

RECOGNITION OF CONCRETE OBJECTS USING PATTERN RECOGNITION TECHNIQUES OF ARTIFICIAL NEURAL NETWORKS

*R. QAMAR, A.R. BAIG, M.A. FAROOQ, S. ABAR and T. AHMED

Computer Division, PINSTECH, P. O. Nilore, Islamabad, Pakistan

(Received March 8, 2002 and accepted in revised form November 1, 2002)

A software "Pattern Recognition System" has been developed to recognize the objects for which this system is trained. An "Intel PC Camera" is used to take the BMP images of the concerned objects. A neural network using Back propagation algorithm is developed as a recognition engine. The BMP images of the objects serve as the input patterns for this engine. Once the neural network is trained for the patterns of the objects, the software is able to detect as well as recognize them.

Keywords: Neural networks, Mult-layer perception, Back propagation, Sabel algorithm, YUV color system, Sigmoid function

1. Introduction

The term pattern recognition encompasses a wide range of information processing problems of great practical significance [1]. The patterns we encounter can fall into two categories: abstract and concrete. Examples of abstract items include ideas and arguments. Recognition of such patterns, termed conceptual recognition, belongs to another branch of artificial intelligence and is beyond the scope of this paper. Examples of concrete items include characters, symbols, pictures, biomedical images, three-dimensional physical object, target signatures, speech waveforms, electrocardiograms, electroencephalograms and seismic waves [2].

Now-a-days two most general and most natural frameworks to formulate the solution to pattern recognition problems are statistical and Artificial Neural Network. In statistical pattern recognition features of patterns are extracted. On the basis of these features patterns are divided into different classes. Artificial neural networks have many remarkable features, such as ability to process noisy, sparse or incomplete data and high fault tolerance. Due to their inherent parallelism they also have the ability to respond in real time. With such attractive features, neural networks can be applied to a wide variety of problems, such as patterns recognition, grouping similar patterns, or speech recognition.

2. Description of work

Whole work can be divided into four stages:

2.1. Constructing the neural network

In this stage different neural networks were studied. The neural network MLP (Multi-layer Perceptron) was employed in this work. An MLP consists of one input layer, one or more intermediate or "Hidden" layers and one output layer. The neurons of each layer are fully connected to the neurons of the next layer as shown in Fig. 1. Software package of neural network has been written in Visual Basic.

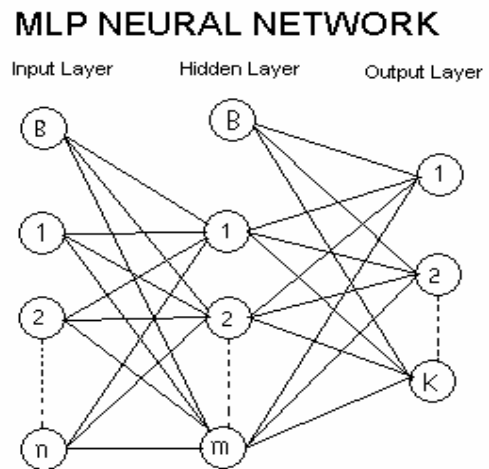


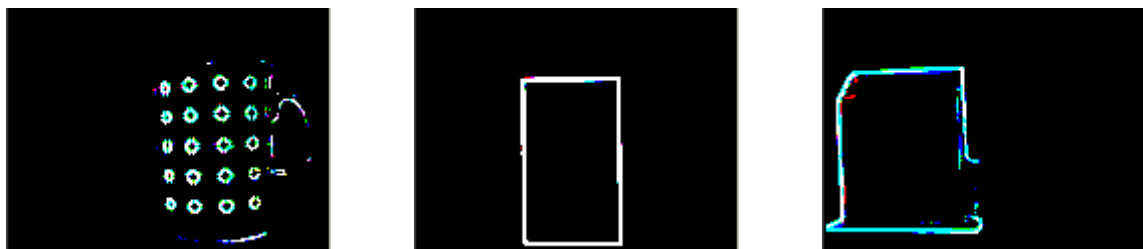
Figure 1. In this MLP Neural Network one Input layer, one Hidden layer and one Output layer is shown.

* Corresponding author : rashid@pinstech.org.pk



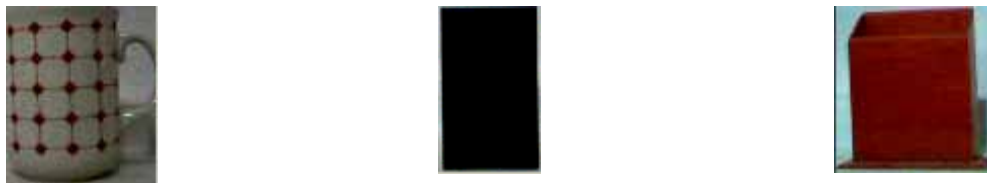
Image of Cup Image of Pen Stand Image of Paper Rack

Figure 2. BMP Images of three objects, for which Recognition System is trained.



Edges of Cup Edges of Pen Stand Edges of Paper Rack

Figure 3. Edges of BMP images



Extracted Cup from whole image Extracted Pen Stand from whole image Extracted Paper Rack from whole image

Figure 4. In this figure only those portions of BMP images are show in which these objects lie.



Resized Cup Resized Pen Stand Resized Paper Rack

Figure 5. BMP images of selected objects after resizing are shown.

2.2. Data preprocessing for input data file.

An Intel PC camera has been used to save the snap shorts in BMP binary files. Original images of three objects are shown in Fig 2.

A Sobel edge detector algorithm [3] was modified for colour images and used to detect the edges of the objects shown in Fig. 2. After applying this algorithm the above images were converted to the images shown in Fig. 3.

Using these edges a subroutine is written to extract original object from the whole images as shown in Fig. 4.

Since the sizes of these extracted images are different, as shown in Fig 4, a subroutine is written to resize the images as shown in Fig. 5 (40 x 40 pixels).

Now these 1600 (40 x 40) pixels for each image are converted from RGB color system to YUV color system.

Like YIQ model YUV model is also used in commercial color TV broadcasting. Basically, YUV is a recoding of RGB for transmission efficiency and for maintaining compatibility with monochrome TV standards. In fact Y component of the YUV

system stands for luminance or brightness, and U and V for chrominance. Actually U and V are the difference between a color and a chosen reference color of the luminous intensity [9, 10].

$$Y = (0.299 \times \text{red} + 0.587 \times \text{green} + 0.114 \times \text{blue}) / 255 [3].$$

$$U = (-0.148 \times \text{red} - 0.287 \times \text{green} + 0.437 \times \text{blue}) / 112 [4].$$

$$V = (0.615 \times \text{red} - 0.515 \times \text{green} - 0.100 \times \text{blue}) / 157 [4].$$

In some other text U and V are given as:

$$U = -0.169 \times \text{red} - 0.332 \times \text{green} + 0.5 \times \text{blue} + 128 [9] [10].$$

$$V = 0.5 \times \text{red} - 0.419 \times \text{green} - 0.0813 \times \text{blue} + 128 [9] [10].$$

2.3. Training of neural network

Once neural network was constructed and input data files made, training phase of Pattern Recognition Neural Network System started.

In training Back Propagation algorithm is used. The mathematical base for the back – propagation algorithm is gradient descent method. The gradient of function gives the direction in which function increases more rapidly [3, 5]. Algorithm consists of two passes through the different layers of network, a forward pass and a backward pass. In the forward pass, an input pattern is applied to the input nodes (1600 input nodes in our case) and its effect is propagated forward, layer-by-layer. The activation function used at each hidden and output node is given as :

$$f(x) = \tanh(x)$$

It is a sigmoid function ranging from -1 to 1.

During the backward pass the error signal is propagated backward through the network against the direction of the links, and the weights are adjusted, from last hidden layer to the first hidden layer. Here the weight correction is proportional to gradient of error function.

Target Output values for training of Neural Network are defined as:

If image of Cup is the input to the system then Target Output values are 1, 0, 0 i.e. Cup= 1, Pen Stand= 0, Paper Rack= 0.

Similarly for Pen Stand Target Output values are 0, 1, 0 i.e. Cup= 0, Pen Stand= 1, Paper Rack = 0.

And for Paper Rack Target Output values are 0, 0, 1 i.e. Cup= 0, Pen Stand= 0, Paper Rack= 1.

2.4 Comparison with the reported work

In the above mentioned paper [11], hardware neural network architecture to recognize the images has been demonstrated. The system is based on the combination of a two-dimensional amorphous silicon photoconductor and a liquid-crystal spatial light modulator array (64 x 64). Overall processing time is 250 micro seconds. In this paper 80% correct results are claimed.

In many other neural networks for pattern recognition we studied, feature extraction technique is used [12-15]. But from the above mentioned paper, we picked an idea of giving pixel's gray level as input to train our Pattern Recognition System.

We have developed a software program for Pattern recognition using neural network instead of its hardware setup. To get images we used "Intel PC Camera". We used YUV colour model instead of RGB colour model. We also reduced our input image size from 120 x 160 pixels to 40 x 40 pixels, which reduced the processing time. But our processing time is in the range of milli seconds. We got 93% to 100% correct results as shown in Table 1.

3. Results and Discussion

Neural Network is trained with 150 training patterns, 50 patterns for each object. Other 75 testing patterns, 25 patterns for each object, are used to test the performance of trained "Pattern Recognition System".

In Table 1 four different input types are given for which neural network is trained. Epochs (Iterations) and system error is optimized so that maximum number of testing patterns, out of 75, must be recognized. Optimized epochs for U and V type are almost double form Red pixel and Y type of input data, but respective system error for U and V type is slightly greater than Red pixel and Y type. It means the training time of neural network using U and V type is much greater than Red pixel and Y type. On the other hand if one trained the system using U or V type of input data file, one can get almost 100% correct result as given in Table 1.

Table 1. All four types of input, their errors, and results are shown below.

Input type	Red Pixel	Y	U	V
Training Patterns	150	150	150	150
No. of Epochs	600	900	1800	1500
System Error	0.019	0.010	0.023	0.023
Testing Patterns	75	75	75	75
Results	70 93%	72 96%	75 100%	75 100%

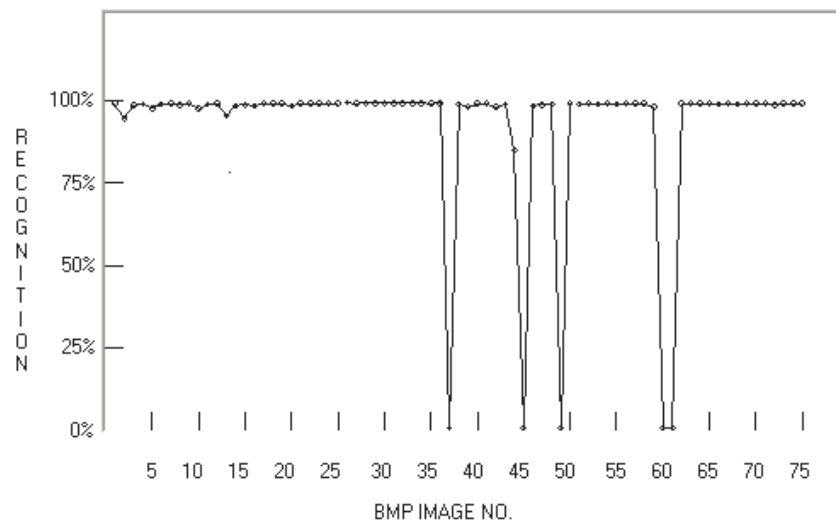


Figure 6. In this figure results of 75 BMP images are shown using Red pixels of the image as input to the Recognition System. Five out of 75 BMP images are not recognized, whose image numbers are 37, 45, 49, 60 and 61.

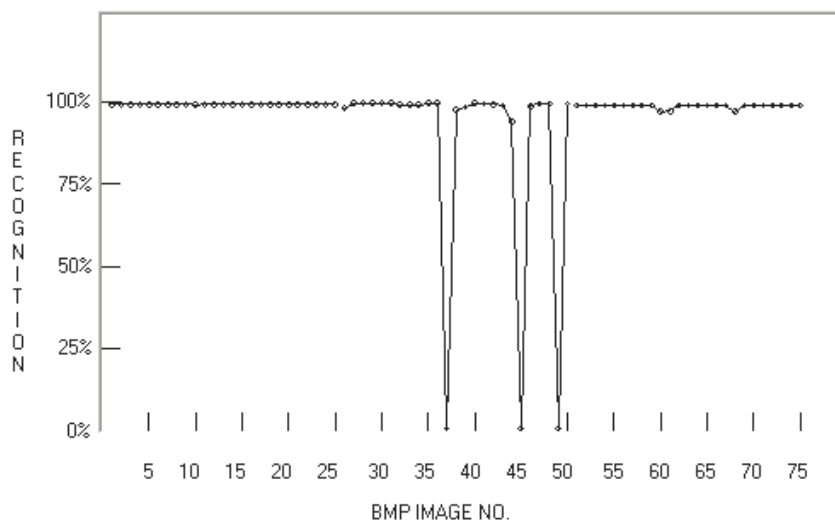


Figure 7. In this figure results of 75 BMP images are shown using Y value of images as input to the Recognition System. Three out of 75 BMP images are not recognized, whose image numbers are 37, 45 and 49.

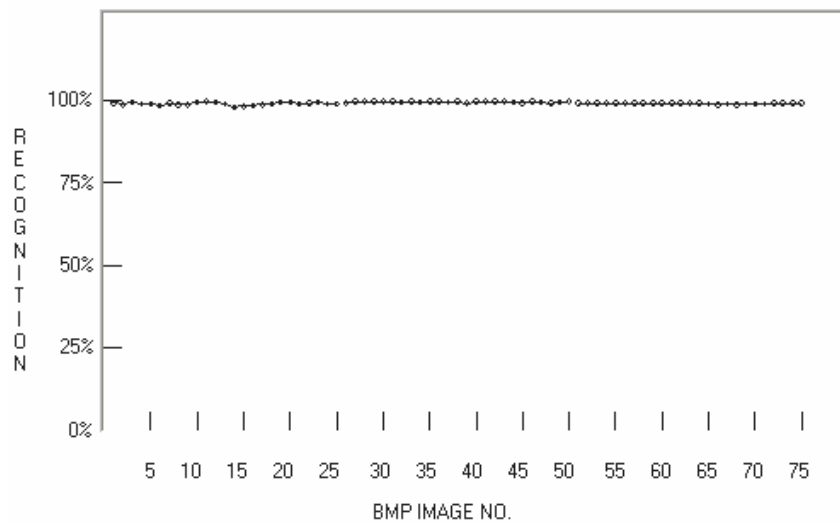


Figure 8. In this figure results of 75 BMP images are shown using U value of images as input to the Recognition System. All 75 BMP images are recognized correctly.

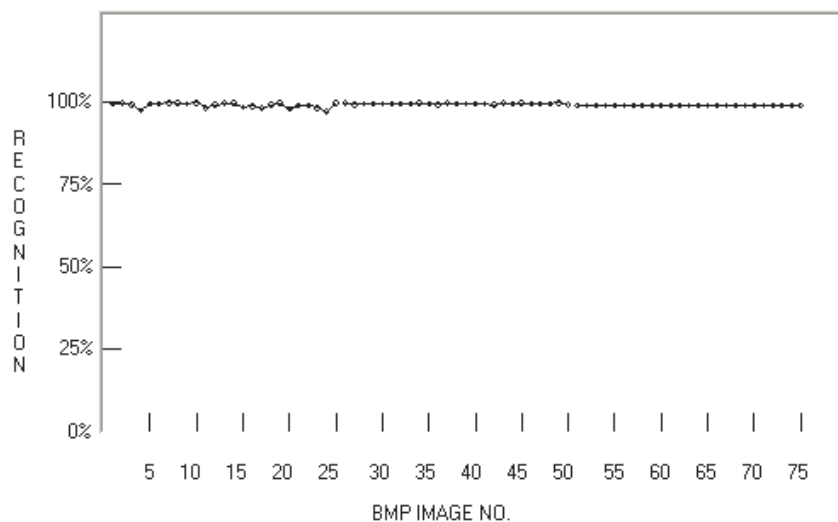


Figure 9. In this figure results of 75 BMP images are shown using V value of images as input to the Recognition System. All 75 BMP images are recognized correctly.

Detailed results in graphics form are given below.

By examining results given in figures from 6 to 9 one can find that U and V types of input data files are the best to recognize the patterns. In other words for sensitive research work U and V types are recommended.

4. Conclusion

We have developed a "Pattern Recognition System" using Neural Network. It is an initial stage of our project named "Automatic Target Recognition" (ATR). Our next attempt will be to

recognize the given target using movies. It will be help full for defense related application of PAEC.

Acknowledgements

The authors would like to thank Mr. Javaid-ur-Rehman, Head: Computer Division, PINSTECH to start such useful research work in Computer Division.

References

- [1] Sing-Tze Bow, Pattern Recognition, Marcel Dekker (1984).
- [2] M. Christopher Bishop, Neural Networks for Pattern Recognition, Oxford (1997).

- [3] L. Pitas, Digital Image Processing Algorithms, Prentice Hall, International (UK) Ltd. (1993).
- [4] A. Rauf Baig, R. Segquier and G. Vaucher, A Spatio-Temporal Neural Network Applied to Visual Speech Recognition, Proceedings of ICANN '99, UK
- [5] K.R. Castleman, Digital Image Processing, Prentice Hall, Inc. (1996).
- [6] J. Sklansky, Pattern Recognition Introduction and Foundations, Dowden Hutchinson & Ross (1973).
- [7] S. Brown, Visual Basic 5, BDB Publication (1997).
- [8] S. Holzner, Visual Basic 6 Programming Black Book, Coriolis 2000.
- [9] Pattern Recognition Neural Networks Computer Division, Joe Maller, Fx Script Reference : RGB and YUV Color http://www.joemaller.com/fcp/fxscript_yuv_color.shtml.
- [10] B. Charles Owen, RGB <---> YUV Conversion Formulas <http://www.cse.msu.edu/~cbowen/docs/yuvtorgb.html>.
- [11] Richard G. Stearns, Applied Optics **34** (1995) 2595.
- [12] A.A. Khan., et al., Applied Optics **33** (1994) 5476.
- [13] J.W. Shual and Z. X. Chen, R. T. Liu. and B. X. Wu, Applied Optics **34** (1995) 6764.
- [14] S.W. Lee and H.H. Song, IEEE Transaction on Neural Network **8** (1997) 331.
- [15] A. A. Khan, A. A. Rizvi and M. S. Zubairy, Applied Optics **33**(1994) 5467.

