A Decoupled Image Space Approach to Visual Servo Control of a Robotic Manipulator

Robert Mahony Dep. of Eng., Australian Nat. Uni., ACT, 0200 Australia. email: mahony@ieee.org Tarek Hamel Cemif-SC FRE-CNRS 2494, 40 rue du Pelvoux, 91020 Evry France. email: thamel@iup.univ-evry.fr Francois Chaumette Irisa/Inria Rennes, Campus Universitaire de Beaulieu, 35042 Rennes, France. email: Francois.Chaumette@irisa.fr

Abstract

An image-based visual servo control is presented for a robotic manipulator. The proposed control design addresses visual servo of 'eye-in-hand' type systems. Using a novel representation of the visual error based on a spherical representation of target centroid information along with a measure of the rotation between the camera and the target, the control of the position and orientation is decoupled. A non-linear gain introduced into the orientation feedback kinematics prevents the target image from leaving the visual field. Semi-global convergence of the closed loop system is proved.

1 Introduction

Visual servo algorithms have been extensively developed in the robotics field over the last ten years [7, 19]. Most visual servo control has been developed for serial-link robotic manipulators with the camera typically mounted on the end-effector [12]. Visual servo systems may be divided into two main classes [15]: Position-based visual servo (PBVS) involves reconstruction of the target pose with respect to the robot and results in a Cartesian motion planning problem. Problems with this approach include the necessity of an accurate 3D model of the target, sensitivity to camera calibration errors, poor robustness of pose estimation and the tendency for image features to leave the camera field of view during the task. Imagebased visual servo (IBVS) treats the problem as one of controlling features in the image plane, such that moving features to a goal configuration implicitly results in the task being accomplished [7]. Feature errors are mapped to actuator inputs via the inverse of an image Jacobian matrix. Features can be derived from image features such as points, lines and circles. Issues associated with optimal selection of features and partitioned control where some task degrees of freedom are controlled visually and others by position or force [13, 2] have been considered in depth. IBVS avoids many of the robustness and calibration problems associated with PBVS, however, it has its own problems [3]. Foremost in the classical approach is

the need to determine depth information for each feature point in the visual data. Various approaches have been reported including estimation via partial pose estimation [15], adaptive control [17] and estimation of the image Jacobian using quasi-Newton techniques [11, 18]. More recently, there has been considerable interest in hybrid control methods whereby translational and rotational control are treated separately [15, 6]. Hybrid methods, however, share with the PBVS approach a difficulty to keep all image features within the camera's field of view. Contemporary work [16, 4] aims to address these issues.

In this paper we propose an image-based visual servo control for a robotic manipulator. The model considered is that of an 'eye-in-hand' type configuration, where the camera is attached to the end effector of the robotic arm. The approach taken is based on recent work by the authors [9] in which an image based error is introduced for which the dynamics have certain passivity-like properties. The visual error representation used is based on a spherical representation of target centroid information. Using this representation it is possible to express the position and orientation kinematics of the visual error as a partially decoupled system. An exponentially stable feedback law is proposed for stabilization of the position. The feedback law for the orientation kinematics is modified via a simple non-linear gain to ensure that the target image never leaves the field of vision of the camera.

2 Image Based Errors

In this section an image based error measurement for a visual servo task is introduced.

Let $\mathcal{I} = \{E_x, E_y, E_z\}$ denote a right-hand inertial frame. Let the vector $\xi = (x, y, z)$ denote the position of the focal centre of the camera in the inertial frame. Let $\mathcal{A} = \{E_1^a, E_2^a, E_3^a\}$ be a (right-hand) body fixed frame for the camera. The orientation of the camera is given by the rotation $R : \mathcal{A} \to \mathcal{I}$.

Let P represent a stationary point target visible to the

camera expressed in the camera frame. The image point observed by the camera is denoted p and is obtained by rescaling onto the image surface S of the camera (cf. Figure 1)

$$p = \frac{1}{r(P)}P.$$
 (1)

Following the approach introduced in [10] we consider a camera with a spherical image plane. Thus¹ r(P) = |P|.

Let $V \in \mathcal{A}$ denote the linear velocity and $\Omega \in \mathcal{A}$ denote the angular velocity of the camera both expressed in the camera frame. The dynamics of an image point for a spherical camera of image surface radius unity are (see [9, 8])

$$\dot{p} = -\Omega \times p - \frac{\pi_p}{r(P)}V,\tag{2}$$

where $\pi_p = (I_3 - pp^T)$ is the projection $\pi_p : \mathbb{R}^3 \to T_p S^2$, the tangent space of the sphere S^2 at the point $p \in S^2$.

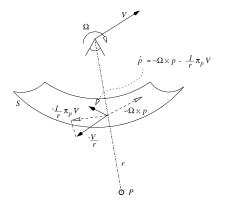


Figure 1: Image dynamics for spherical camera image geometry.

For a given target the centroid is defined to be

$$q := \sum_{i=1}^{n} p_i \in \mathbb{R}^3 \tag{3}$$

where the target consists of n points $\{P_i\}$. Using centroid information is an old technique in visual servo control [1, 14, 21]. Among the advantages of using the target centroid as a visual feature one has that it is not necessary to match observed image points directly to desired features and that the calculation of a image centroid is a highly robust procedure. These properties are expected to ensure strong practical robustness of the proposed control design. Note that the centroid defined above contains more information than a classical definition. This may be seen from the fact that the norm of q is related to the average depth of the target points. Consider the case of a target comprising a finite number of point targets. Recalling Eq. 2 it may be verified that

$$\dot{q} = -\Omega \times q - QV, \tag{4}$$

where

$$Q = \sum_{i=1}^{n} \frac{\pi_{p_i}}{r(P_i)}.$$
(5)

The visual servo task considered is that of positioning a camera relative to a stationary (in the inertial frame) target. In addition to visual information additional inertial information is explicitly used in the error formulation. In particular, it is assumed that the orientation of the camera with respect to the target is known *a priori* in the inertial frame. For example, one may wish to position the camera above the target or to the north of the target. Note that the centroid of an image contains no information relative to the orientation that can be used to position the camera. If purely centroid information is used then additional image or inertial information must be incorporated into the error to fully specify the camera orientation.

Formally, let $b \in \mathcal{I}$ denote the desired inertial direction for the visual error. The norm of b encodes the effective depth information for the desired limit point. Define

$$b^* := R^T b \in \mathcal{A}$$

to be the desired target vector expressed in the camera fixed frame. Since $b^* \in \mathcal{A}$ it inherits dynamics from the motion of the camera

$$\dot{b}^* = -\Omega \times b^*.$$

The image based error considered is the difference between the measured centroid and the target vector expressed in the camera frame

$$\delta := q - b^*. \tag{6}$$

The image error kinematics are

$$\dot{\delta} = -\Omega \times \delta - QV \tag{7}$$

To regulate the full pose of the camera, it is necessary to consider an additional error criterion for the orientation. Note that any rotation of the camera that alters the observed image centroid may be controlled using centroid information. However, the image centroid is invariant under rotation around the axis of the centroid. Thus, the best that can be achieved using the centroid feature is to orient the camera such that the image centroid lies in a desired direction. To control the remaining degrees of freedom one must use additional information; either from the image or a prior inertial direction (linearly independent from b (see section 3.1).

¹The notation |x| represents the norm of any vector $x \in \mathbb{R}^n$ $(|x| = \sqrt{x^T x}).$

In general it is desired that the image centroid is close to the focal axis of the camera. This is particularly important if there is a risk that the image may leave the active field of view of the camera during the servo manoeuvre. For the sake of simplicity we choose to stabilise the orientation in order that the image centroid converges to the focal axis of the camera. Choose the camera frame \mathcal{A} such that the focal axis of the camera is co-linear with the third coordinate axis. Let $\{e_1, e_2, e_3\}$ denote the unit vectors of the coordinate frame \mathcal{A} . The visual error used for the camera orientation is the angle θ between the centroid and the focal axis

$$\cos \theta = \frac{q^T e_3}{|q|} = q_0^T e_3 \quad \text{and} \quad \sin \theta = |q_0 \times e_3|, \quad (8)$$

where the normalized centroid is defined to be $q_0 := q/|q|$. In practice, it is most convenient to work directly with the cosine of θ . Define

$$\alpha := \cos \theta = q_0^T e_3.$$

Thus, the goal of the serve algorithm is to drive $\delta \to 0$ asymptotically and stabilize $\alpha \to 1$.

3 Kinematic Control Design

In this section a Lyapunov control design is given for the kinematic control of the visual error based on the visual errors introduced in Section 2.

The kinematics of an image based visual servo system are the first order dynamics of the visual features q and α . The derivative of α is

$$\dot{\alpha} = \frac{d}{dt}\cos\theta = \frac{e_3^T \dot{q}}{|q|} - \frac{|q|e_3^T q}{|q|^2} = -\Omega^T (q_0 \times e_3) - \frac{1}{|q|} e_3^T \pi_{q_0} QV.$$
(9)

Note that angular velocity Ω is three dimensional, however, only the component in the direction $q_0 \times e_3$, that rotates the focal axis in the direction of the observed centroid, contributes to the first order dynamics of α . Thus, the kinematics considered are given by Eqn's 7 and 9. The inputs to these equations are V and Ω .

Define a storage function S

$$S = \frac{1}{2} |\delta|^2 \tag{10}$$

Taking the time derivative of S and substituting for Eq. 7, yields

$$\dot{S} = -\delta^T Q V \tag{11}$$

Note that Eq. 11 is independent of the angular velocity Ω .

The matrix Q > 0 is not exactly known, however, it is known to be positive definite. Thus, a choice $V = \delta$ is sufficient to stabilise S_1 . The kinematic control chosen for Eq. 11 is

$$V = k_{\delta}\delta \tag{12}$$

Lemma 3.1 Consider the system defined by Eq. 7 and let $k_{\delta} > 0$ be a positive constant. Assume that the image remains in the camera field of view for all time. Then, the closed loop system Eq. 7 along with control Eq. 12 exponentially stabilises the visual error δ .

Proof: Recalling the derivative of S and substituting the control input V by its expression, it yields

$$\dot{S} = -k_{\delta}\delta^T Q\delta$$

Since Q is a positive definite matrix, classical Lypunov theory guarantees that δ converges exponentially to zero.

To control the kinematics of the visual feature α the storage function considered is

$$T := \frac{1}{2}(1 - \cos\theta)^2 = \frac{1}{2}(1 - \alpha)^2.$$
(13)

Following a classical Lyapunov control design the natural choice of Ω is

$$\Omega = -k_{\alpha} \left(q_0 \times e_3 \right). \tag{14}$$

Thus,

$$\dot{T} = (1 - \alpha) \left(\Omega^T(q_0 \times e_3) + \frac{1}{|q|} e_3^T \pi_{q_0} QV \right)$$
 (15)

$$= -k_{\alpha}(1+\alpha)(1-\alpha)^{2} + \frac{(1-\alpha)}{|q|}e_{3}^{T}\pi_{q_{0}}QV.$$
 (16)

Lemma 3.2 Consider the rotational dynamics defined by Eq. 9 with the control law Eq. 14. Let V be an exponentially stable signal

$$|V| \le Be^{-\kappa t}$$

for positive constants $B, \kappa > 0$. Assume that $|q| > \epsilon > 0$ is bounded away from zero and that $||Q|| < \lambda < \infty$ is bounded. Then $\alpha \to 1$ exponentially. As a consequence $\theta \to 0$ exponentially.

An unfortunate possibility with the above control design is that the target image may leave field of view of the camera during the evolution of the closed loop system. To avoid such a situation it is necessary to modify the orientation control design. In particular, if the closed loop system evolves such that the image is tending to leave the field of view, high gain dynamics are introduced in the orientation kinematics to keep the target in the image plane. For a given initial condition it is possible to choose a gain $k_{\alpha} > 0$ sufficiently large such that the exponential stability of the orientation dynamics dominates the camera attitude kinematics and the camera points towards the target centroid regardless of the linear dynamics of the camera. This approach is undesirable as it leads to high gain input over the full dynamics of the system and reduces system robustness as a consequence. In this paper a non-linear gain (linked to a barrier function potential) is used to limit the divergence of the image centroid from the focal axis of the camera.

Let the maximum allowable value for θ be denoted $\theta_m > 0$. The constraint $\theta < \theta_m$ limits the possible orientation of the camera to a positive cone around the centroid. Set $\alpha_m = \cos(\theta_m)$. The approach taken is to choose the control to be

$$\Omega := -\frac{k_{\alpha}}{(\alpha - \alpha_m)} \left(q_0 \times e_3 \right). \tag{17}$$

The barrier term $1/(\alpha - \alpha_m)$ ensures that the control input for the orientation dynamics dominates any finite perturbation due to the linear velocity V if the orientation approaches the limit. Recalling the definition of the storage function T, one has

$$\dot{T} = -\frac{k_{\alpha}(1+\alpha)}{(\alpha-\alpha_m)}(1-\alpha)^2 + \frac{(1-\alpha)}{|q|}e_3^T\pi_{q_0}QV$$

Lemma 3.3 Consider the rotational dynamics defined by Eq. 9. Let the control law Eq. 17. Let V be an exponentially stable signal

$$|V| \le Be^{-\kappa t}$$

for positive constants $B, \kappa > 0$. Let $\theta_m > \theta(0)$ ($\alpha(0) > \alpha_m$). Assume that $|q| > \epsilon > 0$ is bounded away from zero and that $||Q|| < \lambda < \infty$ is bounded. Then $\alpha > \alpha_m$ for all time and $\alpha \to 1$ exponentially.

Proof: First it is necessary to show that T remains bounded. Indeed, from Eq. 13 it follows that $T \leq 2$.

Introducing bounds on q and Q, the derivative of T may be bounded by the following expression

$$\dot{T} \le -\frac{k_{\alpha}(1+\alpha)}{(\alpha-\alpha_m)}(1-\alpha)^2 - \frac{4\lambda}{\epsilon}Be^{-\kappa t}$$

Taking the limit of the above bound on \dot{T} when $\alpha \to \alpha_m$, it yields

 $\dot{T} \leq 0$

This proves that $\alpha > \alpha_m$ for all time.

The exponentially decreasing perturbation does not destroy the classical Lyapunov convergence result that forces $\dot{T} \rightarrow 0$ exponentially and guarantees that $\alpha \rightarrow 1$ exponentially.

3.1 Stabilisation of remaining orientation with reference to visual data.

In this subsection an auxiliary control is proposed that will regulate the remaining degree of freedom in the attitude of the camera using an additional error criteria.

To control the remaining orientation of a camera via an error in image space it is necessary to consider error criteria that depends on another vector direction fixed in the camera frame. A key observation for the proposed control design is that such an error is chosen and minimized after the regulation of the visual errors δ and θ respectively. Thus, the best that can be achieved is to orient the camera within the set δ , $\theta = 0$ and the auxiliary error must be chosen with this in mind. The control objective considered is to align e_1 as closely as possible with new visual 'feature' vector computed from image measurements

$$q^a = \sum_{i=1}^{i=n} a_i p_i \tag{18}$$

where a_i are a set of real constants. It is not required that $a_i > 0$ and differences between observed points may be used to generate the visual feature q^a vector.

Define σ to be the error between the visual feature direction q_0^a $(q_0^a = \frac{q^a}{|q^a|})$ and the camera frame direction e_1 .

$$\sigma = q_0^a - e_1 \tag{19}$$

Taking the derivative of σ , yields

$$\dot{\sigma} = -\Omega \times \sigma + \frac{1}{|q^a|} \pi_{q_0^a} Q_a V - e_2 \Omega^3 + e_3 \Omega^2 \qquad (20)$$

Let U be the storage function for the error criterion σ .

$$U = \frac{1}{2} |\sigma|^2$$

Deriving U and recalling the derivative expression of σ , yields

$$\dot{U} = -\sigma^T \left(-\frac{1}{|q^a|} \pi_{q_0^a} Q_a V - e_3 \Omega^2 + e_2 \Omega^3 \right)$$

Choose the control for Ω^3 to be

$$\Omega^3 = k_\sigma e_2^T \sigma \tag{21}$$

The derivative of U becomes now,

$$\dot{U} = -\sigma_1^T Q_\gamma V + \sigma_1^T e_3 \Omega^2 - k_\sigma \sigma_1^T e_2 e_2^T \sigma_1 \qquad (22)$$

Lemma 3.4 Consider the remaining rotational dynamics defined by Eq. 19 with the control law Eq. 21. Let V and Ω^2 be exponentially stable signals

$$|\Omega^2|, \ |V| \le Be^{-\kappa t}$$

for positive constants $B, \kappa > 0$. Assume that $|q^a| > \epsilon > 0$ is bounded away from zero and that $||Q_a|| < \lambda < \infty$ is bounded. Then $\sigma \to 0$ exponentially.

Sketch of Proof Similar to the case encountered in Lemma 3.2 the Lyapunov function U need not be monotonically decreasing and stabilizing part of the right-hand of Eq. 22 is not negative definite. Nevertheless, using the fact that both V and Ω^2 are exponentially stable (cf. Lemmas 3.1 and 3.2) then a result proving convergence of $e_2^T \sigma$ to zero is possible. Lemma's 3.2 and 3.4 provide a complete characterisation of the convergence of the orientation of the camera. This can be seen by confirming that Lemma 3.2 guarantees that the e_3 axis of the camera frame is aligned in the direction of the centroid q. The camera may still be rotated around the e_3 axis without changing the convergence result obtained, however, Lemma 3.4 ensures convergence of q^a into the plane given by span $\{e_1, e_3\}$. Since e_3 is guaranteed to lie in the direction of the centroid q then e_1 must lie in the unique direction

$$e_1 \in \left(I - \frac{(qq^T)}{|q|^2}\right) q^a.$$

 \triangle

4 Experimental results

In this section experimental results are presented. Our aim is to validate the proposed image-based approach, showing, in real conditions, the convergence and robustness.

The control law has been tested on a six d.o.f. Cartisian robot AFMA (at Irisa/Inria-Rennes). The camera was mounted on the end-effector of the robot. In this experiment a basic stabilization task is considered. The target is the four black marks on the vertices of a planar square. Each vertex is of 0.04m of length. The signals available are the normalized co-ordinates of the center of the four marks observed by the camera, denoted $\{(u_1, v_1), (u_2, v_2), (u_3, v_3), (u_4, v_4)\}$. These coordinates are transformed, in the computer, into normalized spherical coordinates (x_i, y_i, z_i) .

The desired target vector b^* is chosen such that the camera set point is located at 0.2m above the square. Figure 2 shows the initial and goal image appearance of the target. The end-effector of the robot is moved according to the control laws Eqn's 12, 17 and 21.

In this experiment, the barrier term is obtained from the initial condition; it is chosen to be $(\alpha_m = 0.9 \cos \theta(0))$. The following control gains are used, $k_{\alpha} = k_{\sigma} = 0.07$. The control gain k_{δ} is chosen to be a constant matrix $k_{\delta} = 0.07Q^{*-1}$. This choice is used to avoid the conditioning problem of the matrix Q and to decouple the dynamics of the components of δ . We have proved that QQ^{*-1} is positive definite matrix, however, due to space limitations the proof is omitted and may be obtained on application to the authors. The constants a_i used

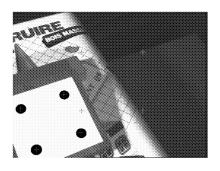


Figure 2: Initial image and desired appearance of the target.

in the definition of the variable σ have been fixed to [1, 1, -1, -1]. The algorithm performance is satisfactory : the pixel trajectories of the image points (cf. Figures 4 and 3) show a good convergence of the difference between the extracted and desired features. The convergence and stability of the positioning process is shown in Figure 5.

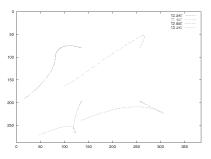


Figure 3: Image point trajectories

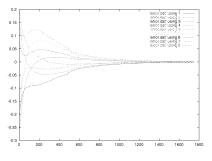


Figure 4: Error on the coordinates of the image features.

5 Concluding remarks

In this paper a novel representation of the visual error, based on a spherical representation of target centroid information, along with a Lyapunov control design is used to decouple the position and orientation servo control design for an image based visual servo system. In contrast to other recent work [5] the proposed servo control design does not need separate depth of target estimates. Moreover, the framework used leads to a simple barrier

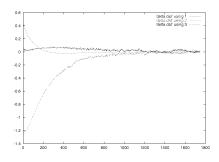


Figure 5: Exponential evolution of the error δ .

function technique that is proposed to guard against the observed target leaving the field of view of the camera without otherwise effecting the performance of the algorithm.

6 Acknowledgments

The authors wish to thank Peter Corke for numerous helpful discussions.

References

[1] R. L. Andersson. A Robot Ping-Pong Player: Experiment in Real-Time Intelligent Control. MIT Press, Cambridge, MA, USA, 1988.

[2] A. Castano and S. Hutchinson. Visual compliance: Task directed visual servo control. *IEEE trasnactions on Robotics and Automation*, 10(3):334–341, June 1993.

[3] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In *The Confluence of Vision and Control*, LNCIS, No. 237 pp. 66-78, 1998.

[4] P. I. Corke and S. A. Hutchinson. A new partitioned approach to image-based visual servo control. In *Proceedings of the International Symposium on Robotics*, Montreal, Canada, May 2000.

[5] P. I. Corke and S. A. Hutchinson. A new hybrid image-based visual servo control scheme. IEEE Int. Conf. on Decision and Control, CDC'2000, Sydnet, Australia, December 2000.

[6] K. Deguchi. Optimal motion control for imagebased visual servoing by decoupling translation and rotation. In *Proceedings of the International Conference on Intelligent Robots and Systems*, pages 705–711, 1998.

[7] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Transactions on Robotics and Automation*, 8(3):313–326, 1992.

[8] C. Fermuller, Y. Aloimonos. Observability of 3D motion. Int. Journal of Computer Vision, 37(1):43-64, June 2000.

[9] T. Hamel and R. Mahony. Robust visual servoing for under-actuated dynamic systems. In *Proceedings* of the Conference on Decision and Control, CDC'2000, Sydney, N.S.W., Australia, 2000.

[10] T. Hamel and R. Mahony. Visual servoing of an under-actuated dynamic rigid-body system: An image based approach. Provisionally accpeted for publication in IEEE Transactions on Automation and Robotics, Submitted June 2000, 2001.

[11] K. Hosada and M. Asada. Versatile visual servoing without knowledge of true jacobian. In *Proceedings* of the *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 186–193, Munich, Germany, 1994.

[12] S. Hutchinson, G. Hager, and P. Cork. A tutorial on visual servo control. *IEEE Transactions on Robotics* and Automation, 12(5):651–670, 1996.

[13] K. P. Khosla, N. Papanikolopoulos, and B. Nelson. Dynamic sensor placement using controlled active vision. In *Proceedings of IFAC 12th World Congress*, pages 9.419–422, Sydney, Australia, 1993.

[14] M. Lei and B. K. Ghosh. Visually guided robotic motion tracking. In *Proceedings of the thirteenth Annual Conference on Communication, Control and Computing*, pages 712–721, 1992.

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

[16] G. Morel, T. Liebezeit, J. Szewczyk, S. Boudet, and J. Pot. Explicit incoporation of 2D constraints in vision based control of robot manipulators, volume 250 of Lecture Notes in Control and Information Sciences, pages 99– 108. Springer-Verlag, New York, USA, 1999. Edited by P. Corke and J. Trevelyan.

[17] N. Papanikolopoulos, P. K. Khosla, and T. Kanade. Adaptive robot visual tracking. In *Proceedings of the American Control Conference*, pages 962–967, 1991.

[18] J. A. Piepmeier. A dynamic quasi-newton method for model independent visual servoing. Phd, Georgia Institute of Technology, Atlanta, USA, July 1999.

[19] R. Pissard-Gibollet and P. Rives. Applying visual servoing techniques to control of a mobile hand-eye system. In *Proceedings of the IEEE International Conference on Robotics and Automation, ICRA'95*, pages 166–171, Nagasaki, JAPAN, 1995.

[20] C. Samson, M. Le Borgne, and B. Espiau. *Robot Control: The task function approach*. The Oxford Engineering Science Series. Oxford University Press, Oxford, U.K., 1991.

[21] B. Yoshimi and P. K. Allen. Active, uncalibrated visual servoing. In *Proceedings of the IEEE International Conference on Robotics and Automation, ICRA'94*, pages 156–161, San Diago, CA, USA, 1994.