

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,800

Open access books available

142,000

International authors and editors

180M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# 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-effector 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  $\Sigma_o$ .

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

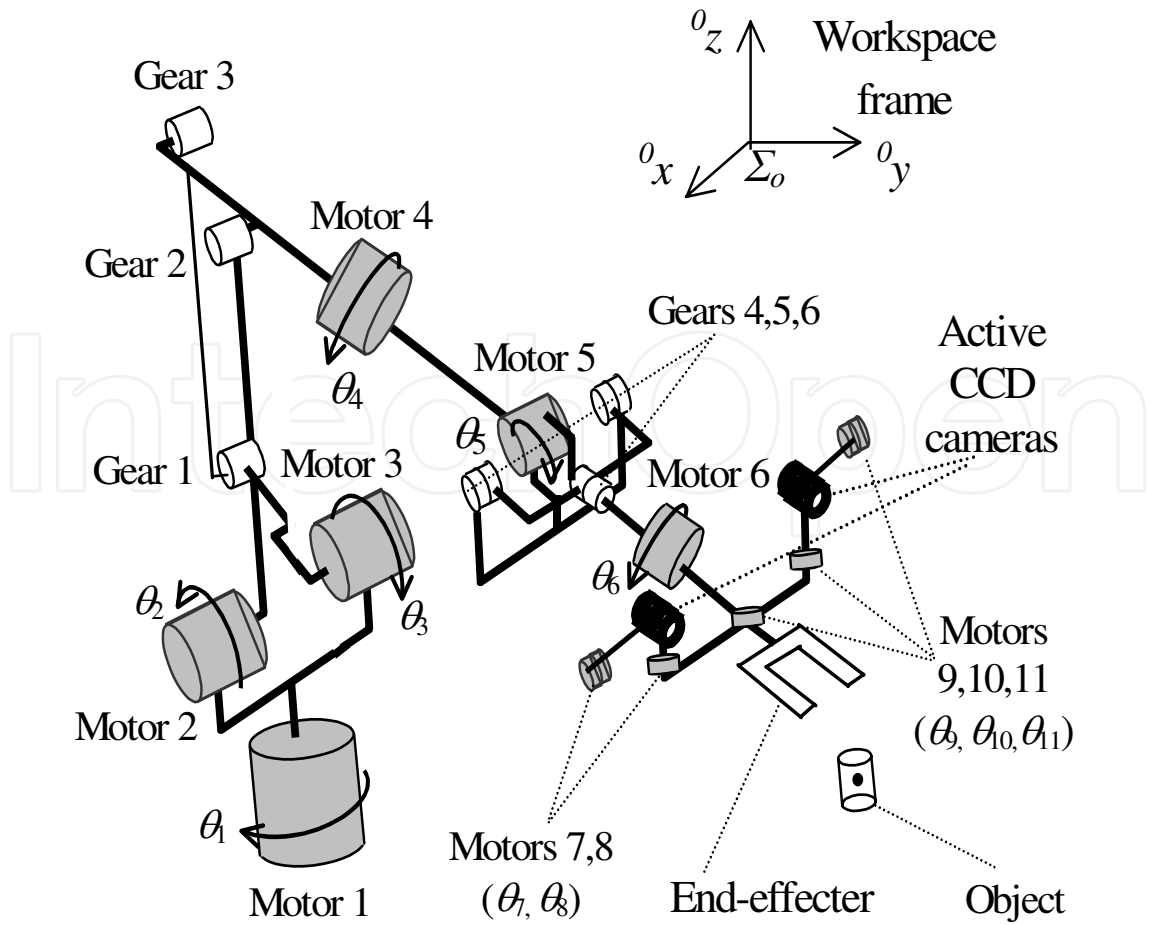


Figure 1. A robot system with active vision

### 3. Many-to-One Mapping Relationships

Figure 2 shows the projective relationships between the active stereo vision system and the object in  $\Sigma_o$ . Let  $p_i$  ( $i=1,2,3,\dots$ ) be a feature point on the object. When the active stereo vision system tracks  $p_i$  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  $Q_i=[\theta_{i7},\theta_{i8},\dots,\theta_{i11}]^T$ ,  $P_i=[x_{oi}, y_{oi}, z_{oi}]^T$  and  $V_i=[{}^l u_{i1}, {}^l v_{i1}, {}^r u_{i1}, {}^r v_{i1}]^T$  be a joint angle vector of the active stereo vision system tracking  $p_i$ , a spatial representation vector of  $p_i$  in  $\Sigma_o$  and an image coordinate vector of  $p_i$  on the left and image planes, respectively. It is known from Fig. 2 that when  $p_1$  and  $p_2$  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.  $P_1$  and  $P_2$  are not changed from that obtained by tracking on  $p_1$  and  $p_2$ , respectively, despite the image coordinate vectors  $V_1$  and  $V_2$  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),  $V_i$  can be expressed as follows:

$$V_i=\varphi(Q_i,p_i), \quad (i=1,2,3,\dots), \quad (1)$$

where  $\varphi$  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,  $P_i$  is specified as

$$P_i=\psi(V_i,Q_i), \quad (i=1,2,3,\dots), \quad (2)$$

where  $\psi$  is a nonlinear many-to-one mapping function, which denotes that the combinations of  $Q_i$  and  $V_i$  correspond to  $P_i$ . On the other hand, it is known from Fig.2 that any combination of  $Q_i$  and  $V_i$  should map to the same  $P_i$ , because  $p_i$  is stationary feature.

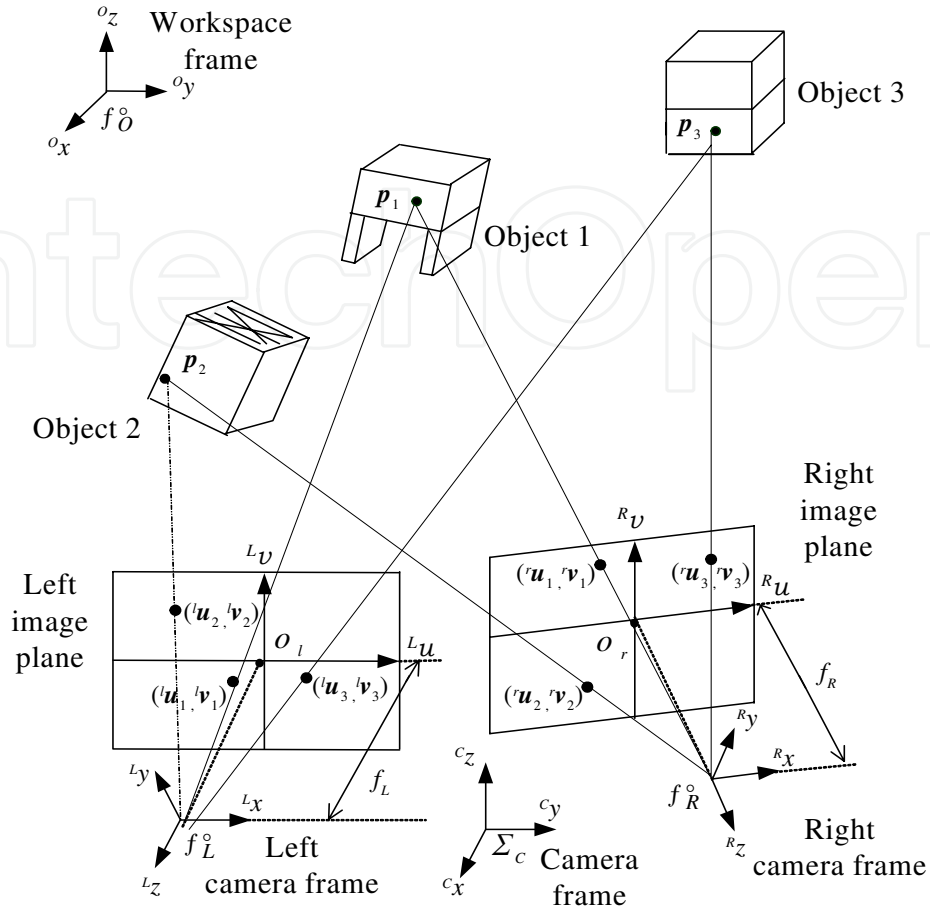


Figure 2. Projection and fixation relationships

## 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  $p_i$ ,  $V_i$  is obtained for  $p_i$  and  $Q_i$ , we have from Equation (2),

$$P_i = \psi(V_i, Q_i), \quad (3)$$

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

$$P_i = \psi(V_{O_i}, Q_{O_i}), \quad (4.a)$$

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

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

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

therefore,  $Q_{O_i}$  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  $\Sigma_C$  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

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

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

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

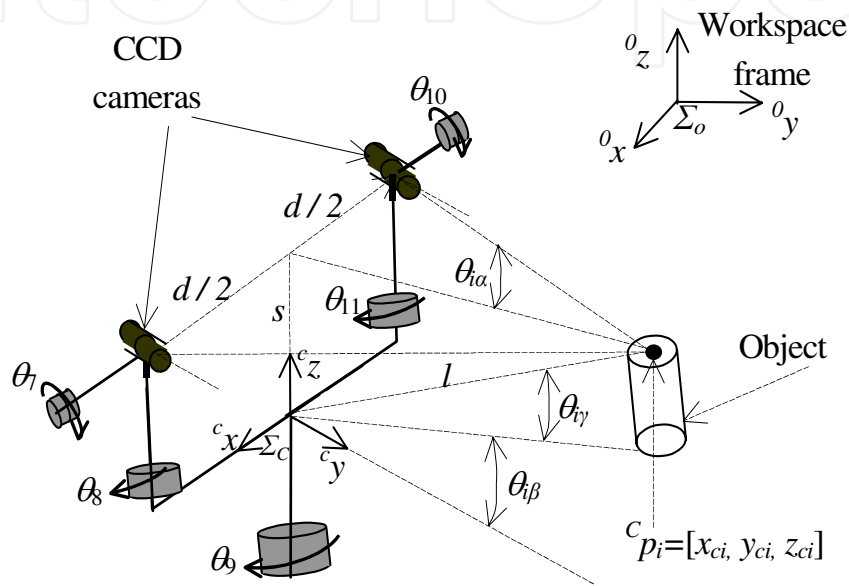


Figure 3. Configuration of the vision system

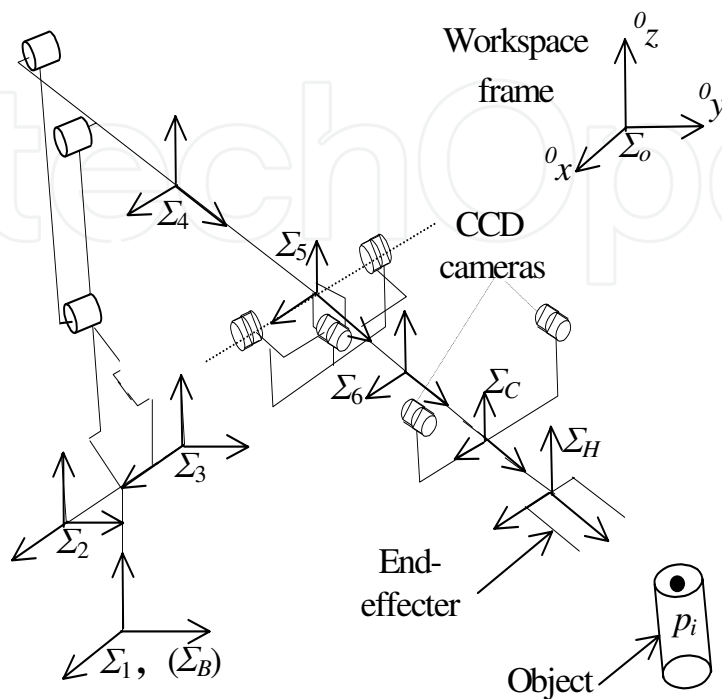


Figure 4. Frames of the active vision system

Figure 4 gives the joint coordinate frames of the robot joint system. In Fig. 4, let  $\Sigma_j(j=1,2,\dots,6)$ ,  $\Sigma_B$ ,  $\Sigma_H$ ,  $\Sigma_C$  be a coordinate frame of  $i$ th robotic joint, base frame, coordinate frame of the end-effector, camera frame,  ${}^B P_i = [x_{bi}, y_{bi}, z_{bi}]^T$ ,  ${}^H P_i = [x_{hi}, y_{hi}, z_{hi}]^T$ ,  $P_i = [x_{oi}, y_{oi}, z_{oi}]^T$  be an Euclidian coordinate vector of  $p_i$  in  $\Sigma_B$ ,  $\Sigma_H$ ,  $\Sigma_O$ , respectively.

By homogeneous transformation relationship,  ${}^B p_i$  can be specified by

$${}^H P_i = {}^H H_C \bullet {}^C P_i, \quad (7.a)$$

$${}^B P_i = {}^B H_6 \bullet {}^6 H_H \bullet {}^H P_i, \quad (7.b)$$

$${}^O P_i = {}^O H_B \bullet {}^B P_i. \quad (7.c)$$

where  ${}^H H_C$ ,  ${}^6 H_H$ ,  ${}^B H_6$ ,  ${}^O H_B$  are the homogeneous matrixes from  $\Sigma_C$  to  $\Sigma_H$ ,  $\Sigma_H$  to  $\Sigma_6$ ,  $\Sigma_6$  to  $\Sigma_B$ ,  $\Sigma_B$  to  $\Sigma_O$ , respectively.

According to the robotic forward kinematics  $A[\theta_r(t)] \in R^{6 \times 1}$ , we obtain

$${}^O P_{Hi} = A[\theta_r(t)], \quad (8)$$

$$J[\theta_r(t)] = \partial A[\theta_r(t)] / \partial \theta_r(t), \quad (9)$$

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

where  ${}^O P_{Hi}$  is the original coordinates of  $\Sigma_H$  in  $\Sigma_O$ ,  $J[\theta_r(t)] \in R^{6 \times 1}$  is a Jacobian matrix of the end-effector,  $\theta_r(t) \in R^{6 \times 1}$  is a reference joint angle vector of the end-effector. Therefore, we have

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

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

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

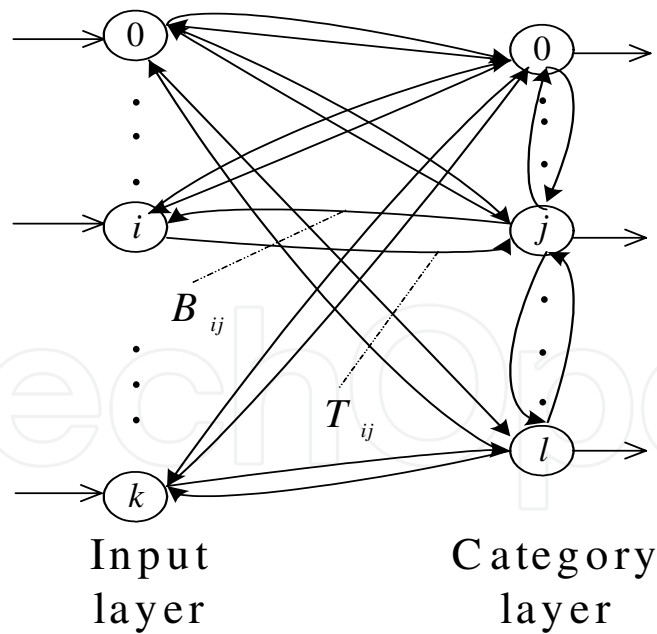
$$\text{or } \theta_r(k+1) = \theta_r(k) + T J^{-1}[\theta_r(k)] \bullet {}^O \dot{P}_{Hi}, \quad (12.b)$$

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

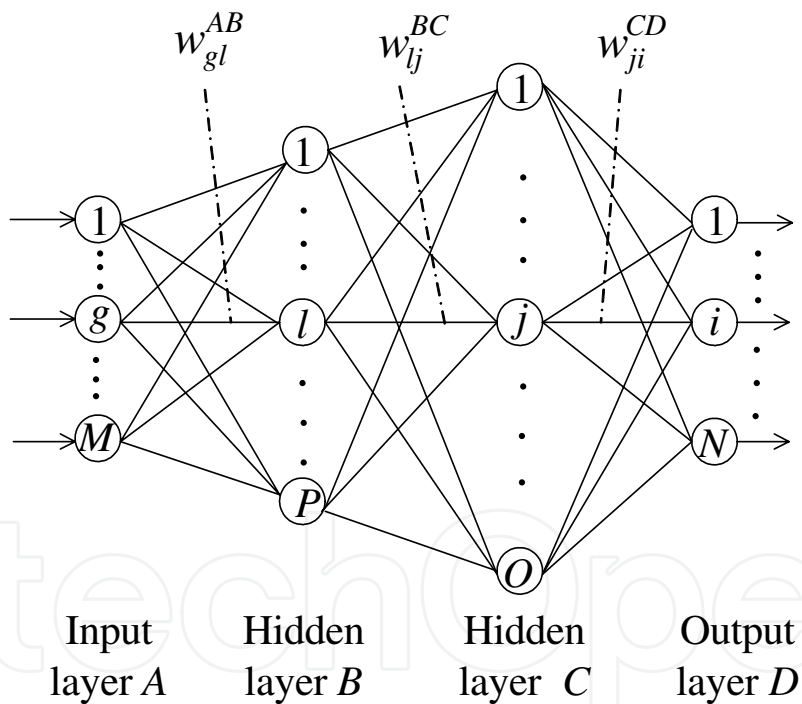
## 5. Visual Robot Learning Control System

### 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  $\psi$  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.

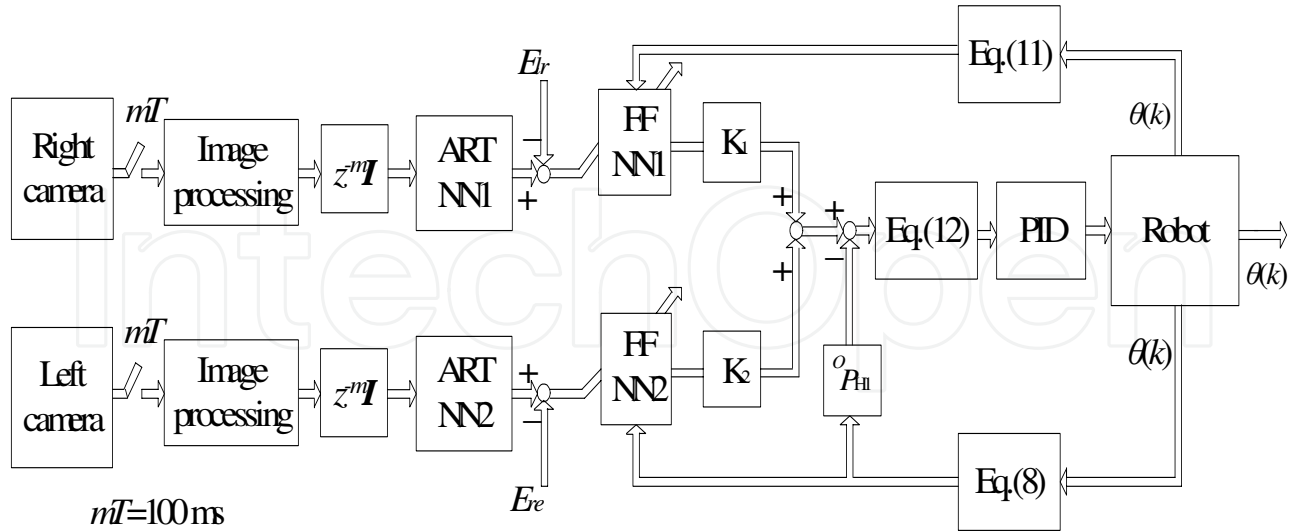


Figure 6. A Robotic Learning Control System

In Fig. 6,  $K_1$  and  $K_2 \in R^{6 \times 6}$  are the coefficient matrixes which were specified empirically.  $E_{le}$  and  $E_{re} \in R^{6 \times 6}$  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-effector.

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  $p_i$  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-effector 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 Networks*, (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.

#### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Nan-Feng Xiao and Saeid Nahavandi (2005). An Active Stereo Vision-Based Learning Approach for Robotic Tracking, Fixating and Grasping Control, Cutting Edge Robotics, Vedran Kordic, Aleksandar Lazinica and Munir Merdan (Ed.), ISBN: 3-86611-038-3, InTech, Available from:

[http://www.intechopen.com/books/cutting\\_edge\\_robotics/an\\_active\\_stereo\\_vision-based\\_learning\\_approach\\_for\\_robotic\\_tracking\\_fixating\\_and\\_grasping\\_control](http://www.intechopen.com/books/cutting_edge_robotics/an_active_stereo_vision-based_learning_approach_for_robotic_tracking_fixating_and_grasping_control)

**INTECH**  
open science | open minds

#### **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](http://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

© 2005 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen