Enlargement of Stable Region in Visual Servo

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 enlarge stable region of feature-based visual servoing. Potential is defined as the norm of image feature error and stable region is a downward convex region of the potential surface that includes reference position. The proposed scheme generates relay images that interpolate initial and reference image features and artificial potential is defined by using relay images. The artificial potentials 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.**

I. 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 presented. 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), thus the robustness of feature-based method is not fully obtained. Morel et al. proposed 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 by using potential. The potential is the norm of the feature error and the task is defined as the minimization problem of the potential. The stable region is 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 configuration show the validity of the potential switching control scheme. Experiments on 6 DOF robot exhibit the effectiveness of image interpolation based on affine transformation.

II. Feature-based Visual Servo

A. 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. 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 $[X_i\ Y_i\ Z_i]^T = {}^c p_{ri} = {}^c R_w(p_{oi} - p_c)$ where ${}^c R_w$ 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

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

B. Control Law

A typical control law is the steepest decreasing law of (1) given by [1], [6], [7], [8], [9]

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

where $J^\dagger$ is the generalized inverse of J and J is defined by

$$J = \frac{\partial \xi}{\partial {}^c p_c}\frac{\partial {}^c p_c}{\partial q} = \begin{bmatrix} J_1 \\ \vdots \\ J_n \end{bmatrix} {}^c J_r. \tag{3}$$

In this definition, ${}^c J_r$ is the robot Jacobi matrix expressed in the camera coordinate system and J_i is the image Jacobian for i th feature point defined by

$$J_i = \begin{bmatrix} -f/Z_i & 0 & x_i/Z_i & x_i y_i/f & -(x_i^2+f^2)/f & y_i \\ 0 & -f/Z_i & y_i/Z_i & (y_i^2+f^2)/f & -x_i y_i/f & -x_i \end{bmatrix}. \tag{4}$$

III. Global Minimization

A. Equilibria

Increasing the number of feature points increases the sensitivity of the visual servo system [10]. However, if the set of feature points is redundant, then there may exist undesired equilibria.

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

$$\frac{\partial V}{\partial q} = -2J^T(\xi_d - \xi) \tag{5}$$

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 $\partial V/\partial q = 0$ and $\dot{q} = 0$ in (2). Thus these features are equilibria of the potential and visual servo stops at these equilibria.

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

B. Examples

B.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. This is the case of 1 DOF arm that rotates around Y axis with a camera attached on the tip looking at the object on the rotational axis (Figure 1). Let the generalized coordinate be $q = \theta$ where θ is the rotation angle of the arm and $\theta = 0$ when the arm is upright position. Then we have ${}^cJ_r = [d\ 0\ 0\ 0\ 1\ 0]$. 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 $q_{lm} = 1.029$. For the features at the local minima $\xi_{lm} = \xi(q_{lm})$ the feature error is $\xi_d - \xi_{lm} = [-1.854, 0.0, 8.911, 0.0, -2.659, 0.0]^T$ and image Jacobian is $J = [20.780, 0.0, -2.591, 0.0, -23.175, 0.0]^T$. The error vector falls into KerJ^T.

B.2 2 DOF Rotation

Let us consider a 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 generalized coordinates of the camera is $q = [Z_c\ \theta]^T$ where Z_c is the camera height and θ is the rotation angle. Then the robot Jacobi matrix becomes

$${}^cJ_r = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{6}$$

Let the reference image be the one obtained at $q_d = [1000, 0]^T$. For the camera position in the range $700 \le Z_c \le 2000, -1.5\pi \le \theta \le 0$, 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,

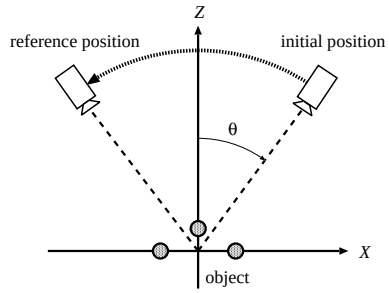

Fig. 1. Camera motion (circular)

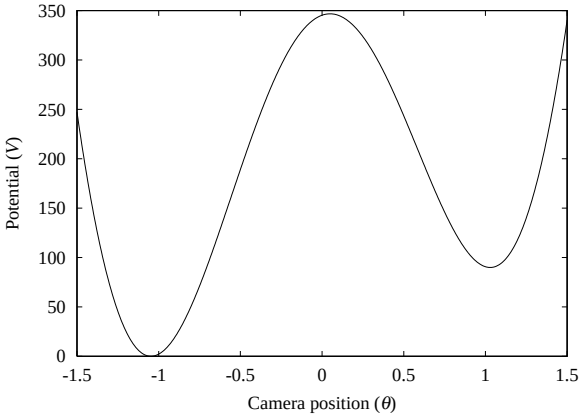

Fig. 2. Potential plot (circular)

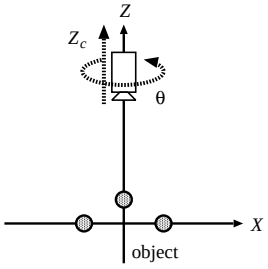

Fig. 3. Camera motion (2D rotation)

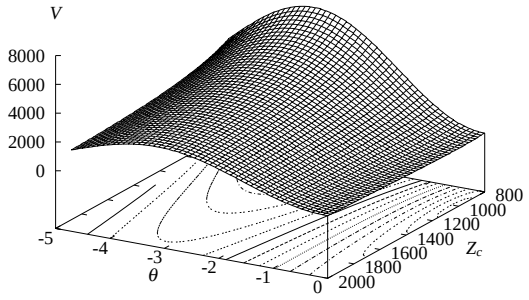

Fig. 4. Potential plot (2D rotation)

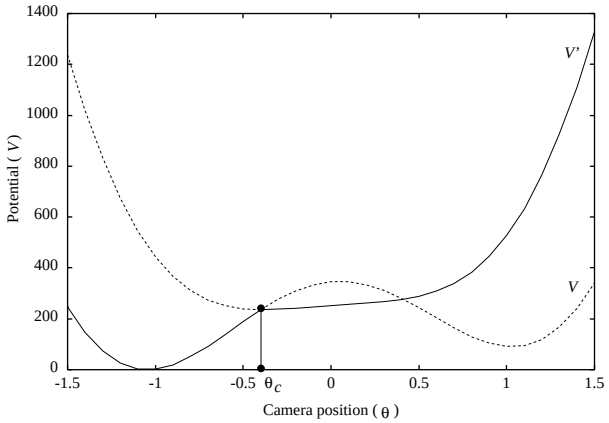

Fig. 5. Patched potential

the camera slowly rotates but it quickly goes away from the object.

IV. Potential Switching

A potential switching control scheme is proposed to enlarge the stable region. It patches artificial potential surfaces over the undesirable downward convex region. For example we show the procedure 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 q_c. Then we patch V' over the unstable region to obtain a larger stable region. Figure 5 shows the patched potential. If V' is globally downward convex, the stability becomes global. If V' has local minima, we can repeat the same procedure.

Selection of q_c and generation of V' are important. In feature-based visual servoing, q_c can not be specified directly and V' should be generated via image features. Thus we use initial and reference images. These images are interpolated and artificial image features ξ_i are generated. Then artificial potential is defined by $V_i(q) = (\xi_i - \xi(q))^T(\xi_i - \xi(q))$. The artificial image ξ_i is called relay image. Let the number of relay images be N then we define the 0-th relay image be the initial image and $N+1$-st relay image be the reference image. Then camera is controlled by feature-based visual servo (2) with the relay image as the reference image. If the current features converges to the neighborhood of the relay image then the reference image is switched to the next relay image. This control scheme is called potential switching.

The image interpolation consists of two steps. The first step is to find the relationship between initial and reference images. We use affine transformation to approximate the relationship. The affine transformation is divided into rotation and translation with transformation by using N-th root of the affine parameters. Since image transformation is usually generated by large camera motion, the second step, magnification, is required. Use of magnified reference image causes approach motion of camera and it is effective for removing local minima. To explain how magnification works we first give an example of 1 DOF case and then introduce the affine interpolation method.

A. Magnification

The image magnification method 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

$$J^T(\xi_c - \xi) = bf(\theta),$$
$$f(\theta) = -s(\frac{c_d}{2} + \frac{c_0}{2} - c) + ac(\frac{s_d}{2} + \frac{s_0}{2} - s) \qquad (7)$$

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 $c = c_d/(1-a)$ [11]. 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 $\xi_r = (1-r)\xi_0 + r\xi_d$ $(0 \le r \le 1)$. Let the magnification ratio be γ then the magnified image becomes

$$\xi_i = \gamma\xi_r = \gamma(1-r)\xi_0 + \gamma r\xi_d \qquad (8)$$

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 $\gamma(0) = 1, \gamma(1) = 1, \gamma(1/2) > (1-a)/c_d$.

B. Affine Transformation

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

$$\xi_0 \quad = \quad A\xi_d + T + \delta \qquad (9)$$

where

$$A = \begin{bmatrix} \tilde{A} & & 0 \\ & \ddots & \\ 0 & & \tilde{A} \end{bmatrix}, \quad T = \begin{bmatrix} \tilde{T} \\ \vdots \\ \tilde{T} \end{bmatrix}. \qquad (10)$$

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

$$\tilde{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \quad \tilde{T} = \begin{bmatrix} T_x \\ T_y \end{bmatrix} \qquad (11)$$

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 to minimize $\|\delta\|$. Once the affine parameters are found, then the rotation matrix $\tilde{R}$ between initial and reference features is estimated by decomposing $\tilde{A}$ as

$$\tilde{A} = \tilde{R}\tilde{B}, \qquad \tilde{R} = \begin{bmatrix} \cos(\arg(\lambda)) & -\sin(\arg(\lambda)) \\ \sin(\arg(\lambda)) & \cos(\arg(\lambda)) \end{bmatrix} \qquad (12)$$

where λ is the eigenvalue of $\tilde{A}$ and $\tilde{B}$ is transformation matrix. Decomposition $A = RB$ is 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. It is straightforward to find affine transformation

$$\Psi_{\Theta_i}(\xi_i) = A_i\xi_i + T_i \qquad (13)$$

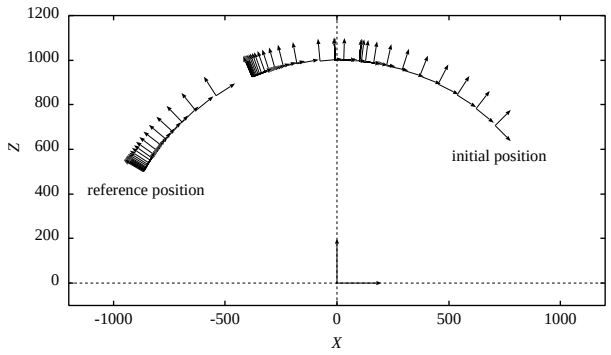

Fig. 6. Camera position

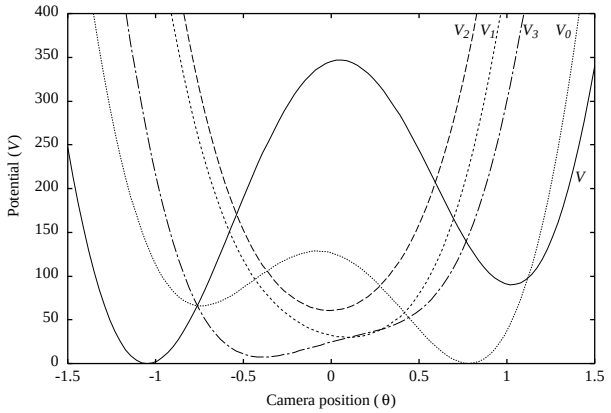

Fig. 7. Potentials for interpolated images

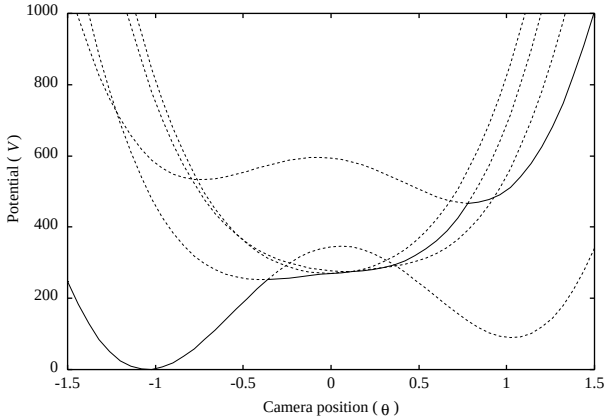

Fig. 8. Patched potential

that satisfies

$$\xi_{i-1} = \Psi_{\Theta_i}(\xi_i), \quad \xi_{N+1} = \xi_d \tag{14}$$

where Ψ_{Θ_i} for $i = 1, ldots, N/2$ are rotation and Ψ_{Θ_i} for $i = N/2+1, ldots, N$ are translation and transformation. We also include magnification in the transformation and translation part.

V. Simulation

A. 1 DOF Circular Motion

Figure 6 is the simulation result using magnification method for the 1 DOF circular motion example (Figure 1). The the initial position is $\theta_0 = \pi/4$ and reference position is $\theta_d = -\pi/3$. Three relay images are used and 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. The object is placed at the origin of the world coordinates. The arrows above the object show the camera motion. A couple of orthogonal arrows correspond to X and Z axes of the camera coordinate system (the optical axis is $-Z$ direction). At the initial position first relay image is set as the reference image and then the camera is controlled by visual servo and the camera converges to the potential minimum of the first relay image ($\theta = 0.12$). If the image error becomes smaller than a threshold then the reference image is switched to the second relay image. Similar process is repeated four times and the camera converges to the desired position. For $i = 1, 2, 3$ the potentials V_i of the relay images ξ_i are plotted in Figure 7. V_0 is the potential of the initial image. The minimum values of V_i for $i = 1, 2, 3$ are not zero. This means that there is no camera position that generates the corresponding images ξ_i. However, the potentials are globally convex because of magnification. This fact exhibits the importance of magnification. The patched potential is shown in 8. The positions on which the arrows are crowded in Figure 6 correspond to the positions of potential minima for relay images.

B. 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 9 and 10. The arrows show the X and Z axis of the camera. Figure 9 is the result without potential switching. As expected by the potential plot the camera moves upward without rotation. However by using three relay images generated by affine transformation, the camera converges to the reference position (Figure 10). The figure is magnified and view angle is changed to show the camera motion clearly. While rotating the camera moves up and down slightly. This is due to the shape of surface of each artificial potential.

VI. Experiment

Experiments are carried out with Mitsubishi RV-E2 robot with a CCD camera mounted on the hand (Figure 11). 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]. Four feature points lies on the floor and make a 32 mm $\times$ 40 mm square. Fifth feature points is at the object center but has 5 mm height.

A. 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 $(X, Y, Z, \phi, \theta, \psi)^T = (500, -153, 512, 1.57, 0.35, 1.57)$ and the reference position is $(500, 208, 457, 1.57, 0.57, -1.57)$. The object images at the initial and reference positions are depicted in Figure 12. Ten relay images (five rotation and five translation-transformation) are generated by affine transformation. These images are also depicted in Figure 12.

Figure 13 shows the experimental result without potential switching. The arrows represent the X and Z axes of the camera coordinate system. The solid lines are the X axes and the dash-dot lines are the Z axes. The camera does not reach the reference position. At the final configuration the image feature error is not zero and the error falls in $\mathrm{Ker} J^T$.

Figure 14 is the result with potential switching. The artificial potential is switched when the norm of feature error becomes

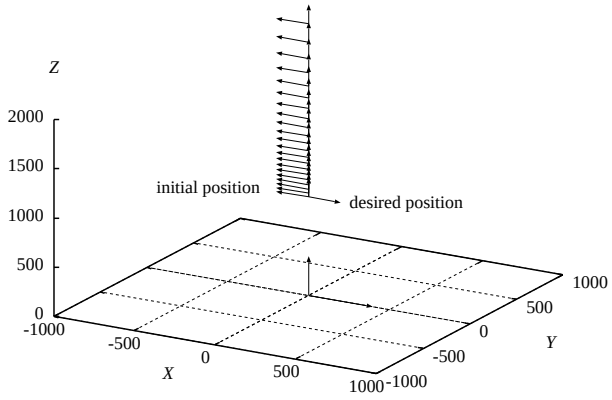

Fig. 9. Camera position

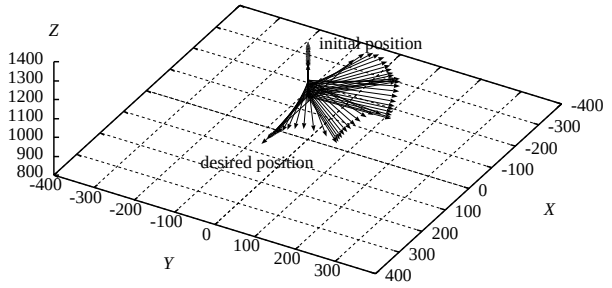

Fig. 10. Camera position

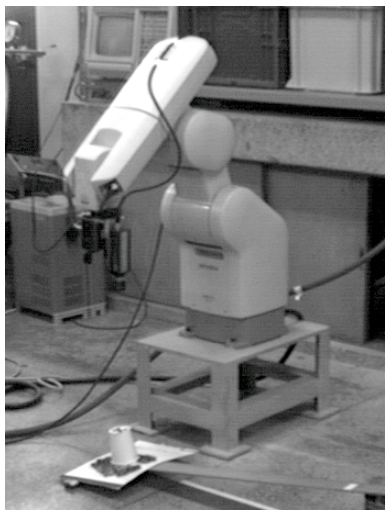

Fig. 11. Experimental System

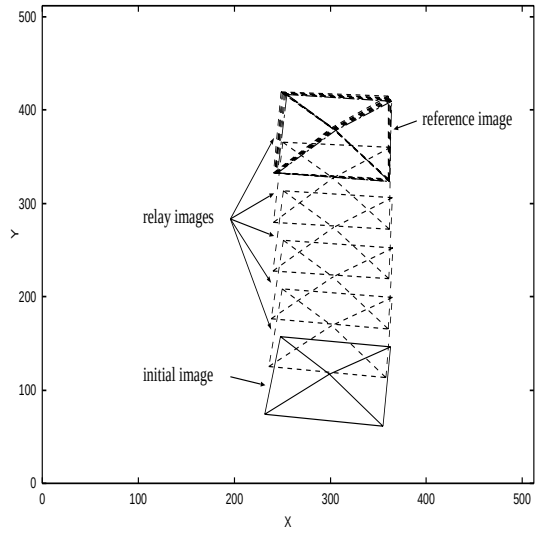

Fig. 12. Initial and reference images

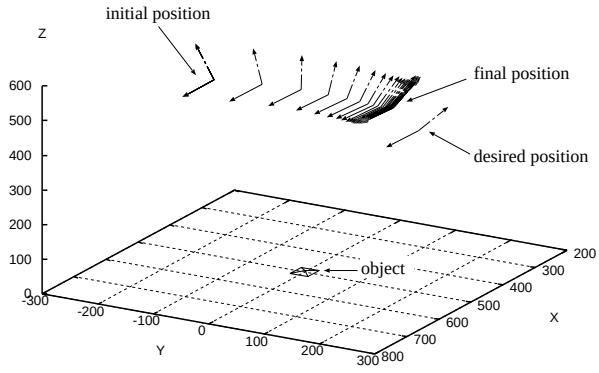

Fig. 13. 3D plot of camera (without switching)

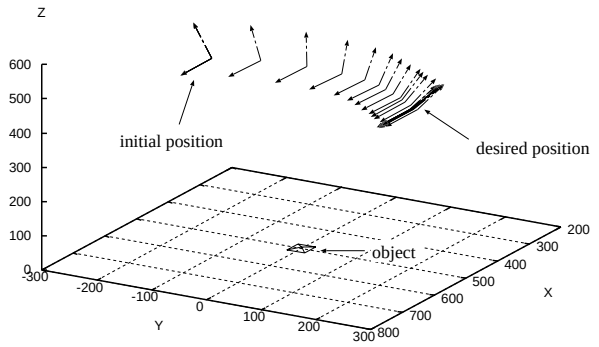

Fig. 14. 3D plot of camera (with switching)

smaller than 120 [pixel^2]. The camera reaches the reference position.

B. 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 $(432.6, -14.3, 483, 0, 0, 3.14)$. To show the effect of translation and rotation separation, small translation is also added. Figure 15 shows the initial and reference images. Ten relay images are also plotted.

The result without potential switching is given in Figure 16. As we have seen in the simulation (Figure 9), the camera moves upward.

Figure 17 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 time. The optimal number of relay image is not trivial and will be our future research subject.

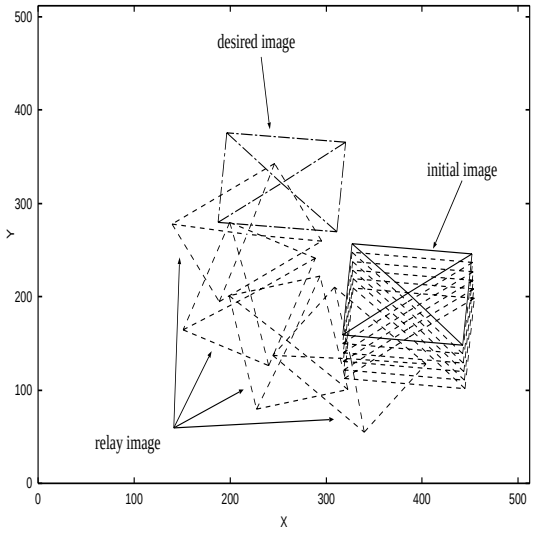

Fig. 15. Initial and reference images

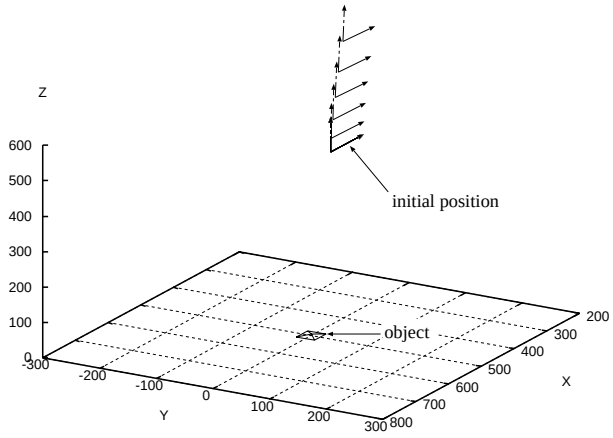

Fig. 16. 3D plot of camera (without switching)

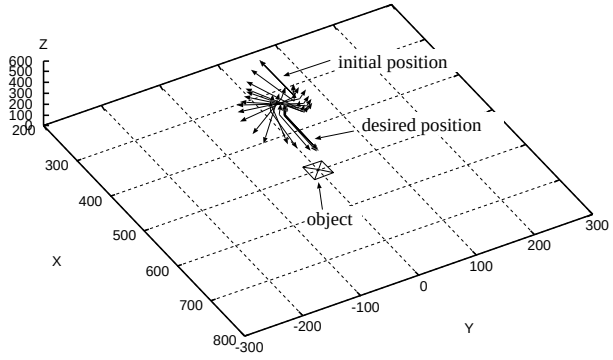

Fig. 17. 3D plot of camera (with switching)

VII. Conclusion

This paper considers the potential and stable region problems in visual servo. Examples have exhibited that local minima exist and the stable region is not large even for very simple camera motion. 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*, vol. RA-3, no. 5, pp. 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*, London, 1998, pp. 66–78.

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

[4] E. Malis, F. Chaumette, and S. Boudet, "2-1/2-D visual servoing," *IEEE Trans. Robotics and Automation*, vol. 15, no. 2, pp. 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*, pp. 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*, Sacramento, Calif., 1991, pp. 2248–2253.

[7] N. Papanikolopoulos, P. K. Khosla, and T. Kanade, "Vision and control techniques for robotic visual tracking," in *IEEE Int. Conf. Robotics and Automation*, Sacramento, Calif., 1991, pp. 857–864.

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

[9] W. Jang and Z. Bien, "Feature-based visual servoing of an eye-in-hand robot with improved tracking performance," in *IEEE Int. Conf. Robotics and Automation*, Sacramento, Calif., 1991, pp. 2254–2260.

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

[11] K. Hashimoto and T. Noritsugu, "Potential switching control in visual servo," in *IEEE Int. Conf. Robotics and Automation*, San Francisco, Calif., 2000.