An Active Stereo Vision-Based Learning Approach for Robotic Tracking, Fixating and Grasping Control

Nan-Feng Xiao & Saeid Nahavandi

1. Introduction

Vision-based robotic tracking, fixating and grasping control depends on many environmental factors in an unknown environment. The robot control systems lack robustness, and the calibration of the CCD cameras is very slow and tedious in the existing methods.

Although the binocular cameras can solve some of these problems, it is necessary to rely on the time consuming and complicated 3-D reconstruction algorithms (8)-(9). Therefore, it is necessary to develop a more effective vision-based robotic tracking, fixating and grasping method, and use the robotic learning ability to improve the tracking, fixating and grasping in the unknown environment.

This chapter presents an active stereo vision-based learning approach for robotic tracking, fixating and grasping. First, the many-to-one functional mapping relationships are derived to describe the spatial representations of the object in the workspace frame. Then, ART_NN and FF_NN are used to learn the mapping relationships, so that the active stereo vision system guides the end-effecter to track, fixate and grasp the object without the complicated coordinate transformation and calibration. Finally, the present approach is verified by simulation.

2. Visual Tracking, Fixating and Grasping

Active vision can easily realize selective attention and prevent an object to go out of the view fields of the cameras, therefore the active stereo vision-based robotic tracking, fixating and grasping can achieve greater flexibility in an unknown environment.

Figure 1 shows an active stereo vision-based robotic system for the tracking, fixating and grasping. The CCD cameras have 5 DOF, the robot has 6 DOF, which constitute an 11 DOF tracking, fixating and grasping system.

Because the active CCD cameras and the robot can move independently or together, the active CCD cameras can observe freely an object in Σ_o.

According to the visual feed back information, the robot can track, fixate and grasp the object autonomously.

Source: Cutting Edge Robotics, ISBN 3-86611-038-3, pp. 784, ARS/plV, Germany, July 2005 Edited by: Kordic, V.; Lazinica, A. & Merdan, M.

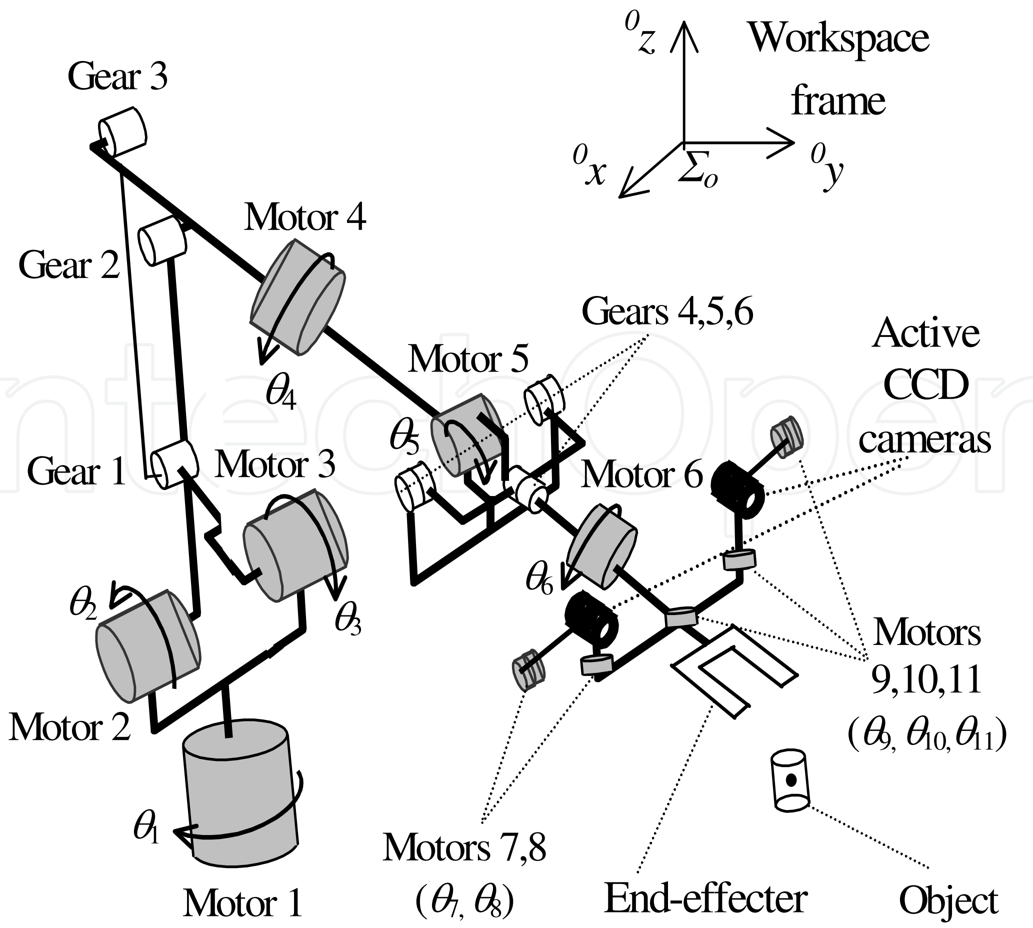


Figure 1. A robot system with active vision

3. Many-to-One Mapping Relationships

Figure 2 shows the projective relationships between the active stereo vision system and the object in Σ_o. Let p_i (i=1,2,3,…)be a feature point on the object. When the active stereo vision system tracks pi and its image coordinates are registered in the centers (o_l, o_r) of the left and right image planes of the two cameras respectively, the active stereo vision system is known as fixation on p_i.

Let $Q_i=[\theta_{i7},\theta_{i8,\ldots,}\theta_{i11}]^T$, $P_i=[x_{oi},\ y_{oi},\ z_{oi}]^T$ and $V_i=[{}^lu_{i1,}{}^lv_{i1,}{}^ru_{i1,}{}^rv_{i1}]^T$ be a joint angle vector of the active stereo vision system tracking p_i, a spatial representation vector of p_i in Σ_o and an image coordinate vector of p_i on the left and image planes, respectively. It is known from Fig. 2 that when p1 and p2 are visible to the CCD cameras, Q_2 and P_2 of p_2 identified by the active stereo vision system tracking on p_2 should be different from Q_1 and P_1. If another joint angle vector Q_3 is obtained by tracking p_3, p_1 and p_2 are still visible. P1 and P2 are not changed from that obtained by tracking on p_1 and p_2, respectively, despite the image coordinate vectors V1 and V2 change on the image planes of the CCD cameras. Therefore, there exist many combinations of Q_i and V_i which correspond to the same P_i, which means that P_i is invariant to the changing Q_i and V_i.

According to the projective geometry (7), Vi can be expressed as follows:

$$V_i=\varphi(Q_i,p_i),\quad (i=1,2,3,\ldots), \tag{1}$$

where φ is a nonlinear projective function which maps the object and the joint angles on the left and right image planes of the CCD cameras. Therefore, Pi is specified as

$$P_i=\psi(V_i,Q_i),\quad (i=1,2,3,\ldots), \tag{2}$$

where ψ is a nonlinear many-to-one mapping function, which denotes that the combinations of Qi and Vi correspond to Pi. On the other hand, it is known from Fig.2 that any combination of Qi and Vi should map to the same Pi, because pi is stationary feature.

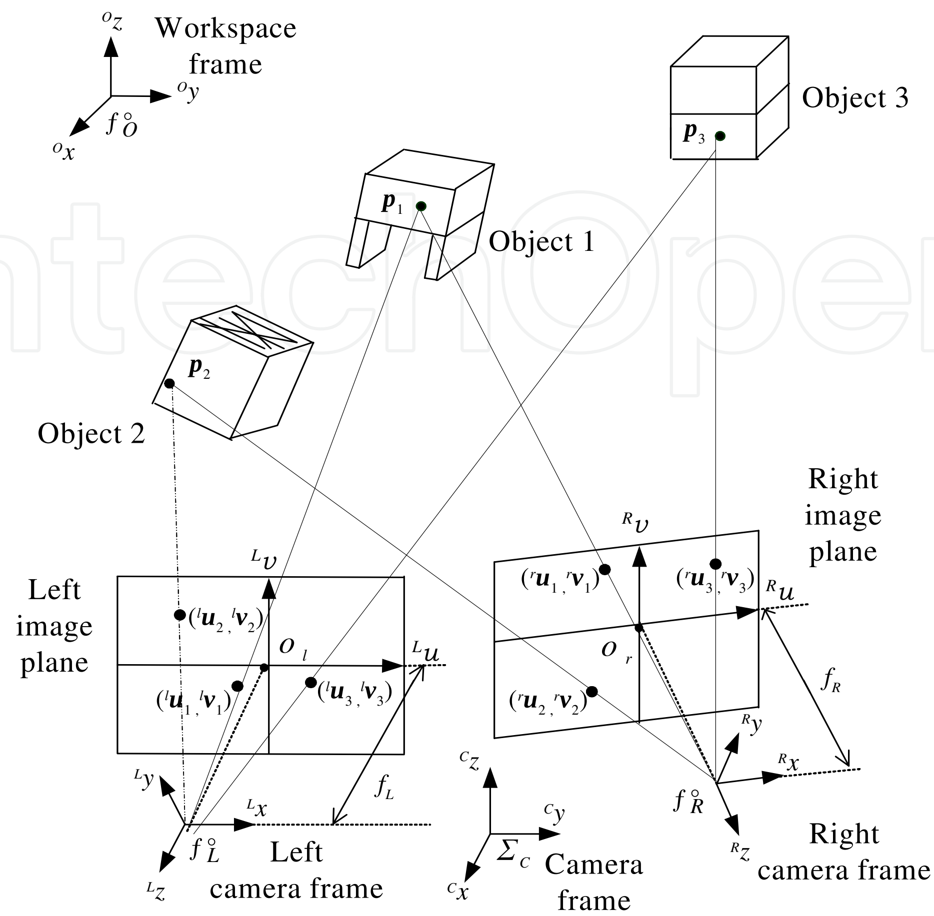


Figure 2. Projection and fixation relationships

4. Active Vision-Based Robot Control

4.1 Tracking and Fixating Control

It is known from Fig. 2 that when the active stereo vision system tracks pi, Vi is obtained for pi and Qi, we have from Equation (2),

$$P_i=\psi(V_i,Q_i), \tag{3}$$

where the active stereo vision system tracks p_i**.** When the active stereo vision system fixates p_i and the coordinate vector V_{Oi} of p_i corresponds to the centers (o_l, o_r) of the image planes, then the desired joint angle vector Q_{Oi} which is necessary to bring V_{Oi} to V_{Oi} can be computed as follows:

$$P_i=\psi(V_{Oi},Q_{Oi}), \tag{4.a}$$

$$\text{or}\quad Q_{Oi}=\psi^{-1}(V_{Oi},P_i), \tag{4.b}$$

where ψ and ψ^{-1} are invertible functions which can be used to control the fixation on p_i, respectively. Because Pi has invariance, Q_{Oi} can be computed by combining Equation (3) with Equation (4),

$$Q_{Oi}=\psi^{-1}[V_{Oi},\psi(V_i,Q_i)], \tag{5}$$

therefore, Q_{Oi} is used to control the active vision system to fixate p_i**.**

4.2 Grasping Control

Figure 3 shows the configuration parameters of the active stereo vision system. Let d, s, l be the distance, height of the CCD cameras and diameter of the sphere coordinate system. Let $\theta_{i\alpha}$, $\theta_{i\beta}$, $\theta_{i\gamma}$ be the configuration angles of the active stereo vision system and the spatial coordinates of p_i in Σ_C be ${}^{C}P_i=[x_{ci}, y_{ci}, z_{ci}]^T$, respectively. When the active stereo vision system fixates p_i, ${}^{C}P_i$ can be computed by the triangular geometry relationships in Fig. 3

$$x_{ci} = l\cos(\theta_{i\gamma})\sin(\theta_{i\beta}), \tag{6.a}$$

$$y_{ci} = l\cos(\theta_{i\gamma})\cos(\theta_{i\beta}), \tag{6.b}$$

$$z_{ci} = l\sin(\theta_{i\gamma}). \tag{6.c}$$

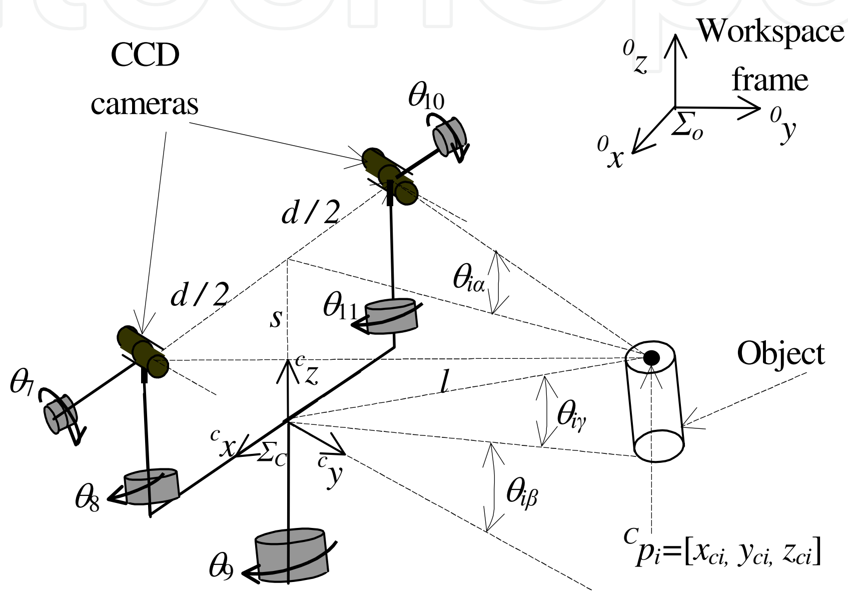


Figure 3. Configuration of the vision system

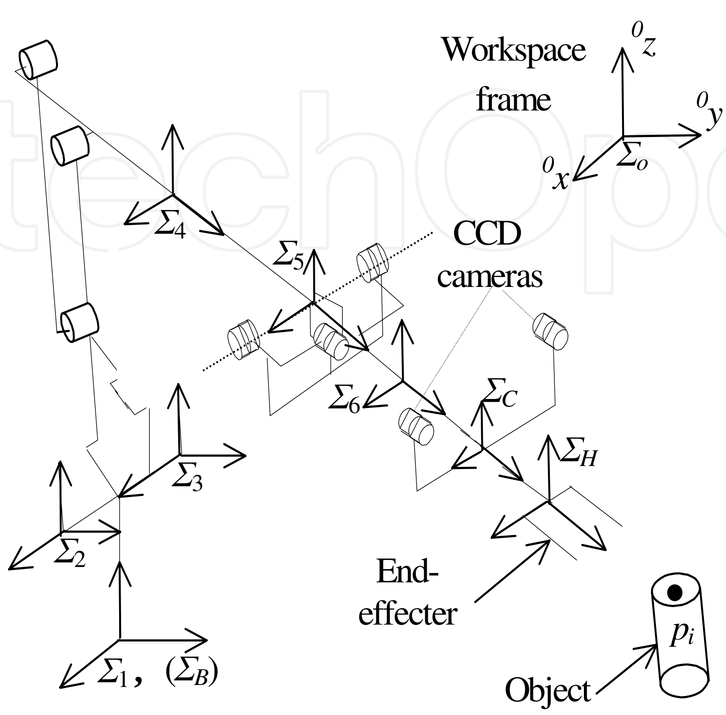


Figure 4. Frames of the active vision system

Figure 4 gives the joint coordinate frames of the robot joint system. In Fig. 4, let Σ_j(j= 1,2,…,6), Σ_B, Σ_H, Σ_C be a coordinate frame of ith robotic joint, base frame, coordinate frame of the end-effecter, camera frame, ${}^BP_i=[x_{bi}, y_{bi}, z_{bi}]^T$, ${}^HP_i=[x_{hi}, y_{hi}, z_{hi}]^T$, $P_i=[x_{oi}, y_{oi}, z_{oi}]^T$ be an Euclidian coordinate vector of p_i in Σ_B, Σ_H, Σ_O, respectively.
By homogeneous transformation relationship, ${}^Bp_i$ can be specified by

$$ {}^HP_i={}^HH_C \bullet {}^CP_i, \tag{7.a}$$

$$ {}^BP_i={}^BH_6 \bullet {}^6H_H \bullet {}^HP_i, \tag{7.b}$$

$$ {}^OP_i={}^OH_B \bullet {}^BP_i. \tag{7.c}$$

where ${}^HH_C$, ${}^6H_H$, ${}^BH_6$, ${}^OH_B$ are the homogeneous matrixes from Σ_C to Σ_H, Σ_H to Σ_6, Σ_6 to Σ_B, Σ_B to Σ_O , respectively.
According to the robotic forward kinematics $\Lambda[\theta_r(t)]\quad R^{6\times1}$, we obtain

$$ {}^OP_{Hi}=\Lambda[\theta_r(t)], \tag{8}$$

$$ J[\theta_r(t)]=\partial\Lambda[\theta_r(t)]/\partial\theta_r(t), \tag{9}$$

$$ {}^O\dot{P}_{Hi}=J[\theta_r(t)]\dot{\theta}_r(t), \tag{10}$$

where ${}^OP_{Hi}$ is the original coordinates of Σ_H in Σ_O, $J[\theta_r(t)]\quad R^{6\times1}$ is a Jacobian matrix of the end-effecter, $\theta_r(t)\quad R^{6\times1}$ is a reference joint angle vector of the end-effecter. Therefore, we have

$$ \dot{\theta}_r(t)=J^{-1}[\theta_r(t)]\bullet {}^O\dot{P}_{Hi}. \tag{11}$$

When the sampling period of the robot joint control system T is every minute, it is suitable that using $\dot{\theta}_r(k)=[\theta_r(k+1)-\theta_r(k)]/T$ to replace $\dot{\theta}_r(k)$ at time t=kT. Therefore, Equation (11) is discreted by

$$ [\theta_r(k+1)-\theta_r(k)]/T = J^{-1}[\theta_r(k)]\bullet {}^O\dot{P}_{Hi}, \tag{12.a}$$

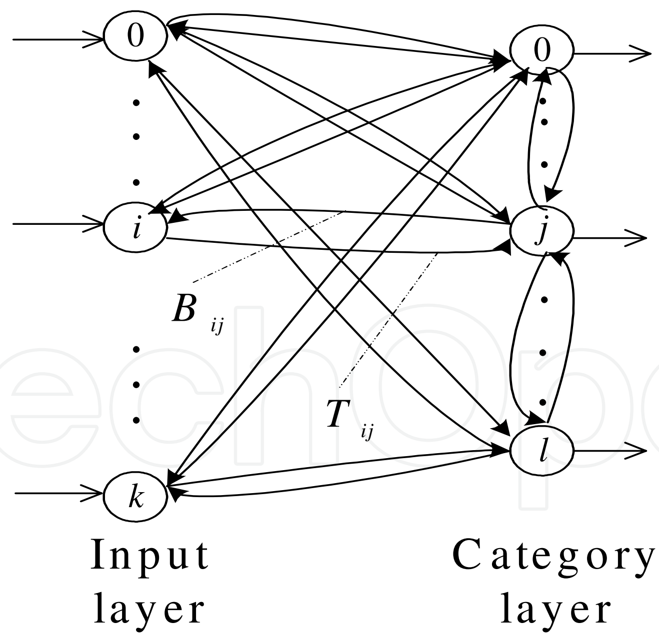
$$ \text{or } \theta_r(k+1)= \theta_r(k)+TJ^{-1}[\theta_r(k)]\bullet {}^O\dot{P}_{Hi}, \tag{12.b}$$

where ${}^O\dot{P}_{Hi}(k)=[{}^oP_{Hi}(k)-{}^oP_{Hi}(k-1)]/T$, $\theta_r(k+1)$ can be used to control the robot joint angles. When ${}^OP_{Hi}={}^OP_i$, the end-effecter can grasp the object.

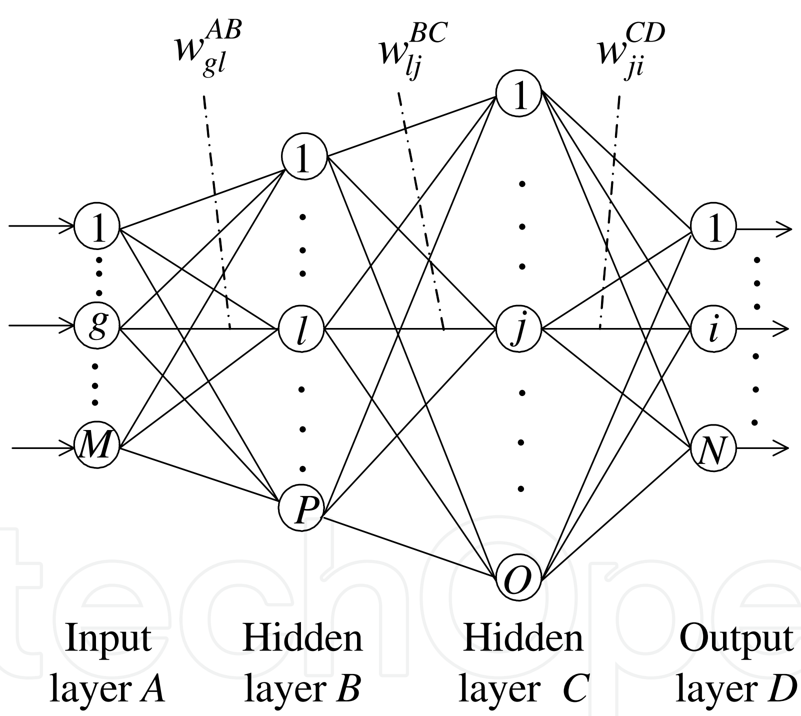
5. Visual Robot Learning Control System

5.1 Visual Learning Control System

In order to obtain the nonlinear many-to-one mapping function, ART_NN are combined with FF_NN to learn ψ defined in Equation (3). The architecture of ART_NN, FF_NN and the vision-based robot control system based on ART_NN and FF_NN are showed in Figs. 5 and 6, respectively.



(a) Architecture of ART_NN



(b) Architecture of FF_NN

Figure 5. Architectures of ART_NN and FF_NN

5.2 Learning of ART_NN and FF_NN

In Fig. 5, T_{ij}, B_{ij}, w_{gl}^{AB}, w_{lj}^{BC} and w_{ji}^{CD} are weights. The self-adaptive resonance algorithms for ART_NN and the learning algorithms for FF _NN are omitted. In Fig. 6, the ART_NN require two types of inputs V_i and Q_i, where V_i corresponds to the image coordinates of p_i on the image planes of the CCD cameras, and Q_i is the joint angle coordinates corresponding to the CCD cameras tracking p_i. The ART_NN clusters V_i into classes within the category layer. The class number in each category layer depends on a vigilant parameter which is a real number between zero and one.

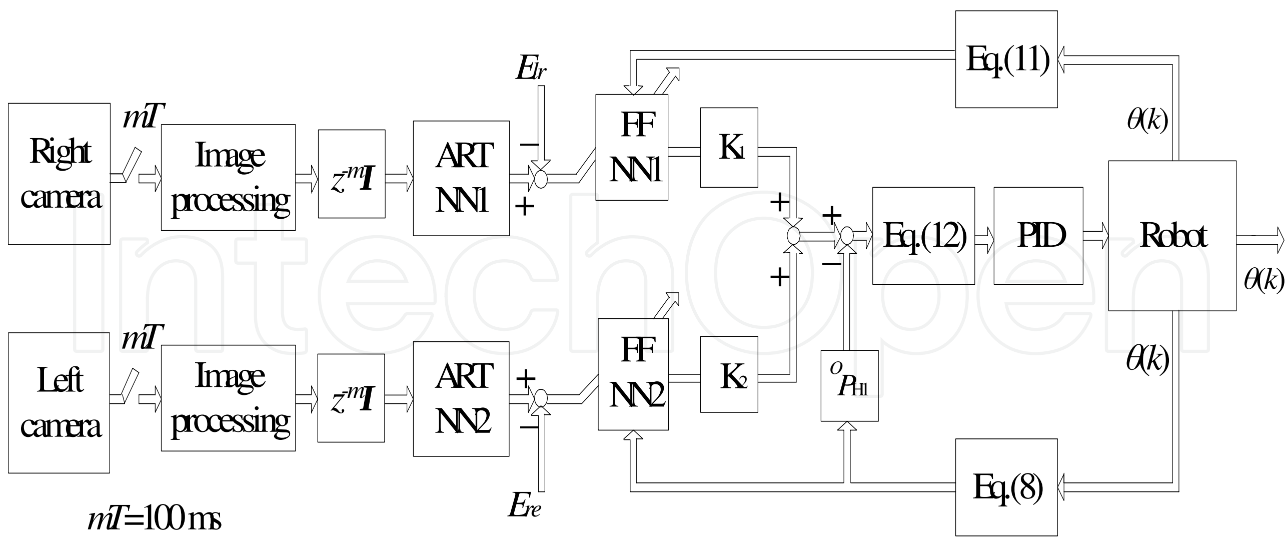


Figure 6. A Robotic Learning Control System

In Fig. 6, $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$?$R^{6\times6}$ are the coefficient matrixes which were specified empirically. $\boldsymbol{E}_{le}$ and $\boldsymbol{E}_{re}$?$R^{6\times6}$ are the differences between the two learning trials, respectively, and the PID controller is used to obtain the joint servoing control with high accuracy.

6. Simulations

To evaluate the validity of the active stereo vision-based robotic tracking, fixating and grasping in the unknown environment, the simulations are carried out using the models of the active stereo vision system installed in the end-effecter.
For controlling the robot to track, fixate and grasp the object, first, ART_NN1 and ART_NN2, FF_NN1 and FF_NN2 learn the many- to-one functional mapping relationships by generating 10000 random pairs of V_i and Q_i signals corresponding to p_i. The ART_NN1 created 500 classes for the inputs from the right CCD camera, and the ART_NN2 also created 500 classes for the inputs from the left CCD camera.
The spatial coordinates of pi are computed by using in the tracking, fixating and grasping control loop.
The simulation results denote that the errors for all of the three components of the spatial representation converged to within 2% of its dynamic range. These results show that the learning of ART_NN and FF_NN is fast, convergent and the end-effecter can also arrive at the position of the object.

7. Conclusions

The following conclusions are drawn from the above experiments:
(1) There exist many-to-one mapping relationships between the joint angles of the active stereo vision system and the spatial representations of the object in the workspace frame.
(2) ART_NN and FF_NN can learn the mapping relationships in an invariant manner to the changing joint angles. The vision and joint angle signals of the active vision system corresponding to the object correspond to the same spatial representation of the object.

(3) The present approach was evaluated by simulation using the models of an 11 DOF active stereo vision system and the simulation confirms that the present approach has high robustness and stability.

8. Acknowledgement

This research was funded by the initial foundation of China Education Ministry for Scholars Returned China from abroad (project no.: 2002247) and the Intelligent Systems research initiative of Deakin university.

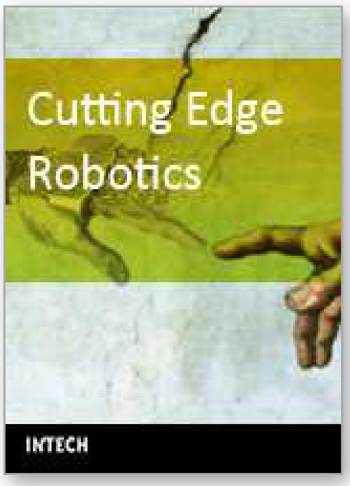
9. References

Bernardino A., Santos-Victor, Bin- ocular Tracking: Integrating Perception and Control, IEEE Transactions on Robotics and Automation, (1999-12), Vol. 15, No. 6, pp. 1080 – 1093.

Sumi, Kawai, Yoshimi, 3-D Object Recognition in Cluttered Environments by Segment-Based Stereo Vision, International Journal of Computer Vision, (2002-1), Vol. 46, No. 1, pp. 5-23.

Srinivasa N., Rajeev S., SOIM: A Self-Organizing Invertible Map with Applications in Active Vision, IEEE Transactions on Neural Net- works, (1997-5), Vol. 8, No. 3, pp. 758-773.

Barnes N.M., Liu Z.Q., Vision guided circumnavigating autonomous robots, *International Journal of Pattern Recognition and Artificial Intelligence*, (2000), Vol. 14, No. 6, pp. 689-714.



Cutting Edge Robotics

Edited by Vedran Kordic, Aleksandar Lazinica and Munir Merdan

ISBN 3-86611-038-3
Hard cover, 784 pages
Publisher Pro Literatur Verlag, Germany
Published online 01, July, 2005
Published in print edition July, 2005

This book is the result of inspirations and contributions from many researchers worldwide. It presents a collection of wide range research results of robotics scientific community. Various aspects of current research in robotics area are explored and discussed. The book begins with researches in robot modelling & design, in which different approaches in kinematical, dynamical and other design issues of mobile robots are discussed. Second chapter deals with various sensor systems, but the major part of the chapter is devoted to robotic vision systems. Chapter III is devoted to robot navigation and presents different navigation architectures. The chapter IV is devoted to research on adaptive and learning systems in mobile robots area. The chapter V speaks about different application areas of multi-robot systems. Other emerging field is discussed in chapter VI - the human- robot interaction. Chapter VII gives a great tutorial on legged robot systems and one research overview on design of a humanoid robot.The different examples of service robots are showed in chapter VIII. Chapter IX is oriented to industrial robots, i.e. robot manipulators. Different mechatronic systems oriented on robotics are explored in the last chapter of the book.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821