International Journal of Computer Architecture and Mobility
(ISSN 2319-9229) Volume 1-Issue 6, April 2013

Vision Based Hand Gesture Recognition using Artificial Neural Network

Preeti.S.Ratnaparkhi[#1], Devesh.D.Nawgaje[*2]

Department of E & TC

Shri Sant Gajanan Maharaj College of Engineering,Shegaon
preetiratnaparkhi08@gmail.com
dnawgaje@gmail.com

***Abstract*— We propose a method of hand gesture recognition consisting of three phase feature extraction, neural training, testing. This project is limited the hardware part to computer and webcam, we just need to consider the software and programming parts. We use the vision based approach for gesture recognition. The aim of gesture recognition researches is to create system that can easily identify gestures, and use them for device control, or convey some formations without using the normal input devices which are the keyboard and mouse. There are several approaches for gesture recognition using MATLAB. Artificial Neural networks are flexible in a changing environment. System should be able to get static image through the webcam and perform the classification. The implemented system should able to perform classification correctly. Canny edge detection technique is used to find the boundary of hand gesture in image.**

Keywords: ***Gesture Recognition, Artificial Intelligence, classification, Histogram, Segmentation***

I. INTRODUCTION

In the recent year framework of interactive, intelligent computing, an efficient human computer interaction is assuming utmost importance in our daily lives. Human gesture typically constitutes the space of motion expressed by the body, face or hand. Among these, hand gesture is the most expressive & most frequently used. Gesture can also be defined as a meaningful physical movement of the fingers, hands & arm or non-verbal interaction among people. Gesture recognition can be termed as an approach in this direction.

Gestures can be static (posture or certain pose) which require less computational complexity [6] or dynamic (sequence of postures) which are more complex but suitable for real time environments [2] [3]. Different methods have been proposed for acquiring information necessary for recognition gestures system [4][5]. Some methods used additional hardware devices such as data glove devices and color markers to easily extract comprehensive description of gesture features [4]. Other methods based on the appearance of the hand using the skin color to segment the hand and extract necessary features [4], these methods considered easy, natural and less cost comparing with methods mentioned before [4].

The use of hand gestures provides an attractive alternative to cumbersome interface devices for human-computer interaction (HCI). There are two approach for gesture recognition vision based and glove based. Vision based interfaces are feasible and popular at this moment because the computer is able to communicate with user using webcam. This means, user able to give command to the computer by just showing some actions in front of the webcam without typing keyboard and clicking mouse button. Canny edge detection technique is used to find the boundary of hand gesture in image.This paper gives the overall detail information of gesture recognition procedure.

The application of gesture system on interactive applications produces many challenges. The first and important challenge is the response time which should be fast [1]. There should be no noticeable time between user gesture movement and computer replies [1]. The designed computer vision algorithms should be reliable and work for different ethnic people [1] especially when the color of human is changed comparing with white and black people. One more challenge which is the cost challenge, the gesture system needs special hardware such as the camera and sensors as necessarly, those special hardware will be the replacement of the existing hardware devises which may considered as low cost [1] such as the keyboard and mouse, but the gesture system with these new devices will be more worthwhile for wire-less communication.

II. RELATED WORK

Xingyan [7] applied gesture recognition using fuzzy C-means algorithm, He has used image processing methods to transform the raw image into the feature vector. The feature vector is created by applying segmentation using HSV color model and then he has reduced the noise, the feature vector of the image is thirteen parameters long. The first feature is the aspect ratio of the hand's bounding box as decided by the authors in [7]. The last 12 features are values representing

coarse parameters of the image, where each grid cell is the mean gray level value in the 3 by 4 block division of image. Each of the 12 values calculated by the mean value of 3 by 4 partitions, which represent the mean of the brightness value. The classification phase is applied using a recognition algorithms based on the Fuzzy C-Means (FCM) algorithm, he had used samples taken from 6 persons with 6 gestures each and achieves a time of 2-4 seconds with 86 % recognition rate.

P. Garg [6] used the vision based technique for gesture recognition. For vision based technique basically no specialized hardware is used, Only single camera is used. Supervised feed-forward neural net based training and back propagation algorithm for classifying hand gestures. Feature extraction is necessary, and it is done very accurately. This method gives the recognition rate 94.6%.

Fuzzy expert systems has also been investigated for gesture recognition[8] based on analyzing complex features of the signal like the doppler spectrum. The disadvantage of these methods is that the classification is based on the separability of the features, therefore two different gestures with similar values for these features may be difficult to classify.

Method [7] used 3D Imaging data glove & 3D electromagnetic sensor. The advantage of this method is that it is a very high precision technique but the disadvantage is that Forces the user to carry a load of cables which are connected to the computer and hinders the ease and naturalness of the user interaction. The accuracy of this method is 91.4.

Gonzalo Bailador[8] Continuous time Recurrent Neural Network is used. This approach is based on the idea of creating specialized signal predictors for each gesture class. Thes forecast future acceleration values from current ones. The errors between the measured acceleration of a given gesture and the predictors are used for classification. This approach is modular and allows for seamless inclusion of new gesture classes. These predictors are implemented using Continuous Time Recurrent Neural Networks (CTRNN). On the one hand, this kind of networks exhibits rich dynamical behaviour that is useful in gesture recognition and on the other, they have a relatively low computational cost that is interesting feature for real time systems.

A new method [9] for hand gesture recognition that is based on a hand gesture fitting procedure via a new Self-Growing andSelf Organized Neural Gas(SGONG) network is proposed. Initially, the region of the hand is detected by applying a color segmentation technique based on a skin color filtering procedure in the YCbCr color space. Then, the SGONG network is applied on the hand are a so as to approach its shape. Based on the output grid of neurons produced by the neural network, palm morphologic characteristics are extracted.The secharacteristics,in accordance with powerful finger features,allow the identification of the raised fingers.Finally,the hand gesture recognition is a ccomplished through alike lihood-based classification technique.The proposed system has been extensively tested with success.

III. METHODOLOGY

In order to be implemented in real-time systems, the hand gesture detection algorithm for automatic picture taking has to meet certain criteria, such as fast processing time and low computation complexity. Our system is composed of four main segments in order to recognize the input gesture, these stages are summarized below

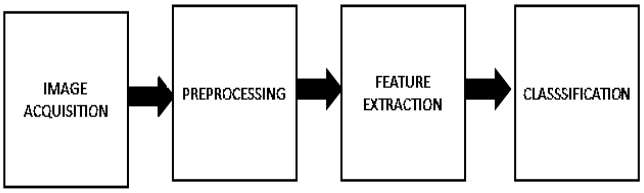

Fig.1 Gesture Recognition System

A. Image Database

The starting point of the project was the creation of a database with all the images that would be used for training and testing. The image database can have different formats. Images can be either hand drawn, digitized photographs or a 3D dimensional hand. Photographs were used, as they are the most realistic approach. . Images came from two main sources. VariousASL databases on the Internet and photographs taken with a digital camera. This meant that they have different sizes, different resolutions and sometimes almost completely different angles of shooting. Images belonging to the last case was very few but they were discarded, as there was no chance of classifying them correctly. This meant that they have different sizes, different resolutions and sometimes almost completely different angles of shooting. Images belonging to the last case was very few but they were discarded, as there was no chance of classifying them correctly.

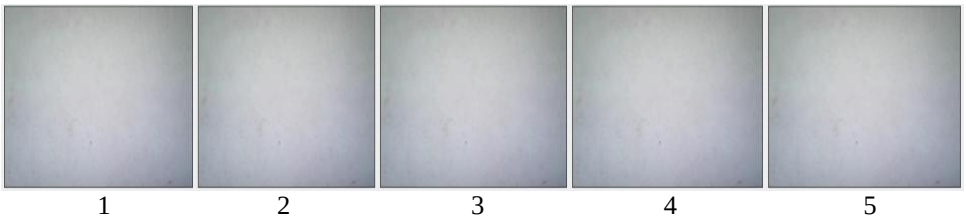

Fig. 2 Five background database

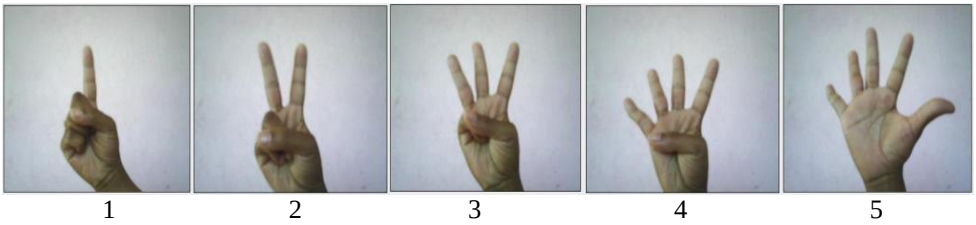

Fig. 3 Five hand gesture database

B. Image Preprocessing

Preprocessing means to extract the meaning from raw data. The web-cam captures the input and using an image differencing technique, the sequence of (x, y) coordinates representing the gesture is determined. This raw set of (x, y) coordinates will have to be preprocessed before it can be fed into the trained neural net for classification. One of the major limitations of neural nets is that they require a fixed number of inputs. Preprocessing must ensure that this condition is met. This means that a gesture with an inadequate number of inputs must not be passed onto the neural classifier or it must be 'enlarged' in an appropriate manner to meet the size requirement. A gesture that is too long must be sampled appropriately to fit the exact number of inputs in the neural classifier.

The resultant processed input can now be passed into the classifier. Yet further preprocessing can be performed. Preprocessing can also be used to extract further 'meaning' from the raw data and then passing the interpreted data onto the neural classifier. This has the general effect of improving the performance of neural nets. The *n* input sequence of (x, y) coordinates is preprocessed into a vector sequence, which is then passed into the trained neural net for classification. The general effect of this is improved gesture recognition performance. as compared to using raw (x, y) coordinates. Scaling can also be introduced to improve performance.

Segmentation is done to convert gray scale image into binary image, so that we can have only two objects in image one is hand and other is background. Segmentation is done to segment the hand area & isolate it from the background. Basically there are two methods for segmentation. These are HSV model based technique and Thresholding technique.HSV model based technique deals with the color pigment of the human skin. Thresholding technique depends on threshold value of probability. If the probability of a pixel is greater than or equal to threshold value, represents skin color. If Condition not satisfied does not represent skin color. Skin color pixels represents white and the other ones represents black

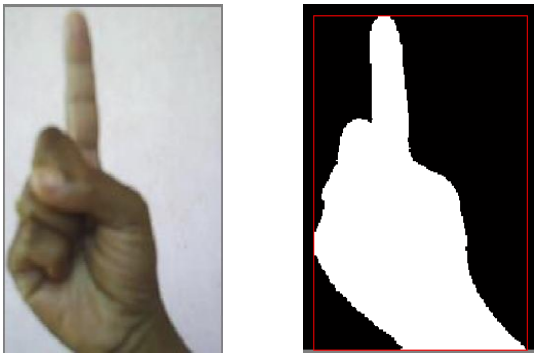

Fig. 4 Original image Binary image

C. *Feature Extrction*

Good segmentation process leads to perfect features extraction process and the latter play an important role in a successful recognition process.Feature extraction is very important in terms of giving input to a classifier .Features vector of the segmented image can be extracted in different ways according to particular application. Various methods have been applied for representing the features can be extracted. Some methods used the shape of the hand such as hand contour and silhouette while others utilized fingertips position, palm center, etc.

Our prime feature is local contour sequence (L.C.S) .In feature extraction first we have to find edge of the segmented

and morphological filtered image. Canny edge detector is used to find the edge which leads us to get boundary of hand in image. In image processing finding edge is fundamental problem because edge defines the boundaries of different objects. Edge can be defined as sudden or strong change in the intercity or we can say sudden jump in intensity from one pixel to other pixel. By finding the edge in any image we are just reducing some amount of data but we are preserving the shape.

Fig. 5 Fingertip edge detection

D Histogram

We want gestures to be the same regardless of where they occur with the images boarders. To achieve this we will ignore position altogether, and tabulate a histogram of how often each orientation element occurred in the image. Clearly, this throws out information and some distinct images will be confused by their orientation histograms. In practice, however, one can choose a set of training gestures with substantially different orientation histograms from each other. One can calculate the local orientation using image gradients. I used two 3 – tap x and y derivative filters. The outputs of the x and y derivative operators will be dx and dy. Then the gradient direction is a tan (dx, dy). I had decided to use the edge orientation as the only feature that will be 0presented to the neural network. The reason for this is that if the edge detector was good enough it would have allowed me to test the network with images from different databases. Another feature that could have been extracted from the image would be the gradient magnitude using the formula below [13]

$$\sqrt{dx^2 + dy^2}$$

This would lead though to testing the algorithm with only similar images. Apart from this the images before resized should be of approximately the same size. This is the size of the hand itself in the canvas and not the size of the canvas. Once the image has been processed the output will be a single vector containing a number of elements equal to the number of bins of the histogram.

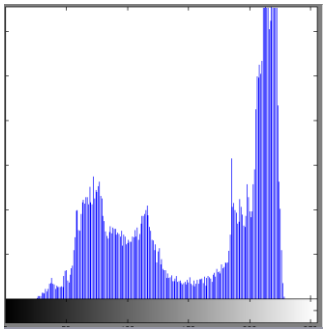

Fig. 6 Histogram

E. Neural Training

The set of input is passed through the trained Neural Network which classifies the gesture into one of several predefined classes that can be identified by the system. After modeling and analysis of the input hand image, gesture classification method is used to recognize the gesture. There many types of neural network are available. If classification is accurate then testing gives perfect result. Feed forward neural network is very simple and gives very good result. Radial basis function neural network[8] used for interpolation in multidimensional space. Neural Network can be used for both static and dynamic gesture recognition Artificial Neural Networks are one of the technologies that solved a broad range of problems in an easy and convenient manner. The working concept of Artificial Neural Networks (ANNs) is similar to human nervous system, hence it has synonym with the word neural networks. Artificial neuron is called perceptron. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques.

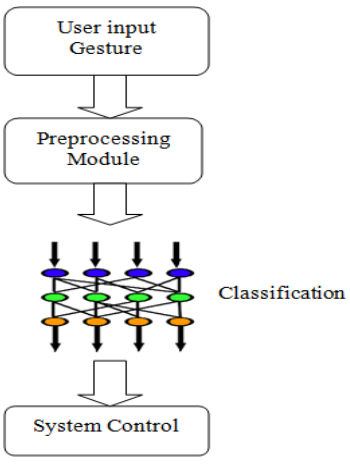

Fig. 7 One-hidden-layer feed forward neural network

Once the network weights and biases are initialized, the network is ready for training. The multilayer feedforward network can be trained for function approximation (nonlinear regression) or pattern recognition. The training process requires a set of examples of proper network behavior network inputs p and target outputs t. The process of training a neural network involves tuning the values of the weights and biases of the network to optimize network performance, as defined by the network performance function net perform function. The default performance function for feedforward networks is mean square error the average squared error between the network outputs a and the target outputs t. The output will compare with the *Target vector*. After that, if there is an error, the Perceptron network will re-adjust the weights value until there is no error or minimized and then it will stop. Each pass through the input vectors is called epoch. Neural network training and Performance graph is as shown below

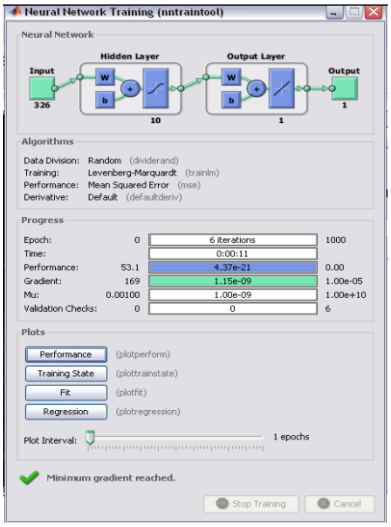

Fig. 8 Neural network training

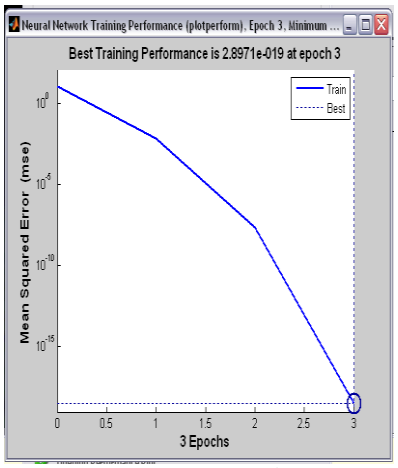

Fig. 9 Neural Network performance

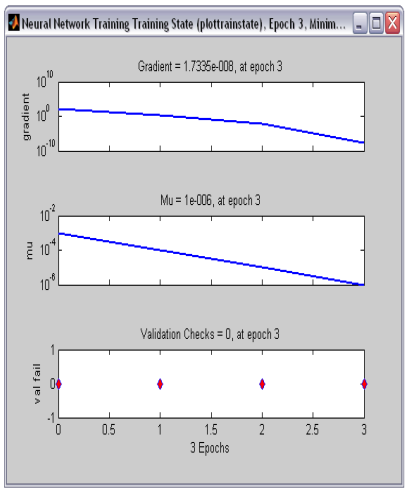

Fig. 10 Neural Network training state graph

IV. TESTING & VALIDATION

Now in the last phase we take the back snap and then test the hand gesture. By giving the one by hand gesture, the system matches the one hand gesture with previous taken database one. We set the (x,y) co-ordinate at the start button and if gesture is recognize successfully then cursor moves automatically at the start button. So 1 is recognize successfully. The accuracy of this system is very good as compared to the other method.

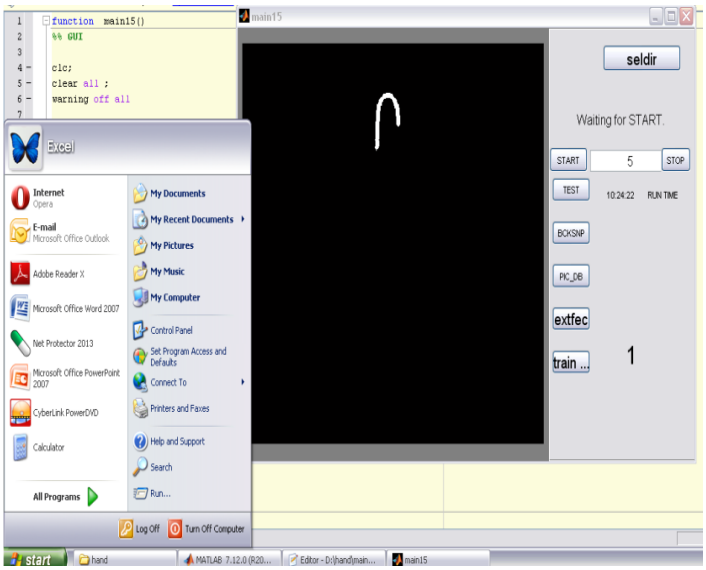

Fig. 11 Result of gesture recognition

V. CONCLUSION

Hand gesture recognition becomes important Interactive human-machine interface and virtual environment. We have captured the hand gesture images against a uniform background under normal illumination. The idea is to make the classification of gestures as key problem, not the segmentation. Artificial neural networks are emerging as the technology of choice for many applications, such as pattern recognition, gesture recognition, prediction, system identification, and control.ANN provides good and powerful solution for gesture recognition in MATLAB.

The major goal of this project is to develop a system that will aid in the interaction between human and computer through the use of hand gestures as a control commands.

REFERENCES

[1] K. Symeonidis. "Hand Gesture Recognition using Neural Networks", Master Thesis, School Of Electronic And Electrical Engineering, 2000

[2] Xingyan Li. (2003). "Gesture Recognition Based on Fuzzy C-Means Clustering Algorithm", Department of Computer Science. The University of Tennessee Knoxville.

[3] S. Mitra, and T. Acharya. (2007). "Gesture Recognition: A Survey" IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 37 (3), pp. 311- 324, doi: 10.1109/TSMCC.2007.893280.

[4] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). "A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012 173 Networks," IEEE Proceedings of the 9th International Conference on Neural Information Processing, Singapura.

[5] Joseph J. LaViola Jr., (1999). "A Survey of Hand Posture and Gesture Recognition Techniques and Technology", Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization, USA.

[6] P. Garg, N. Aggarwal & S. Sofat, (2009). "Vision Based Hand Gesture Recognition," World Academy of Science, Engineering and Technology, Vol. 49, pp 972-977.

[7] Y. Su, C. R. Allen, D. Geng, D. Burn, U. Brechany, G. D. Bell, and R. Rowland, "3D motion system ("data-gloves"): Application for Parkinson's disease," IEEE Trans. Instrum. Meas., vol. 52, no. 3, pp. 662–674, Jun. 2003.

[8] Real time gesture recognition using Continuous Time Recurrent Neural Networks by Gonzalo Bailador & Daniel Roggen & Gerhard Tröster & Gracián Triviño

[9] E. Stergiopoulou & N. Papamarkos. (2009) "Hand gesture recognition using a neural network shape fitting technique,"Elsevier Engineering Applications of Artificial Intelligence, Vol. 22, No. 8, pp 1141–1158, doi: 10.1016/j.engappai.2009.03.008

[10] L. Fausett, "Fundamentals of Neural Networks, Architectures, Algorithms, and Applications", Prentice-Hall, Inc. 1994, pp-3.

[11] E. Stergiopoulou & N. Papamarkos. (2009) "Hand gesture recognition using a neural network shape fitting technique," Elsevier Engineering Applications of Artificial Intelligence, Vol. 22, No. 8, pp 1141–1158, doi: 10.1016/j.engappai.2009.03.008

[12] Mahmoud E., Ayoub A., J¨org A., & Bernd M., (2008) "Hidden Markov Model-Based Isolated and Meaningful Hand Gesture Recognition", World Academy of Science, Engineering and Technology, Vol. 41, pp 393-400.

[13] W. Freeman, and M. Roth, "Orientation Histograms for Hand Gesture Recognition", Technical Report 94-03a, Mitsubishi Electric Research, Laboratories Cambridge Research Center, 1994

[14] Aditya Ramamoorthy et al. "Recognition of dynamic hand gestures ", page 1-13. Department of Electrical Engineering IIT New Delhi-110016 India.

[15] Kouichi M. & Hitomi T. (1999) "Gesture Recognition using Recurrent Neural Networks, " ACM SIGCHI conference on Human factors in computing systems: Reaching through technology (CHI'91), pp 237-242, doi: 10.1145/108844.108900

[16] Dionisio. C.R.P., Cesar. R.M. and Jr.: A Project for Hand Gesture Recognition", Proc. of IEEE Symposium on Computer Graphics and Image Processing, pp. 345, 2000.

[17] Manar Maraqa & Raed Abu-Zaiter, (2008) "Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks," IEEE First International Conference on the Applications of Digital Information and Web Technologies, (ICADIWT 2008), pp 478-48. Doi: 10.1109/ICADIWT.2008.4664396

[18] Matthew A. Turk and Alex P. Pentland, " Face recognition using eigenfaces", IEEE Society Conference on Computer Vision and Pattern Recognition, pages 586–591, Lahaina, Maui, Hawaii, June 3-6 1991.