

Comparative Study of AI Based Gesture Recognition

Preeti.S.Ratnaparkhi and Devesh D. Nawgaje

Abstract: Gesture recognition plays a very vital role in Human Computer Interaction. In this paper a survey of recent hand gesture recognition systems is presented. This research paper gives the overview of different methods for gesture recognition. A comparative study included in this paper with focusing on different AI technique, research advantages and drawbacks are provided as well. This research paper gives the overview of AI for gesture recognition. It also describes the process of gesture recognition using AI.

Keywords: Gesture Recognition, Artificial Intelligence, Human Computer Interaction, Classification.

I. INTRODUCTION

With the development of information technology in our society, one can expect that computer systems to a larger extent will be embedded into our daily life. People use the computer either at their work or in their spare time. The communicating with computers at this moment are limited to mouse, keyboard, track ball, web-cam, light pen and etc. But still Interaction with computers are not comfortable experience. People should communicate with computer with body language. Hand gesture recognition becomes important Interactive human-machine interface and virtual environment. A primary goal of gesture recognition research is to create a system which can identify specific human gestures and use them to convey information or for device control. Interaction between humans comes from different sensory modes like gesture, speech, facial and body expressions [2]. Gesture recognition is the process by which gestures made by the user are made known to the system. [13] Gesture recognition is also important for developing alternative human-computer interaction modalities [14]. It enables human to interface with machine in a more natural way.

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. We develop *gesture recognition techniques which in turn results in developing a low cost interface device for interacting with objects in virtual environment using hand gestures.*

Preeti.S.Ratnaparkhi is staying Behind Wankhede Petrol Pump Malkapur, and and Devesh D. Nawgaje is with ²Department of E & TC ,SSGMCE Shegaon, Shri Sant Gajanan Maharaj College of Engineering, Shegaon, preetiratnaparkhi08@gmail.com, dnawgaje@gmail.com.

II. GESTURE RECOGNITION

Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures via mathematical algorithms. Gestures can originate from any bodily motion or state but commonly originate from the face or hand.

A. TYPES OF GESTURE RECOGNITION

Gesture recognition is mainly divided into following 3 types

- Face Recognition
- Facial expression Recognition
- Hand gesture recognition

These 3 types again divided into 4 subtypes

- 1) Static Gesture Recognition - Static gesture have less computational complexity.
- 2) Dynamic Gesture Recognition - Dynamic gesture have more computational complexity.
- 3) On-line Gesture Recognition- Direct manipulation gestures.
- 4) Off-line gesture Recognition- Those gestures that are processed after the user interaction with the object.

III. Block Diagram of Gesture Recognition System

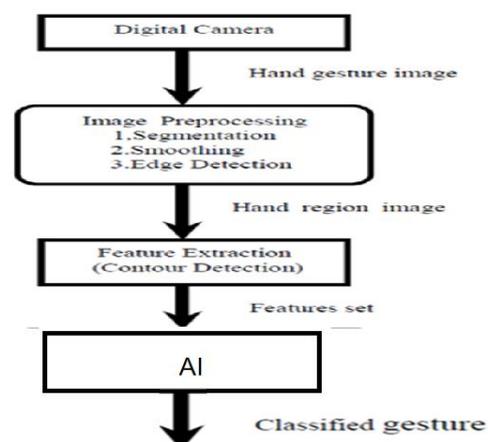


FIGURE 1 Basic block diagram of Gesture Recognition system

Collect the input: The starting point of the project was the creation of a database with all the images that would be used for training and testing. Photographs were used, as they are the most realistic approach. Capture the images using digital camera.

Image Preprocessing: Next most important step is image preprocessing. Basically preprocessing means to extract the meaning from raw data. When the Gesture is determined,

this raw image 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 input gestures are used for both training and testing purposes this phase is divided down into the following sub phases.

Segmentation
Smoothing
Edge Detection

Segmentation: The main objective of image segmentation is to extract various feature of the image which can be merged or split in order to build objects of interest on which analysis & interpretation can be performed. 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

Edge Detection: Edge detection is the process of finding the meaningful transition in an image. The purpose of edge detection is to identify areas of an image where large change in intensity occurs. Intensity jump between the connected pixel is called edge and non-edge areas do not has any intensity difference. Edge detection preserves the important information in the images that represents the edges. Edge detection is usually done with local linear gradient operator .We use the Canny edge detector By finding the edge in any image we are just reducing some amount of data but we are preserving the shape.

Feature Extraction: After the preprocessing of the image and segmenting of the hand gesture, a black-white image is created and represented the hand pose inset; the feature extraction phase will start. Feature Extraction seeks to identify inherent characteristics, features of objects. These characteristics are used to describe the object, or attribute of the object, prior to the subsequent task of classification. Feature extraction operates on two-dimensional image *arrays* but produces a *list* of descriptions, or a ' feature vector'. For posture recognition, (static hand gestures) features such as fingertips, finger directions and hand's contours can be extracted. But such features are not always available due to self-occlusion and lighting conditions. Feature extraction is a complex problem, and often the whole image or transformed image is taken as input. Features are thus selected implicitly and automatically by the recognizer [9]. In this paper we select the hand contour as a good feature to describe the hand gesture shape. Contour detection process consists of two steps: first find the edge response at all points

in an image using gradient computation and in the second step modulate the edge response at a point by the response in its surround [10] and extract intersection point. second step modulate the edge response at a point by the response in its surround [10] and extract intersection point.

Artificial Intelligence: Artificial intelligence is used for classification purpose. Under AI we can use Neural Network, Fuzzy logic and Genetic algorithm, Neuro Fuzzy. But as there are many types of neural are available so different technique can be possible by using different neural network. Fuzzy rule-based methods are also used for gesture recognition. Under fuzzy mostly fuzzy clustering algorithm is used [3]. Fuzzy C-Means Clustering Algorithm shows good performance with complex background using fuzzy [3].

IV.EXISTING TECHNIQUE FOR GESTURE RECOGNITION

- Data gloves based system
- Vision based system

A. DATA GLOVES BASED SYSTEM

Data-Gloves approaches are based on the use of sensors devices, which digitize the human hand and finger movements in input parameters for a virtual reality simulation system. These methods employs mechanical or optical sensors Attached to a glove that transforms finger flexions into electrical signals to determine the hand posture [6]. Using this method the data is collected by one or more data- glove instruments which have different measures for the joint angles of the hand and degree of freedom (DOF) that contain data position and orientation of the hand used for tracking the hand [7]. However, this method requires the glove must be worn and a wearisome device with a load of cables connected to the computer, which will hampers the naturalness of user-computer interaction [5].



FIGURE 2 Data glove based system

B. VISION BASED SYSTEM:

There are few technologies already using vision based analysis system. Example: Arabic Sign Language Recognition[13]. Vision based approaches use image capturing devices. In this way a more natural interaction is achieved. a vision based gesture recognition using a simple system connected with a web camera. When the camera

capture the image of hand gesture, the system extract the human hand region which is the region of interest (ROI) using the intensity color information. The system is obtained the motion velocity and the direction by tracking the center of gravity (COG) of the hand region, which provides the speed of any conducting time pattern.[8]. Vision based analysis, is based on the way human beings perceive information about their surroundings, yet it is probably the most difficult to implement in a satisfactory way. Several different approaches have been tested so far.

One is to build a three-dimensional model [18] of the human hand. The model is matched to images of the hand by one or more cameras, and parameters corresponding to palm orientation and joint angles are estimated. These parameters are then used to perform gesture classification.

Second one to capture the image using a camera then extract some feature and those features are used as input in a classification algorithm for classification [19].

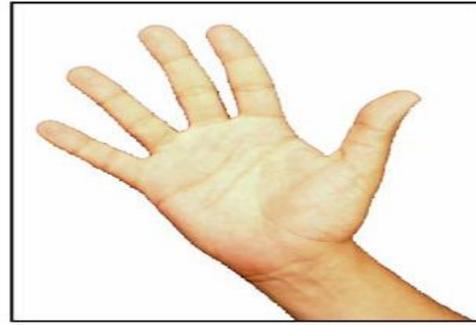


FIGURE 3 Vision based system

C. COMPARISON BETWEEN GLOVE BASED & VISION BASED METHOD

Referenc e	Method used	Advantage	Disadvantage	%of Accurac y
22	3D Imaging data glove & 3D electromagnetic sensor is used.	High precision technique .	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.	91.4
4	Web Camera is used. Supervised feed-forward neural net based training and back propagation algorithm for classifying hand gestures.	No specialized hardware is used. Only single camera is used	Feature extraction is necessary	94.6

D. COMPARISON BETWEEN AI TECHNIQUES.

Fuzzy set theory most widely used theory for soft computing, which deals with the design of flexible information processing systems, with applications in control systems, decision making, expert systems etc. Fuzzy can also be used in gesture recognition system. Fuzzy can be used for both data glove and vision based method. Mostly fuzzy is used for data glove approach. When we used the data glove method using fuzzy so to deal with the uncertainties in the data provided by the data glove, an approach based on interval fuzzy logic is used [20]. Under Fuzzy basically Fuzzy C Means algorithm is used to classify the hand gesture since it has good speed in recognizing gestures with sufficient accuracy for real-time operation[3]. To transform an expressed gesture into a meaningful statement can be a computationally intensive task and may not be easy to achieve in real-time Fuzzy Logic [21] can be a solution. Fuzzy Logic can help to find (and evaluate) strict aspects which describe a gesture’s behavior good enough to identify it. Neural Network can also be used for classification purpose. 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 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. All these neural network are used in different gesture recognition system. So each network has its own advantage and disadvantage. Some are listed below

- Feed forward Neural Network
- Radial Basis Function (RBF) Neural

- Kohonen Self-organizing Neural Network
- Learning Vector Quantization Neural Network
- Continuous Time Recurrent Neural Network
- Self-Growing and Self-Organized Neural Gas (SGONG) network
- Elman Recurrent Neural Network

Refer ence	Method used	Advantage	Disadvantage	% of accuracy
8	Continuous time Recurrent Neural Network is used. Use of signal predictors to recognize gestures. Tri-axial accelerometer is used to capture the gesture	Low computational cost. Use inexpensive accelerometer	Classification based on separability. Recognition accuracy depend on segmentation Signal Predictor are fast, simple and modular	94
11	Shape fitting technique is used. Self-Growing and Self-Organized Neural Gas (SGONG) network is used for classification. Skin color filtering is used for color segmentation .	The exact shape of the hand was obtained which led to good feature extraction	Input images include exactly one hand Gestures are made with the right hand only The arm must be Vertical The palm is facing the camera, and the image background is plain	90.45
16	Feed forward neural network & back propagation algorithm is used. Sigmoid activation function used. Implemented application used for video gaming	Perform accurate classification. Structure is very simple Low computational complexity	Feature extraction is very lengthy procedure.	96.44
15	Back propagation Neural Network used for postures recognition and Elman Recurrent Neural Network For gesture recognition. Thresholding is used for segmentation. 13 data item for postures/ 16 data For gestures are taken.	Simple and active, and successfully can recognize a word and alphabet. Automatic sampling, and augmented filtering data improved the system performance.	Required long time for Learning	90
12	Hidden Markov Model (HMM) is used. Gaussian Mixture Model (GMM) Used for skin color detection. For features extraction, the orientation between the centroid points of current frame and previous frame were determined by vector	Recognized both isolated and meaningful gestures for Arabic numbers.	Recognition limited to numbers only.	89

	quantization.			
23	Radial Basis Function neural network is used. Histogram based thresholding algorithm is used. A localized contour sequence (LCS) based feature is used here to classify the hand gestures.	This network is faster Rotation invariant gesture recognition is possible	Algorithm used is very complex	99.6
3	Fuzzy C-Means Clustering algorithm is used. Implemented application is used in mobile remote. The system compares sequential images to track an object	Good performance with complex background	Performance decreases when distance greater than 1.5m between user & camera	85.83
6	Gesture recognition based on the local brightness of each block of the gesture image. Colored segmentation is done using HSV	Efficient and intuitive interaction between the human and the computer The system successfully recognized static and dynamic gestures	The database samples, variation in scale, translation and rotation led to a misclassification	91

V.CONCLUSION

In this paper various methods are discussed for gesture recognition, these methods include from Neural Network, HMM, fuzzy c-means clustering. This paper gives detailed comparison of artificial intelligence based gesture recognition system. We studied both neural network based and fuzzy logic gesture recognition system. Here we also mention advantages and disadvantages of different method. ANN provides good and powerful solution for gesture recognition. Feed forward neural network is very simple and gives very good result. Learning Vector Quantization Neural Network ration results in good classification scheme. For dynamic gestures HMM tools are perfect and have shown its efficiency especially for robot control . The single hand gesture recognition system works successfully for real-time static hand gesture recognition. From literature survey it is observed that neural network gives more accuracy as compared to the fuzzy logic approach.

REFERENCES

[1] Luigi Lamberti & Francesco Camastra, (2011) “Real-Time Hand Gesture Recognition Using a Color Glove”, Springer 16th international conference on Image analysis and processing: Part I (ICIAP'11), pp 365-373.
 [2] Ismail, N. A., & O’Brien, A., (2008), “ Enabling Multimodal Interaction in Web-Based Personal Digital Photo Browsing”, Proceedings of the International Conference on Computer and Communication Engineering , Kuala Lumpur, Malaysia, May 13-15, pp. 907-910.
 [3] Xingyan Li. (2003). “Gesture Recognition Based on Fuzzy C-Means Clustering Algorithm”,Department of Computer Science. The University of Tennessee Knoxville.

[4] P. Garg, N. Aggarwal & S. Sofat, (2009). “Vision Based Hand Gesture Recognition,” World Academy of Science, Engineering and Technology, Vol. 49, pp 972-977.
 [5] S. Mitra, and T. Acharya. (2007, May). “Gesture Recognition: A Survey” IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 37 (3), pp. 311- 324, available: doi: 10.1109/TSMCC.2007.893280.
 [6] M. M. Hasan & P. K. Mishra, (2011) “HSV Brightness Factor Matching for Gesture Recognition System”, International Journal of Image Processing (IJIP), Vol. 4, No. 5, pp 456-467.
 [7] Joseph J. LaViola Jr., “A Survey of Hand Posture and Gesture Recognition Techniques and Technology,” Master Thesis, NSF Science and Technology Center for Computer Graphics and Scientific Visualization, USA, 1999.
 [8] Real time gesture recognition using Continuous Time Recurrent Neural Networks by Gonzalo Bailador & Daniel Roggen & Gerhard Tröster & Gracián Triviño
 [9] K.Murakami, H.Taguchi: Gesture Recognition using Recurrent Neural Networks. In CHI 91 Conference Proceedings (pp. 237-242). ACM. 1991.
 [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] Daniel Thalmann, October 2002. “Gesture Recognition Motion Capture, Motion Retargeting, and Action Recognition”, EPFL – VRlab.

- [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.
- [19] J. Edward Jackson, "A Users Guide to Principal Components", Wiley Series in Probability and Mathematical Statistics, A Wiley-Interscience Publication, 1st edition, 1991.
- [20] Hand Gesture Recognition in an Interval Fuzzy Approach by B.R.C. BEDREGAL1, DIMAP, Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil G.P. DIMURO2, A.C.R. COSTA3, PPGINF, Universidade Católica de Pelotas, Pelotas, RS, Brazil. In *TEMA Tend. Mat. Apl. Comput.*, 8, No. 1 (2007), 21-31.
- [21] **H.-J. Zimmerman**, *Fuzzy Set Theory – and its Applications*, 2nd edition, Kluwer-Nijhoff Publishing, 1990
- [22] 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.
- [23] A Static Hand Gesture Recognition Algorithm Using K-Mean Based Radial Basis Function Neural Network by Dipak Kumar Ghosh & Samit Ari.