Applying Visual Servoing Techniques to Control a Mobile Hand-Eye System

Roger Pissard-Gibollet and Patrick Rives
INRIA
2004 Route des Lucioles
06902 Sophia Antipolis Cedex - France
rives@sophia.inria.fr, pissard@sophia.inria.fr

Abstract

In this paper, we discuss both theoretical and implementation issues of a vision based control approach applied to a mobile robot. After a brief presentation of the visual servoing framework used, we point out some key problems associated to its application in the case of nonholonomic mobile robots. To overpass these problems, we propose to use a hand-eye system mounted on the mobile robot in order to introduce redundancy with respect to the task. The second part of the paper presents our first results obtained on our experimental testbed.

1 Introduction

It is now wellknown that nonholonomic systems like wheeled mobile robots lead to specific problems of control. As a consequence of the Brockett's theorem [1], it has been proved that, a smooth pure state-space feedback law capable to stabilize a nonholonomic mobile robot around a fixed point in the configuration space does not exist. Recently, some authors started investigating alternative control techniques, little studied up to now, which consist in using *time-varying* feedback laws [2], or discontinuous or piecewise smooth controllers. For more details about these techniques of control, the reader is referred to [3]. Unfortunately, to date, very little results exist dealing with the robustness and accuracy aspects of such schemes with regard to the perturbations due to the imperfect modeling of the robot and to the errors in the odometry measurements. Moreover, these control laws are designed in the configuration space, and do not take into account sensory data from their environment. In this paper, we propound another kind of approach. We assume that, for correctly completing a task, we need a low level of task specification which explicitely integrates the interaction between the robots and its local environment. Therefore, we introduced a framework [4] linked to the sensor based control approach. This framework was originally designed for manipulator arms using a camera mounted on the end effector (cf. [5], [6] for more details). In this paper, we point out the difficulties to transpose these techniques to the case of nonholonomic mobile robots. In the first part, we briefly recall theoretical aspects of visual servoing which allow to relate visual constraints expressed in the image frame and geometric constraints expressed in the configuration space. In the second part, using recent results [7], we propose a general way to compute kinematic model of a mobile robot available for different mechanical structures. Then, we show that, due to the presence of nonholonomic constraints, visual servoing schemes performing well for holonomic hand-eye systems, cannot be directly transposed. To overpass this problem, we propose to add judicious degrees of freedom to the sensor in such a way that the dimension of the configuration space of the whole mechanism is equal or greater than the dimension of the control input vector. This approach was applied to an experimental testbed [8] including a mobile platform carrying a pan-and-tilt camera. Experimental results on the visual tracking of ellipses are presented. The trajectories produced with a very simple velocity control scheme for different starting configurations are shown.

2 Local tasks based on visual servoing techniques

2.1 Modeling the interaction camera-environment

Let us consider a sensor (S) with its related frame F_S moving with respect to a reference frame F_0. (S)

interacts with a part of the environment with an associated frame F_T (figure (1)). The pose (location and attitude) of (S) is an element $\bar{r}$ of SE_3 which is a six dimensional differential manifold. Now, let us consider that the useful sensor output is a k-dimensional vector $\underline{s} = (s_1, \cdots, s_k)^T$. Then, to be used as feedback in a sensor based control scheme, each component s_j has to define a mapping from SE_3 to $I\!R$ such that s_j is only dependant of the relative pose between the sensor frame F_S and the frame F_T attached to the environment.

$$s_j = s_j(F_S, F_T) = s_j(\bar{r}) \in I\!R$$

Moreover, we will assume $s_j(.)$ to be C^2. Let us note that if the sensor is motionless with regard to a static target, s_j remains constant. Due to the properties of the mapping from SE_3 to $I\!R$, the differential of s_j is a mapping from se_3 to $I\!R$. It can be expressed as a screw product :

$$\dot{s}_j = \frac{\partial s_j}{\partial \bar{r}} \bullet \frac{d\bar{r}}{dt} = H_j \bullet T_{ST}$$

where H_j is called the *interaction screw*, T_{ST} is the velocity screw between F_S and F_T, and $\bullet$ represents the screw product.

Let us consider the bilinear form L^T associated to the interaction screw H; with an obvious breach of notation, we obtain, in the sensor frame :

$$\dot{s}_j = L_j^T . T_{ST}, \qquad (1)$$

L_j^T can be viewed as a *Jacobian* of the elementary sensor output s_j with regard to the relative displacement between the camera and the environment.

Important remark: L_j^T characterizes the interactions between the sensor and its environment. It is independant of the way to physically perform the motion of the sensor (handled by a robot arm or embarked on a mobile robot).

For example, in the case of vision sensors, the simplest interaction matrix links the variation of the coordinates (X_i, Y_i) of the pixels, to the relative motion between the camera and the 3D corresponding points in the scene. This is the wellknown optical flow relations.

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} = \begin{bmatrix} -1/z & 0 & X/z & XY & -(1+X^2) & Y \\ 0 & -1/z & Y/z & 1+Y^2 & -XY & -X \end{bmatrix} . \begin{bmatrix} V_x \\ V_y \\ V_z \\ \Omega_x \\ \Omega_y \\ \Omega_z \end{bmatrix}$$

This result was generalized to a lot of geometrical primitives like points, lines and circles. A closed form of the interaction screw can be obtained for the set of 3D geometrical primitives which projects onto the image frame under the form of a general parametric curve [5].

2.2 Relation between constraints in the configuration space and constraints expressed in the sensor space

As a straightforward consequence of the definition of s_j (i.e. a mapping from SE_3 to $I\!R$), we have the following properties :

Property 1: For a given geometrical primitive characterised by a vector of elementary visual signal $\underline{s} = [s_1 \ldots s_k]^T$, the mapping is *injective*, ie, it is possible to relate a certain configuration $\bar{r}^d$ between F_S and F_T to a certain value $\underline{s}^d$ of the vector $\underline{s}$.

Property 2: In function of the value $\underline{s}^d$, it may exist motions of the camera which leave the value of $\underline{s}$ unchanged. For example, in the figure (2), if the velocity screw T_{ST} belongs to the plane constituted by the optical center and the 3D line, no change appears in the image frame during the motion.

Given the Interaction Matrix L^T, the velocity screws T^* which leave $\underline{s}$ invariant during the motion are solutions of :

$$\dot{\underline{s}} = L^T . T^* = 0$$

The motions T^* belong to the subspace S reciprocal to the subspace spanned by $\{L_1^T \ldots L_k^T\}$ a basis of which is defined by $Ker(L^T)$.

This result allows us to state that: Imposing $\dot{\underline{s}} = 0$ is equivalent to introduce *geometrical constraints* in the configuration space SE_3 in order to achieve a *Virtual Linkage* between the sensor and its environment.

We define the *class* N of the virtual linkage as the number of degrees of freedom of the sensor leaving $\underline{s}$ unchanged.

$$N = dim(Ker(L^T)) \text{ with } rank(L^T) = 6 - N$$

2.3 Using sensor signals in a robust closed loop control scheme

Using the general formalism introduced in [4], we can express a robot task in term of the regulation of an output function $\underline{e}(\bar{r}, t)$. We consider that the task is perfectly achieved during the time interval $[0, T]$ iff :

$$\underline{e}(\bar{r}, t) = 0, \ \forall t \in [0, T]$$

Applying this formalism to the sensory based task, we showed in previous papers [6] that a generic form of $\underline{e}(\bar{r},t)$ can be built using the sensor output $\underline{s} = [s_1 \dots s_k]^T$ in a following way :

$$e(\bar{r},t) = \begin{pmatrix} s_1(\bar{r},t) - s_1^d(t) \\ \vdots \\ s_k(\bar{r},t) - s_k^d(t) \end{pmatrix}$$

It was also shown that a very simple *gradient based approach* embedded in a velocity control scheme is sufficient to ensure an exponential convergency to 0 of the task function $\underline{e}(\bar{r},t)$ by using the following desired velocity screw T^d as control input :

$$T^d = -\lambda . L^{T^+}_{\underline{s}=\underline{s}^d} . (\underline{s}\,(\bar{r}) - \underline{s}^d) \tag{2}$$

where λ is a positive gain and $L^{T^+}_{\underline{s}=\underline{s}^d}$ is the pseudo inverse of the interaction matrix computed at the desired value $\underline{s}^d$.

Such a scheme, using vision sensors, has been successfully implemented [5] for a camera mounted on a 6 dof manipulator end effector. In this case, the desired velocity screw T^d was effectly realized by using the inverse Jacobian of the manipulator :

$$\dot{q}^d = J^{-1}(q).T^d \tag{3}$$

Transposing such a scheme to nonholonomic mobile robots is not straightforward.

3 Modeling and control of a mobile manipulator

Applying vision based control theory to mobile robots provides an interesting framework to implement reactive robotic tasks. However, we have to take some care due to the presence of nonholonomic constraints. These constraints imply that the dimension of the configuration space is greater than the number of degrees of freedom that can really be controlled. Consequently, a simple control scheme based on a gradient descent method cannot ensure a convergence of the task function when the camera is rigidly attached to a mobile cart.

One way to overpass this problem consists in adding some degrees of mobility to the camera by mounting it on a motorized device like a manipulator. Then, considering the whole mechanical system constituted by the cart and its manipulator like a single kinematic chain, we will show that the mapping between the configuration space (d.o.f of the cart + d.o.f of the manipulator) and the pose (attitude and location) of the camera is almost everywhere regular (excepted in some isolated configurations). Also, it becomes possible to fully control the motion of the camera without being limited by the cart nonholonomic constraints. Doing so, we shall be able to execute visual servoing tasks directly defined in the camera frame, in a similar way to the case of holonomic systems. Obviously, in such an approach, the trajectory of the cart will not be explicitely controlled. However, by using some trivial geometric considerations, we shall be able to bound the sweeping surface occupied by the cart during its displacement.

In this section, we propose a general approach to compute the kinematic model of a wheeled mobile manipulator, then we apply it to a robot constitued by a two-wheel driven cart with a camera mounted on a three d.o.f. manipulator described in figure (1).

3.1 Kinematic model for wheeled mobile robots

Recent works [7] provide a general framework to compute kinematic models of wheeled mobile robots. Let us briefly recall this framework.

The robot posture is specified by the vector : $\underline{\xi} = (x\, y\, \theta)^T$ where $(x\, y)^T$ are the coordinates of the reference point M in F_0 and θ the orientation of F_M with an associated rotation matrix $R(\theta)$ (see figure (1)).

Four types of wheels was defined :

- N_f fixed wheels,
- N_c centered oriented wheel,
- N_{oc} off-centered oriented wheel,
- N_{sw} swedish wheels.

where $N = N_f + N_c + N_{oc} + N_{sw}$ is the whole number of robot's wheels. Each class of wheel is characterized by a set of constant parameters (for example distance l between M and the center of wheel) and time varying parameters : φ for the rotation of the wheel around its horizontal axis, β_c (resp. β_{oc}) for the direction of a (resp. off) centered oriented wheel. The whole set coordinates $\underline{z} = \begin{pmatrix} \underline{\xi} & \underline{\beta_c} & \underline{\beta_{oc}} & \underline{\varphi} \end{pmatrix}^T$ constitutes the *configuration coordinates* of a mobile robot.

With this parametization, the pure rolling and non slipping constraint can be expressed under a general matrix form :

$$\left\{ \begin{array}{l} J_1(\underline{\beta_c}, \underline{\beta_{oc}}) R(\theta) \dot{\underline{\xi}} + J_2 \dot{\underline{\varphi}} = 0 \\ C_1(\underline{\beta_c}, \underline{\beta_{oc}}) R(\theta) \dot{\underline{\xi}} + C_2 \dot{\underline{\beta_{oc}}} = 0 \end{array} \right. \tag{4}$$

where J_1, J_2, C_1, C_2 are matrices whose rows derive from constraint equations :

$$C_1(\underline{\beta_c}, \underline{\beta_{oc}}) = \begin{pmatrix} C_{1f} \\ C_{1c}(\underline{\beta_c}) \\ C_{1oc}(\underline{\beta_{oc}}) \end{pmatrix} \quad , C_2 = \begin{pmatrix} 0 \\ 0 \\ C_{2oc} \end{pmatrix} \tag{5}$$

Let us consider the $(N_f + N_C)$ first constraints, (4) can be written :

$$\begin{cases} C_{1f} R(\theta)\dot{\underline{\xi}} & = 0 \\ C_{1c}(\underline{\beta_c}) R(\theta)\dot{\underline{\xi}} & = 0 \end{cases} \tag{6}$$

These constraints imply that the vector $R(\theta)\dot{\underline{\xi}}$ belongs to the null space of the following matrix $C^*(\beta_c) = (C_{1f}, C_{1c}(\beta_c))^T$. The rank of this matrix allows to compute the *degree of mobility* of a mobile robot and its classification.

The state space representation can be easily derived from these constraints :

$$R(\theta)\dot{\underline{\xi}} \in Ker[C_1^*(\underline{\beta_c})] \Leftrightarrow \exists \underline{\eta}(t) \; \dot{\underline{\xi}} = R(\theta)^T \Sigma(\underline{\beta_c})\underline{\eta} \tag{7}$$

where $\Sigma(\underline{\beta_c})$ forms a basis of a $Ker(C_1^*(\underline{\beta_c}))$ and $\underline{\eta}$ can be interpreted as control inputs. We can augment (7) by adding control input on $\dot{\underline{\beta_c}}$ resulting:

$$\begin{cases} \dot{\underline{\xi}} = R(\theta)^T \Sigma(\underline{\beta_c})\underline{\eta} \\ \dot{\underline{\beta_c}} = \underline{\zeta} \end{cases} \tag{8}$$

Finally, we can express the kinematic model by relating the robot configuration vector $\underline{z}$ to the control inputs $\dot{\underline{q}}_{cart} = (\underline{\eta}\,\underline{\zeta})^T$:

$$\dot{\underline{z}} = S(\underline{z})\, \dot{\underline{q}}_{cart} \tag{9}$$

with:

$$S(\underline{z}) = \begin{pmatrix} R^T(\theta)\Sigma(\beta_c) & 0 \\ 0 & I \\ -C_{2oc}^{-1} C_{1oc}(\beta_{oc})\Sigma(\beta_c) & 0 \\ -J_2^{-1} J_1(\beta_c, \beta_{oc})\Sigma(\beta_c) & 0 \end{pmatrix} \tag{10}$$

From this general approach and for a given mechanical structure, we can compute the direct kinematic model in order to have instantaneous velocities (in translation and rotation) of F_M:

$$\begin{pmatrix} \vec{V_M} \\ \Omega_{F_M/F_O} \end{pmatrix} = \begin{pmatrix} J_{MV} \\ J_{M\dot{\theta}} \end{pmatrix} \dot{\underline{q}}_{cart} \tag{11}$$

3.2 Kinematic model for manipulators

The position $\underline{r}$ of a frame F_S linked to an arm effector which only involves holonomic links parametrized by a vector $\underline{q}_{arm}$ can be expressed by the geometrical model:

$$\underline{r} = \Phi(\underline{q}_{arm}) \tag{12}$$

Classically, we use Denavit-Hartenberg parametrization to compute such a model.

The derivation of (12) provides the kinematic model:

$$\dot{\underline{r}} = J_{arm}(\underline{q}_{arm})\, \dot{\underline{q}}_{arm} \tag{13}$$

with $\dot{\underline{r}} = (\vec{V_{S/FM}}\; \Omega_{F_S/F_M})^T$.

3.3 Kinematic model for mobile manipulators

Using the fundamental kinematic relation, we can express camera kinematic screw:

$$\begin{cases} \vec{V_S} = \vec{V_M} + \Omega_{F_M/F_O} \wedge \vec{MS} + \vec{V_{S/F_M}} \\ \Omega_{F_S/F_O} = \Omega_{F_M/F_O} + \Omega_{F_S/F_M} \end{cases} \tag{14}$$

Let us consider all the mobile manipulator d.o.f represented by the vector: $\underline{q} = (\underline{q}_{cart}\, \underline{q}_{arm})^T$. Using (11), (13) and the previous equations (14), we can compute the jacobian matrix of the complete mechanism system:

$$J = \left(\begin{array}{c|c} J_{M\dot{\theta}} \wedge \vec{MC} + J_{MV} & J_{arm} \\ J_{M\dot{\theta}} & \end{array} \right) \tag{15}$$

3.4 Application to a mobile manipulator

The mobile manipulator, described in figure (1), has five dof: $\underline{q} = (q_d\, q_g\, q_1\, q_2\, q_3)^T$. The previously described general approach is used to compute its kinematic model J. The details of the computation can be found in [9].

Computing the determinant of J provides :

$$Det(J) = -\frac{d l_2 r^2 cos(q_2)}{2R} \tag{16}$$

It means that a singular configuration appears only when the robot reaches its limit configurations ($q_2 = \frac{\pi}{2} + k\pi$) which corresponds to a vertical position of the second segment (l_2). The condition $d \neq 0$ and $l_2 \neq 0$ implies that there is no redundancy between the cart rotation axis and q_1. Consequently, no singularity occurs from the nonholonomic constraints and J is alway a full rank matrix when the manipulator does not reach its limit configurations.

4 Experiments

In order to validate and illustrate our approach, we have implemented a task for positioning the camera of the mobile manipulator in front of a target. The selected $3D$ target is a cone represented in figure (3). It projects on to the image frame as two ellipses corresponding to the two circles bounding the cone. The detail of the parametrization of ellipses, the computation of the interaction matrix are fully described in [10]. The positioning task consists in aligning the cone axis with the optical axis of the camera and in tuning the distance between the camera and the target along with this axis (see for notations figure (3)). From each ellipse (i=1,2) on the image, we extract five parameters $(x_{ci}, y_{ci}, \mu_i^{1,1}, \mu_i^{2,0}, \mu_i^{0,2})$.

For this robotic task, we must control three d.o.f of the camera frame : two translations along $\vec{z}$ and $\vec{x}$ axis and one rotation around $\vec{y}$. Among the ten parameters available from the two ellipses, we choose the following visual signal vector :

$$\underline{s} = (x_{c1}\, x_{c2}\, \zeta)^T \text{ with } \zeta = \frac{R_1}{\sqrt{\mu_1^{2,0} + \mu_1^{0,2}}} \tag{17}$$

Using this vector, the interaction matrix can be computed for the desired final value:

$$L^T_{s=s^d} = \begin{pmatrix} -\frac{1}{Z} & 0 & 0 & 0 & -1-\left(\frac{R_1}{Z}\right)^2 & 0 \\ -\frac{1}{Z+H} & 0 & 0 & 0 & -1-\left(\frac{R_1}{Z+H}\right)^2 & 0 \\ 0 & 0 & \frac{\lambda}{R_1} & 0 & 0 & 0 \end{pmatrix} \tag{18}$$

Computing $Ker(L^T_{s=s^d}) = (0\,1\,0\,1\,0\,1\,)^T$ allows us to verify that $\underline{s}$ is fully consistent with the robotic task. It shows that only three d.o.f of the robot (two translation and one rotation) will be controlled using this feature. The interaction matrix (18) is used to compute the control input as defined in (2) and (3).

We have developped a highly versatile testbed in order to validate vision based approaches in real experiments. The robotic system (fig.4) uses a cart-like mobile robot carrying a hand-eye system with two d.o.f [8]. The on board computer architecture is built around a VME backplane which can accept until 12 VME boards. The robot control is performed by a Motorola MVME 167 under VxWorks and a 6 axes custom-made servocontrol board. To solve critical time constraints at the image processing level, we have developped an original machine vision [11] characterized by its modularity and its real time capabilities. Its architecture is based both on VLSI chips for low level processing and on multi DSP's processors for more elaborated processings which allows to perform image processing at video rate.

As shown in figure (6), the mobile manipulator moves adequatly to reach the goal position indepently of the different starting configurations. Figure (5) shows the initial and the final view of the target corresponding to the desired location of the camera. Figure (7) shows the exponential convergency of the task function and the behaviour of the control inputs during the experiment.

5 Summary and future work

We proposed here a set of formalisms to build visual servoing tasks allowing to explicitly control mobile manipulators in closed loop with regard to their environment. Then, we validated our approach on our experimental testbed.

We can formalize simple sensor based tasks (e.g. following a wall, going in front of a door,...) in terms of *virtual linkage*. Therefore, this work has to be pursued in order to enlarge the set of elementary visual tasks available on our experimental system.

A fruitful way to implement more complex robotic tasks is currently investigated: visual events (e.g. reaching the end of a wall or detecting obstacle,...) are taken into account in a judicious way to sequence elementary tasks.

References

[1] R. W. Brockett :*Asymptotic Stablity and Feedback Stabilisation*, Geometric Control Theory, Progress in Math., vol. 27, Birkhauser, pp 181-208, 1983.

[2] C. Samson : *Time varying feedback stabilisation of a car like wheeled mobile robot*, The Int. Journal of Robotics Research, vol 12, n 1, pp 55-64, Feb. 1993.

[3] C. Canudas de Wit, H. Khennouf, C. Samson, O. J. Sordalen : *Recent trends in mobile robots* , pp 121-154, Editor Yuan F. Zheng, World Scientific Series in Robotics and Automated Systems, Vol 11, 1994.

[4] C. Samson, B. Espiau, M. Le Borgne: *Robot control: the task function approach*, Oxford University Press, 1990.

[5] F. Chaumette : *La relation vision-commande: théorie et applications à des tâches robotiques*, Ph-D Thesis, University of Rennes I, France, July 1990.

[6] Espiau B., Chaumette F., Rives P. : *A New Approach to Visual Servoing in Robotics*, IEEE Trans. on Robotics and Automation, Vol. 8, No. 3, June 1992

[7] G. Campion, G. Bastin, B. d'Andrea-Novel : *Structural properties and classification of kinematic and dynamic models*, Proceedings IEEE Int. Conf. on Robotics and Automation, Atlanta, May 1993.

[8] P. Rives, R. Pissard-Gibollet, K. Kapellos : *Development of a Reactive Mobile Robot Using Real Time Vision*, 3rd Int. Symp. on Experimental Robotics, Kyoto, Oct. 1993.

[9] R. Pissard-Gibollet : *Conception et Commande par Asservissement Visuel d'un Robot Mobile*, Ph-D Thesis, Ecole des Mines de Paris, December 1993.

[10] P. Rives, H. Michel : *Visual Servoing Based on Ellipse Features*, Proc. of the Spie Int. conf. on "Intelligent Robots and Computer Vision", pp 356-367, Boston, Sept. 1993.

[11] P. Rives, J.J. Borrelly, J. Gallice, P. Martinet : *A Versatile Parallel Architecture for Vision Based Applications*, Workshop on Computer Architecture for Machine Perception, New Orleans, Dec 15-17, 1993

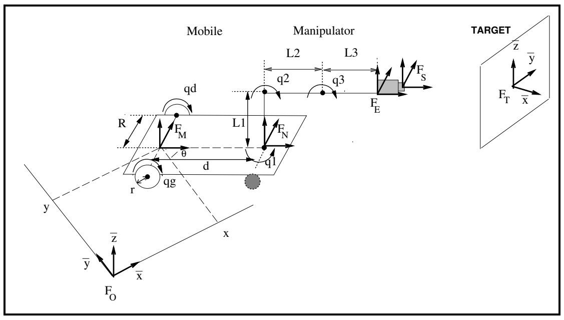

Figure 1: Frames and Notations

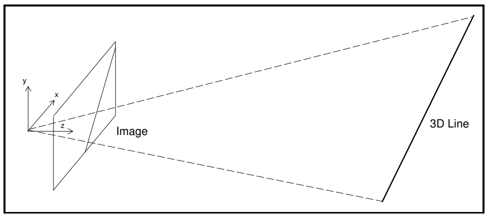

Figure 2: Virtual linkage

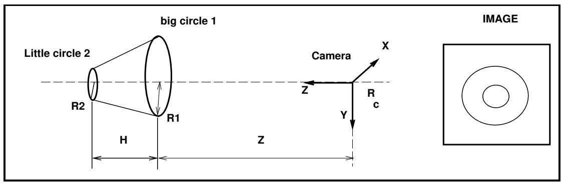

Figure 3: Nominal position

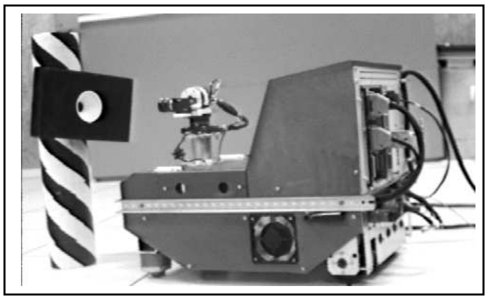

Figure 4: Task Positioning in front of a target

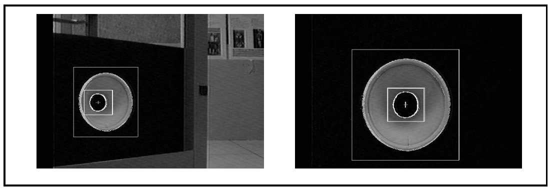

Figure 5: Initial and final view of the target

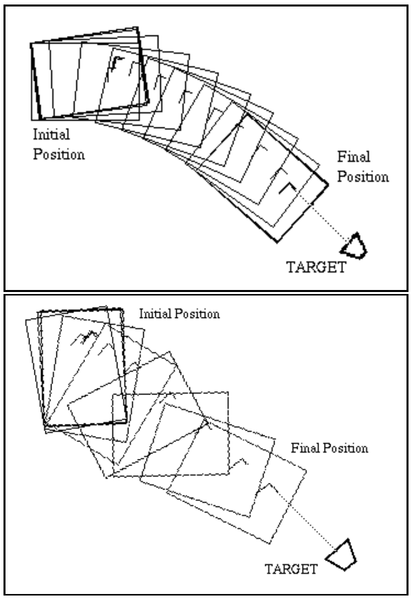

Figure 6: Robot Trajectories for two initial positions

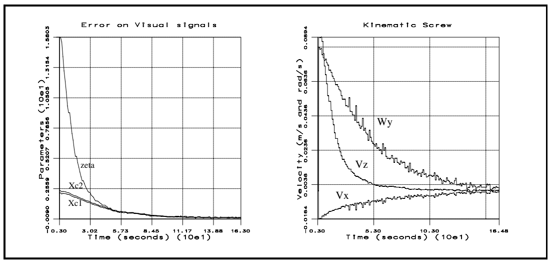

Figure 7: Error on visual signals and Kinematic screw