SLIC based Hand Gesture Recognition with Artificial Neural Network

Harpreet Kaur
PG Scholar
Department of Computer Science & Engineering
GZSCCCET, Bathinda, India

Er. Jyoti Rani
Assistant Professor
Department of Computer Science & Engineering
GZSCCCET, Bathinda, India

Abstract

In the past years, the recognition of gesture feature has been glamour attention as a natural human. The communication system can build the human relationships. The mode of communication will be verbal and non-verbal. The non-verbal communication is helpful for deaf people and also helpful in surgeries, remote control etc. There is no need of peripheral device to interact with the computer. This paper mainly focuses on recognition of hand gesture on the basis of simple linear iterative clustering and the implementation is done with the artificial Neural Network. Initially, trained the data by creating a cluster of hand region using SLIC and extract the hand region of interest. After training phase the system is tested on 10 gesture classes and shows the 99.0% average accuracy and 0.219 seconds Recognition time and finally drawn the confusion matrix.

Keywords: Artificial Neural Network, Hand Gesture Recognition, Simple Linear Iterative Clustering, Superpixel, Human Computer Interaction

I. INTRODUCTION

The Human relationships can build by the communication system. It enables easy to exchange ideas, feelings, emotions, and endearment to other human beings. The communication plays a very vital role. A message is being formulated and transmitted to the person or persons whether we are talking, walking, playing, sitting, or even sleeping. Communication can be in verbal or non – verbal form [1]. The verbal or spoken communication is the combination of speaking and listening. In this communication, people can express their views by face to face, telephone, radio or television and other media. In the non-verbal mean, a person can express their thoughts by body language, gestures (Swearing at someone, smiling, nodding, leaning forward), posture (Standing in a certain way to express emotions. Example-eye contact and Facial expressions like Frowns, Smiling. In today's era only 35% verbal communication and 65% non-verbal communication occurs in between the two persons as shown in fig.1. The Gun Shoot in the air is the example of non-verbal communication [2].

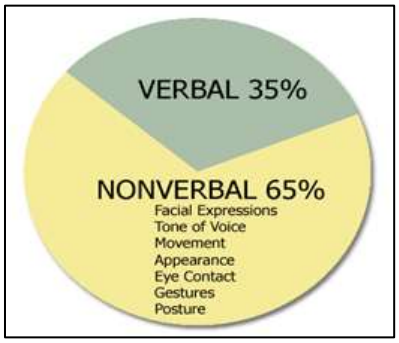

Fig. 1: Partition of Verbal and Non-Verbal Communication

The word 'Gesture' means movement of the hand, arm, body, and head or face that is too expressive of an idea, opinion, emotions etc. The word 'Recognition' means to detect and identifying the objects in the digital medium.

Fig. 2: Non-Verbal Communication

This paper focuses on Hand gesture recognition the gesture recognition is in the form of Static and dynamic in nature. Generally the name suggests the static gesture recognition appeals that at the time of gesture period the hand position still remains constant. Static gesture based on the shape and flexural angles of the hand and dynamic refers that at the time of

gesturing period hand position changes continually [3]. The shape of a hand and flexural angles of fingers are included in dynamic gestures and it also includes orientation of hand.

The gesture technology permits the users to control or manipulate devices in a more inherent manner, such as TV, Microwave, and AC remote operates with hand movements. Gesture has the diversity in hand movement which is produced by the hand. The term 'Posture' focuses on the shape of hand, but 'Gesture' focuses on the movement of the hand. Human Computer Interaction (HCI), Robotics, T.V interaction, Sign Language recognition is widely applications used in the recognition of hand gesture [4].

Section I describes the Introduction of Hand Gesture Recognition System. Section II describes the Methodology; Section III defines the various Experimental Results. Section IV describes the Applications and finally, Section V describes the conclusion and Future work.

II. METHODOLOGY

In this section we will discuss the step by step methodology. Fig 3 shows the flow diagram of proposed methodology.

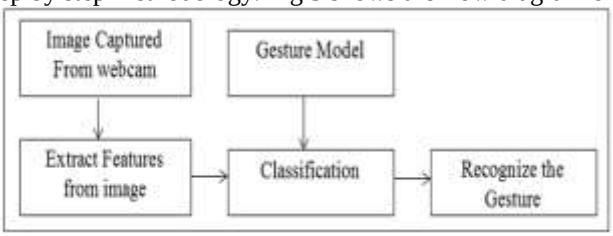

Fig. 3: Block Diagram of proposed methodology

In a block diagram (fig. 3) firstly captured the images from a webcam. A set of features is extracted from every frame captured and extracted features are used as input in a classification step. The gesture model tends to go to the classification model for the purpose of compared the features of hand [5]. Finally gesture recognition performs and it requires high definition cameras and to recognize the hand signals by using some algorithms.

A. Simple Linear Iterative Clustering approach

In the computer vision applications the concept of superpixels are becoming popular in use. The high quality segmentations will be produced by this technique. To enforce the superpixel compactness the simple linear iterative clustering algorithm is used. In this system, the clustering is performed with the similar behaviour of the image i.e color tone or intensity value. The segmentation of pixels in an image to form the clusters is known as Superpixels and each superpixel is nearly equal to the size of pixels. Suppose N pixels of an image are segmented into k superpixels then each superpixel is nearly equal to N/K pixels of an image[7].

B. Artificial Neural Network

In artificial neural network, neurons consist of a simple artificial nodes and processing elements are connected together to form a network. ANN mimics a biological neural network and solves the difficult problems which are not easily solved by conventional computer or the humans. It is capable of machine learning and it requires the training. They contain the basically three layers i.e. input, hidden and the output layer. The inputs are passed having some weights reach to the next node and doing the sum of all weights [8]. This process is continues until the data reach to the output layer where the outcomes are predicted by the model.

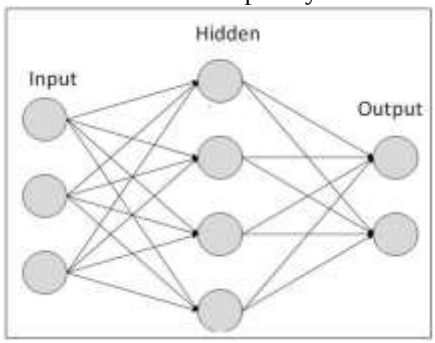

Fig. 4: Different layers in Artificial Neural Network

The comparison of actual output and predicted output is done in supervised learning technique. In this technique if the actual output is equal to the predicted output then weights are doesn't change. But in other case, if the predicted output is up or down from the actual result then it is back to the system and weights are adjust accordingly [9]. This is also known as back-propagation and then final results will be produced.

C. *Algorithm for Hand Gesture Recognition system*

- Step 1: the images from the training dataset are loaded into memory. These images are pre-processed and apply the normalisation.
- Step 2: the depth map of the image is used to segment out the hand region using thresholding at certain distance and SLIC based region segmentation is performed over them.
- Step 3: the features from hand shapes are computed in order to form a feature matrix. This feature matrix is used to train Artificial Neural Network.
- Step 4: the artificial neural network is trained and this artificial neural network is used to test the system in next step.
- Step 5: a test set is used to test the accuracy of the system.
- Step 6: When the testing phase is correctly done, compute the confusion matrix and the recognition time graph will be drawn.

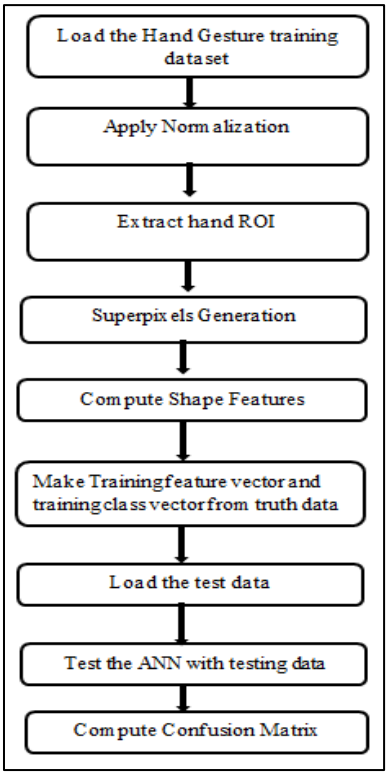

Fig. 5: Algorithm of Hand Gesture Recognition

III. EXPERIMENTAL RESULTS

The following system were implemented using MATLAB and tested on an Intel core i5 with 8GB Ram running on window 7. The datasets has been obtained in both controlled and non-controlled forms. The dataset defines the gestures that represent the number from 0 to 9.

Fig. 6: Training hand Gesture Samples (0-9)

A. *Dataset*

The dataset has been obtained on a constant background which can provide maximum accuracy. The system is tested with 10 different gestures naming from 0 to 9.

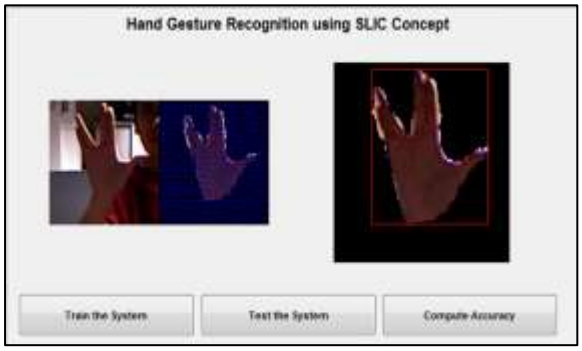

Fig. 7: Training phase of the HGR system

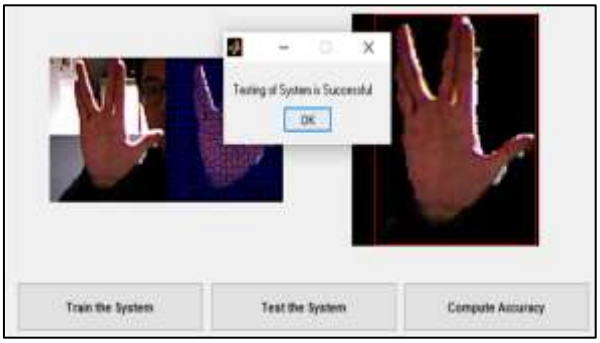

Fig. 8: Testing phase of the HGR system

B. Confusion Matrix

It is a contingency table which describes the correct or incorrect classifications used to compute performance of classification algorithm in machine learning. The results from confusion matrix show a total of 99.0 % average accuracy of the system.

C. Recognition Time

It is the time taken to recognize the each gesture in a proposed system and it calculates the average recognition time to measure the each gesture is 0.219 seconds.

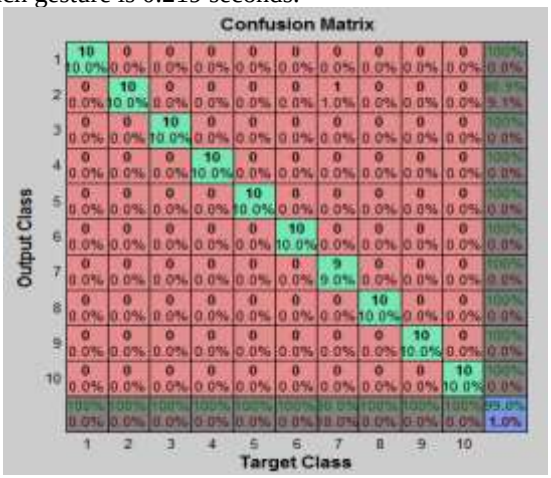

Fig. 9: Confusion Matrix

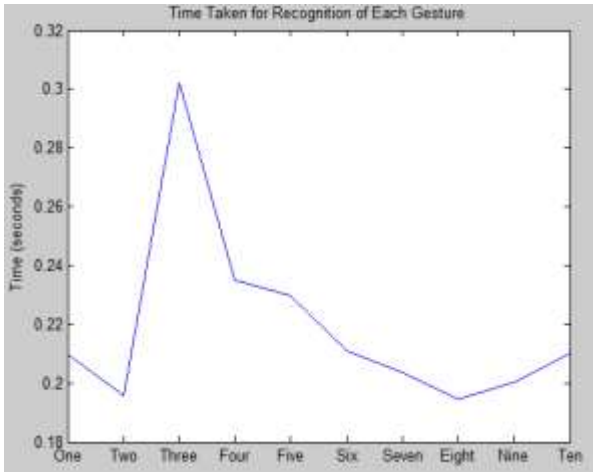

Fig. 10: Time Taken to Recognize the Gesture

IV. APPLICATION PROSPECT

The applications for recognition of the hand gesture system can be stated as follows:

- In the use of Virtual environment interaction
- In the use of Sign language
- Gaming Zone, medicine, robotics, computer science, remote control etc. [5].

A. Candy Crush Game

The Candy crush saga is a popular match puzzle (minimum three). The main objective is to form one formation, i.e. either of 3, 4 or 5 with matching candies in horizontal or vertical way. Stripping candies is just like a bomb, wrapping candies clearing the nearby 8 candies and color bomb removes the formation with the same color of candies. In Fig.11 to represent 3 commands different hand gesture will be employed. When user selects a square one formation can make [15].

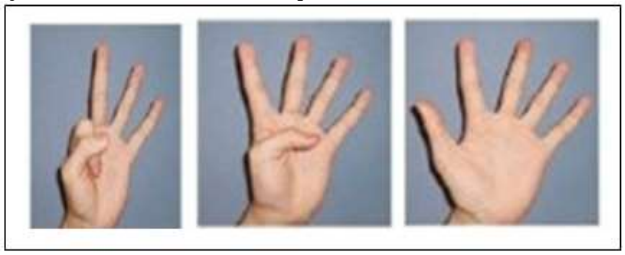
Fig. 11: Candy Crush Saga adopted the 3 Gesture Commands

V. CONCLUSION AND FUTURE WORK

In the field of 3D gaming, Surgeries, robotics the Hand gesture recognition is applicable. It is just like a non-verbal Communication. There is a need for research due to an increasing number of applications. In this paper the segmentation is performed on the basis of simple linear iterative clustering. The features are computed from the hand shapes in the form of feature matrix. These feature matrixes will be trained by the use of artificial neural network and after training the testing of feature matrix is done with the help of ANN. It provides the recognition results in all 0 to 9 types of gestures. The effectiveness of proposed system is explained by the experiments on normal test dataset. When the training dataset and tested dataset are matched correctly then compute the confusion matrix with 99.0% accuracy is achieved and recognition time is 0.219 sec to recognize the hand gesture. ANN provides the efficient algorithm to recognize the hand gestures and it is better for real life Human Computer Interaction. In future, system is upgraded to support dynamic gestures and recognize the gestures from different plane of angles.

REFERENCES

[1] C. Wang, Z. Liu, and S.C. Chan, "Superpixel-Based Hand Gesture Recognition With Kinect Depth Camera," Trans. Multimed., vol. 17, no. 1, pp. 29–39, 2015.
[2] L. Chen, F. Wang, H. Deng, and K. Ji, "A Survey on Hand Gesture Recognition," 2013 Int. Conf. Comput. Sci. Appl., pp. 313–316, 2013.
[3] P. K. Pisharady and M. Saerbeck, "Recent methods and databases in vision-based hand gesture recognition: A review," Comput. Vis. Image Underst., vol. 141, pp. 152–165, 2015.
[4] H. Cheng, L. Yang, and Z. Liu, "A Survey on 3D Hand Gesture Recognition," IEEE Trans. Circuits Syst. Video Technol., vol. PP, no. 99, p. 1, 2015.
[5] M. K. Ahuja and A. Singh, "Static vision based Hand Gesture recognition using principal component analysis," 2015 IEEE 3rd Int. Conf. MOOCs, Innov. Technol. Educ., pp. 402–406, 2015.
[6] D. Q. J. Dqg, V. Iurp, L. Pdun, O. Kdqg, H. W. Lv, K. Kinect, H. G. Recognition, Z. Fdq, E. H. Fodvvlilhg, E. S. Fodvvlilhuv, and V. Dv, "using Kinect Depth Camera," pp. 5–8, 2015.
[7] J. S. Sonkusare, N. B. Chopade, R. Sor, and S. L. Tade, "A Review on Hand Gesture Recognition System," 2015 Int. Conf. Comput. Commun. Control Autom., pp. 790–794, 2015.
[8] S. Atsushi, "Hand Gesture based TV Control System," 19th Korea -Japan Jt. Work. Front. Comput. Vis., pp. 121–126, 2013.
[9] Y. Zhu, Z. Yang, and B. Yuan, "Vision based hand gesture recognition," 2013 Int. Conf. Serv. Sci. ICSS 2013, vol. 3, no. 1, pp. 260–265, 2013.
[10] J. Chastine, N. Kosoris, and J. Skelton, "A study of gesture-based first person control," Proc. CGAMES 2013 USA - 18th Int. Conf. Comput. Games AI, Animat. Mobile, Interact. Multimedia, Educ. Serious Games, pp. 79–86, 2013.
[11] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust part-based hand gesture recognition using kinect sensor," IEEE Trans. Multimed., vol. 15, no. 5, pp. 1110–1120, 2013.
[12] M. A. Moni and A. B. M. Shawkat Ali, "HMM based hand gesture recognition: A review on techniques and approaches," Proc. - 2009 2nd IEEE Int. Conf. Comput. Sci. Inf. Technol. ICCSIT 2009, pp. 433–437, 2009.
[13] D.-L. Dinh, J. T. Kim, and T.-S. Kim, "Hand Gesture Recognition and Interface via a Depth Imaging Sensor for Smart Home Appliances," Energy Procedia, vol. 62, pp. 576–582, 2014.
[14] J. L. Raheja, M. Minhas, D. Prashanth, T. Shah, and A. Chaudhary, "Robust gesture recognition using Kinect: A comparison between DTW and HMM," Optik (Stuttg)., vol. 126, no. 11–12, pp. 1098–1104, 2015.
[15] Z. Ren, J. Meng, and J. Yuan, "Depth Camera Based Hand Gesture Recognition and its Applications in Human-Computer-Interaction," IEEE Int. Conf. Inf. Commun. Signal Process., no. 1, pp. 3–7, 2011.