

Fusion of Multi-Vision of Industrial Robot in MAS-Based Smart Space

Hexi Li ^{1,a}, Juanjuan Zhang ¹

¹ School of Computer Science, Wuyi University, Jiangmen, Guangdong Province, China

Abstract. The paper presents a fusion method of multi-vision of industrial robot in a smart space based on multi-agent system(MAS), the robotic multi-vision consists of top-view, side-view, front-view and hand-eye cameras, the moving hand-eye provide vision guidance and give the estimation of robot position, other three cameras are used for target recognition and positioning. Each camera is connected to an agent based on an image-processing computer that aims at analyzing image rapidly and satisfying the real-time requirement of data processing. As a learning strategy of robotic vision, a back-propagation neural network(BPNN) with 3-layer-architecture is first constructed for each agent and is independently trained as a classifier of target recognition using batch gradient descent method based on the region features extracted from the images of target samples(typical mechanical parts), and then the outputs of trained BPNNs in MAS-based smart space are fused with Dempster-Shafer evidence theory to form a final recognition decision, the experimental results of typical mechanical parts show that fusion of multi-vision can improve the robotic vision accuracy and MAS-based smart space will contribute to the parallel processing of immense image data in robotic multi-vision system.

1 INTRODUCTION

At present, the majority of industrial robot still rely on manual teaching in practical application, adding vision to robot is becoming a trend of intelligent robot, and the robotic vision with monocular or binocular camera has widely used in intelligent robot system, whereas one or two cameras can not deal with complex vision task such as workpiece recognition and positioning, visual servoing, robot tracking, multi-robot cooperating. Using multi-cameras is an effective method to deal with the complex vision task of industrial robot[1-3], and the subsequent problem of multi-vision is immense image data processing that may be not undertaken by one computer. In this paper, a multi-agent system(MAS) based on computer network is adopted for the immense image data processing of multi-vision task, in which each agent is a computer connected to a camera with independent image processing capability, the MAS forms a smart space surrounding the robot and provides the parallel processing of multi-vision information, there are some models for the communication and cooperation of multi-agents in MAS[4-5], a blackboard system of artificial intelligence is used for this purpose, and we put emphasis on the learning and fusing strategy of multi-vision in industrial robot application.

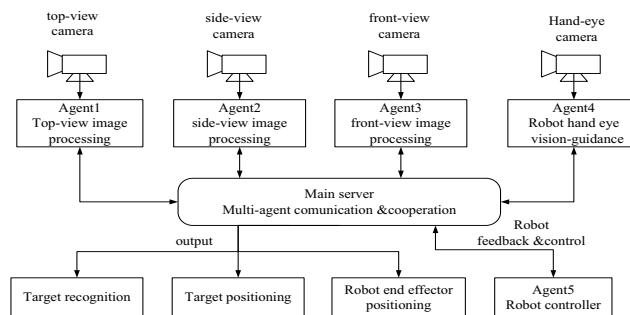


Fig.1 multi-vision of industrial robot in MAS-based smart space

2 Learning Strategy of Multi-Agent Based on BPNN

It is necessary to design a learning strategy for each agent oriented to robotic vision, BPNN has been widely used for many years since its adaptability and flexibility[6-7], therefore it is chosen as the learning strategy of agents in MAS-based robotic vision system in this paper. As an input of BPNN, each image feature can be characterized by converting its pixel matrix into a vector, resizing its normal sizes to small sizes for reducing the dimensionality of feature space. A typical BPNN structure is as follows:

^a Corresponding author: jmlihexi@163.com

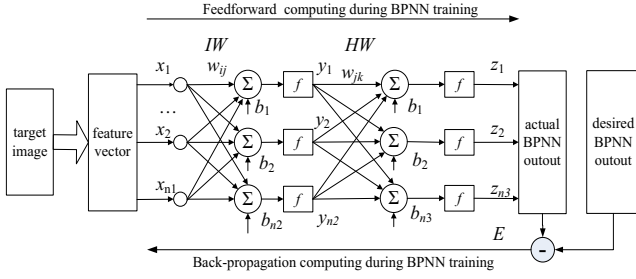


Fig. 2 BPNN architecture of robotic vision

For each BPNN, an input vector of input layer, $x=\{x_1, x_2, \dots, x_{n1}\}^T$, stands for a target image which piles up along column direction, $y=\{y_1, y_2, \dots, y_{n2}\}^T$ is an output vector of hidden layer, $z=\{z_1, z_2, \dots, z_{n3}\}^T$ is an output vector of the BPNN, subscript n_1, n_2 and n_3 stand for the neuron number of input, hidden and output layer respectively. IW is a matrix of weight between input layer and hidden layer, HW is a matrix of weight between hidden layer and output layer. In order to improve learning efficiency, the batch training with N target images is used during BPNN training, and all computations of BPNN is completed with matrix operation instead of element operation. Training procedures of BPNN for target image recognition is as follows:

2.1 Feedforward computing

All columns of each target image are piled up into a vector X , N images of different target are arranged to form a matrix $X_{n_1 \times N}$, the matrix will be extended to a new matrix $X_{n_1+1 \times N}$ with first row of -1 value for the sake of threshold of neurons, the sum of weighted X is given by:

$$v_1 = IW_{n_2 \times n_1+1} \times X_{n_1+1 \times N}$$

$$= \begin{bmatrix} v_{11} & \dots & v_{1N} \\ v_{21} & \dots & v_{2N} \\ \vdots & \dots & \vdots \\ v_{n_2+1} & \dots & v_{n_2N} \end{bmatrix} = \begin{bmatrix} iw_{11} & \dots & iw_{1n_1+1} \\ iw_{21} & \dots & iw_{2n_1+1} \\ \vdots & \dots & \vdots \\ iw_{n_2+1} & \dots & iw_{n_2n_1+1} \end{bmatrix} \begin{bmatrix} -1 & \dots & -1 \\ x_{11} & \dots & x_{1N} \\ \vdots & \dots & \vdots \\ x_{n_11} & \dots & x_{n_1N} \end{bmatrix} \quad (1)$$

The output Matrix Y of hidden layer is computed with transfer function $f=1/[1+\exp(-v)]$ as follows:

$$Y_{n_2 \times N} = \frac{1}{1 + \exp(-v_1)} = \begin{bmatrix} y_{11} & \dots & y_{1N} \\ y_{21} & \dots & y_{2N} \\ \vdots & \dots & \vdots \\ y_{n_21} & \dots & y_{n_2N} \end{bmatrix} \quad (2)$$

Similar to the hidden layer, for the output layer we have

$$v_2 = HW_{n_3 \times n_2+1} \times Y_{n_2+1 \times N}$$

$$= \begin{bmatrix} v_{11} & \dots & v_{1N} \\ v_{21} & \dots & v_{2N} \\ \vdots & \dots & \vdots \\ v_{n_3+1} & \dots & v_{n_3N} \end{bmatrix} = \begin{bmatrix} hw_{11} & \dots & hw_{1n_2+1} \\ hw_{21} & \dots & hw_{2n_2+1} \\ \vdots & \dots & \vdots \\ hw_{n_3+1} & \dots & hw_{n_3n_2+1} \end{bmatrix} \begin{bmatrix} -1 & \dots & -1 \\ y_{11} & \dots & y_{1N} \\ \vdots & \dots & \vdots \\ y_{n_21} & \dots & y_{n_2N} \end{bmatrix} \quad (3)$$

The output Matrix Z of output layer are computed with transfer function $f=1/(1+\exp^{-v})$ as follows.

$$Z_{n_3 \times N} = \frac{1}{1 + \exp(-v_2)} = \begin{bmatrix} z_{11} & \dots & z_{1N} \\ z_{21} & \dots & z_{2N} \\ \vdots & \dots & \vdots \\ z_{n_31} & \dots & z_{n_3N} \end{bmatrix} \quad (4)$$

2.2 Back-propagation computing

Let D stands for a desired output matrix of the network, from feedforward computing result, the error of the BPNN output is given by:

$$E = Z - D \quad (5)$$

The local gradient of output layer δZ can be computed by Eq.6, here denotation $*$ stands for the element-by-element product of the matrices.

$$\delta Z = E * Z * (1 - Z) \quad (6)$$

Deleting first column, the weight matrix of hidden layer, HW , is changed into

$$HW_1 = \begin{bmatrix} hw_{12} & hw_{13} & \dots & hw_{1n_2+1} \\ hw_{22} & hw_{23} & \dots & hw_{2n_2+1} \\ \dots & \dots & \dots & \dots \\ hw_{n_22} & hw_{n_23} & \dots & hw_{n_2n_2+1} \end{bmatrix} \quad (7)$$

Local gradient of hidden layer δY can be given by:

$$\delta Y = Y * (1 - Y) * (HW_1^T * \delta Z) \quad (8)$$

Update the weight of hidden layer using batch gradient descent method with learning rate r as follows:

$$HW = HW + (r/N) * (\delta Z * Y^T) \quad (9)$$

Update the weight of input layer using batch gradient descent method with learning rate r as follows:

$$IW = IW + (r/N) * (\delta Y * X^T) \quad (10)$$

2.3 Judgement of BPNN training end

The sum of squared errors between the actual output of the BPNN and desired output is computed using Eq. 11 and is used to evaluate training performance of BPNN based on batch gradient descent method.

$$E = \frac{1}{2N} \sum_{j=1}^N \sum_{i=1}^{n_3} e_j^2 \quad (11)$$

E is a measure to the performance of trained BPNN based on batch gradient descent method, BPNN training process will end while E being less than a given threshold value E_t , for example, $E_t=0.0001$.

3 Fusion of Multi-Vision Based on D-S evidence theory

Considering the fusion of top-view, side-view and front-view image from multi-vision sensors, three BPNNs need to be constructed respectively, each network is composed of input layer, hidden layer and output layer[6,7], their

outputs are fused together with Dempster-Shafer (D-S) theory to produce a final decision, as shown in Fig.3.

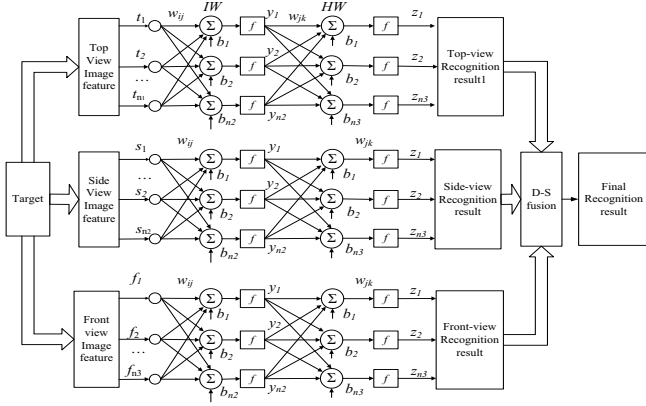


Fig.3 Fusion of Multi-vision of industrial robot

D-S evidence theory is a useful decision level fusion[8], considering the importance of different images, the D-S evidence theory will be modified by adjusting the weight of basic probability assignment (BPA) for fusion of outputs of three BPNNs in the paper. Define a complete (exhaustive) set Ω , frame of discernment, which consists of mutual exclusive element, $m : 2^\Omega \rightarrow [0,1]$, A is a subset that contains at least two elements of Ω , called as compound hypothesis, the mass function m (BPA) satisfies: (1) $m(\phi)=0$; (2) $0 \leq m(A) \leq 1, \forall A \subset \Omega$; (3) $\sum m(A)=0$. Here ϕ is an empty set.

Let m_1, m_2 and m_3 be BPA functions corresponding to the output of the three BPNN respectively, the focal element of m_1 consists of A_1, A_2, \dots, A_l , the focal element of m_2 consists of B_1, B_2, \dots, B_m , and focal element of m_3 consists of C_1, C_2, \dots, C_n , the evidence m_1, m_2 and m_3 with focal elements $A_i (i=1 \sim l)$, $B_j (j \sim m)$ and $C_k (k=1 \sim n)$ respectively can be combined to yield a new mass function $m(D)$ as follows by a combination rule.

$$m(D) = \begin{cases} \frac{\sum_{A_i \cap B_j \cap C_k = D} m_1(A_i) m_2(B_j) m_3(C_k)}{1 - K} & D \subseteq \Omega \text{ and } D \neq \phi \\ 0 & D = \phi \end{cases} \quad (12)$$

$$K = \sum_{A_i \cap B_j \cap C_k = \phi} m_1(A_i) m_2(B_j) m_3(C_k) < 1 \quad (13)$$

K measures the degrees of conflict between the different bodies of evidence. In this paper, the six target classes will be discussed, suppose class space $\Omega = \{ \text{link rod, piston, gear, bolt, hexagon nut, other} \}$, the output vectors of the three BPNNs of top-view, side-view and front-view vision can be written as follows respectively,

$$\begin{cases} ta = [ta(1) \ ta(2) \ \dots \ ta(6)]^T \\ tb = [tb(1) \ tb(2) \ \dots \ tb(6)]^T \\ tc = [tc(1) \ tc(2) \ \dots \ tc(6)]^T \end{cases} \quad (14)$$

Hence, the BPA of the three BPNNs in Eq.12 can be given by:

$$m_1(A_i) = \frac{ta(i)}{\sum_{p=1}^6 ta(p)}, \quad m_2(B_j) = \frac{tb(j)}{\sum_{p=1}^6 tb(p)}, \quad m_3(C_k) = \frac{tc(k)}{\sum_{p=1}^6 tc(p)} \quad (15)$$

4 Experimental Results of Multi-vision Fusion of Robot in MAS-Based Space

Target classes = {link rod, piston, gear, bolt, hexagon nut, other}, test samples are acquired from the captured images of mechanical part from top-view, side-view and front-view direction, the region feature are extracted through morphological pre-processing such as dilation and erosion, as shown in Fig.4.

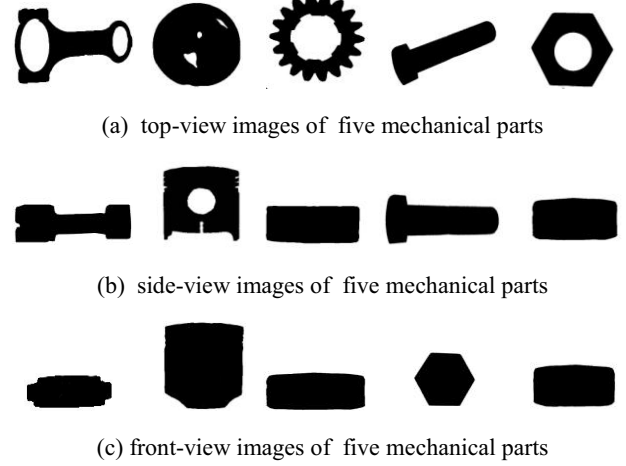


Fig.4 Binary images of mechanical parts obtained from different directions

Using these typical samples which include basic shapes, the three BPNNs are trained according to the above BPNN training procedures, Fig.5 displays the training results for top-view, side-view and front-view images of mechanical parts using region feature vector with the goal of performance $E_t=0.001$.

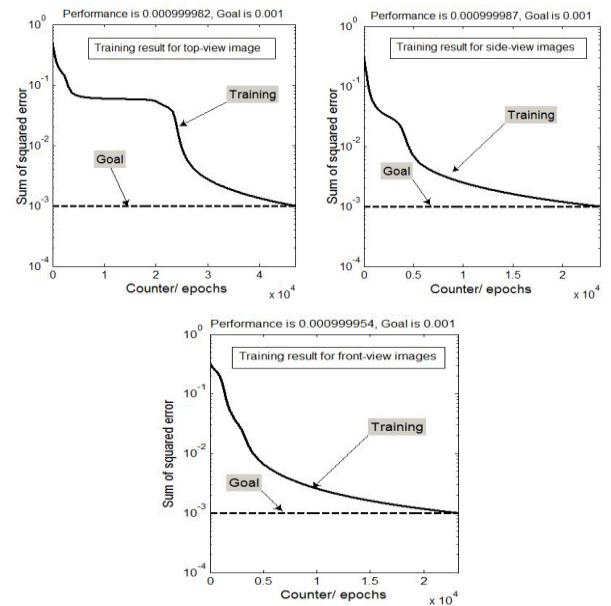


Fig.5 Training results of BPNN for top-view, side-view and front-view images

In order to compare the recognizing accuracy between multi-vision fusion and single vision based on BPNN, 160 samples of mechanical part is used, in which 120 samples are training samples and 40 samples as verifying

samples, as shown in Table 1. The test results in Table 1 shows that the recognizing accuracy of multi-vision fusion proposed in the paper is higher than one of single BPNN, it helps to improve the accuracy of robotic vision.

Table 1 Comparison of recognition results between single BPNN and BPNN-fusion using D-S theory

	Test results			
	Fusion of BPNNs	(top-view) Single BPNN	(side-view) Single BPNN	(front-view) Single BPNN
Training samples	120	120	120	120
Verifying samples	40	40	40	40
Fault	1	3	7	6
Error ratio	2.5%	7.5%	17.5%	15%

5 Conclusions

In MAS-based smart space, multi-vision fusion of industrial robot has been completed using D-S evidence theory, the BPNN classifier of each agent is trained independently to recognize the various targets (mechanical parts) from different directions, and the region features of top-view, side-view and front-view image are inputted to their own BPNN respectively, the outputs of three BPNNs are fused with D-S theory to produce a final recognition decision. The experimental results of target recognition for a variety of mechanical parts have shown that recognizing accuracy of the proposed fusion method in the paper is higher than conventional one of single BPNN method and is robust to degraded target image, and MAS-based smart space will contribute to the parallel processing of immense image data in robotic multi-vision system.

6 Acknowledgment

This work is financially supported by the Guangdong natural science foundation (No. S2012010010265) and Jiangmen science and technology bureau (No.20140060117111).

References

1. Weber B. , Zeller P. and Kuhnlenz K. Multi-camera based real-time configuration estimation of continuum robots, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Jinan, China, 2012, p. 3350 - 3355.
2. Boochs F. , Schutze R. , Simon C. , Marzani F. and et al. Increasing the accuracy of untaught robot positions by means of a multi-camera system, International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2010, p.1-9.
3. Kermorgant O. and Chaumette F. Multi-sensor data fusion in sensor-based control: Application to multi-camera visual servoing, IEEE International Conference on Robotics and Automation (ICRA), 2011, p. 302-308.
4. Hou X. ,Wang Y. and Wang H. Research on coordination CSCM based on multi-agent system. Journal of system &management. 2008, vol.17, no.5, p.548-552.
5. Bulka B. , Gaston M. and Desjardins M. Local strategy learning in networked multi-agent team formation, Autonomous Agents and Multi-Agent System, 2007, vol.15, no. 1, p.318-324.
6. Basheer I. A. and Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods, 2000, vol.43, no.1, p. 3-31
7. Cheng Hua Li and Soon Cheol Park. Combination of modified BPNN algorithms and an efficient feature selection method for text categorization. Information Processing and Management, 2009 vol.45, no.3, p. 329-340.
8. Guan J. W. and Bell D. A. Evidence Theory and its Applications, North-Holland, New York, 1991