MULTI-VIEW TECHNIQUE FOR 3-D ROBOTIC OBJECT RECOGNITION SYSTEM USING NEURO-FUZZY METHOD

*M. Khusairi Osman, Mohd. Yusoff Mashor and Mohd. Rizal Arshad

Control and ELectronic Intelligent System (CELIS) Research Group, School of Electrical & Electronic Engineering,
Universiti Sains Malaysia, Engineering Campus,
14300 Nibong Tebal, Seberang Perai Selatan, Pulau Pinang, Malaysia.
Tel: +604-5937788, Fax: +604-5941023, *E-mail:muhd_khusairi@yahoo.com

Abstract: The recognition of objects is one of the most challenging goals in robotic vision system. The problems increase when the process of recognition involves three dimensional (3-D) objects. To overcome this problem, many researchers have proposed their own solution. In this paper, a multiview 3-D object recognition system based on neuro-fuzzy system is proposed. The system consists of two stages. The first stage is feature extraction of object. In practice, there are many techniques for feature extraction such as moment invariants, Fourier transform coefficients, M-transform etc. We use moment invariants for feature extraction. The second stage is the process of recognition. We use Multiple Adaptive Network based Fuzzy Inference System (MANFIS) as a classifier for this stage. This method has been successfully tested using five different objects. High recognition rate was obtained using the proposed method.

Keywords: Robot vision**,** 3-D object recognition, moment invariants, neuro-fuzzy system.

I. Introduction

An object recognition system finds an object in the real world from an image of the world, using object models which are known priori [9]. The process of recognition is one of the hardest problems in computer vision. Although human can perform object recognition effortlessly and instantaneously, an algorithmic description of this task for implementation on machine has been very difficult especially in case of 3-D objects. In robotics application such as object grapping or manipulation, efficient 3D object recognition will assist in faster identification and localization process for real-time dynamic arm motion control.

In general most model based object recognition system considers the problem of recognizing objects from the image of a single view [2][5][22]. However, a single view may not contain sufficient features to recognize the object. In addition, it required complex feature sets and this make the recognition process time consuming [7][22]. To overcome this problem, modeling 3D object recognition using multiple 2D views was proposed by some researchers.   It summarised the set of possible 2D appearances of a 3D object. Some of the early studies such

as use of aspect graph was proposed by Koenderink and van Doorn [13]. An aspect graph represents all stable 2D views of a 3D object. However, the extraordinarily large in size and complexity of aspect graphs for even simple object has hindered the use of this method. Edelman and Bulthoff [4] found a strong and stable correlation between recognition performance and viewpoint variation and suggest object representations by multiple viewpoint specifically 2D representations. Murase and Nayar [17] and Nene [18] developed a parametric eigenspace method to recognize 3D objects directly from their appearance. This technique however is not robust to occlusion and do not provide indication on how to optimize the size of the database with respect to the types of objects considered for recognition and their respective eigenspace dimensionality.

Recently, some papers have proposed an effective recognition algorithm using neural networks. The advantage of this model is the ability to learn from a training data set and perform a prediction of the other data set.   Lin et. al. [15], Nasrabadi and Li [19] used Hopfield neural networks for their 3D object recognition system. Compare with conventional 3D object recognition, it provides a more general and parallel implementation paradigm. Lu et al. [16] recognized 3D objects using a back-propagation algorithm which has been commonly used in pattern recognition applications. Other works using neural networks such as Foresti and Pieroni [6] used neural tree (NT), Ham and Park [7] used hidden Markov model-based system combined with neural networks and Carpenter and Ross [3] used ART-EMAP networks. However, the use of neuro-fuzzy system; a combination of neural network and fuzzy system, is not widely used in 3D object recognition. In this paper, we used a type of neuro-fuzzy system called Multiple Adaptive Network based Fuzzy Inference System (MANFIS) to perform 3D object recognition.

II. System Overview

In this section, a methodology for image acquisition and data extraction is presented. A 3D object recognition system using multiple views was developed. The system aims to recognize 3D objects which are stand alone, separated and are independent to each other. The possible object and camera set-up for the proposed system is

illustrated in Fig. 1.

Object to be recognized must be placed at the centre of the turntable. Three B/W CCD cameras are used to capture the images simultaneously from different viewpoints (different angle). These cameras are fixed at the same height (y co-ordinate), at 45° from the centre of turntable. The cameras must have same focal length and distance from the centre of turntable. The angle that separated camera 1 and camera 2, camera 2 and camera 3 are fixed at 45°. We assumed that the location of camera 1 as a reference point (scene at 0^0). For the first condition, camera 1 views the scene at 0°, camera 2 views the scene at 45° and camera 3 views the scene at 90°. Fig. 6 shows an example of images taken from the cameras at the reference points. Next, the object will be rotated 5° clockwise to get the second condition. At this condition, camera 1 views the scene at 5°, camera 2 views the scene at 50° and camera 3 views the scene at 95°. For image acquisition process, each object will be rotated 360° and images will be captured for each 5° rotation. Hence, for each object, we will have 72 conditions after a complete 360° rotation. Captured images are then digitized by the DT3155 frame grabber from Data Translation Inc. and set to the pre-processing and feature extraction stage. In this study, moment invariants are used as a feature as it is invariants with position, orientation and scale changes. The algorithm has been commonly used in pattern recognition because it explains geometrical properties of an object.
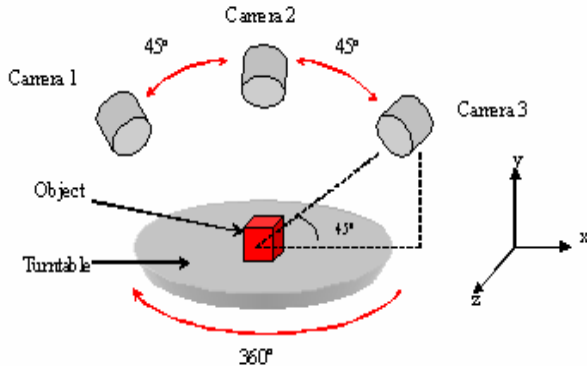
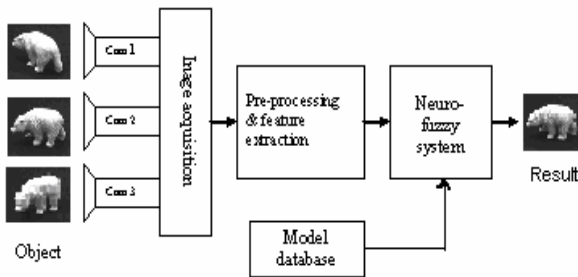

Figure 1: Image acquisition set-up

Figure 2: System configuration

Furthermore it takes short processing time as the algorithm is simple. Some works using moment descriptions and its properties can be found in [7, 14, and 21]. All the features extracted from various viewpoints will be presented as an input for the recognition stage. Fig. 2 depicts the overall proposed system. The invariance properties of moments of 2-D and 3-D shapes have received considerable attention in recent years. They are useful as they define a simply calculated set of region properties that can be used for shape classification and part recognition. Hu [8] derived a set of invariants based on combinations of regular moments using algebraic invariants. These invariants are invariant under change of size, translation and rotation. In this work, the first moment invariant is selected to be used as suggested in [23].

III. Neuro-fuzzy system

Neuro-fuzzy system is a combination of neural network and fuzzy system in such a way that neural network learning algorithms, is used to determine parameters of the fuzzy system [20]. ANFIS is a neuro-fuzzy model proposed by Jang [11]. The structure of ANFIS with five layers is shown in Fig. 3. x and y are the inputs for ANFIS. Note that the input layer is not calculated as an ANFIS layer.
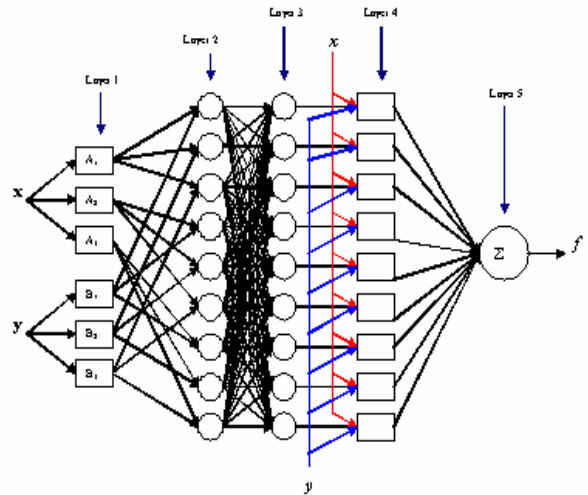

Figure 3: ANFIS Architecture

For learning rule of ANFIS, hybrid learning algorithm [4,5] which combines the gradient descent and least-squares method is used to find a feasible set of parameters. Table 1 shows the hybrid learning procedure for ANFIS. Further information can be obtained from [10, 11, 12, and 20]

Table 1: Two passes in the hybrid learning procedure for ANFIS

-	Forward Pass	Backward Pass
Premise Parameters	Fixed	Gradient Descent
Consequent Parameters	Least Squares Estimate	Fixed

However, ANFIS itself only suitable for single output system. For a system with multiple outputs, ANFIS will be placed side by side to produce a Multiple ANFIS (MANFIS) [12]. The number of ANFIS required depends on the number of required output. Fig. 4 shows a MANFIS with five outputs. Since the input data remains the same for each ANFIS, they also have the same initial parameter such as initial step size $?$, membership function (MF) type and number of MF.
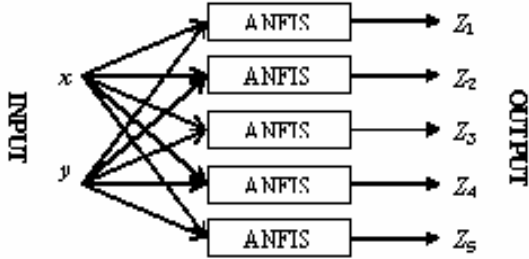

Figure 4: MANFIS with five output

IV. Experimental results

In order to examine the performance of this system, we have selected five 3D objects for this recognition. Some examples are shown in Fig. 5. As we mentioned earlier, each object will has 72 conditions. We choose odd condition (1, 3, 5, …71) as a training data and the even condition (2, 4, 6,…72) for the testing, so that the views of the testing images have never appeared in the training process at all. Hence, for five objects, we will have 180 data for training set and 180 testing data set.

MANFIS with five outputs was used to perform this task. We have analyzed the MANFIS performance using different initial parameter set. To find the best, first, we run our system using MF=2 with initial step size, $?$ =0.01, 0.05, 0.10, 0.25, and 0.35. Increasing the initial step size value will increase the learning rate for the ANFIS. However, if the step size is set too large (i.e. 0.35), the system will fail to learn properly. Table 2 summarized the system performance.

Table 2: System performance using MF=2 with different step size

Step size $?$	Maximum accuracy (%)
0.01	82.78
0.05	76.11
0.10	82.22
0.25	67.78
0.35	70.00

We also analyzed the system performance with the number of MF=3 and 4. MF=5 and above are not suitable for the analysis since the number of data is smaller than the number of adjustable parameters in the network. Table 3 and 4 summarized the results for each number of MF.

Table 3: System performance using MF=3 with different step size

Step size $?$	Maximum accuracy (%)
0.01	83.33
0.05	83.33
0.10	79.44
0.25	77.78

The results show that selecting a proper number of MF and initial step size value will affect the system performance. The system produces the best result at MF=4, $?$ =0.10 with 84.44% recognition accuracy. However, MF=2 is adequate to perform a good and fast recognition with a slightly less accuracy at 82.78%.

Table 4: System performance using MF=4 with different step size

Step size $?$	Maximum accuracy (%)
0.01	77.78
0.05	81.11
0.10	84.44
0.25	56.67

V. Conclusion

A multiple view 3D robotic object recognition system using neuro-fuzzy system is proposed in this paper. Our experiments show that 3D objects can be modeled and represented by a set of multiple 2D views. In addition, it does not require complex feature sets for 3D object modeling, thus improve processing time for feature extraction stage. Our experiments also proved that neuro-fuzzy system can perform well in 3D object recognition task although we are using simple feature. While we use simple feature for the purpose of illustration, one may use or combine other feature such as edge, Zernike moment, texture, corner etc to improve the performance of this system. Future work will be the comparison of the approach with other neural networks and/or neuro-fuzzy and actual implementation of the system in a robotic arm object handling and motion planning applications.

VI. References

[1] Bamieh, B. and De Figueiredo, R. A General Moment Invariants/attributed Graph Method for Three Dimensional Object Recognition From a Single View. *IEEE Journal of Robotics and Automation.* 2(1):31-41. 1986.

[2] Besl, P. J. and Jain, A. C. Three-dimensional Object Recognition, *ACM Computer Survey*. 17:76-145. 1985.

[3] Carpenter, G. A. and Ross, W. D. ART-EMAP: A Neural Network Architecture for Object Recognition by Evidence Accumulation. *IEEE Transactions on Neural Networks*. 6(4):805-818. 1995.

[4] Edelman, S. Y.; Bulthoff, H. H.; Tarr, M. J., How Are

Three-Dimensional Objects Represented In The Brain?. *Technical Report* , MIT. 1994.

[5] Elsen, I.; Kraiss, K. –F.; and Krumbeigel, D. Pixel Based 3D Object Recognition with Bidirectional Associative Memories. International Conference on Neural Netwoks. 3:1679-1684, 1997.

[6] Foresti, G. L. and Pieroni, G. G. 3D Object Recognition By Neural Trees. Proc. International Conference on Image Processing, 3:408-411, 1997.

[7] Ham, Y. K. and Park, R. –H. 3D Object Recognition In Range Images Using Hidden Markov Models And Neural Networks. *Pattern Recognition*. 32:729-742, 1999.

[8] Hu, M. K. Visual Pattern Recognition By Moment Invariants. *IRE Transaction on Information Theory*. 8(2):179-187. 1962.

[9] Jain, R.; Kasturi, R; and Schaunck, B.G. *Machine Vision*: McGraw-Hill. 1995.

[10] Jang, J. –S. R. Fuzzy Modeling Using Generalized Neural Networks And Kalman Filter Algorithm. Proc. of the Ninth National Conference on Artificial Intelligence (AAAI-91). 762-767.1991.

[11] Jang, J. –S. R., ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Trans. On Systems, Man and Cybernetics*. 23(3):665-685. 1993.

[12] Jang, J. –S. R.; Sun, C. –T.; and Mizutani, E., *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*: Prentice Hall. 1997.

[13] Koenderink, J. J. and van Doorn, A. J. Internal Representation Of Solid Shape With Respect To Vision. *Biological Cybernetics*. 32(4):211-216, 1979.

[14] Koker, R.; Oz, C.; Ferikoglu, A. Development Of A Vision Based Object Classification System For An Industrial Robotic Manipulator. The 8th IEEE International Conference on Electronics, Circuits and Systems. 3: 1281 -1284, 2001.

[15] Lin, W. –C.; Liao, F. –Y.; Tsao, C. –K. and Lingutla, T. A Hierarchical Multiple-View Approach To Three-Dimensional Object Recognition. *IEEE Trans. On Neural Networks*. 2(1):84-92, 1991.

[16] Lu, M. C.; Lo, C. H. and Don, H. S. 3D Object Identification And Pose Estimation. *Intelligence of Engineering System Through Artificial Neural Networks*. ASME Press, 1991.

[17] Murase, H. and Nayar, S. K, Visual Learning And Recognition Of 3d Objects From Appearance. *International Journal of Computer Vision.* 14:5-24. 1995.

[18] Murase, H., Nayar, S. K, and Nene, S. A. Real-Time 100 Object Recognition System. Proc. IEEE International Conference on Robotic and Automation. 1996.

[19] Nasrabadi, N. M. and Li, W. Object Recognition By A Hopfield Neural Network. *IEEE Trans. System, Man and Cybernetic*. 21(6):1523-1535. 1991.

[20] Nauck, D.; Klawonn, F.; and Kruse, R., *Foundations of Neuro-fuzzy Systems*. John Wiley & Sons. 1997.

[21] Ngan, K.N.and Kang, S.B.; 3-D Object Recognition Using Fuzzy Quaternions. Proc. IEEE Communications, Speech and Vision. 139(6):561-568. 1992.

[22] Roh, K. S.; You, B. J. and Kweon, I. S. 3D Object Recognition Using Projective Invariant Relationship by Single View. Proc. Of the IEEE Int. Conf.on Robotic and Automation. 3394-3399. 1998.

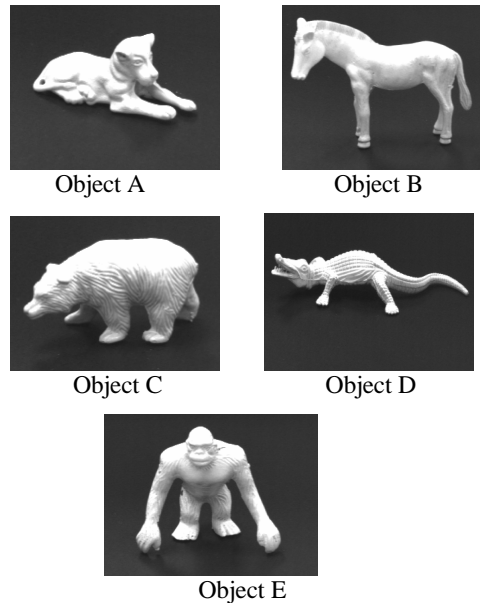[23] Vernon, D. *Machine Vision: Automated Visual Inspection and Robotic Vision*: Prentice Hall. 1991.

Object A   Object B

Object C   Object D

Object E

Figure 5: Example of objects used in the experiment
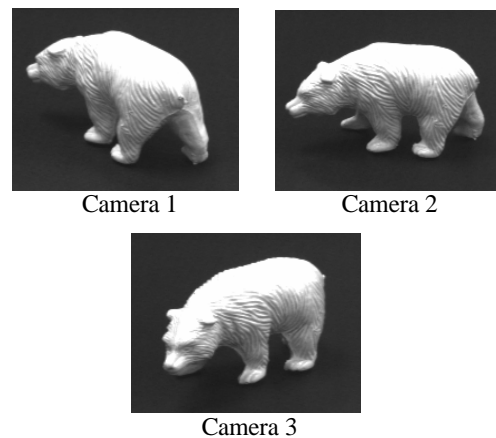

Camera 1   Camera 2

Camera 3

Figure 6: Image scene from different view at reference point