  $\|\mathbf{a}\| > 1$ ;
- 2) the parameters  $(\mathbf{a}, \mathbf{t})$  are not admissible as  $\|\mathbf{p} - \mathbf{p}^*\|_\infty > \delta$ .

Instead, whenever  $w$  is nonnegative, an admissible parameter  $\mathbf{a}$  is simply given by  $(a_1, a_2, a_3, \sqrt{w_1})^T$  as in (34), and moreover all admissible parameters  $\mathbf{a}$  are recovered by letting  $a_1, a_2, a_3$  vary in the unit sphere (observe in fact that there is no loss of generality in taking the positive square root of  $w_1$  because  $\mathbf{a}$  and  $-\mathbf{a}$  provide the same rotation matrix). Then, lower bounds of  $s_r(\delta)$  and  $s_t(\delta)$  can be obtained as

$$s_r^-(\delta) = \max_{a_1, a_2, a_3, \mathbf{t}} \psi_r \quad (35)$$

and

$$s_t^-(\delta) = \max_{a_1, a_2, a_3, \mathbf{t}} \psi_t \quad (36)$$

which can be solved by using, for instance, the simplex algorithm, which can handle the non-differentiability of the functions  $\psi_r$  and  $\psi_t$ . The found solutions may be non-optimal since these functions are non-convex, nevertheless any found solution it is guaranteed to be a lower bound of the sought worst-case error since it is obtained for an admissible camera pose. Lastly, let us observe that an initialization for (35)–(36) is simply given by  $(0, 0, 0, \mathbf{0}_3)^T$  which is admissible for any  $\delta$ .

### C. Error measures

In this section we describe some norms  $\mu$  and  $\nu$  which can be used in the proposed approach to define the errors  $s_r(\delta)$  and  $s_t(\delta)$ . Let us consider first the rotation. Then, a possibility is to select

$$\mu(\mathbf{R}) = \theta \quad (37)$$

where  $\theta \in [0, \pi]$  is the angle in the representation of  $\mathbf{R}$  via exponential coordinates. Let us observe that  $\theta$  is not a polynomial function of  $\mathbf{a}$  and hence (37) cannot be directly used in (27). However, this problem can be simply solved by defining  $\mu$  as

$$\mu(\Omega(\mathbf{a})) = a_1^2 + a_2^2 + a_3^2 \quad (38)$$

which indirectly provides (37) because

$$a_1^2 + a_2^2 + a_3^2 = \left(\sin \frac{\theta}{2}\right)^2 \quad (39)$$

(see for instance [16] for details about this relationship).

Let us consider now the translation, for which one can similarly select

$$\nu(\mathbf{t}) = \|\mathbf{t}\| \quad (40)$$

which provides the standard euclidean measure of the translation. Again, this is not a polynomial function of  $\mathbf{t}$  and hence cannot be directly used in (27), nevertheless one can simply use  $\|\mathbf{t}\|^2$  which is polynomial and indirectly provides (40).

Another useful possibility is to select

$$\nu(\mathbf{t}) = \|t_i\| \quad (41)$$

for some  $i = 1, 2, 3$ , which provides a bound of the translation along the  $i$ th axis, and which can be obtained similarly to the technique used for (40). Several other norms can be similarly defined.

#### IV. EXAMPLES

In this section we present some examples of the proposed approach. The the upper bounds  $s_r^+(\delta), s_t^+(\delta)$  in (30)–(31) and the lower bounds  $s_r^-(\delta), s_t^-(\delta)$  in (35)–(36) have been computed by using Matlab. Their computational time is less than 10 seconds on a standard PC.

##### A. Example 1

Let us consider the situation shown in Figure 1a where a camera is observing four dices. The chosen object points are the centers of the eight large dots. The screen size is  $640 \times 480$  pixels, and the camera intrinsic parameters are

$$\mathbf{A} = \begin{pmatrix} 500 & 0 & 320 \\ 0 & 500 & 240 \\ 0 & 0 & 1 \end{pmatrix}. \quad (42)$$

Figure 1b shows the corresponding camera view. The problem is to estimate the worst-case location error introduced by image points matching, i.e. the errors  $s_r(\delta)$  and  $s_t(\delta)$  in (9)–(10) where  $\delta$  represents the total image error in (14).

To this end, let us select the norms  $\mu(\mathbf{R}) = \theta$  and  $\nu = \|\mathbf{t}\|$  as in (37) and (40). Then, we compute the upper bounds  $s_r^+(\delta), s_t^+(\delta)$  in (30)–(31) and the lower bounds  $s_r^-(\delta), s_t^-(\delta)$  in (35)–(36) for some values of  $\delta$ . We find the values shown in Table I.

$\delta$ [pixels]	$s_r^-(\delta)$ [deg]	$s_r^+(\delta)$ [deg]	$s_t^-(\delta)$ [mm]	$s_t^+(\delta)$ [mm]
0.5	0.122	0.352	0.384	0.605
1.0	0.247	0.705	0.785	1.21
1.5	0.370	1.06	1.25	1.82
2.0	0.491	1.41	1.78	2.42

TABLE I

EXAMPLE 1. UPPER AND LOWER BOUNDS OF THE WORST-CASE LOCATION ERRORS FOR THE OBJECT POINTS IN FIGURE 1 ACCORDING TO THE MEASURES  $\mu(\mathbf{R}) = \theta$  AND  $\nu = \|\mathbf{t}\|$  FOR SOME VALUES OF  $\delta$ .

##### B. Example 2

Let us consider the real image shown in Figure 2 where twelve object points, indicated by the “+” marks, are observed. The 3D points have been estimated by acquiring a second image from another location, and then performing a standard object reconstruction via the camera pose estimated

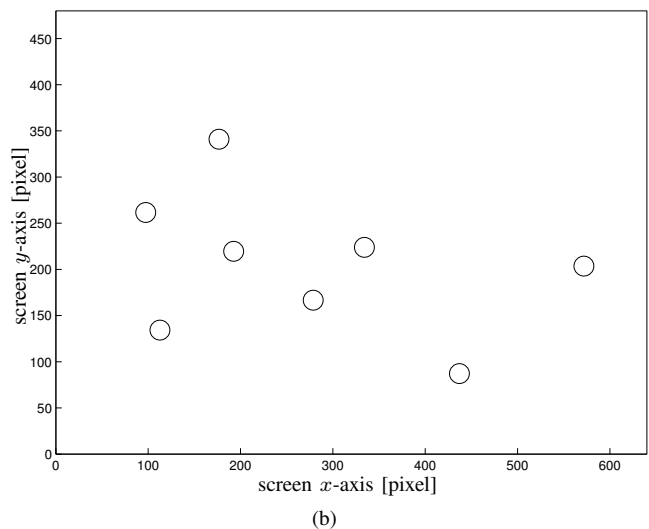
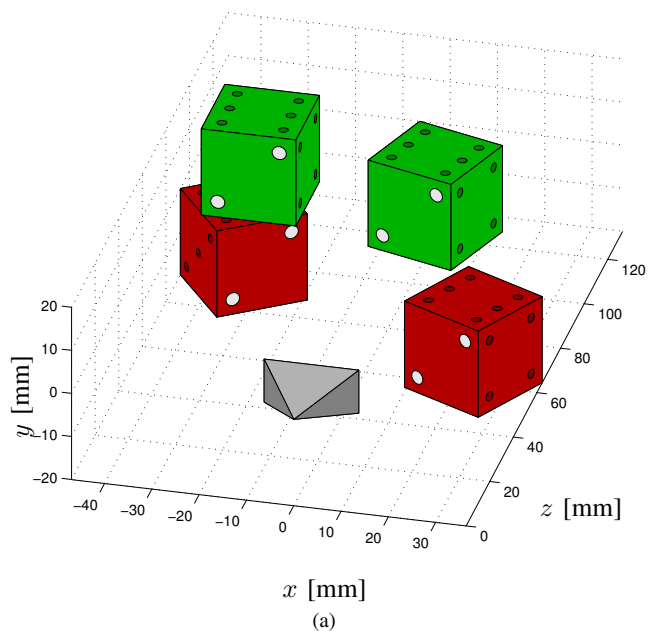


Fig. 1. Example 1. (a) 3D scene with the camera observing three dices. (b) Corresponding camera view: the used object points are the centers of the 8 large dots of the dices.

through the essential matrix with calibrated camera and with the knowledge of the distance among two points in the 3D space (in order to estimate the norm of the translation). The screen size is  $659 \times 493$  pixels.

Let us select the norms  $\mu(\mathbf{R}) = \theta$  and  $\nu = \|\mathbf{t}\|$ . We find the values shown in Table II.

#### V. CONCLUSION

We have addressed the estimation of the worst-case location error introduced by image points matching. Specifically, we have proposed some strategies for computing upper bounds and lower bounds of such an error according to several possible measures for certain image noise intensity and camera-object configuration. The computation of the upper bounds is based on LMI, while the computation of



Fig. 2. Example 2. Real image with twelve observed points.

$\delta$ [pixels]	$s_r^-(\delta)$ [deg]	$s_r^+(\delta)$ [deg]	$s_t^-(\delta)$ [mm]	$s_t^+(\delta)$ [mm]
0.5	0.060	1.46	0.148	0.295
1.0	0.119	2.91	0.301	0.591
1.5	0.180	4.36	0.505	0.892
2.0	0.240	5.82	0.658	1.20

TABLE II

EXAMPLE 2. UPPER AND LOWER BOUNDS OF THE WORST-CASE LOCATION ERRORS FOR THE OBJECT POINTS IN FIGURE 2 ACCORDING TO THE MEASURES  $\mu(\mathbf{R}) = \theta$  AND  $\nu = \|\mathbf{t}\|$  FOR SOME VALUES OF  $\delta$ .

the lower bounds is obtained through barrier functions. These bounds provide an admissible region of the sought worst-case location error.

The proposed strategies allows one to investigate an important issue of visual servoing schemes. Indeed, in such schemes a robot end-effector is positioned by matching some object features in the image, which are unavoidably affected by image noise. Therefore, such a positioning is unavoidably affected by errors, and the bounds proposed in this paper allows one to investigate this performance limitation.

Future work will be devoted to derive procedures for establishing tightness of the obtained bounds, and to analyze the propagation of the location error introduced by image points matching in closed-loop control systems.

#### ACKNOWLEDGEMENT

The authors would like to thank the Associate Editor and Reviewers for their time and comments.

#### REFERENCES

[1] B. Thuilot, P. Martinet, L. Cordesses, and J. Gallice, "Position based visual servoing: keeping the object in the field of vision," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Washington, D.C., 2002, pp. 1624–1629.

[2] K. Hashimoto, T. Kimoto, T. Ebine, and H. Kimura, "Manipulator control with image-based visual servo," in *Proc. IEEE Int. Conf. on Robotics and Automation*, 1991, pp. 2267–2272.

[3] E. Malis, F. Chaumette, and S. Boudet, "2 1/2 D visual servoing," *IEEE Trans. on Robotics and Automation*, vol. 15, no. 2, pp. 238–250, 1999.

[4] P. I. Corke and S. Hutchinson, "A new partitioned approach to image-based visual servo control," *IEEE Trans. on Robotics and Automation*, vol. 17, no. 4, pp. 507–515, 2001.

[5] P. Oh and P. Allen, "Visual servoing by partitioning degrees-of-freedom," *IEEE Trans. on Robotics and Automation*, vol. 17, no. 1, pp. 1–17, 2001.

[6] N. J. Cowan and D. E. Chang, "Geometric visual servoing," *IEEE Trans. on Robotics*, vol. 21, no. 6, pp. 1128–1138, 2005.

[7] E. Malis, "Visual servoing invariant to changes in camera-intrinsic parameters," *IEEE Trans. on Robotics and Automation*, vol. 20, no. 1, pp. 72–81, 2004.

[8] F. Schramm and G. Morel, "Ensuring visibility in calibration-free path planning for image-based visual servoing," *IEEE Trans. on Robotics*, vol. 22, no. 4, pp. 848–854, 2006.

[9] O. Tahri and F. Chaumette, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Trans. on Robotics*, vol. 21, no. 6, pp. 1116–1127, 2005.

[10] G. Chesi, K. Hashimoto, D. Prattichizzo, and A. Vicino, "Keeping features in the field of view in eye-in-hand visual servoing: a switching approach," *IEEE Trans. on Robotics*, vol. 20, no. 5, pp. 908–913, 2004.

[11] G. Chesi and A. Vicino, "Visual servoing for large camera displacements," *IEEE Trans. on Robotics*, vol. 20, no. 4, pp. 724–735, 2004.

[12] Y. Mezouar and F. Chaumette, "Path planning for robust image-based control," *IEEE Trans. on Robotics and Automation*, vol. 18, no. 4, pp. 534–549, 2002.

[13] H. Zhang and J. Ostrowski, "Visual motion planning for mobile robots," *IEEE Trans. on Robotics and Automation*, vol. 18, no. 2, pp. 199–208, 2002.

[14] J. Park and M. Chung, "Path planning with uncalibrated stereo rig for image-based visual servoing under large pose discrepancy," *IEEE Trans. on Robotics and Automation*, vol. 19, no. 2, pp. 250–258, 2003.

[15] B. Allotta and D. Fioravanti, "3D motion planning for image-based visual servoing tasks," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Barcelona, Spain, 2005.

[16] G. Chesi and Y. S. Hung, "Global path-planning for constrained and optimal visual servoing," *IEEE Trans. on Robotics*, vol. 23, no. 5, pp. 1050–1060, 2007.

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

[18] K. Hashimoto, "A review on vision-based control of robot manipulators," *Advanced Robotics*, vol. 17, no. 10, pp. 969–991, 2003.

[19] F. Chaumette and S. Hutchinson, "Visual servo control, part I: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, pp. 82–90, 2006.

[20] —, "Visual servo control, part II: Advanced approaches," *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, pp. 109–118, 2007.

[21] V. Kyrki, D. Kragic, and H. Christensen, "Measurement errors in visual servoing," in *Proc. IEEE Int. Conf. on Robotics and Automation*, New Orleans, Louisiana, 2004, pp. 1861–1867.

[22] G. Chesi and Y. S. Hung, "Image noise induced errors in camera positioning," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 8, pp. 1476–1480, 2007.

[23] G. Chesi, A. Tesi, A. Vicino, and R. Genesio, "On convexification of some minimum distance problems," in *5th European Control Conf.*, Karlsruhe, Germany, 1999.

[24] G. Chesi, A. Garulli, A. Tesi, and A. Vicino, "Solving quadratic distance problems: an LMI-based approach," *IEEE Trans. on Automatic Control*, vol. 48, no. 2, pp. 200–212, 2003.

[25] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia: SIAM, 1994.

[26] G. Chesi, A. Garulli, A. Tesi, and A. Vicino, *Homogeneous Polynomial Forms for Robustness Analysis of Uncertain Systems*. Springer (in press).

[27] G. Chesi, A. Garulli, A. Vicino, and R. Cipolla, "Estimating the fundamental matrix via constrained least-squares: a convex approach," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 397–401, 2002.

[28] G. Chesi, "Camera displacement via constrained minimization of the algebraic error," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 370–375, 2009.