Potential problems and switching control for visual servoing

Koichi Hashimoto
Okayama University

Toshiro Noritsugu
Okayama University

Proceedings of the 2000 IEEE/RSJ
International Conference on
Intelligent Robots and Systems

Potential Problems and Switching Control for Visual Servoing

Koichi Hashimoto and Toshiro Noritsugu

Department of Systems Engineering
Okayama University
3-1-1 Tsushima-naka, Okayama 700-8530 JAPAN

Abstract

This paper proposes a potential switching scheme that enlarges stable region of feature-based visual servoing. Relay images that interpolate initial and reference image features are generated by using affine transformation. Artificial potentials defined by the relay images are patched around the reference point of the original potential to enlarge the stable region. Simulations with simplified configuration and experiments on a 6 DOF robot show the validity of the proposed control scheme.

1 Introduction

Feature-based visual servo is a robot control scheme based on the information from cameras. Typical configuration uses single hand-mounted camera and the image features of object and environment are controlled directly in the image plane. Robot is controlled so that the current image features converge to the reference image features [1]. Since the feedback loop is closed in external sensor level, modeling error in kinematic (internal sensor) level will not affect the stability. Also feature-based visual servo is robust against image noise because it does not require 3D position estimation of the object which is fairly sensitive to image noise. However, the stable region of the feature-based visual servo is local. The reason is the nonlinearity and singularity of the mapping from the object image to the joint angle.

To analyze the structure of joint-image mapping Chaumette discussed the singularity. A sufficient condition is given and an example of unpredictable camera motion is presented [2]. Also some comments regarding this condition and selection of the control law are given. Cowan and Koditschek [3] proposed a globally stabilizing method using navigation function for a planar camera motion. The method is limited for a very simplified case, but it gives a complete solution for global stabilization. On the other hand, Malis, Chaumette and Boudet recently proposed a 2-1/2 D visual servoing which incorporates both the image features and the camera orientation parameters into the controlled variables [4]. This method also gives global stability but it is not purely feature-based (it requires depth computation). Morel et al. propose another 2-1/2 D visual servoing that keeps the whole feature points in the camera's field of view [5].

This paper considers the stability problems in feature-based visual servoing. Potential is defined as the norm of the feature error. Then the stable region is represented by the downward convex region of the potential surface that includes the reference point. The reason of unpredictable camera motion presented by Chaumette is visualized with some examples. Also locality of stability of the feature-based visual servo is visualized. To enlarge the stability region, a potential switching method is proposed. The proposed scheme generates relay images by interpolating the initial and reference object image features. Interpolation-magnification method and affine transformation method are proposed to generate the relay images. Then the artificial potential is formed as a function of relay images. Next, to enlarge the stable region the artificial potentials are patched around the reference point of the original potential. Simulations with 1 DOF and 2 DOF configurations show the validity of the potential switching control scheme. Experiments on 6 DOF robot exhibit the effectiveness of image interpolation based on affine transformation.

2 Feature-based Visual Servo

2.1 Formulation

We assume that the camera is mounted on the robot hand and the hand is controlled by observing the image feature points on an stationary object. Let the generalized coordinates of the camera be q and the position and orientation vector of the camera be p_c.

0-7803-6348-5/00/$10.00

Let the position and orientation vector of the object be p_o and the number of visible feature points be n. Let the ith feature point be p_{oi} and the relative position and orientation vector between the camera and the object be $[X_i\ Y_i\ Z_i]^T = {}^cp_{ri} = {}^cR_w(p_{oi} - p_c)$ where ${}^cR_w$ is the rotation matrix from the world coordinate system to the camera coordinate system. Let the feature vector of the ith feature point in the image coordinates be $[x_i\ y_i]^T = \xi_i$ and define the feature vector as $\xi = [\xi_1^T \cdots \xi_n^T]^T$.

Feature-based visual servo works so that the current image features converge to the reference features. Let q_d be the reference robot configuration and $\xi_d = \xi(q_d)$ be the reference features, then the visual servo problem is formulated as a potential minimization problem with the potential function being

$$V(q) = (\xi_d - \xi(q))^T(\xi_d - \xi(q)) \tag{1}$$

2.2 Control Law

A typical control law is the steepest decreasing law of (1) given by [1, 6, 7]

$$\dot{q} = J^{\dagger}(\xi_d - \xi) \tag{2}$$

where J is defined by $J = \frac{\partial \xi}{\partial {}^cp_c}\frac{\partial {}^cp_c}{\partial q} = J_f {}^cJ_r$, ${}^cJ_r$ is the robot Jacobi matrix expressed in the camera coordinate system and J_f is the image Jacobian.

3 Global Minimization

3.1 Equilibrium Points

Increasing the number of feature points increases the sensitivity of the visual servo system [8]. However, if the set of feature points is redundant, then there may exist undesired equilibria.

The partial derivative of the potential $V(q)$ is

$$\frac{\partial V}{\partial q} = -2J^T(\xi_d - \xi) \tag{3}$$

Let the dimensions of q and ξ be m and $2n$, respectively. Suppose $2n > m$ then J becomes tall. For all q in the neighborhood of q_d, the necessary and sufficient condition for local stability ($\xi \to \xi_d$ then $q \to q_d$) is rank$J = m$. Even if this condition is satisfied, there exists $2n - m$ linearly independent error vectors $\xi_e = \xi_d - \xi$ that belong to KerJ^T. For these error vectors, we have $\partial V/\partial q = 0$. Thus these features are equilibria of the potential.

Since the mapping from q to ξ_e depends on the robot kinematics and robot-object configuration, it is not easy to discuss existence of local minima without specifying the configuration. Thus we give two simple examples and draw potential plots.

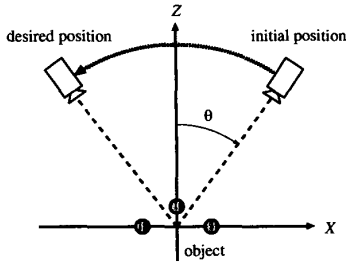

Figure 1: Camera motion (circular)

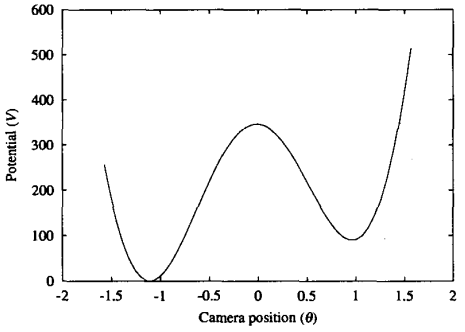

Figure 2: Potential plot (circular)

3.2 Examples

3.2.1 1 DOF Circular Motion

Assume that the object is a triangle whose vertexes are $p_{o1} = [-B\ 0\ 0]^T, p_{o2} = [0\ 0\ H]^T, p_{o3} = [B\ 0\ 0]^T$ and the camera has 1 DOF; the camera moves in X-Z plane; the distance from the origin is constant d; and the optical axis always go through the origin (Figure 1). Let the generalized coordinate be $q = \theta$ where θ is the rotation angle. Suppose that the reference is the features obtained at $\theta_d = -\pi/3$ and the parameters are as follows: $B = 100, H = 20, d = 1000$. Then the potential plot is given by Figure 2. A local minimum exists. For initial position $q_0 < q_a$ where $q_a = 0.05$ is the local maximum, the camera position converges to the local minima q_{lm}. For the features at the local minima $\xi_{lm} = \xi(q_{lm})$ the feature error $\xi_d - \xi_{lm} \neq 0$ and it falls into KerJ^T.

3.2.2 2 DOF Rotation

Let us consider another 2 DOF case, one DOF is translation along the Z axis and the other DOF is the rotation around the Z axis (Figure 3). The optical axis always coincides with the Z axis. The general-

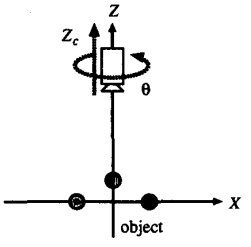

Figure 3: Camera motion (2D rotation)

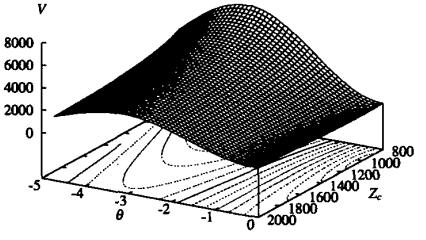

Figure 4: Potential plot (2D rotation)

ized coordinates of the camera is $q = [Z_c\ \theta]^T$ where Z_c is the camera height and θ is the rotation angle. Let the reference image be the one obtained at $q_d = [1000, 0]^T$. For the camera position in the range $700 \leq Z_c \leq 2000, -1.5\pi \leq \theta \leq 0$, the potential is plotted in Figure 4. The plot for $\theta > 0$ is symmetric with respect to the plane $\theta = 0$. For the initial value $q_0 = [1000, -\pi]^T$, the initial image is symmetric to the reference with respect to the image center. This initial point is an unstable equilibrium for the change of θ and the surface is monotonically decreasing for the change of Z_c. Thus the camera goes upward straightly without rotation. This coincides with the Chaumette's observation [2] but our potential plot is more informative. If the initial point is not strictly point symmetric, the camera slowly rotates but it quickly goes away from the object.

4 Potential Switching

The procedure of potential switching is shown with the 1 DOF example. First, we chose an adequate point θ_c in the stable region and generate an artificial potential V' that has global minima at θ_c. Then we patch V' over the unstable region to obtain a larger stable region. If V' is globally downward convex, the stability becomes global. If V' has local minima, we can repeat the same procedure. In this section we discuss two points: selection of θ_c and generation of V'. In feature-based visual servo θ_c can not be selected directly. Thus we use interpolated image and V' is defined as the potential for the interpolated image.

4.1 Magnification

The image interpolation is presented for the 1 DOF circular motion case (Figure 1). For simplicity, assume that the imaging model is weak perspective projection and let $\theta_0 = -\theta_d$ and $\theta_d < 0$. First, let us adopt the averaged image $\xi_c = (\xi_d + \xi_0)/2$ as a relay image. Then we have

$$J^T(\xi_c - \xi) = bf(\theta),$$
$$f(\theta) = -s(\frac{c_d}{2} + \frac{c_0}{2} - c) + ac(\frac{s_d}{2} + \frac{s_0}{2} - s) \quad (4)$$

where $a = \frac{H^2}{2B^2}, b = \frac{2f^2B^2}{d^2}, c = \cos(\theta), s = \sin(\theta), c_d = \cos(\theta_d), s_d = \sin(\theta_d)$. By substituting $\theta_0 = -\theta_d$ into the above equation, we can see that the potential has equilibria at $\theta = \theta_d$, $\theta = 0$ and $c = c_d/(1-a)$. Thus for $a < 1$, the averaged image is not adequate for relay image.

Next, we magnify the interpolated image around the object center. Since the object center is always projected to the image center, the interpolated image is $\xi_r = (1-r)\xi_0 + r\xi_d$ $(0 \leq r \leq 1)$. Let the magnification ratio be γ then the magnified image becomes

$$\xi_i = \gamma\xi_r = \gamma(1-r)\xi_0 + \gamma r\xi_d \quad (5)$$

When $r = 1/2$, the solutions of $f(\theta) = 0$ are $\theta = \theta_d$, $\theta = 0$ and $c = \frac{\gamma}{1-a}c_d$. Thus if $\gamma > (1-a)/c_d$, then the artificial potential V_i with relay image ξ_i does not have local minima and the camera will converge to the minimum of V_i. Since the global minimum of V_i coincides with the local maxima of V, the camera can 'climb up' to the local maxima.

To investigate the characteristics of the potential for $r \neq 1/2$, we compute $f'(\theta)$ and we have $f'(0) < 0$ for $r < 1/2$ and $f'(0) > 0$ for $r > 1/2$. Thus the global minimum of the potential exists in the region $\theta > 0$ for $r < 1/2$. Also for $r > 1/2$, it is in $\theta < 0$. If we change r from $1/2 - \epsilon$ to $1/2 + \epsilon$, then the global minimum of the potential changes from negative to positive. Thus using these images, the camera goes across the local maxima and falls down to the global minimum by using ξ_d. To carry out the visual servo task, the final image should be ξ_d, thus γ may be a continuous function of r that satisfies $\gamma(0) = 1, \gamma(1) = 1, \gamma(1/2) > (1-a)/c_d$.

4.2 Affine Transformation

The interpolation method is useful for initial and reference features that do not include rotation. Since the relation between initial and reference features is approximated by affine transformation, the eigenvalue of the transformation matrix is used to estimate the rotation angle. Also, the relay images are generated by the affine transformation.

Let ξ_0 and ξ_d be the initial and reference features. Then the affine transformation between them is given by

$$\xi_0 = A\xi_d + T + \delta \tag{6}$$

where

$$A = \begin{bmatrix} \tilde{A} & & 0 \\ & \ddots & \\ 0 & & \tilde{A} \end{bmatrix}, \quad T = \begin{bmatrix} \tilde{T} \\ \vdots \\ \tilde{T} \end{bmatrix}. \tag{7}$$

The sub-matrices $\tilde{A}, \tilde{T}$ are the affine transformation for each feature point

$$\tilde{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \quad \tilde{T} = \begin{bmatrix} T_x \\ T_y \end{bmatrix} \tag{8}$$

and δ represents the approximation error. The affine parameters $\Theta = (a_1, \cdots, a_4, T_x, T_y)$ are computed by Newton-Raphson method so as to minimize $\|\delta\|$. Once the affine parameters are found, then the rotation angle between initial and reference features is estimated by decomposing $\tilde{A}$ as

$$\tilde{A} = \tilde{R}\tilde{B}, \qquad \tilde{R} = \begin{bmatrix} \cos(\arg(\lambda)) & -\sin(\arg(\lambda)) \\ \sin(\arg(\lambda)) & \cos(\arg(\lambda)) \end{bmatrix} \tag{9}$$

where λ is eigenvalue of $\tilde{A}$. Decomposition $A = RB$ can be done similarly.

Let the number of relay images be N, then the relay images ξ_i $(i = 1, \ldots, N)$ are generated by using N-th root of A. Let the N-th root of R and B be R' and B' respectively. Then it is straightforward to find affine transformation

$$\Psi_{\Theta_i}(\xi_i) = A_i\xi_i + T_i \tag{10}$$

that satisfies

$$\xi_{i-1} = \Psi_{\Theta_i}(\xi_i), \quad \xi_{m+1} = \xi_d. \tag{11}$$

5 Simulation

5.1 1 DOF circular motion

For the 1 DOF circular motion example (Figure 1) with $\theta_d = -\pi/3, \theta_0 = \pi/4$, a simulation result using interpolation method is given in Figure 5. The magnification ratio is $\gamma(r) = 1 + \sin(r\pi)$, $r = 1/4, 1/2, 3/4$. The coordinate system at the center of the figure is the world coordinates and at the origin of the world coordinates the object is placed. The arrows above the object show the camera coordinate system. Since the motion is in X-Z plane, only X and Z axes are plotted (the optical axis is $-Z$ direction). The camera position converges to the desired position. The potentials V_i for the interpolated images ξ_i are plotted in Figure 6. The positions on which the arrows are crowded in Figure 5 correspond to the positions of potential minima for interpolated images. The potentials work effectively to pull the camera up to the local maxima of V.

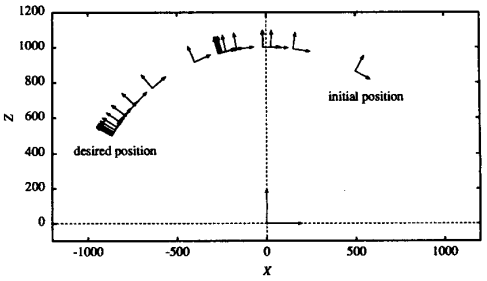

Figure 5: Camera position

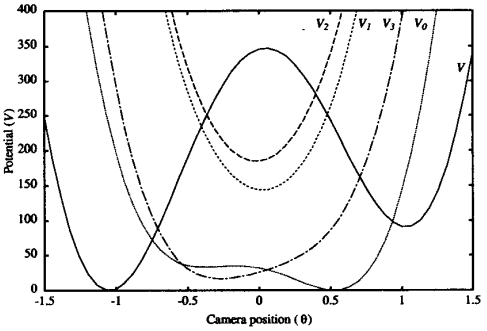

Figure 6: Potentials for interpolated images

5.2 2 DOF Rotation

Simulation result of 2 DOF rotational motion (Figure 3) with $Z_d = 1000, \theta_d = 0, Z_0 = 1000, \theta_0 = \pi$ is given in Figures 7 and 8. The arrows show the X and Z axis of the camera. Figure 7 is the result without potential switching. As expected by the potential plot the camera moves upward without rotation. However

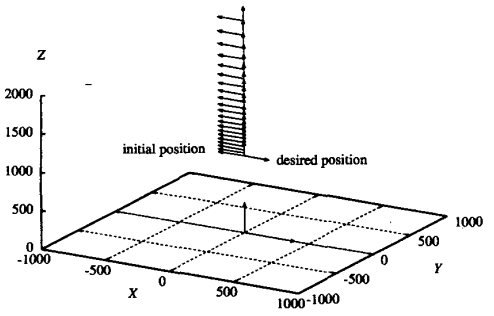

Figure 7: Camera position

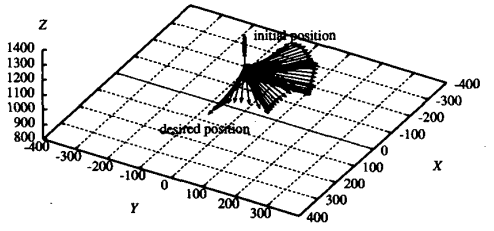

Figure 8: Camera position

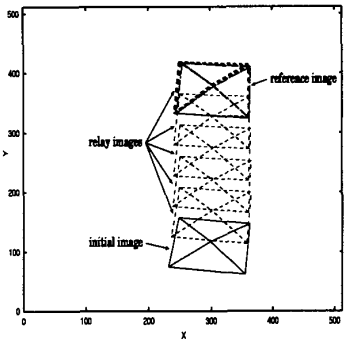

Figure 9: Initial and reference images

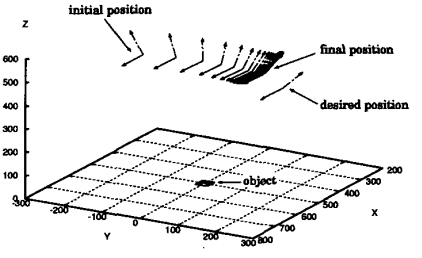

Figure 10: 3D plot of camera (without switching)

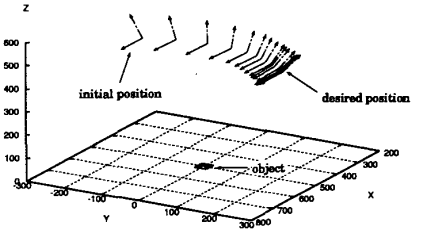

Figure 11: 3D plot of camera (with switching)

by using three relay images generated by affine transformation, the camera converges to the reference position. Figure 8 is an magnified plot with higher view angle. While rotating camera position moves up and down slightly. This is due to the shape of surface of each artificial potential.

6 Experiment

Experiments are carried out with Mitsubishi RV-E2 robot with a CCD camera mounted on the hand. The origin of the world coordinate system is located at the base of the robot. The object has five feature points and placed in front of the robot. The position of the object center is $(500, 0, 0)$ [mm] and the feature points makes a 32 mm square and a feature point with 5mm height is added at the object center.

6.1 6 DOF Translation

First experiment is translation from right hand side to left hand side over the object. The initial position of the camera is $(X, Y, Z, \phi, \theta, \psi) = (500, -153, 512, 1.57, 0.35, 1.57)$ and the reference position is $(500, 208, 457, 1.57, 0.57, -1.57)$. The object images at the initial and reference position is depicted in 9. Also eight relay images generated by affine transformation ($N = 10$) are depicted.

6.2 6 DOF rotation

Second experiment is 180 degree rotation above the object. The initial position of the camera is $(469.5, -29.6, 483, -0.12, 0, 0.12)$ and the reference position is $(432.6, -14.3, 483, 0, 0, 0)$. To show the effect of translation and rotation separation, small translation is also added. Figure 12 shows the initial and reference images. Ten relay images are also plotted.

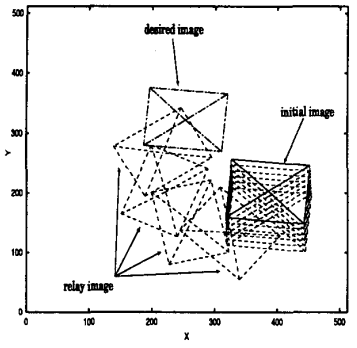

Figure 12: Initial and reference images

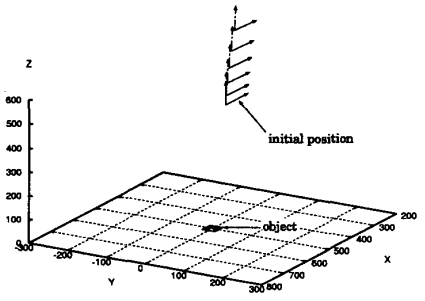

Figure 13: 3D plot of camera (without switching)

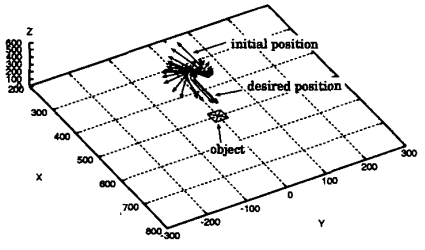

Figure 14: 3D plot of camera (with switching)

The result without potential switching is given in Figure 13. As we have seen in the simulation (Figure 3), the camera moves upward. Figure 14 shows the result with potential switching. The camera rotates to the reference position with small up and down motion. This experiment uses 10 relay images but the number of relay images can be reduced to enhance the convergence speed. The optimal number of relay image is not trivial and will be our future research subject.

7 Conclusion

This paper considers the potential and stable region problems in visual servo. Examples have exhibited existence of local minima and limited region of stability. A potential switching control is proposed to enlarge the stable region. The artificial potential is generated by using relay images. The relay images are generated using magnification and affine transformation. Simulations are carried out to show the validity of the switching control. Experiments on 6 DOF robot are also carried out to see the effectiveness of the switching control.

References

[1] L. E. Weiss, A. C. Sanderson, and C. P. Newman. Dynamic sensor-based control of robots with visual feedback. *IEEE J. Robotics and Automation*, RA-3(5):404–417, 1987.

[2] F. Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In *The Confluence of Vision and Control, D. J. Kriegman, G. D. Hager and A. S. Morse eds., Springer-Verlag*, pages 66–78, London, 1998.

[3] N. J. Cowan and D. E. Koditschek. Planar image based visual servoing as a navigation problem. In *IEEE Int. Conf. Robotics and Automation*, pages 611–617, Detroit, Michigan, 1999.

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

[5] G. Morel, T. Liebezeit, J. Szewczyk, S. Boudet, and J. Pot. Explicit incorporation of 2D constraints in vision based control of robot manipulator. *Experimental Robotics VI*, pages 99–108, 1999.

[6] F. Chaumette, P. Rives, and B. Espiau. Positioning of a robot with respect to an object, tracking it and estimating its velocity by visual servoing. In *IEEE Int. Conf. Robotics and Automation*, pages 2248–2253, Sacramento, Calif., 1991.

[7] K. Hashimoto et al. Manipulator control with image-based visual servo. In *IEEE Int. Conf. Robotics and Automation*, pages 2267–2272, Sacramento, Calif., 1991.

[8] K. Hashimoto and T. Noritsugu. Performance and sensitivity in visual servoing. In *IEEE Int. Conf. Robotics and Automation*, pages 2321–2326, Leuven, Belgium, 1998.