A Study on Feature-Based Visual Servoing Control of Robot System by Utilizing Redundant Feature

Sung Hyun Han*
Division of Mechanical and Automation Engineering, Kyungnam University, Masan, Kyungnam 631-701, Korea
Hideki Hashimoto
Institute of Industrial Science University of Tokyo, 7-22-1, Roppongi, Minato Tokyo 106, Japan

This paper presents how effective it is to use many features for improving the speed and accuracy of visual servo systems. Some rank conditions which relate the image Jacobian to the control performance are derived. The focus is to describe that the accuracy of the camera position control in the world coordinate system is increased by utilizing redundant features in this paper. It is also proven that the accuracy is improved by increasing the number of features involved. Effectiveness of the redundant features is evaluated by the smallest singular value of the image Jacobian which is closely related to the accuracy with respect to the world coordinate system. Usefulness of the redundant features is verified by the real time experiments on a Dual-Arm Robot manipulator made by Samsung Electronic Co. Ltd..

Key Words : Visual Servoing Control, Redundant Feature, Feature-Based Visual Tracking, Real Time Control, Image Jacobian

1. Introduction

Although robots can perform assembly recently, material handling jobs with speed and precision, when compared to human workers robots, are hampered by their lack of sensory perception yet. To address this deficiency considerable research into force, tactile and visual perception has been conducted over the past two decades.

Visual servoing is the fusion of result from many elemental areas including high-speed image processing, kinematics, dynamics, control theory, and real-time computing. It has much in common with research into active vision and structure from motion, but is quite different from the often described use of vision in hierarchical task-level robot control systems. Many of the control and vision problems are similar to those encountered by active vision researchers who are building "robotic heads". However, the task in visual servoing is to control a robot to cope with its environment using vision as opposed to just observing the environment.

Most visual servoing problems can be considered as nonlinear control problems with the gray level of each two dimensional pixel array being an observation. The difficulty of the problems is the size and the nonlinearity. The size of each observation is larger than ten thousand and they have nonlinear interaction with each other. A few researches based on the stochastic models of the two-dimension al observation are found, but most visual servoing schemes use the features of an image as the observation. To manipulate objects with complex shapes, it is important to deal with complex features such as spheres and

* Corresponding Author,
E-mail : shhan@kyungnam.ac.kr
TEL : +82-55-249-2624; **FAX** : +82-55-249-2617
Division of Mechanical and Automation Engineering, Kyungnam University, Masan, Kyungnam 631-701, Korea. (Manuscript **Received** July 29, 2000; **Revised** November 30, 2001)

cylinders. However, the time for extracting complex features will become too long due to limited hardware. Accordingly, a visual servoing scheme which utilizes many features effectively is required. Furthermore, exploiting carefully the information from the features will give robust and accurate control performance.

Sanderson et al. (1987) proposed a feature-based approach and defined the Jacobian of are ideal inverse interpretation which was considered as the infinitesimal change of the relative position and orientation between the camera and the object in the environment.

Newman et al. (1987) proposed an adaptive control law based on a single input single output model and a feature selection criterion was proposed. The criterion addressed a selected feature should be used to control an actuator, where the number of the selected features is equal to the number of the actuators. Feddema et al. (1987; 1989) also studied the selection method of the features to make the Jacobian well-conditioned. Real time experiment on gasket tracking showed that the proper selection of features is necessary to minimize the effect of image noise.

Papanikolopoulos et al. (1991; 1993) experimentally examined many control algorithms including Proportional-Plus-Integral, pole assignment and linear quadratic gaussian. Some adaptive control schemes were also examined in Hashimoto (1991). These approaches do not consider to use the redundant features which are defined as the features of more than the degrees of freedom of a robot manipulator.

Chaumette et al. (1991) and Espiau et al. (1992) derived the interaction matrix, and introduced the concept of task function. Chaumette (1994) extended the task function approach to the complex features. Jand and Bien (1991) mathematically defined the "feature", and derived the feature Jacobian matrix. The authors derived the image Jacobian, and used its generalized inverse and PD control to generate the hand trajectory. These schemes are based on the generalized inverse of the Jacobian. Redundant features could be used. However, the parameters for improving the control performance were too limited and the controllability of the redundant features was not discussed.

The authors proposed a linearized dynamic model of the visual servo system and linear quadratic control scheme for redundant features. The controllability problem was discussed but the performance improvement by utilizing the redundant features was not presented.

This paper presents how the control performance of the feature-based visual servoing system is improved by utilizing redundant features. Effectiveness of the redundant features is evaluated by the smallest singular value of the image Jacobian which is closely related to the accuracy in the world coordinate system. Usefulness of the redundant features is verified by the real time control experiments. To illustrate the accuracy of the redundant visual servo system, real time experiments on the Dual-Arm robot with eight joints are carried out. Translational and rotational step response with three, four and five features are examined in the experiments.

2. System Modeling and Formulation

The object image moves with the joint angle to the object image, which is composed of the kinematic model and the camera model as shown in Fig. 1. Suppose that a camera is mounted on the robot hand and the object does not move. The kinematic model is a map from the joint angle to a position of the camera. Since the camera is on the robot hand, the camera position is uniquely defined by the joint angle θ based on the kinematic structure of the robot. The camera model is a map from the position of the camera to the image of the object. The object image is generated by the perspective of the relative position between the camera and the object. The perspective projection is a map between two different representations of the position of the object, i.e., the representations in the camera coordinate system $[XYZ]^T$ and in the image plane $[xy]^T$.

The perspective projection with f being the focal length of the lens is given as

$$[x \quad y]^T=[X \quad Y]^T(f/z) \tag{1}$$

Suppose that there are feature points, namely $p_i=[X_iY_iZ_i]^T$ $(i=1, \cdots, n)$, on an object and the corresponding positions in the image plane are $\xi_i=[x_iy_i]^T$ $(i=1, \cdots, n)$. Assume that the shape and size of the object are known and constant (i.e., the object is a rigid body). Then ξ_i's become functions of the joint angle θ. Let us define $2n$ a dimensional feature vector as $\xi\cong[\xi_1^T\cdots\xi_n^T]^T$. Then the system model for n feature points is defined as the map, $\psi : \mathbf{R}^m \rightarrow \mathbf{R}^{2n}$, from the joint angle θ to the feature vector ξ,

$$\psi(\theta)\cong\xi \tag{2}$$

where m is the number of joints of the robot.

Since the task must be carried out in the nonsingular region of the robot, the nonsingular region is called the operation region $\mathrm{M}_\theta\subset\mathbf{R}^M$. We restrict the robot motion in the operation region. Thus the robot Jacobian J is invertible in the working area. It is useful to introduce the feature manifold , which is defined as

$$\mathrm{M}=\{\,\xi\in\mathbf{R}^{2n} : \xi=\psi(\theta),\ \theta\in\mathrm{M}_\theta\,\} \tag{3}$$

The features on the feature manifold is called the admissible features. If the features are admissible, then the robot Jacobian is invertible by definition. In Eq. (3), θ represents the joint angle.

Differentiation of the system model yields

$$\dot{\xi}=J\dot{\theta} \tag{4}$$

where the $2n\times m$ matrix J is defined as

$$J\cong\begin{bmatrix} J_{im}^{(1)} \\ \vdots \\ J_{im}^{(n)} \end{bmatrix} {}^cJ_{ARM} \tag{5}$$

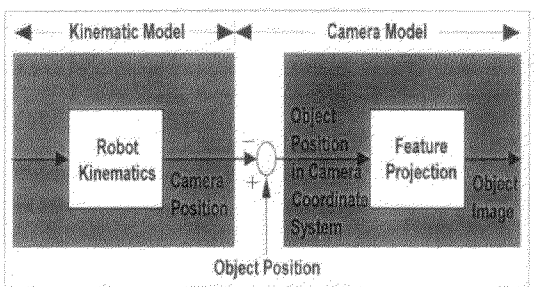

Fig. 1 System modeling

The matrix $J_{im}^{(i)}$ is given by Feddema (1987) and Papaniklolpoulos (1991)

$$J_{im}^{(i)}\cong\begin{bmatrix} -\frac{f}{z_i} & 0 & \frac{x_i}{Z_i} & \frac{x_iy_i}{f} & -\frac{x_i^2+f^2}{f} & y_i \\ 0 & -\frac{f}{z_i} & \frac{y_i}{Z_i} & \frac{y_i^2+f^2}{f} & \frac{x_iy_i}{f} & -x_i \end{bmatrix} \tag{6}$$

and called the image Jacobian. ${}^cJ_{ARM}$ is the robot Jacobian expressed in the camera coordinate system. Since the vector ${}^cJ_{ARM}\dot{\theta}\in\mathbf{R}^6$ is the linear and angular velocities of the camera expressed in the camera coordinate system, $J_{im}^{(i)}$ becomes the infinitesimal change of the position of the camera. Moreover, $J(i)$ $(\cong J_{im}^{(i)}{}^cJ_{ARM})$ becomes the infinitesimal change of the features according to the infinitesimal change of the joint angles.

The degenerated features are the features for which the extended image Jacobian is of not full rank. The degenerate features should be avoided because the inverse map (the map from ξ to θ) becomes singular. Thus, when the number of joints is m,

$$rank\,J(\theta)=m \quad \forall\theta\in\mathrm{M}_\theta \tag{7}$$

is required for all admissible features. To satisfy this condition, $n\geq m/2$ is an obvious necessary condition, but it is not sufficient for some cases.

For example, consider a general six degree of freedom case $(m=6)$. In this case, $n\geq3$ is necessary. If $n=3$, the camera lies on the cylinder (Fig. 2) which includes the three points and the axis of which is perpendicular to the plane containing these points. For any attitude of the camera, J is singular. Thus $n=3$ is not sufficient and $n\geq4$ is desirable. For the case of $n=4$, we have the following theorem.

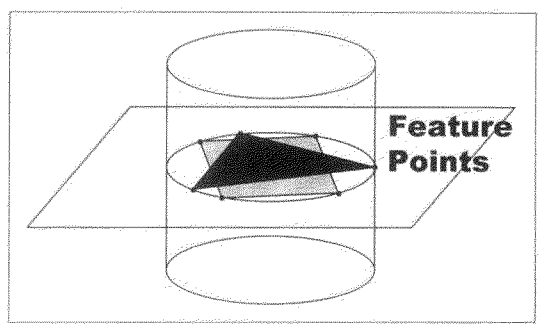

Fig. 2 Singular cylinder

Theorem 1 : Suppose that there are four points on a plane and the corresponding feature vector is admissible. Then the extended image Jacobian is of full rank if any three feature points out of them are not collinear in the image plane.

Proof : Let the plane on which the four points exist be $Z=pX+qY+r$. Then Z_i satisfies $Z_i=pX_i+qY_i+r$ for $i=1, \cdots, 4$. Substituting (1) into this yields

$$\frac{f}{Z_i}=\frac{f-px_i-qy_i}{r} \tag{8}$$

And substituting this into (6) yields

$$J_{im}^{(i)}=M_iN \tag{9}$$

where M_i and N are defined as

$$M_i=\begin{bmatrix} f & 0 & x_i & y_i & 0 & 0 & x_i^2/f & x_iy_i/f \\ 0 & f & 0 & 0 & x_i & y_i & x_iy_i/f & y_i^2/f \end{bmatrix},$$

$$N=\frac{1}{r}\begin{bmatrix} -1 & 0 & 0 & 0 & -r & 0 \\ 0 & -1 & 0 & r & 0 & 0 \\ p & 0 & 1 & 0 & 0 & 0 \\ q & 0 & 0 & 0 & 0 & r \\ 0 & p & 0 & 0 & 0 & -r \\ 0 & q & 1 & 0 & 0 & 0 \\ 0 & 0 & -p & 0 & -r & 0 \\ 0 & 0 & -q & r & 0 & 0 \end{bmatrix} \tag{10}$$

Then we obtain $J=MN^cJ_{ARM}$, where $M=[M_1^T\ M_2^T\ M_3^T\ M_4^T]^T$. It is straightforward to see that

$$\det M=\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}\cdot\begin{vmatrix} 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \\ 1 & x_4 & y_4 \end{vmatrix}\cdot\begin{vmatrix} 1 & x_3 & y_3 \\ 1 & x_4 & y_4 \\ 1 & x_1 & y_1 \end{vmatrix}\cdot\begin{vmatrix} 1 & x_4 & y_4 \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{vmatrix} \tag{11}$$

Thus M is invertible because any three feature points are not collinear. On the other hand, if $p^2+q^2\neq 0$, the first six rows of is linearly independent. If $p=q=0$, the first four and the last two rows are linearly independent. Thus rank $N=6$. Finally, since all features are admissible, ${}^cJ_{ARM}$ is invertible. Therefore, the extended image Jacobian J is of full rank.

3. Analysis of Visual Servoing System

For evaluating the performance of the feature-based visual servo system, it is useful to discuss the ratio of the joint angle error to the feature vector. The following theorem shows that increasing the number of the feature point is an effective way to improve the performance.

Let the joint error be $\Delta\theta=\theta-\theta_d$ and the feature error be $\Delta\xi=\xi-\xi_d$. Define the worst joint/feature error ratio ER, called sensitivity, as follows :

$$ER=\sup_{\|\Delta\xi\|\neq 0}\frac{\|\Delta\theta\|}{\|\Delta\xi\|}=\frac{1}{\beta_{\min}(J)} \tag{12}$$

where $\beta_{\min}(J)$ is the minimum singular value of J. Then the sensitivity ER decreases strictly by increasing the number of non-degenerated features on the object.

Let J_n be the image Jacobian for n feature points and J_{n+1} be the image Jacobian obtained by adding an extra feature point to the already existing feature points.

Then we have

$$J_{n+1}=\begin{bmatrix} J_n \\ J^{(n+1)} \end{bmatrix} \tag{13}$$

where $J^{(n+1)}$ is the $2\times m$ image Jacobian corresponding to the newly added feature point. It is straightforward to see that

$$\beta_{\min}(J_n)\leq\beta_{\min}(J_{n+1}) \tag{14}$$

The equal sign holds only if each row of $J^{(n+1)}$ is linearly dependent on J_n, i.e., only if J_{n+1} is not of full rank. Since we assumed that the features are not degenerated, the equal sign should be dropped. Thus adding extra feature points strictly increases the minimum singular value.

This theorem says that we can reduce the joint angle error by increasing the number of feature points.

Linearizing the model (2) with the feature vector being the state vector yields an uncontrollable model because ξ can not move arbitrarily in $\mathbf{R}^{2n}$. A simple way to avoid this problem is to map $\xi\in$M onto the tangent space of M by using the following transformation.

$$z=J_d^T(\xi-\xi_d) \tag{15}$$

where $J_d(=J(\theta_d))$ is the image Jacobian at the desired point. Note that z and θ are one-to-one

in the neighborhood of θ_d. The dynamics of the feature error on the tangent space of the manifold M is given as

$$\dot{z}=J_d^T J(\theta)\dot{\theta} \tag{16}$$

Thus, for a simple continuous time control law $\dot{\theta}=-Kz$ with a positive definite constant matrix K yields an asymptotic stability if $J_d^T J(\theta)$ is positive definite. It was shown that this condition is satisfied in fairly large region of θ_d.

4. Experiments and Discussions

As shown in Fig. 3, the objects are white boards with three, four and five black marks. Three points are arranged to make an equilateral triangle with an edge length of 120mm. Four points are on corners of a square with an edge length of 120mm. All marks are on a plane except the center one of five points, with a height of 40mm. Dual-Arm robot holds the objects and a camera(Fig. 4). The world coordinate system $\omega_X-\omega_Y-\omega_Z$ is at the base of the Dual-Arm robot. A nominal camera position is almost in front of the plane on which the marks exist. To avoid the singular cylinder(Fig. 2), the optical axis and the normal axis of the object plane are not aligned. The distance is about 1000mm. The features are the x and y coordinates of the center of the image of each mark. Computing their minimum singular values at the reference position gives

$$\begin{aligned}\beta_{\min}&=(J_3)=0.35,\\ \beta_{\min}&=(J_4)=0.65,\\ \beta_{\min}&=(J_5)=3.60.\end{aligned} \tag{17}$$

Thus accuracy of the position control of the camera in the 3D work space will be improved by using 5 features. We carried out many step tests to confirm this observation.

The experiment is on a step motion in vertical axis. The object is moved upward for 120mm (i.e., in ω_Z direction). The camera is controlled to keep the features at the initial positions. Thus the initial values and the reference values are the same. The object motion is considered as a disturbance for the plots of the features in the image plane. On the other hand, the object motion becomes the step change of the reference position for the plots of the camera motion in the world coordinate system. Since Dual-Arm robot has only 6 degrees of freedom, the orientation of the object changed slightly. Thus, the reference orientation is [2.8, 0, −1.8] degrees expressed in the Euler angles, say p, η, ϕ.

Figure 5 has six curves which show the and

Table 1 Specification of dual arm robot

Content		Unit	Spec.	Remark
Workspace	1st Arm	deg	180	
	2nd Arm	deg	450	
	Z Axis	mm	150	
	R Axis	deg	±180	
Maximum Reach		mm	(350+260)	
Payload		kg	2.5	High-speed
Max. Resultant Vel.		m/sec	5.4	1, 2 Axis
Position Repeatability	Plane	mm	0.05	1, 2 Axis
	Z Axis	mm	0.02	
	M Axis	deg	0.05	
Weight		kg	200	
Coincident Control Axis No.		EA	8Axis (4+4)	

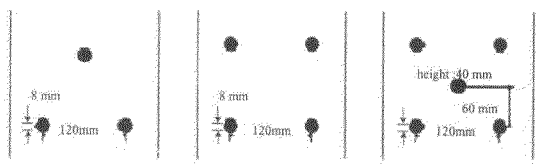

Fig. 3 Configuration of feature points

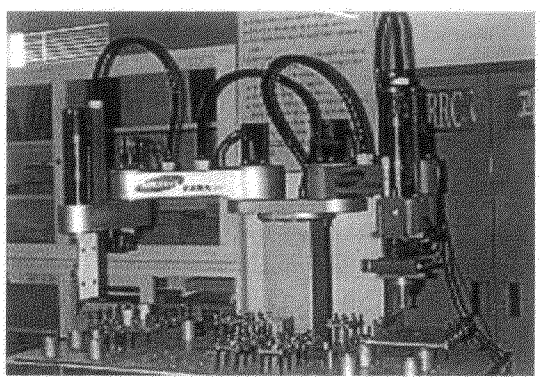

Fig. 4 Experimental setup

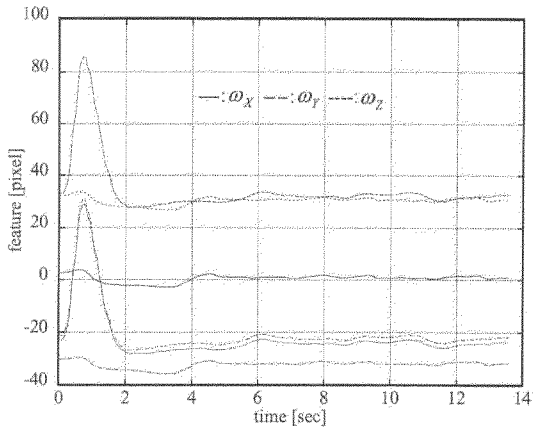

Fig. 5 Response in image plane for 3 points (vertical step)

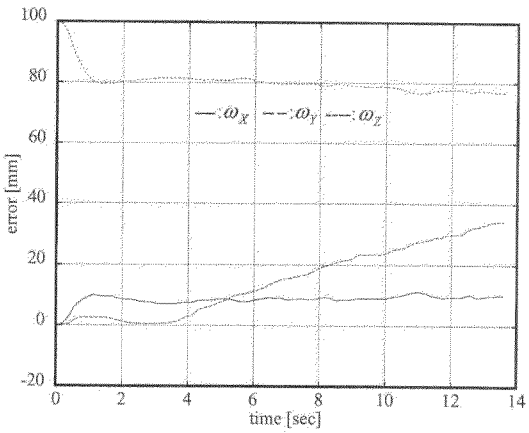

Fig. 8 Error in 3D for 3 points (vertical step)

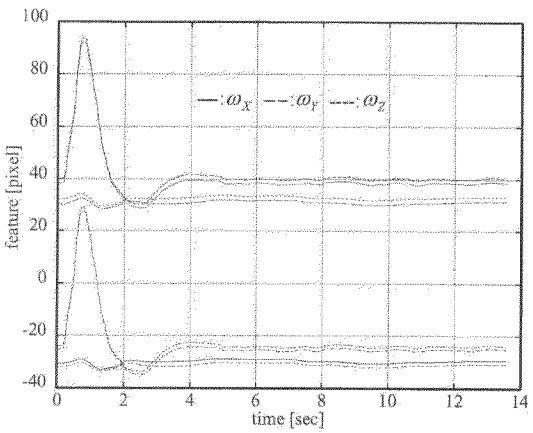

Fig. 6 Response in image plane for 4 points (vertical step)

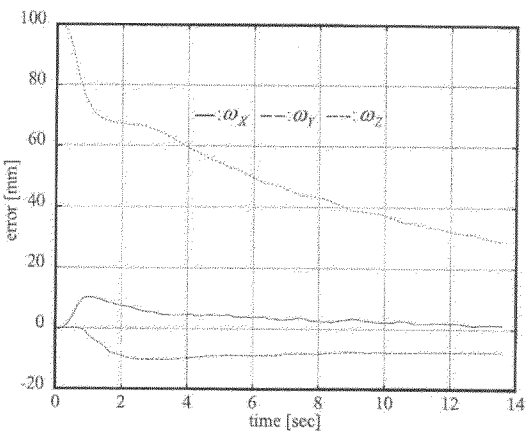

Fig. 9 Error in 3D for 4 points (vertical step)

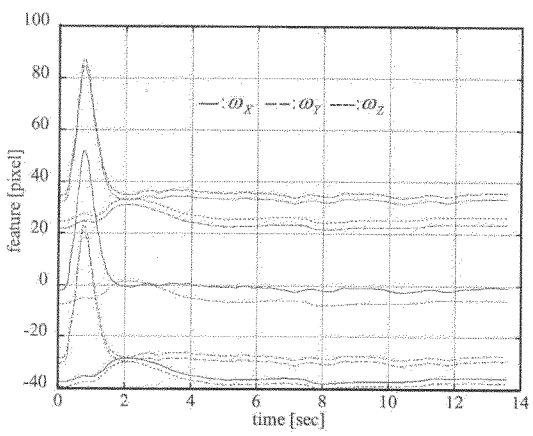

Fig. 7 Response in image plane for 5 points (vertical step)

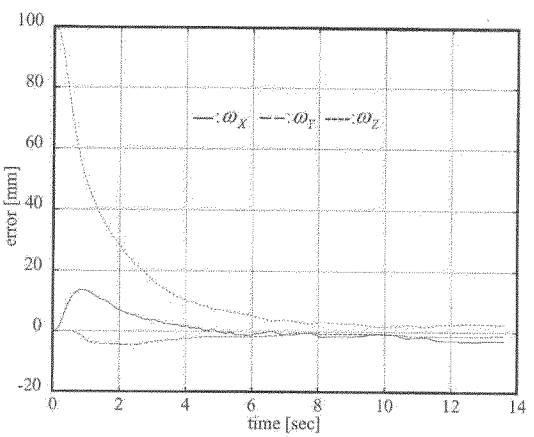

Fig. 10 Error in 3D for 5 points (vertical step)

coordinates of the three feature points in the image plane. The horizontal axis is the time. The largely dicturbed curves are the y coordinates and the others are the x coordinates. They are almost stabilized in two seconds. Figure 7 depicts the image coordinates of the five points. All responses in the image plane are similar to each other.

The plots in Fig. 8 depicts the position errors of the camera for the three feature points (measured in the world coordinate system). The error ω_Y in direction is diverging. However, as shown in Fig. 9, the response of the camera position with four feature points is stabilized. It is sluggish, and it takes more than 20 seconds to stabilize the disturbance. Figure 10 is the response with the five feature points. It is improved very much for both speed and accuracy. The steady state errors are smaller than 5mm for all directions.

5. Conclusions

In this paper, it was presented how the control performance of the feature-based visual servo system is improved by utilizing redundant features. Effectiveness of the redundant features was evaluated by the smallest singular value of the image Jacobian which is closely related to the accuracy in the world coordinate system. It was shown that the accuracy of the camera position control in the world coordinate system was increased by utilizing redundant features. Real time experiments on Dual-Arm robot were carried out to evaluate the improvement of the accuracy and speed by utilizing redundant features. The results verified that the minimum singular value of the extended image Jacobian plays an important role in evaluating performance improvement of the feature-based visual servoing.

References

Allen, P. K., Timcenko, A., Yoshimi, B. and Michelman, P., 1993, "Automated Tracking and Grasping of a Moving Object with a Robotic Hand-Eye System," *IEEE Trans. Robotics and Automation*, Vol. 9, No. 2, pp. 152~165.

Chaumette, F., 1994, "Visual Servoing Using Image Features Defined Upon Geometrical Primitives," in *Proc. 33rd Conf. Decision and Control*, Lake Buena Vista, Florida, pp. 3782~3787.

Chaumette, F., Rives, P. and Espiau, B., 1991, "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., pp. 2248~2253.

Espiau, B., Chaumette, F. and Rives, P., 1992, "A new Approach to Visual Servoing in Robotics," *IEEE Trans. Robotics and Automation*, Vol. 8, No. 3, pp. 313~326.

Feddema, J. T. and Michell, O. R., 1989, "Vision Guided Servoing with Feature-Based Trajectory Generation," *IEEE Trans. Robotics and Automation*, Vol. 5, No. 5, pp. 691~700.

Feddema, J. T., Lee, C. S. G. and Michell, O. R., 1987, "Automatic Selection of Image Features for Visual Servoing of a Robot Manipulator," in *IEEE Int. Conf. Robotics and Automation*, Scottsdale, Ariz, pp. 832~837.

Hashimoto, K., 1991, "Image-Based Dynamic Visual Servo for a Hand-Eye Manipulator," in *MTNS-91*, Kobe, Japan, pp. 609~614.

Hashimoto, K., Ebine, T. and Kimura, K., 1996, "Visual Servoing with Hand-Eye Manipulator—Optimal Control Approach," *IEEE Trans. on Robotics and Automation*, Vol. 12, No. 5, pp. 766~774.

Hashimoto, K., 1991, "Manipulator Control and with Image-Based Visual Servo," in *IEEE Int. Conf. Robotics and Automation, Sacramento*, Calif., pp. 2267~2272.

Hashimoto. K., Ebine, T., Sakamoto, K. and Kimura, H., 1993, "Full 3D visual tracking with nonlinear model-based control," in *American Control Conference*, San Francisco, Calif., pp. 3180~3184.

Jand, W. and Bien, Z., 1991, "Feature-Based Visual Servoing of an Eye-in-Hand Robot with Improved Tracking Performance," in *IEEE Int. Conf. Robotics and Automation*, Sacramento, Calif., pp. 2254~2260.

Longuet-Higgins, H. and Prazdny, K., 1984,

"The Interpretation of a Moving Retinal image," Proc. Royal Soc. London Ser. B, Vol. 208, pp. 385~397.

Papaniklolpoulos, N., Khosla, P. K. and Kanade, T., 1991, "Vision and Control Techniques for Robotic Visual Tracking," in *IEEE Int. Conf. Robotics and Automation*, Sacramento, Calif., pp. 857~864.

Papanikolopoulos, N. and Khosla, P. K., 1993, "Adaptive Robotic Visual Tracking: Theory and Experiments," *IEEE Trans. Automatic Control*, Vol. 38, No. 3, pp. 429~445.

Sung Hyun Han, Man Hyung Lee and Hideki Hashimoto, 2000, "Image-Based Servoing Control of a SCARA Robot" *KSME International Journal in Korea*, Vol. 14, No. 7, pp. 782~788.

Weiss, L. E., Sanderson, A. C. and Newman, C. P., 1987, "Dynamic Sensor-Based Control of Robots with Visual Feedback," *IEEE J. Robotics and Automation*, Vol. RA-3, No. 5, pp. 404~417.