SVD based Object Recognition using Functional Link Neural Network for Bin Picking Application

R Nagarajan, Sazali Yaacob, Paulraj Pandiyan, Hema C R[1]
Artificial Intelligence Research Group
School of Engineering and Information Technology
Universiti Malaysia Sabah
Kota Kinabalu, Malaysia
[1]hemacr@yahoo.com

Shamsudin Amin, Marzuki Khalid
CAIRO, Faculty of Electrical Engineering
Universiti Teknologi Malaysia
Kuala Lumpur
Malaysia

Abstract – **A vision system attached to a bin picking robot must be able to analyze the objects in the bin and must be able to pick an object. In this paper a simple object recognition method is presented using singular value decomposition of the object image and a functional link neural network for a bin picking vision system. The results of the functional link net are compared with a simple feed forward net. The net is trained using the error back propagation procedure. The proposed method seems to be robust for recognition of objects.**

I. INTRODUCTION

Mechanical manipulators are being used increasingly for machine loading, welding, painting and sealing. However they have not been used extensively for applications such as assembly operations. One of the reasons for this is that the manipulators typically just play back a sequence of motions. The blind robot has difficulty dealing with uncertainty in the position of objects [1]. Feeding mechanism and fixtures are needed to present the objects precisely in predefined locations for pick up, which increase the cost of automation of such operations.

The above problem can be tackled with a bin picking system wherein the objects to be picked are placed inside the bin in an unorganized manner. The bin picking system with vision sensors depicts a manual assembly unit were objects to be assembled are placed in bins surrounding the work station. However the vision system attached to the bin picking system should be capable analyzing the objects in the bin as well as identify the object to be picked from the pile irrespective of the pose or orientation of the object.

In this paper we address the problem of recognizing complex industrial objects with different orientations. The method proposed uses the images of industrial objects in various orientation acquired by a digital camera to extract the singular value features. These features are fed as input to a functional link neural network classifier [FLNN] which identifies the objects..

The proposed method using FLNN classifier gave better results when compared with conventional classifier in terms of training time and epochs.

A. Feature Extraction

Feature extraction is the most fundamental and important problem in the field of object recognition [2] and finding the efficient feature is always the key to solving the problem of object recognition

There are four major types of features for object recognition.

1. *Visual Features*: They include edges, contours, textures and regions of an image; these are all visual features of a pixel
2. *Statistical Features*: Histogram features, properties of moments etc. belong to this kind of features of a pixel.
3. *Transform Coefficient Features*: Fourier descriptors have a good ability to describe edges of objects and Fourier transforms are used to extract texture features of images.
4. *Algebraic Features*: They represent intrinsic attributions of an image. Any image can be considered as a matrix; therefore various algebraic transforms or matrix decompositions can be used for algebraic feature extraction of the image.

Eigenvector of a matrix represents algebraic attributes of the matrix and hence can be used for feature extraction. Similarly Singular Value Decomposition [SVD] of a matrix results in singular values which is another algebraic feature of an image.

B. Singular Value Decomposition

The SVD is a widely used technique to decompose a matrix into several component matrices, exposing many of the useful and interesting properties of the original matrix [3]. Any m x n matrix a (m >= n) can be written as the product of an m x n column-orthogonal matrix u, an n x n diagonal matrix with positive or zero elements, and the transpose of an *n* x *n* orthogonal matrix **v**[6]:

$$\mathbf{a} = \mathbf{u}.\mathbf{w}.\mathbf{v}^{t} \quad \dots (1)$$

where

$$\mathbf{w} = \begin{bmatrix} w_1 & 0 & . & 0 & 0 \\ & w_1 & . & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & . & w_{n-1} & 0 \\ 0 & 0 & . & 0 & w_n \end{bmatrix} \quad \dots (2)$$

and

$$\mathbf{u}^t\mathbf{u} = \mathbf{v}\mathbf{v} = \mathbf{I} \qquad \ldots (3)$$

$w_1, w_2, \ldots, w_{n-1} w_n \geq 0$

The diagonal elements of matrix **w** are the singular values of matrix **a** .

The singular values [SV] obtained by the Singular Value Decomposition of an image matrix is an algebraic feature of an image which have the properties such as [2]

1. *Stability*: The SV features do not have large changes when the grey values of an image have small changes.

2. *Invariance*: The SV features are invariant to proportional variance of image intensity in optimal discriminant space. This property is important for description and recognition of objects.

3. *The invariance of SV features to transposition transform.*

4. *The invariance of SV features to rotation transform.*

5. *The invariance of SV features to translation*

6. *The invariance of SV features to mirror transform.*

The above properties of the SV features play an important role in the recognition of objects in different orientation.

C. Functional Link Neural Networks

Functional link enhance the computing power of the neural network, by functional transformation an input pattern is enhanced in its representation. The neural network saves time to learn additional information obtained via such transformation. The functional link can improve the network in both learning capacity and efficiency. The functional link generates a set of linearly independent functions and evaluates these functions with the input pattern as the argument. Suppose each input node encodes a certain feature, applying the functional link to a feature causes it to multiply and generate more features,. As a consequence the expressive power of the network increases, as does its modeling capability. The benefit of the functional link when applied to mathematical modeling is the increased accuracy of mapping through expansion of the basis set. [4].

II. METHODOLOGY

Image Preprocessing

Images of industrial objects are acquired in different orientation using a Sony digital camera. The acquired images are of size 1024 x 1024 pixels and are preprocessed for image enhancement, binarization and edge detection for the extraction of singular value features. The extracted features are fed as input to the FLNN classifier. Figure 5 show the steps involved in the object recognition process. Figure 1 shows the original image of an industrial object which is to be preprocessed to extract the singular value features.

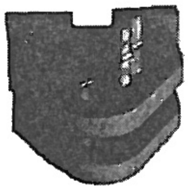

Figure 1 Image of an object before processing

A. Image Enhancement

Image enhancement is done to process the image for noise removal. The object images are resized to 32 x 32 pixels. The image is then passed through a median filter. A median filter is used because these filters are effective in the presence of impulse noise. The image is sharpened to define the edge details.

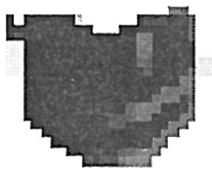

Figure 2 Object image after resizing and noise removal

B. Binarization

Gray level thresholding is applied to the sharpened image to compute the global threshold level to convert the intensity image to a binary image. The binary conversion is done to convert the RGB image to grayscale image. The output binary image has values of 0 for all pixels less than the threshold level and 1 for all pixels greater than the threshold level. A clear definition of the image is obtained from which the edge boundary of the object can be extracted.

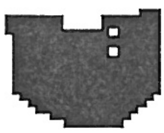

Figure 3 Binary image of the object

C. Edge Detection

Edge boundary of the image is detected by means of a canny edge detector. This method finds edges by looking for local maxima of the gradient of the image. The gradient is calculated using the derivative of a Gaussian filter. The method uses two thresholds, to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges. This method is therefore less likely than the others to be "fooled" by noise, and more likely to detect true weak edges.

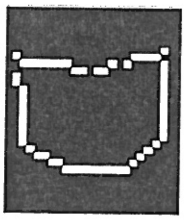

Figure 4 Edge image of the object

D. Feature Extraction

The singular values can be extracted by the decomposition of the edge image matrix using SVD. SVD produces a diagonal matrix, of the same dimension as the edge image and with nonnegative diagonal elements in decreasing order. The diagonal matrix elements otherwise called as singular values represent intrinsic attributes of the image.

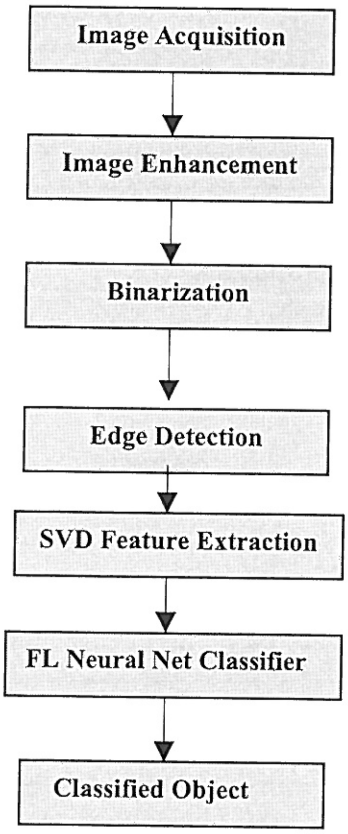

Figure 5 Block Diagram of the Object recognition method

III. NEURAL NETWORK ARCHITECTURE

The Functional Link Neural Network architecture shown in Figure 6 consists of a feed forward neural network with the functional link inputs. 25 singular values are fed to the network as input data in addition to 49 functional link inputs.

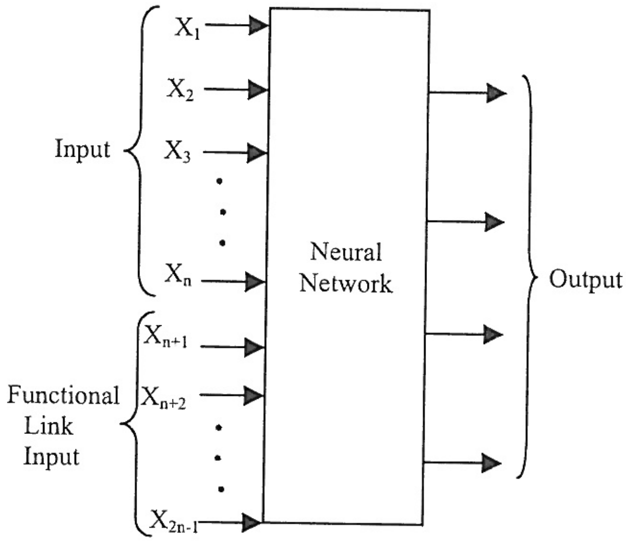

Figure 6 Functional Link Neural Network

The hidden layer is chosen to have 12 neurons and the output consists of 4 neurons. The hidden and input neurons have a bias value of 1.0 and are activated by bipolar activation function. The choice of initial weight will influence whether the net reaches a global minimum of the error and if so how quickly it converges. It is important that the initial weights must not be large, or the initial input signals to each hidden or output unit will be likely to fall in the saturation region. On the other hand if the initial weights are too small, the net input to a hidden or output unit will be close to zero which also causes extremely slow learning [5]. Hence the initial weights for the above network are randomized between –0.5 and 0.5 and normalized. The initial weights that are connected to any of the hidden or output neuron are normalized in such a way that the sum of the squared weight values connected to a neuron is one.

The sum squared error criteria as defined by equation (4) is used as a stopping criteria while training the network [7]. The sum squared tolerance defined in equation (4) is fixed as 0.01. The network is trained by the conventional back propagation procedure. The cumulative error versus epoch plot of the trained neural network is shown in Figure 7. The cumulative error is the sum squared error for each epoch. Equation (4) gives the formula for determining the sum squared error.

$$\text{Sum squared error} = \sum_{p=1}^{p}\sum_{k=1}^{m}(t_{kp} - y_{kp})^2 \qquad \dots (4)$$

where

t_k is the expected output value for the kth neuron

y_k is the actual output value for the kth neuron

m is the total number of output neurons

p is the total number of input neurons

III. EXPERIMENT RESULTS

In the experimental study, four industrial assembly objects are taken. The images of the objects are captured in different orientations. Each object is captured in five to seven such stable poses [orientation] depending on the shape of the object. Thirty one images of these four objects are acquired. The images are pre-processed as detailed in section II. From the edge image matrix the singular value features are extracted through singular value decomposition.

The feature data are normalized using the equation (5)

$$x_{n\_i} = 1.8\left[\frac{x_i - x_{min}}{x_{max} - x_{min}}\right] + 0.1 \qquad , i = 1, 2, 3, \ldots n \quad \ldots(5)$$

where x_i - data to be normalized

x_{n_i} - normalized value of x_i

x_{min} - $\min(x_1, x_2, x_3, \ldots, x_n)$

x_{max} - $\max(x_1, x_2, x_3, \ldots, x_n)$

Out of the thirty two singular values obtained, only first twenty five prominent singular value features of magnitude greater than one are considered for classifying each image. The network is trained for 25 feature data and tested with 31 samples. Figure 7 shows the four objects and samples of some of their orientations. The orientation of the object is such that all prominent edge features are visible to get a complete picture of the shape of the object. This is essential to get a good description of the shape of a particular object. The network is trained for 27 sample images and tested for 31 samples. The FLNN is trained using the error back propagation procedure. The cumulative error versus epoch plot of the trained FLNN is shown in Figure 8

The success rate of the object classification using the FLNN is 90.32 %. These results were compared with a simple feed forward three layer network classifier and it was observed that the FLNN is able to classify the samples in lesser time and minimum number of epochs as against the time of 53 seconds and 9331 epochs of the former net for the same input parameters and training procedure.

The training parameters and test results of the proposed FLNN classifier is shown in Table 1 and the training parameters and test results of the NN classifier is shown in Table 2.

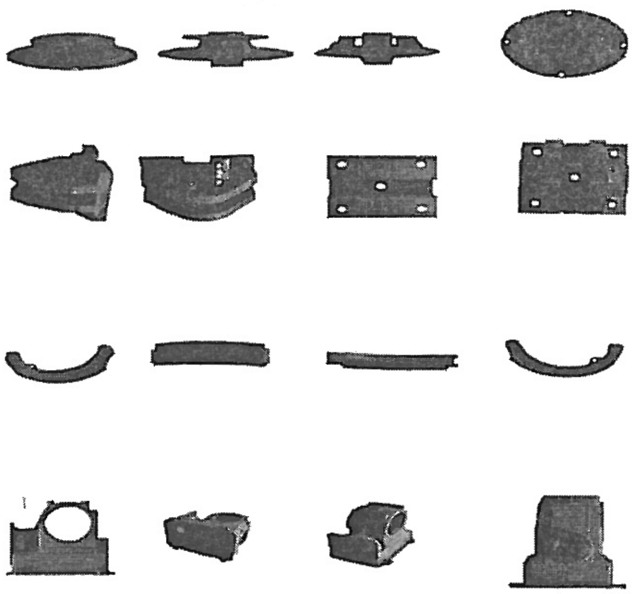

Figure 7 Sample images of the industrial objects with different orientation

TABLE 1
RESULTS OF FLNN CLASSIFIER

Parameters	Values
No. of input neurons	25
No. of Output neurons	4
No. of Functional link input neurons	49
No. of hidden neurons	12
Bias value	1.0
Learning Rate	0.1
Tolerance	0.01
Training time	6 seconds
Percentage Classification	90.32%
Maximum Epoch	5892

TABLE 2
RESULTS OF NN CLASSIFIER

Parameters	Values
No. of input neurons	25
No. of Output neurons	4
No. of hidden neurons	12
Bias value	1.0
Learning Rate	0.1
Tolerance	0.01
Training time	53 seconds
Percentage Classification	89.28%
Maximum Epoch	9331

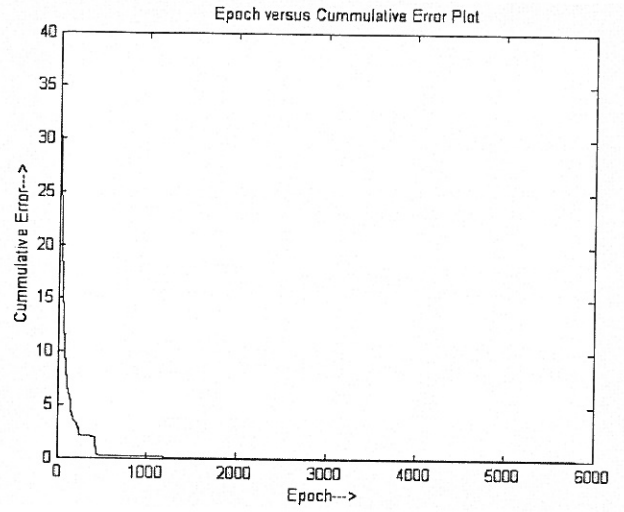

Figure 8 Cumulative Error versus Epoch Plot of FLNN

IV. CONCLUSION

In this paper, an object recognition methodology based on shape classification using the singular value decomposition of the image matrices and FLNN classifier is presented. 27 sample images are used to train the net and 31 samples are used to test the FL neural network. The successful classification percentage is 90.32%. The method proposed was able to classify the objects with lesser training time and epoch in comparison with a simple feed forward net classifier. The experimental results prove that effective recognition of an object with different orientation is possible using a Functional Link Neural Network Classifier.

V. ACKNOWLEDGMENT

The authors thankfully acknowledge the funding received from IRPA through UTM and UMS.

VI. REFERENCES

[1.] Berthold Klaus Paul Horn, *Robot Vision,* The MIT Press, McGraw-Hill Book Company, London 1993, p426

[2] Zi-Quan Hong, Algebraic Feature Extraction of Image for Recognition *IEEE Transaction s on Pattern Recognition* Vol. 24 No. 3 pp 211-219 , 1999

[3] W. Keith Nicholson, *Elementary Linear Algebra.* McGraw-Hill International Edition, Singapore 2001

[4] Li Min Fu, *Neural Networks in Computer Intelligence,* McGraw- Hill International Editions, Singapore 1994.

[5] Laurene Fausett, *Fundamentals of Neural Network Architecture and Applications*, Prentice Hall, USA. 1990.

[6] David S. Watkins, *Fundamentals of Matrix Computations.* John Wiley & Sons 2001.

[7] S.N. Sivanandam, M.Paulraj, *Introduction to Artificial Neural Networks* Vikas Publishing House, India 2003.