Hand Gesture Recognition Using Scale Invariant Feature Transform Technique and Genetic Algorithm

Fatima Sufi[1], Jaskiran Kaur[2], Daljit Kaur[3]

[1] *M.Tech Student, Department of Information Technology, Chandigarh Engineering College, Landran, Punjab, India*

[2, 3] *Assistant Professor, Department of Computer Science and Engineering, Chandigarh Engineering College, Landran, Mohali, Punjab, India*

Abstract - Hand Gesture recognition systems have been considered as the able-bodied and well made systems with the goal of interpreting gestures, making use of suitable algorithms to accomplish the recognition task. So far, various techniques were used to identify the gestures; many have their own advantages but also have few shortcomings like responsiveness, classification, accuracy etc. Hence, gives the need to use such algorithms each of which is unique in its own effective and creative way and try to bring out the best possible results. Our proposed work focuses on extracting the features using Scale-Invariant Feature Transform (SIFT) algorithm, optimizing those features using Genetic Algorithm and finally classifying the gestures using Back Propagation Neural Networks (BPNN). By doing this our system obtains better performance in terms of classification and faster response time or delayed outputs.

Keywords - Gesture interpretation; SIFT Algorithm; Genetic Algorithm; BPNN; Performance

I. INTRODUCTION TO HAND GESTURES

Hand gesture recognition plays an outstanding role in spotting the gestures and acts as a bridge between humans and computers so that computers can come to know human body languages and provide an efficient inter-communication in a consistent manner. The image is actually a function of f(x, y) and takes its magnitude at a location(x,y).Researchers have presented many techniques to monitor the captured gestures. Every gesture recognition system has three main building constituents based on which it carries the recognition process.

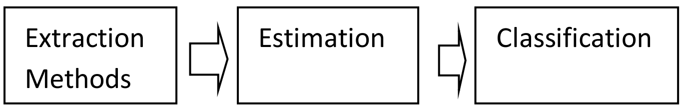

Fig.1: Gesture recognition steps

The very first step in the recognition phase is called segmentation. Segmentation partitions the captured image into discrete parts by bounds. The easiest and beneficial way used for segmentation of hand is the skin color, for it is changeless to size, translation, and rotation features of the hand[1].Once the image is segmented, the next step is extracting the unique properties or features from the segmented image. Finally, gesture classification method is used to interpret the reconstructed gesture with proper classification algorithm. Misclassification of gestures degrades the image quality and leads to poor performance of the system. These systems serve a wide range of applications on various domains including medical systems, virtual environments, smart examination, sign language translation etc.

A. Problems encountered in Hand gesture recognition

During the study of many approaches, it has been found that gesture verification issues many problems and need further improvements in order to get efficient pattern matching. The problems encountered are as: *Variation of illumination conditions*: It means any change in the lighting conditions has a bad impact on the extracted skin region. *Rotation Problem*: It is the problem faced when the hand posture is moved in any order in the scene. *Background Problem*: A complicated background also creates problems when there are other objects around with the hand objects and may have similar skin color[8]. *Scale Problem*: This problem occurs if the hand poses have varying sizes[2].*Rotation problem*: this difficulty arrives when hand is moved in any order in the backdrop. *Translation Problem*: The changing hand positions in differing images also leads to false depiction of the features. *Difficulties in handling the training input and testing output*: Artificial Neural Network (ANN) creates one network on only one layer which causes delay in response time which is the problem of main concern.

Earlier work in the field of hand gesture recognition for feature extraction was performed using Principal Component Analysis (PCA). PCA is a mathematical algorithm and has the trait to determine the Eigen values and Eigen vectors directly in the form of a matrix. However, due to its massive matrices formation it becomes tedious to calculate the Eigen values and Eigen vectors and is relatively sensitive to scaling. Also intelligent techniques like Artificial Neural Networks was used as a mechanism for gesture classification and matching but lead to delayed response time and hampered system performance. Since, hand gesture systems are used for security purpose, it is vital that the system becomes efficient [3, 4].

B. Proposed solution

Our proposed work takes three techniques into consideration. Our goal is to overcome the above limitations and based on it we have used the embedded algorithms namely; Scale Invariant Feature Transform (SIFT), Genetic Algorithm for optimization and Back Propagation Neural Network.

II. TECHNIQUES USED

A. Scale Invariant Feature Transform

SIFT, is an algorithm to find and highlight local features in images [5].By local features we mean that for any object, unique properties also called as the key points are separated from its segmented image. The primary function of SIFT is to test whether the image is segmented properly or not and then to calculate the number of key features from the scalable image and localize them. SIFT works by dividing the input segmented scalable image data into modules and partitions it. For every step SIFT does a deep study to determine number of unique features not only in the centre of the image but also depicts them at the co-ordinates. SIFT is constant to scaling, rotation, orientation problems and has no impact of noise factors. SIFT determines huge amount of features even from a clutter of images, which lessens the occurrence of errors of other feature resembling errors. This detection, extracted from a training image can later be used to point out the object in test images with rest of the objects.

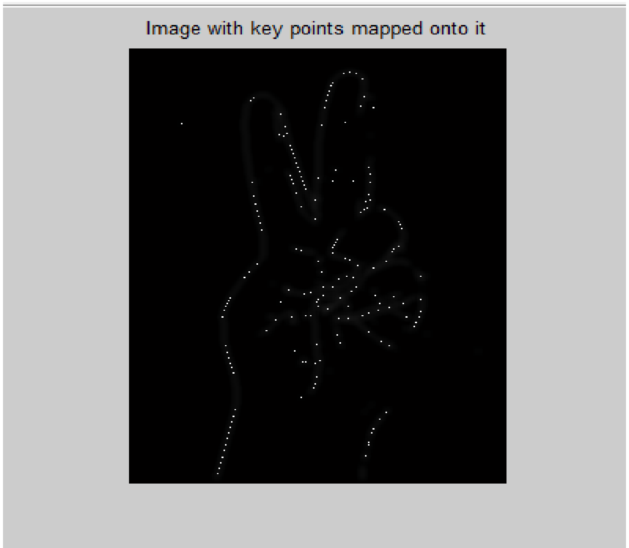

Fig3: key features of segmented image

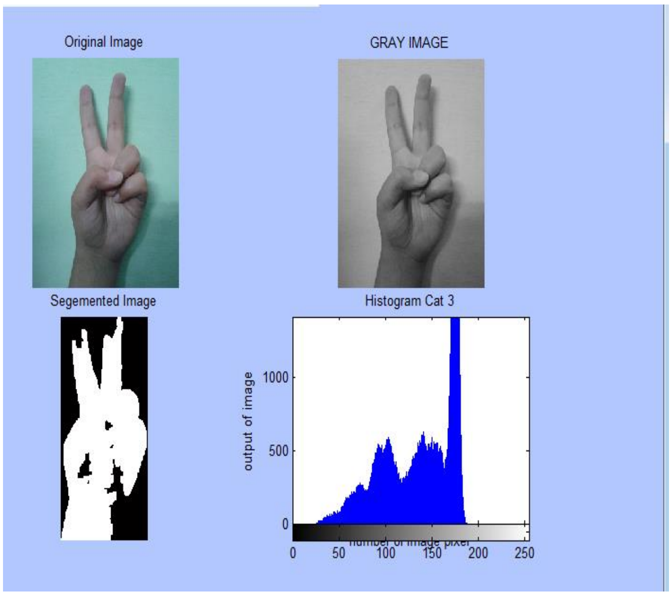

Fig2: segmentation process of hand

B. Genetic Algorithm

A genetic algorithm is an analytical and investigative optimization technique, used when the search space is huge or too complex for analytical examination. It performs a detailed study on extracted features in order to yield better results in an efficient manner. The individuals in population are computed by fitness function and fit values. The operators namely; *Selection*, *Crossover*, *Mutation* carry the remaining task. Selection function selects or initializes the chromosomes in a population that are made to repeat each time producing better chromosomes than before. Crossover is used to vary or interchange the sub parts among chromosomes from one generation to another. Mutation changes the gene values from its original state. A random variable is created in gene pool which lets know whether the specific bit is improved or not. Genetic algorithm is a vigorous search method requiring less information to search efficiently to produce optimized output in an enormous search space and is free of derivatives[6][7].

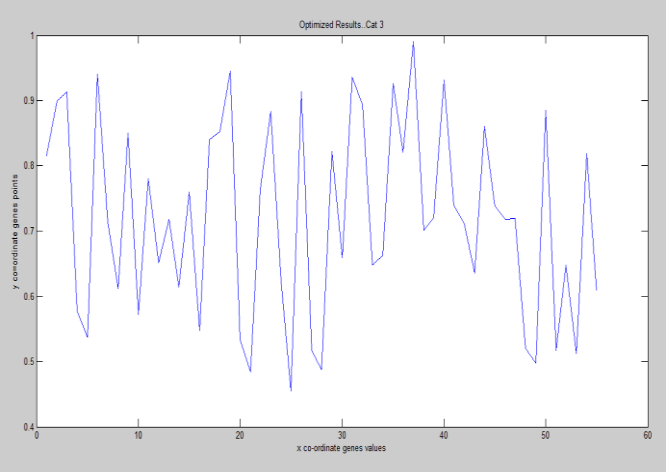

Fig4: optimization graph of Genetic algorithm

C. Back Propagation Neural Network

Back Propagation Neural Network is similar in architecture with Artificial Neural Networks, however differ in processing. BPNN works faster and fastens the response time by sending optimized data in multiple layers unlike neural network which sends data in only a single layer. The error is traced back to previous node. As a result, the delay in input training and testing output can be reduced to a large extent using BPNN which is a feed forward algorithm [9].

III. METHODOLOGY

A. SIFT ALGORITHM

Step1: The SIFT takes the obtained key features in its scenario. The image is looped with Gaussian filters at varying scales, followed by the difference of consecutive Gaussian-blurred image. This is called as the Scale-space extrema detection. The key points generated are nothing but the maximum and minimum of the Gaussian differences represented as(DoG) shown as

$$D(x,y,\sigma)=L(x,y,k_i\sigma) - L(x,y,k_j\sigma), \qquad (1)$$

where $L(x,y,k\sigma)$ is the looping of actual image $I(x,y)$ with the blurred Gaussian image $G(x,y,k\sigma)$ at some scale kσ,and '*' is the looping factor such that

$$L(x,y,k\sigma) = G(x,y,k\sigma) * I(x,y) \qquad (2)$$

Therefore, DoG, is the difference of Gaussian blur images at$k_i\sigma$ and $k_j\sigma$. The procedure is performed using various octaves as shown in fig(6).

Step2: SIFT distills the key points situated on the edges and eliminates the less distinctive key features. This is called as key point localization[10].The area of extremism, z, is given by:

$$Z= - [(d^2 D^(-1))/[dx]^2]^ dD/dx \qquad (3)$$

Step3: This is the key to obtain changeless rotation by providing some orientations based on the computations of Gradient Magnitude and direction belonging to that part. This step is called orientation assigning as in fig(5). For an image sample L(x, y) at 'σ', the gradient magnitude, m(x, y), and rotation,ɵ(x, y) are already evaluated as under

m(x,y)=

$$\sqrt{\mathbf{L(x+1,y)- L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}}$$

ɵ (x, y) = a tan2 (L (x,y+1) – L (x,y-1) – L (x+1,y)–L (x-1,y))

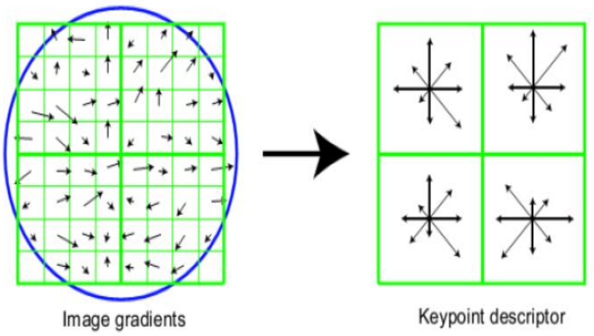

Fig5: Orientation allotment

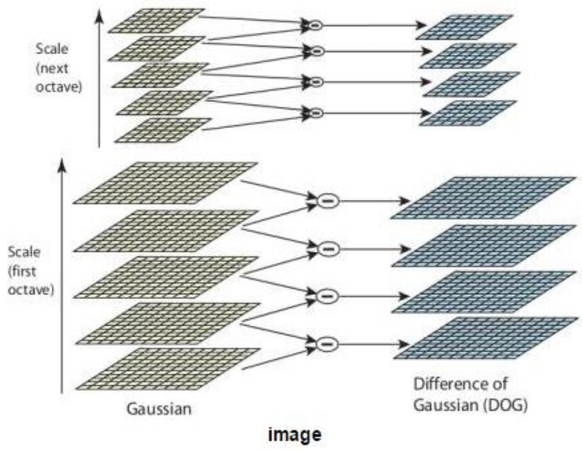

Fig6: Gaussian Pyramid

Step4: This step is carried out for images to get matched and distinguished among nearing features. Key feature descriptors usually make an adaptation for 16 histograms, organized in a 4x4 settlement, representing eight introduction containers, for the attitude compass bearings and the mid-purposes headings.

B. BLOCK DAIGRAM

3.2.1

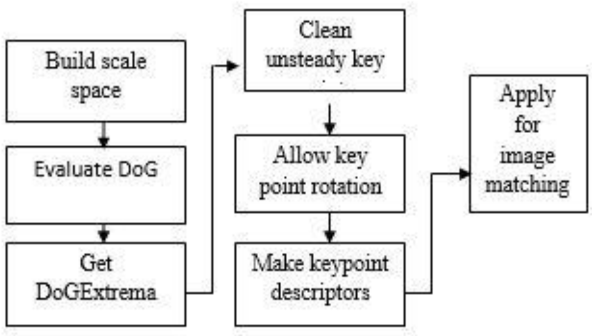

Fig7: Sift algorithm steps

3.2.2

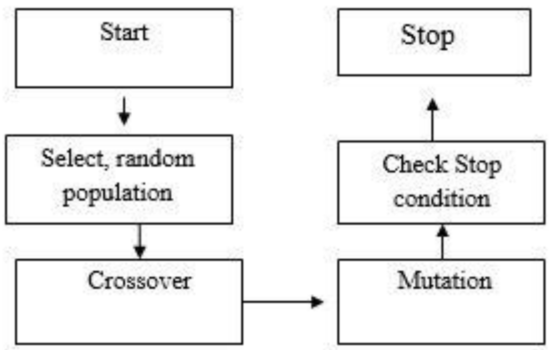

Fig8: Genetic algorithm step

IV. ALGORITHM FOR OVERALL WORKFLOW

Step1: Upload images of many categories
Step2: pre-process and segment those images
Step3: Represent the segmented images graphically using histogram
Step4: Use SIFT to develop key features
Step5: Apply Genetic algorithm to above features
Step6: Select population size and allow for crossover and mutation
Step7: Using fitness function generate fit values
Step8: Classify Using Back propagation Neural Network
Step9: Compute performance of the system

V. CONCLUSION AND FUTURE SCOPE

In this paper we have taken three techniques into consideration viz SIFT algorithm, Genetic algorithm and Back Propagation Neural Network by replacing the already existing techniques. The system was planned with the prime focus on feature separation and optimization so as to achieve accurate gestures. However, in near future Back Propagation neural network can be introduced as a gesture identifier or interpreter to work on the grounds of increased performance.

VI. REFERENCES

[1] N. Ibraheem, M. Hasan, R. Khan, P. Mishra, (2012). “comparative study of skin color based segmentation techniques”, Aligarh Muslim University, A.M.U., Aligarh, India.

[2] Kim-Tien, Nguyen, Nguyen Truong-Thinh, and Trinh DucCuong. "A Method for Controlling Wheelchair Using Hand Gesture Recognition." Robot Intelligence Technology and Applications 2012.Springer Berlin Heidelberg, 2013.961-970.

[3] Jalab, Hamid A., and Herman Omer. "Human computer interface using hand gesture recognition based on neural network." Information Technology: Towards New Smart World (NSITNSW), 2015 5th National Symposium on. IEEE, 2015

[4] Dardas, Nasser H., and Emil M. Petriu. "Hand gesture detection and recognition using principal component analysis." Computational Intelligence for Measurement Systems and Applications (CIMSA), 2011 IEEE International Conference on. IEEE, 2011.

[5] Zhou, Yimin, et al. "Hand Gesture Recognition Based on the Parallel Edge Finger Feature and Angular Projection." Computer Vision-ACCV 2014Workshops. Springer International Publishing, 2014

[6] Justina, J. P., and SangeethaSenthilkumar. "A Survey Analysis on: Vision-Based Hand Gesture Recognition." International Journal 4.1 (2014).

[7] Wu, Chih-Hung, Wei-Lun Chen, and Chang Hong Lin. "Depth-based hand gesture recognition." Multimedia Tools and Applications (2015): 1-22.

[8] S, NAJI. R, ZAINUDDIN. H.A. JALAB. 2012. Skin segmentation based on multi pixel color clustering models. Digital Signal Processing,22, 933-940.

[9] H, AL-JBARA.M, KIAH. H.A, JALAB.2012. Increased capacity ofimage based steganography using artificial neural network. AIP Conf.
Proc, 2012. 20-25.

[10] Li, Zisheng, et al. "Evaluation of medical image registration by using 3D SIFT and phase-only correlation." Abdominal Imaging.Computational and Clinical Applications.Springer Berlin Heidelberg, 2012.255-264.

[11] Sharma, Ram Pratap, and Gyanendra K. Verma. "Human Computer Interaction using Hand Gesture." Procedia Computer Science 54 (2015): 721-727

[12] Rautaray, Siddharth S., and AnupamAgrawal. "Vision based hand gesture recognition for human computer interaction: a survey." Artificial Intelligence Review 43.1 (2015): 1-54