Potential problems in visual servo

Koichi Hashimoto
Okayama University

Kouhei Tanaka
Okayama University

Toshiro Noritsugu
Okayama University

Potential Problems in Visual Servo

Koichi Hashimoto, Kouhei Tanaka and Toshiro Noritsugu
Department of Systems Engineering, Okayama University
3-1-1 Tsushima-naka, Okayama 700-8530 JAPAN
{koichi,kouhei,toshiro}@sys.okayama-u.ac.jp

Abstract

Stability of feature-based visual servo controllers proposed so far is local.[1] *The initial features far from the reference may converge to features different from the reference or even worse they may not converge. In this paper, stability of feature-based visual servo is considered by using potential. The stable region is visualized and artificial potential switching is proposed to extend the stable region.*

1 Introduction

With feature-based visual servo the position of the robot hand is controlled so that the image features of the robot hand converge to the reference image features [9]. It is well known that the stable region of the feature-based visual servo is local. However, the stability region is not well studied because it depends on the kinematic structure of the robot manipulator.

As a related research Chaumette derived a condition for the camera moving unpredictable direction [1]. It gives a sufficient condition and an example of unpredictable motion. 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 [7]. This method is not purely feature-based, but it also gives global stabilization for general setup.

This paper considers the potential problems in feature-based visual servoing. The potential is the norm of the feature error and the task is the minimization of the potential. The reason of unpredictable camera motion is visualized and some detailed examples are given. Also an artificial potential switching controller is proposed and enlargement of stability region is discussed.

[1] except for some really recent ones [7, 3]

2 Feature-based Visual Servo

We assume that the camera is mounted on the robot hand and the hand position is controlled on the basis of the image feature points on an stationary object.

2.1 Formulation

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

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

2.2 Control Law

A typical control law is the steepest decreasing law of (1) given by [9, 2, 8, 4, 6].

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

where J is defined by $J = \frac{\partial \xi}{\partial {}^c p_c} {}^c R_w \frac{\partial p_c}{\partial q} = J_f {}^c R_w J_r$ and J_r is the robot Jacobi matrix and J_f is the image Jacobian.

0-7803-6456-2/00/$10.00 ©2000 IEEE

3 Global Minimization

Increasing the number of feature points increases the sensitivity of the visual servo system [5]. However, if the set of feature points is redundant, then there may exist undesired equilibria.

3.1 Equilibrium Points

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 whether local minima exists or not without specifying the configuration. Thus we give some simple examples.

3.2 Examples

Let us consider one or two degree of freedom (DOF) camera motion and draw the potential plot. Assume that the camera motion is in X-Z plane and the object is a triangle whose vertexes are $p_{o1} = [-B\ 0\ 0]^T, p_{o2} = [0\ 0\ H]^T, p_{o3} = [B\ 0\ 0]^T$.

3.2.1 1 DOF Straight Motion

Suppose that the camera has 1 DOF. Let the camera position be $p_c = [X_c\ Y_c\ Z_c]^T$. The camera can translate in X_c direction but it can not translate nor rotate in other directions (Figure 1). The generalized coordinate is taken as $q = X_c$. Then ${}^cR_w$ becomes unity and ${}^cR_w J_r = [1\ 0\ 0\ 0\ 0\ 0]$. Let $Y_c = 0, Z_c = d, f = -256, 2B = 200, H = 20, d = 1000$, and the reference position be $X_d = -1000$. The potential plot is depicted in Figure 2. Since the depth Z_i $(i = 1, 2, 3)$ is constant, the error vector ξ_e becomes a linear function of X_c and $V(X_c)$ becomes a quadratic function with $V(X_d) = 0$. Thus the only equilibrium is the global minimum. This is the case of trivially redundant features because the camera position is uniquely determined by only one feature point. For the trivially redundant features all the basis of KerJ^T can not be

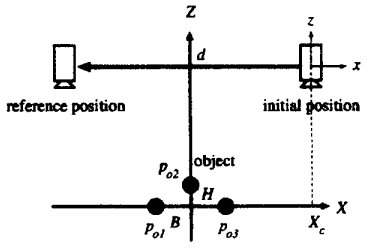

Figure 1: Camera motion (X translation)

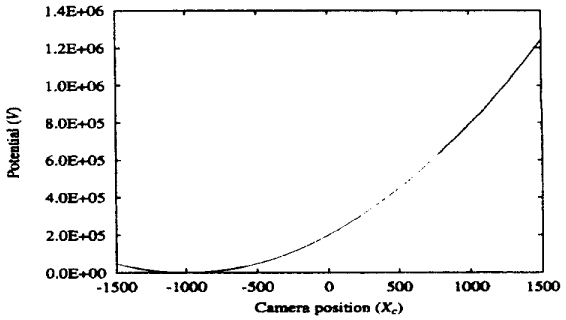

Figure 2: Potential plot (X translation)

generated by any camera motion[2] .

3.2.2 1 DOF Circular Motion

Assume that the camera has 1 DOF; 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 rotational axis. 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 ${}^cR_w J_r = [d\ 0\ 0\ 0\ 0\ 1\ 0]$. Suppose that the reference is the features obtained at $\theta_d = -\pi/3$. Assume that the parameters are as follows: $f = -256, 2B = 200, H = 20, d = 1000$, then the potential plot is given by Figure 4. A local minimum exists because the image obtained at $\theta = \theta_d$ is similar to the reference but the image at $\theta = 0$ is not. This similarity reversal will be solved by setting H large. A sufficient condition for existing local minima will be given in the next section.

3.2.3 2 DOF Straight Motion

Let us consider a 2 DOF case with constant height but the camera can rotate around the Y_c axis, as shown

[2] The basis are $[0\ 1\ 0\ 0\ 0\ 0], [0\ 0\ 0\ 1\ 0\ 0\ 0], [0\ 0\ 0\ 0\ 0\ 1]$ and $[1\ 0\ -1\ 0\ 0\ 0], [1\ 0\ 0\ 0\ -1\ 0]$. The first group requires Y_c motion and the second group requires Z_c motion.

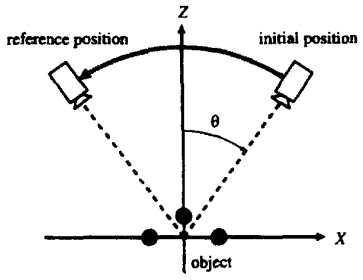

Figure 3: Camera motion (circular)

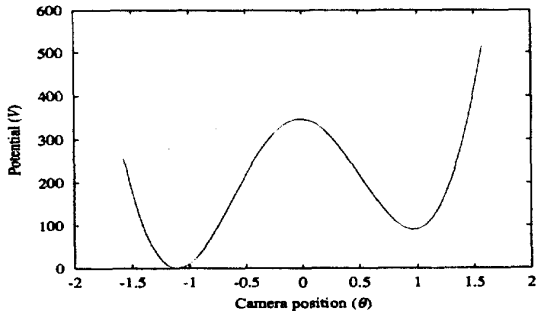

Figure 4: Potential plot (circular)

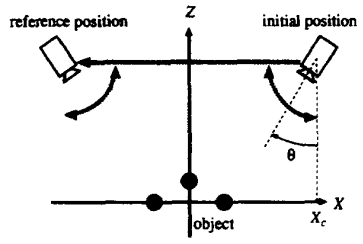

Figure 5: Camera motion (X translation and rotation)

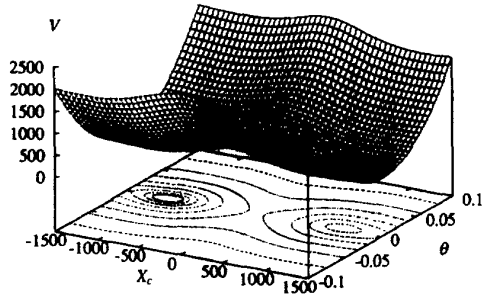

Figure 6: Potential plot (Close up 2D X and q)

in Figure 5. The generalized coordinates are $q = [X_c\ \theta]^T$, where θ is the rotation angle. The Jacobi matrix becomes

$$ {}^cR_w J_r = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad (4) $$

Let the parameters be $Y_c = 0, Z_c = d, f = -256, 2B = 200, H = 20, d = 1000$ and the reference position be $X_d = -1000, \theta_d = \pi/4$. At the reference, the object features are located near the image center. Thus it is easy to conclude that the potential surface has a valley along a curve $\theta = \arctan X_c$. To magnify the bottom of the potential surface, a transformation $\theta' = \theta - \arctan X_c$ is carried out and the potential in the region $-1500 \leq X_c \leq 1500, -0.1\pi \leq \theta' \leq 0.1\pi$ is plotted in Figure 6. The contours on the base plane show the existence of a local minimum. The reason why the local minima exists is the same as the case of 1 DOF circular motion and will be explained in the next section.

3.2.4 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. 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 Jacobi matrix becomes

$$ {}^cR_w J_r = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad (5) $$

Let the reference image be the one obtained at $q_d = [1000, 0]^T$ and assume the parameters are $f = -256, 2B = 200, H = 20$. 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 8. The plot of $\theta > 0$ is symmetric with respect to a 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 back straightly without rotation. This coincides with the Chaumette's observation [1] 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.

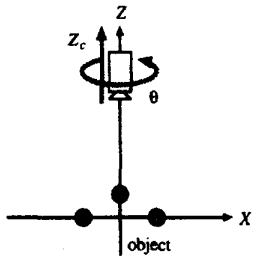

Figure 7: Camera motion (2D rotation)

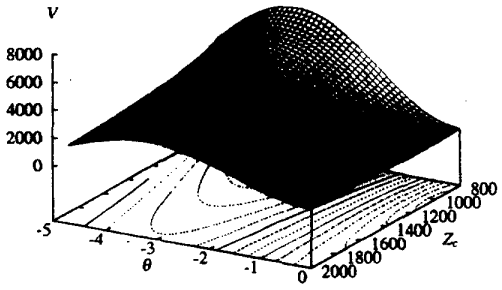

Figure 8: Potential plot (2D rotation)

4 Existence of Local Minima

For the example of Figure 1, the image obtained at $\theta = -\theta_d$ is similar to the reference because the positions in the image plane of the two zero-height feature points are close to the reference. On the other hand, the image at $\theta = 0$ is not very similar to the reference because all of three points are not close to the reference. Thus there exists a local minimum near $\theta = -\theta_d$ and a local maximum near $\theta = 0$. However, the local minima will disappear by increasing the object height H because at $\theta = -\theta_d$ the difference of the feature point with height becomes dominant by increasing the height. This section gives an analysis for existence of local minima for the case of 1 DOF circular motion.

4.1 Gradient of the Potential

Since the optical axis pieces the object center, camera motion is equivalent to the object rotation around the object center. For simplicity we assume $B, H \ll d$ and the weak perspective projection. Then the object position with respect to the camera becomes $X_1 = -Bc, X_2 = -Hs, X_3 = Bc, Y_1 = Y_2 = Y_3 = 0, Z_1 = Z_2 = Z_3 = -d$, where abbreviations $c = \cos(\theta), s = \sin(\theta)$ are used. Then (3) becomes

$$\frac{\partial V}{\partial \theta} = 2b(s(c_d - c) - ac(s_d - s)) \qquad (6)$$

where $a = \frac{H^2}{2B^2}, b = \frac{2f^2B^2}{d^2}, c_d = \cos(\theta_d), s_d = \sin(\theta_d)$. The existence of local minima can be found by evaluating the number of solution of $f(\theta) = s(c_d - c) - ac(s_d - s) = 0$ in the range $\theta \in (-\pi/2, \pi/2)$. After some calculations, we have

$$\begin{aligned} f(\theta) &= \cos(\theta)\cos(\frac{\theta_d + \theta}{2})\sin(\frac{\theta_d - \theta}{2})(g(\theta) + a) \\ g(\theta) &= \tan(\theta)\tan(\frac{\theta_d + \theta}{2}) \end{aligned} \qquad (7)$$

Thus $f(\theta) = 0$ has a solution $\theta = \theta_d$ and the other solutions are, if any, the solutions of $g(\theta) = -a$. It is straightforward to see that $g(\theta) = -a$ is equivalent to

$$h(t) = at_d t^3 + (2 - a)t^2 + t_d(2 - a)t + a = 0 \qquad (8)$$

where $t = \tan(\theta/2)$ and $t_d = \tan(\theta_d/2)$. Thus the number of real solutions of $h(t) = 0$ in the range $t \in (-1, 1)$ determines the existence of local minima. If $t_d = 0$, no solution is in this range. Since $h(t)$ is symmetric about t_d, we consider only for $t_d < 0$.

4.2 Object Height and Local Minima

To find a condition on t_d and a for which equation (8) has solution in range $-1 < t < 1$, one can differentiate $h(t)$ and it is easy to check the followings: for $2/(1 + 3t_d^2) < a < 2$, $h'(t) = 0$ has no real solution ($h(t)$ is monotonically decreasing); for $a < 2/(1 + 3t_d^2)$ or $2 < a$, $h'(t) = 0$ has two real solutions β_1, β_2 where $\beta_1 < \beta_2$; if $h'(t) = 0$ has real solutions and $h(\beta_1) < 0$, $h(t) = 0$ has two solutions in $-1 < t < 1$. However the equation $h(\beta_1) = 0$ becomes a quartic equation and does not have analytical solution. Thus we solve it numerically.

Let $\theta_d = -\pi/3$, $\beta_1 = 0.2950$ and $a = 0.1559$. Then for $a > 0.1559$, that is for $H > 0.5584B$, no local minimum exists.

From the numerical computations we found that the upper limit $\bar{a}$ of a for which local minima exist decreases if θ_d increases to 0, and on the other hand, if θ_d decreases to $-\pi/2$, $\bar{a}$ increases. This fact corresponds to following intuitive: when we look at the object from above, then we can recognize the object orientation easily even if the object do not have features with height. However if we look at the object from very low angle and if the object has no features

with height, then it is difficult to recognize whether we are looking from right side or from left side. Thus when we need global stabilization, the object must have considerably high feature point.

In fact, for our 1 DOF circular motion case, the global stability is achieved if $a > 2/(1+3t_d^2)$. If t_d is unknown, a conservative evaluation is $a > 2$, that is $H > 4B$. Thus, if the object height is larger than twice of base length, global stability is achieved for any reference position.

5 Potential Switching

If the potential has local minima and if the initial position q_0 is near a local minimum, then to converge to the global minimum the camera must cross over the local maximum at q_a. To climb up the local maximum we generate an artificial potential V' which has the minimum at q_a and control the camera position based on V'. If the camera comes sufficiently close to q_a, then we switch the potential to the original V. If there is no local minimum between q_a and the reference q_d, then the camera will converge to q_d. If there is another local minimum then we repeat the same procedure. We call this control scheme as **potential switching**.

5.1 Interpolation

The initial and the reference images are used to generate the artificial potential. The image used to generate the artificial potential is called the relay image. An example of image interpolation is presented for the 1 DOF circular motion case. However this method can be extended to general 6 DOF case because the method is independent of the number of feature points.

For simplicity, assume that the imaging model is weak perspective projection and let $\theta_0 = -\theta_d$ and $\theta_d < 0$. For the first example let us adopt the averaged image $\xi_c = (\xi_d + \xi_0)/2$ as a relay image. Then we have

$$f(\theta) = -s(\frac{c_d}{2} + \frac{c_0}{2} - c) + ac(\frac{s_d}{2} + \frac{s_0}{2} - s),$$
$$J^T(\xi_c - \xi) = bf(\theta) \qquad (9)$$

By substituting $\theta_0 = -\theta_d$ into the above equation, we can see that the potential has equilibria for $\theta = 0$ 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 $\xi_r = (1-r)\xi_0 + r\xi_d$ $(0 \le r \le 1)$. Let the magnification ratio be γ then the magnified image becomes

$$\xi_i = \gamma\xi_r = \gamma(1-r)\xi_0 + \gamma r\xi_d \qquad (10)$$

When $r = 1/2$, the solutions of $f(\theta) = 0$ are $\theta = 0$ 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 minimum of V_i is the local maxima of V, the camera 'climbs 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 $\gamma(0) = 1, \gamma(1) = 1, \gamma(1/2) > (1-a)/c_d$.

6 Simulation

6.1 1 DOF circular motion

For $\theta_d = -\pi/3, \theta_0 = \pi/4$, a simulation result is given in Figure 9. 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 10. The positions on which the arrows are crowded in Figure 9 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.

6.2 2 DOF Straight Motion

For 2 DOF straight motion with $X_d = 1000, \theta_d = -\pi/4, X_0 = 1000, \theta_0 = \pi/3$, simulation result is shown in Figure 11. The magnification ratio is the same as the previous example. The camera converges to the desired position and orientation.

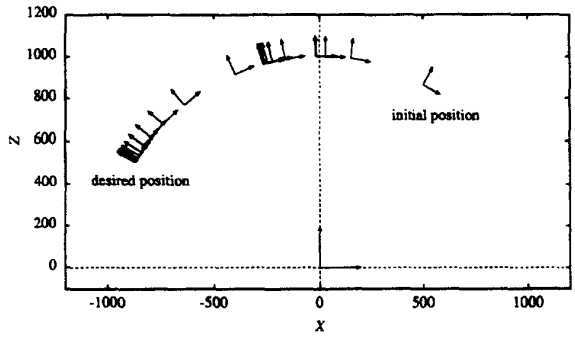

Figure 9: Camera position

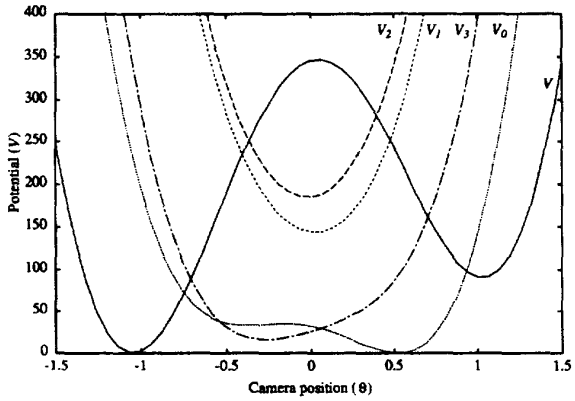

Figure 10: Potentials for interpolated images

7 Conclusion

This paper considered the potential problems in visual servo. Examples have exhibited the existance of local minima and the smallness of stable region even for very simple camera motion. We also have shown that the object height is essential to enlarge the stable region. Moreover, use of switching control with interpolated-magnified relay images is effective to enlarge the stable region. The interpolation method presented here is limited, however by combining with affine transformation the switching control can be generalized to 6 DOF visual servoing.

References

[1] 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.

[2] 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.

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

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

[6] 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*, pages 2254–2260, Sacramento, Calif., 1991.

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

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

[9] 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.

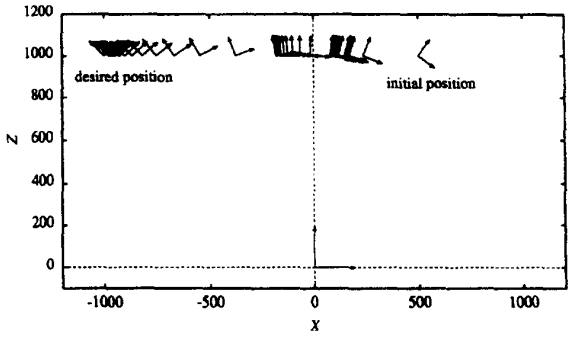

Figure 11: Camera position