ISSN:0975-9646

Priyanka Parvathy Dali et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (2) , 2015, 1111-1114

Interpretation of Hand Gestures Using Neural Networks:A Review

Priyanka Parvathy Dali[#1], Dr. Hema C.R[*2]

[#] *Department of Computer Science & Engineering*
Karpagam University, Coimbatore
[*] *Dean, Faculty of Engineering*
Karpagam University, Coimbatore

Abstract— **Since its emergence from the early 1980s, the field of Human Computer Interaction has moved on and advanced in many significant ways. It has opened up a world in which communication between human and computer has become easier and richer. Among the different modes of interaction, Gestures provide the most natural and convenient way of communication. Hence gesture recognition has been extensively researched and many systems have been implemented. However because of the complexity in understanding gestures, to date, we have not succeeded in devising a perfect gesture recognition system. Gesture recognition involves 4 stages out of which the final gesture recognition stage is the most complicated. Various approaches have been studied in recognizing gestures with varying degrees of success. In this paper, we discuss several studies which employ Artificial Neural Networks as classifiers. The review focuses on those systems that use the Feed Forward Networks as classifiers, mainly Multi-Layer Perceptron Networks, Back Propagation Networks and Radial Basis Function Networks.**

Keywords— **Neural Networks, Human Computer Interaction, Back Propagation Neural Networks, Multi-Layer Perceptron, Radial Basis Function Networks, Gesture Recognition System**

I. INTRODUCTION

Human computer interaction emerged in the 1980's with the primitive text user interfaces or graphical user interfaces, where majority of the input was done using a keyboard and mouse [1]. Though these devices have become very familiar to us, they can create a bottle neck in the effective usage of information flow. They limit the naturalness of interacting with the computer. Computers have started taking roles in all aspects of our life, and hence the ways humans communicate with computers have become a crucial issue. It would be more advantageous if we could communicate with them naturally using gestures, as we do with other humans.

Gestures can be classified as static or dynamic. Static gesture is a posture which does not involve a movement but conveys information, for example hand postures for STOP hand signs. Dynamic gestures are a sequence of postures that involve some movement, as the 'goodbye' or 'come here' hand gestures. "A gesture is a bodily motion that conveys some information." [2]. Gestures can originate from the hand, face or body. Current research works focus more on hand and face gesture recognition as their application domains cover a wide range like Virtual Reality, human robot interaction [3], sign language recognition, game interaction etc.

The primary goal of a gesture recognition research [4] is to create a system to capture, analyze and recognize hand gestures and to use them for device control. A hand recognition system mainly implements 4 stages - data acquisition, gesture modeling, feature extraction and gesture recognition. In the early days data acquisition was accomplished using either color markers [5] or mechanical gloves equipped with various kinds of sensors such as the Data Glove. But these methods required the user to wear a cumbersome glove which hampers the naturalness of interaction. And they were also known for having problems in electro magnetically noisy environments. Later researches were carried out employing the vision based approach [6], [7] where video based non contact interaction techniques were employed. After capturing the images using suitable input devices, the gesture has to be segmented from the background and then it is processed for noise removal, edge detection and normalization. This stage prepares the data for the computational stage for feature extraction. The extracted features are finally evaluated and subjected to gesture matching for recognition. The number of gestures that can be recognized depends on the database that has been pre set.

Several methods like template matching, statistical matching, dictionary lookup, adhoc method, linguistic matching and neural networks have been used for gesture recognition [8]. In this paper we provide a survey on the use of Feed Forward neural networks in gesture recognition. The advantage of using Neural Networks in gesture recognition is its ability to capture the complexity of the gestures. However, Neural Networks have to be extensively tuned to produce desired results. Section 2 describes about Artificial Neural Networks in general. Section 3 shows the comparisons of different researches conducted using Back Propagation Networks and Multi Layer Perceptron Networks. Section 4 outlines the conclusion drawn from the comparisons.

II. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks are biologically inspired; that is, they are composed of elements that perform in a manner that is analogous to the most elementary functions of the biological neuron. They resemble the brain in two aspects:
a) Learning – ANN can modify their behaviour in response to the environment.
b) Inter-neuron connection strengths known as synaptic weights are used to store knowledge.

An artificial neuron [Fig1] is an information processing unit that is fundamental to the operation of neural networks. Each neuron has a set of inputs which represent output of other neurons. Every input is multiplied by a corresponding weight analogous to a synaptic strength, and then all of these weighted inputs are summed to determine the activation level of the neuron. [9]

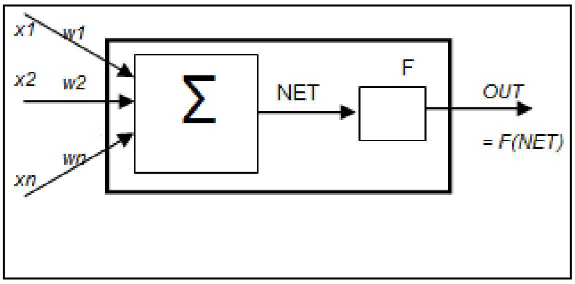

Fig1 : Artificial Neuron with Activation Function

The summation block, which corresponds roughly to the biological cell body, adds all the weighted inputs, producing an output **NET.** This can be stated in vector notation as :

$$\boxed{\mathbf{NET = XW}}$$

F is processed by an activation function to produce the neuron's output signal, which can be represented by the following linear function.

$$\boxed{\mathbf{OUT = K\ (NET)}}$$

where K is a constant, threshold function.

A. Types of Neural Networks

1) Single Layer Perceptron Networks: The power of neural computing comes from connecting neurons into networks. A single layer perceptron (SLP) [10] is the simplest form of a neural network used for the classification of patterns that are linearly separable. SLP is concerned with only a single neuron. The output layer can be expanded to include more neurons, but it is required that the classes are linearly separable for the perceptron to function properly.

2) *Multi Layer Perceptron Networks:* The Multi Layer Perceptron Network (MLP) [10] is a network with one or more hidden layers of computation between the input and output layers. They are usually networks that are fully connected. Several MLP Networks have been studied, which includes the Back Propagation Network (BPN) and the Radial basis Function Network (RBFN).

3) *Back Propagation Networks:* BPN [9] is a multi-layer forward network using extend gradient-descent based delta learning rule (back propagation of errors rule). BPN is a computationally efficient method for altering the weights in a feed forward network, using an activation function. The back propagation requires only that the activation function has to be differentiable. The sigmoid function satisfies this condition. This function can be mathematically expressed as:

$$\boxed{\mathbf{F(x) = 1/1+exp\ (-x)}}$$

4) *Radial Basis Function Networks:* RBFN [10] is a feed forward network which uses a Gaussian potential function as its activation function. The response of such a function is non-negative for all value of x. The function can be expressed as:

$$\boxed{\mathbf{F(x) = exp\ (-x^2)}}$$

The Neural Networks discussed in the above sections fall into the category of supervised training [10], where the network is inputted with a series of sample inputs and then the outputs are compared with the expected responses. The networks are trained until the expected response is achieved, or until the error comes to a minimum. The weights are adjusted according to a learning algorithm.

III. NEURAL NETWORKS BASED CLASSIFIERS

Neural Networks, once trained, are to a certain degree insensitive to minor variations in input. This is a most desirable feature which allows them to deal with the imperfect world. In addition to this, their ability to learn and self-adjust enables us to train ANN’s to produce consistent responses. Different forms of Neural Networks have been tried and tested for gesture recognition applications, giving varied levels of accuracy.

Back propagation (BP) networks are widely used and tested for this purpose. Jiafeng Zhang [11] used BP network to achieve a recognition rate of 83.3% for predefined gestures. Ten gestures were predefined to express different semantics. The hand gestures were first segmented using optical flow vectors from real-time video stream. To compute the optical flow, it was assumed that: the object being imaged is a flat surface, the illumination on the image is uniform and that the reflectance of the object varies smoothly. These assumptions assure that the image brightness or intensity is differentiable. The motion segmentation was represented as a series of 12 vectors that came from consecutive video frames. These vectors were also normalized before becoming part of the training set.

Jinwen Wei [12] achieves a correctness ratio of 95% using BP networks and a binary decision tree for the recognition of 12 hand shapes. Here a touchpad is used to sense the touching of the fingers and a video camera which is placed right above the touchpad is used to capture the images. Each finger can assume 4 statuses: stretching-touching on the touchpad, retracting -touching on the touchpad, stretching-detaching over the touchpad, and retracting-detaching over the touchpad. Three schemes of BP networks were tested and compared in the study.

Hamid [13] uses a supervised feed-forward neural network with back propagation algorithm for recognition. Wavelet network algorithm is used for image segmentation and feature extraction. The hand gestures were captured using a digital camera under natural light settings. Six hand gestures were chosen and a total of 120 images were used for training and another 60 for testing. He managed to achieve a 97% recognition rate with the captured data.

Haitham [14] proposed a system to recognize 6 hand gestures. The input images were captured from a single subject under different lighting conditions. Two methods are compared: the first method extracts the hand contour as a geometric feature. The second method deals with the problem of rotation by using the hand complex moments feature. The first method showed a performance of 70.83% recognition rate and second method showed a better performance of 86.83% recognition rate.

Ednaldo [15] proposes a system with a two level MLP architecture for the recognition of the Brazilian Sign Language (LIBRAS). Hand gestures were captured using a camera using 45 inexperienced people to replicate 18 alphabets. Image segmentation was performed using binarization and edge detection (Canny Algorithm). Signs with similar hand postures were grouped together and applied to a 1-layer and 2-layer MLP. MLP-1L exhibited 84.80% correct recognition for hand postures and the MLP-2L produced a 96% recognition rate.

Vatavu [16] presented a system for interacting with virtual environment. Image is captured using a top-mounted camera that monitors the user's hands on the working desktop, under controlled lighting so as to maintain a good contrast between the user's hands and the background. Hands segmentation is achieved using a simple low cost skin filtering [17]in the HSV colour space on the hue and saturation components. Recognition is performed using a multi-layered Perceptron, organized using a 3-layer structure of 39 neurons. The results show a level of accuracy of 92%.

A Radial Basis Function Neural Network (RBFNN) is proposed by Dipak [18]. Segmentation of the hand gesture is done using a histogram based thresholding. Then the segmented gesture is rotated to make it rotation invariant. To remove the background noise morphological filtering techniques are employed. A localized contour sequence (LCS) is selected as a feature set of the hand gesture. A k-mean based radial basis function neural network (RBFNN) is used here for classification of hand gestures. The experiment is conducted on 500 train images and 500 test images of 25 static hand gestures. They were able to achieve 99.6% classification accuracy with the proposed method.

Comparisons between the chosen methods have been accomplished taking into consideration certain important factors, and this is presented in Table 1.

TABLE I
FEATURE COMPARISON OF NEURAL NETWORK CLASSIFIERS

Journal Reference	Image Capture	Gesture Modeling	Feature Extraction	Type of NN	Activation function	No. of gestures: Image data set	Recognition Rate
[13]	Digital camera under natural lighting.	Wavelet network algorithm	Wavelet network algorithm	BP Network	Binary-sigmoid. Linear transfer function was used for the output layer.	6 hand gestures: 120 images for training and 60 images for testing	97%
[12]	Touchpad to sense finger touch and video camera.	2-D filtering and edge detection.	Relative height of fingers obtained by analyzing scan lines.	BP Network	Sigmoid function	10 basic gestures: 144 images to train	95%
[14]	Digital camera from one subject under different lighting condition.	Thresholding algorithm Median filter. Sobel operator for edge detection.	Hand contour detection algorithm. General features represented as matrix.	BP Network	Binary sigmoid function.	6 gestures: 30 images for training and 84 images with scaling, translation	86.38%
[15]	Image captured using video camera in a dark background.	Binarization	Edge Detection	MLP Network	Hyperbolic tangent function.	100 images of each hand posture and 50 to test the data.	90.7%
[11]	Real time video stream with flat surface and uniform lighting.	Optical flow vector	Vector representation of motion.	BP Network		12 Hand shapes	83.3%
[16]	Top mounted video camera which auto controls the brightness and exposure settings.	Skin Filtering in HSV space.		MLP Network		4 hand postures: 152images for training 67 images for testing.	92%
[18]		Histogram based threshold algorithm. Rotation Morphological filtering technique	Localised Contour Sequence	RBFN	Gaussian Function	25 static gestures: 500 training images and 500 test images	99.6%

IV. Conclusions

In this paper we have reviewed different gesture recognition systems that employ Neural Networks as classifiers for gesture recognition. The advantage of ANN classifier over other classifiers is its self-learning ability on-line. This enables the network to decrease the error ratio with the increase in training samples. All papers that have been reviewed have used Feed Forward Networks, mainly the Back propagation Network, Radial Basis Function Neural Network and the Multi Layer Perceptron Network. As can be seen, the different systems produce different recognition rates, even in cases where the type of Neural Network used is same. The success of the gesture recognition stage depends largely on the successful modeling and feature extraction stages. So n order to have a robust classification, it is very important to choose the features properly and to present the features to the classifier in an appropriate way. The overall efficiency of the system depends on the right combination of the various stages.

References

[1] P. Garg, N. Aggarwal and S.Sofat, "Vision Based Hand Gesture Recognition," *World Academy of Science, Engineering and Technology,* pp. 821 - 826, 2009.

[2] Wikipedia, "Gesture recognition," [Online]. Available: http://en.m.wikipedia.org/wiki/Gesture_recognition.

[3] J. L. Raheja, R. Shyam and P. B. P. Umesh Kumar, "Real-time Robotic Hand Control Using Hand Gestures," *IEEE,* pp. 12-16, 2010.

[4] A. R. Sarkar, G. Sanyal and S.Majumder, "Hand Gesture Recognition Systems: A Survey," *International Journal of Computer Applications,* vol. 71, no. 15, pp. 26-37, 2013.

[5] R. Y.Wang and J. Popovic, "Real-time Hand Tracking with a Color Glove," *ACM Transactions on Graphics,* 2009.

[6] G. Murthy and R. Jadon, "A Review of Vision Based Hand Gestures Recognition," *International Journal of Information Technology and Knowledge Management,* vol. 2, no. 2, pp. 405-410, 2009.

[7] N. A. Ibraheem and R. Z.Khan, "Vision Based Gesture Recognition Using Neural networks Approaches: A Review," *International Journal of Human Compter Interaction,* vol. 3, no. 1, 2012.

[8] V. I. Pavlovic, R. Sharma and T. S.Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction : A Review," *IEEE Transactions on Pattern Analysis And machine Intelligence,* vol. 19, pp. 677 - 695, 1997.

[9] P. D.Wasserman, Neural Computing : Theory and Practice, New York: Van Nostrand Reinhold, 1989.

[10] S. N. Sivanandam, S. .Sumathi and S.N.Deepa, Introduction to Neural Networks Using MATLAB 6.0, Delhi: Mc Graw Hill, 2006.

[11] J. Zhang, F. Zhang and M. Ito, "Image processing based remote control with robot arm simulator," *SICE,* pp. 2344 - 2348, 2009.

[12] H. Q. J. G. Jinwen Wei and Y. C. Jinwen Wei, "The Hand Shape Recognition of Human Computer Interaction with Artificial Neural Network," *IEEE,* p. 3809, 2009.

[13] H. A. Jalab, "Static Hand Gesture Recognition for Human Computer Interaction," *Information Technology Journal,* pp. 1265 - 1271, 2012.

[14] H. Hasan and S.Abdul-Kareem, "Static Hand Gesture Recognition using Neural Networks," *Springer,* 2012.

[15] E. B.Pizzolato, M. d. S. Anjo and G. C.Pedroso, "Automatic Recognition of Finger Spelling for LIBRAS based on a Two-Layer Architecture," *ACM,* pp. 969-973, 2010.

[16] R. D. Vatavu, C. Chaillou, G. L, P. .S and D. .S, "Visual Recognition of Hand Postures for Interacting with Virtual Environments," in *Int. Conference on Development and Application Systems.* , 2006.

[17] C. M. Sharma and S. Saxena, "A Context-aware Approach for Detecting Skin Colored Pixels in Images," *International Journal of Computer Applications,* vol. 71, pp. 8-13, 2013.

[18] D. K. Ghosh and S. Ari, "A Static Hand Gesture Recognition Algorithm Using K-Mean Radial Basis Function Neural Network," *IEEE,* 2011.