

Hand Gesture Recognition based on Invariant Features and Artificial Neural Network

Gurwinder Kaur* and Gourav Bathla

Chandigarh University, Gharuan - 140413, Punjab, India; gurwinder.kaur646@gmail.com, gouravbathla@gmail.com

Abstract

Objectives: Its objective is to develop a system which can recognize specific human hand gesture features from images and use these features to convey information for machine such as HCI (Human Computer Interaction). **Method:** this paper represents that, we propose a new method that supports the hand gesture recognition system in the static form, using SIFT feature extraction with feed forward neural network using MATLAB. We use SIFT technique to extract the invariant features of gesture sign. We have developed a gesture recognition system using feed forward neural networks which could recognize a finger alphabet of different types of symbols and where each gesture specifies a word. **Findings:** In the process of hand gesture recognition system there are many challenges addressed as: Illuminance conditions such that a small change in the lighting conditions effects badly on extraction process like color from which may produce misclassification problem. In the various scenes when the hand region rotates in any direction then the rotation problem get arises. When there is complicated backgrounds like there are other objects in the image with hand poses, this may refer to the background problems. Scale or Size problem, this problem may occur when the hand poses have not same size in the gestured image. At last the transformation problem, like from different different images it's difficult to represent features of various hand positions. **Improvements:** To remove theses hindrances we have to develop an algorithm that can effectively recognize the hand gestures. Hand gestures is one of the techniques used in security purposes and robot communication.

Keywords: Data Acquisition Toolbox, Feed Forward Neural Network, Hand Gesture Recognition, Image Processing, SIFT (Scale-Invariant Feature Transform), Sign Classification

1. Introduction

Communication is an essential prerequisite for survival and collaboration. Actually, distinctive communication ways are utilized for connection, for example, dialect, eyes, body development, outward appearance, hand motion and stances. A gesture or pose is a type of interaction correspondence made by the parts of body and utilized rather than face to face correspondence by speaking. A communication via gestures is a dialect which utilizes signals rather than sound to pass on importance consolidating hand-poses, introduction and development of the hand, arm or body parts, outward appearances and lips-designs. Gesture based communication is a visual dialect and comprises of 3 noteworthy parts¹: finger-signs: used to represent words letter by letter, word sign vocabulary: utilized for the greater part of correspondence, non-man-

ual elements: outward appearances and tongue, mouth and body position. Gesture based communication is one type of correspondence for the hearing and discourse debilitated.

1.1 Human Computer Interaction (HCI)

Communication through signing acknowledgment is a multidisciplinary research zone including design acknowledgment, PC vision, regular dialect preparing and phonetics². In addition for Human Computer Interaction (HCI), when contrasted with the customary association methodologies, for example, console, mouse, pen and so forth, and vision based hand cooperation is more characteristic and effective.

Hand motions fall into two classifications, to be specific static and dynamic³. Some hand signals likewise have

*Author for correspondence

both static and dynamic components, as in gesture based communications. Static hand signals are portrayed by the hand stance which are dictated by a specific finger thumb-palm design and spoke to by a solitary picture. Dynamic hand motions are then again described by the underlying and last stroke movement of a moving signal.

1.2 Various Techniques

There was a technique of hand gesture recognition system for recognition of Arabic sign language's alphabets and convert these alphabets into voice correspondingly to make Arabian deaf people to get in interaction with normal people. The technique was proposed to capture a colour image and convert that image in to YCbCr, because it provides an effective and accurate method to extract skin regions from coloured images under various illuminance changes⁴.

Research demonstrates different sorts of framework and techniques which have been created for gesture based communication acknowledgment. In⁵ proposed a novel and continuous framework for collaborating with a computer game by means of hand signals utilizing SIFT and SURF highlight extraction. The hand signal acknowledgment framework includes both uniform and complex foundation. Numerous works consolidate skin shading division for sectioning hand from a mind boggling foundation. In a⁶ proposed framework for perceiving 26 hand signals for American Sign Language letters in order to from a mind boggling foundation utilizing the Euclidean separation measure. A hand motion acknowledgment framework for ASL utilizes back proliferation neural system⁷. To group the hand signals for letters in order in the wake of portioning the hand from the information picture utilizing the Lab shading areas and the classify of tops and separating points as the components the neural network system⁸ is utilized. The pose division system utilizing foundation subtraction is given as a part of⁹ for extraction of the features from the picture with simple and complicated foundation.

A powerful framework¹⁰ extracting key points and interpretation is actualized for perceiving motion signs for utilization Artificial Neural Network. A unique strategy for example acknowledgment is introduced in¹¹ for perceiving 36 unique signals utilizing SIFT highlights with PCA and coordinating techniques.

Despite the fact that numerous strategies are there in recognition field, not much work has been done in

acknowledgment of Indian Sign Language. A¹² Proposed framework for perceiving Indian Sign Language (ISL) motions utilizing Eigen value weighted Euclidean separation as an arrangement system. Gabor filters are convolved with pictures to get the components and the feature points are reduced using PCA¹³. Utilizing the decreased Gabor highlights, signals are perceived by SVM. To prepare the neural system in acknowledgment of Persian Sign Language¹⁴, Discrete Wavelet Transform elements are utilized.

Different color models are utilized as a part of the division of hand in the signal acknowledgment framework. The perceptual color space is utilized for hand motion acknowledgment¹⁵ for remote robot control applications. Whereas the finger spelling is subject to the particular dialect of a nation. Finger spelling¹⁶ is utilized as a part of Indian gesture based communication, American Sign Language, Chinese gesture and British Sign Language based communication, Persian Sign Language, and Arabic gesture based communication, Malaysian communication via gestures for acknowledgment of communication through signing.

Neural Networks can be able to determine the effective parameters of ground motions with sufficient correlation. Neural Networks were used in many fields of sciences such as mathematics, engineering, medical science, etc., for prediction, function approximation, classification of data¹⁷.

2. SIFT (Scale Invariant Feature Transform)

The most dependable feature extraction method that is used in classification system on Finger-kunuckle-Print (FKP)¹⁸ is SIFT (Scale Invariant Feature Transform). SIFT is a picture descriptor for picture based coordinating and acknowledgment created by David Lowe (1999, 2004). This descriptor and in addition related picture descriptors are utilized for an expansive number of purposes in PC vision identified with point coordinating between various perspectives of a 3-D scene and perspective based article acknowledgment. The SIFT descriptor is invariant to interpretations, turns and scaling changes in the picture space and powerful to direct point of view changes and brightening varieties.

Tentatively, the SIFT descriptor has been ended up being extremely helpful practically speaking for picture coordinating and question acknowledgment under true

conditions. The SIFT descriptor included a strategy for distinguishing interest focuses from a dark level picture at which measurements of nearby inclination bearings of picture intensities were amassed to give an abridging portrayal of the neighbourhood picture structures in a neighbourhood around every interest point, with the goal that this descriptor ought to be utilized for coordinating relating interest focuses between various pictures. Later, the SIFT descriptor has likewise been connected at thick frameworks (thick SIFT) which have been appeared to prompt better execution for assignments, for example, object arrangement, surface characterization, picture arrangement and biometrics. Figure 1. shows the working of SIFT.

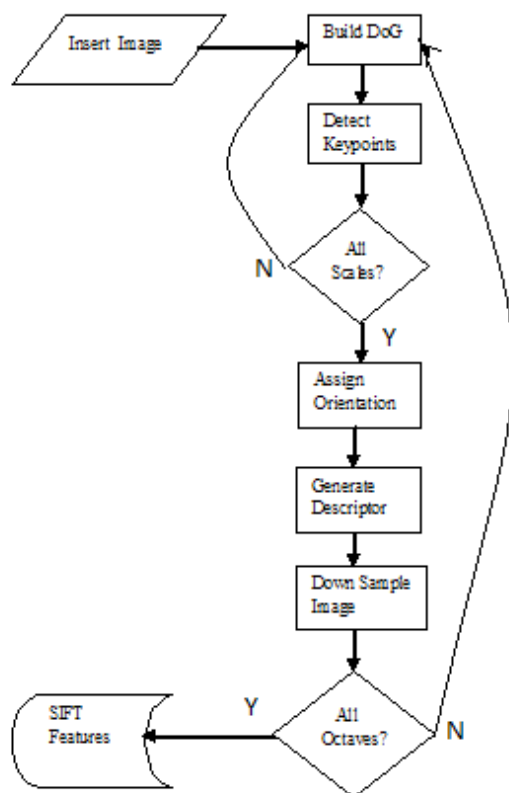


Figure 1. SIFT Feature Extraction Method.

3. Neural Network (NN)

The basic aim of neural network is to work like human brain works. The neural network system contains a number of neurons and their performance is similar to the human brain's neuron structure. The two main basic

properties of neural networks are: mapping based on experimental data (ability and potency of generalize ability) and another one, similar or paralleled structure ability¹⁹. There are various types of neural networks, the feed forward neural network with two hidden layers (4 and 17 neurons for the economic model) had the best results and it can be used to estimate the energy ratio with high precision²⁰, but the Back Propagation Neural Network is commonly in use. Two types of structure has been found in neural network model:

- Cyclic;
- Acyclic

To train the Multilayer FNN, the normal Back propagation is normally applied. The linear and nonlinear outputs are correspondingly given by:

The net input has given by:

$$n1^{k1+1}(i) = \sum_{j=1}^{s1k1} w1^{k1+1}(i,j) a1^{k1}(j) + b1^{k1+1}(i)$$

The unit i is given by

$$a1^{k1+1}(i) = f1^{k1+1}(n1^{k1+1}(i))$$

This recurrence relation is executed at the final layer

$$-F1^{M1}(n1^{M1})(t1_{q1} - a1_{q1})$$

The algorithm is described below:

1. Initialise weights
2. Initialize inputs and each input unit is represented as $(Ci, i = 1 \dots m)$.
3. Each hidden unit is represented as $(Xj, j = 1, \dots, a)$ and weighted input Signals are represented below,
 $Xin = B0j + \sum_{mi=1} Ci Bij$
 Where $B0j$: hidden unit j bias.
 Bij : Weight between output and input unit.
 $Xj = g(Xin)$
4. Each output unit gets its output as show below;
 $Uimk = Eok + \sum_{aj=1} Xj Ejk$
5. Output unit having activation function is shown below:
 $Uk = g(uimk)$
6. Finding of back propagation error.
7. Each output unit gets its output by changing error values.
 $\delta k = (yk - lk)g'(uimk)$
8. Calculation of weight function.
 $\Delta Ejk = \partial \delta k Xj$
9. Calculation of bias correction term

$$\Delta E_{ok} = \partial \delta_k$$

10. Update weights and biases:

11. Each output unit (Y_k , $k = 1, \dots, m$) updates its bias and weights ($j = \dots, p$):

$$E_{jk}(\text{new}) = E_{jk}(\text{old}) + \Delta E_{jk}$$

12. Test stopping condition.

4. Problem Formulation

In the process of hand gesture recognition system there are many challenges addressed as:

Illuminance conditions such that a small change in the lighting conditions effects badly on extraction process like color from which may produce misclassification problem. Scale or Size problem, this problem may occur when the hand poses have not same size in the gestured image. At last the transformation problem, like from different different images its difficult to represent features of various hand positions.

5. Proposed Methodology

Firstly image acquisition will be done then on uploaded image ROI image will be extracted then in next step feature extraction using SIFT will be done and apply Genetic algorithm on key points for optimization. Then training has been done by using optimized key points

and in the last step image classification will be done using neural network. In addition to this feature reduction is also being done on the basis of GA. Figure 2 represents the over view of the proposed work and proposed algorithms for these processes are shown in Figure 3.

5.1 Phases of Recognition System

There are two phases: training phase and testing phase. The various steps involved in the process of recognizing hand gestures are described as:

5.1.1 Training Stage

1. Upload Gesture Image from training data set.
2. Edge detection to find the hand gesture image boundaries.
3. Feature extraction using SIFT.
4. Feature reduction using genetic algorithm (GA).
5. Learning using Feed forward back propagation neural networks.

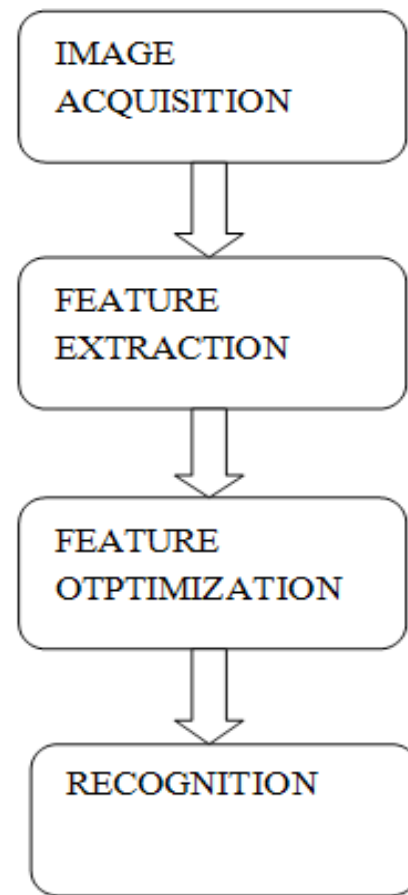


Figure 2. Block diagram of Hand gesture recognition system.

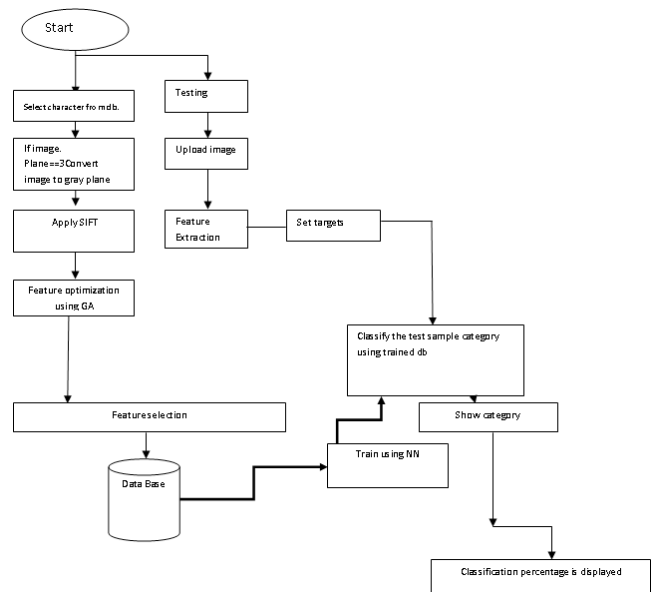


Figure 3. Proposed Model.

5.1.2 Testing Stage

1. Upload Gesture Image from testing data set.
2. Upload test image of SIFT, GA and NN.
3. Parameter evaluation in MATLAB.

5.1.3 Classification using NN

1. Set Training set as 1:5,1 and 6:10,1 and so on
2. Declare groups.
3. Start newff and assign epochs=50.

Classification using NN shown in Figure .4

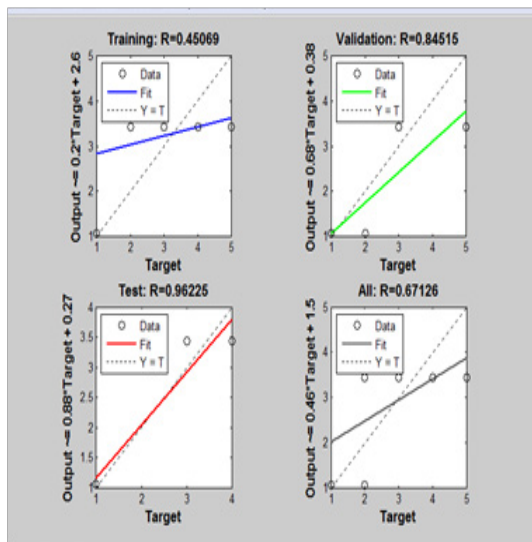


Figure 4. Classification using NN.

6. Experimental Result

The whole implementation is being done in MATLAB using various parameters like FAR, FRR and accuracy.

Table 1 represents the accuracy of purposed work. The proposed work is implemented for .jpg image format. In proposed work already built in toolbox of neural network in utilised. Firstly images will be trained then images will be tested using proposed algorithm.

Table 1. Accuracy table

S.No	Alphabets	FAR	FRR	Accuracy value (%)
1	A	.1396	.1326	96.20
2	B	.1896	.1802	95.01
3	C	.2396	.2276	95.02
4	V	.3395	.3326	93.20
5	5	.2896	.2791	94.20

6.1 Computation Parameters

Accuracy: It is the ratio of corrected detected image samples to un- corrected image samples.

$$Accuracy(x) = \left(\frac{\text{sum of correct classification}}{\text{total number of classification}} \right) \times 100$$

FAR: It is ratio of total false accepted image sample to the total number of corrected image samples. It is generally calculated in percentage form.

$$FAR(x) = \left(\frac{\text{no. of correctly classified instances of class.}}{\text{no. of instances classified as belonging to class.}} \right)$$

FRR: It is ratio of total false rejected image sample to the total number of corrected image samples. It is generally calculated in percentage form.

$$FRR(x) = \left(\frac{\text{no. of noncorrectly classified instances of class.}}{\text{no. of instances classified as belonging to class.}} \right)$$

From the Table-1, it has been evaluated that average accuracy of each alphabet came out to be 94.8%. So it is concluded that this proposed method has giving us promising results. Table 2 describes the comparison between various techniques.

Table 2. Comparison Table

Technique	Accuracy (%)
Bayesian Sensing Hidden Markov Model ²¹	88.84
Depth Map ²²	90
SIFT+ BPNN	94.8

7. Conclusion

Hand gesture recognition is very useful in military fields is very helpful. So, correct gesture recognition must be done in order to get good rate of accuracy. In proposed work a system of gesture recognition based on SIFT and NN is built in which good features are extracted using SIFT and good classification rate has been achieved by NN in terms of FAR, FRR and accuracy parameter in MATLAB environment with the accuracy rate of 94.8%.

8. Future Work

The results can be improved using LDA. Linear Discriminant Analysis is utmost commonly utilized as

dimensionality lessening method in the pre-processing stage for machine learning applications in addition to design-classification. The main objective is to project a specific dataset on top of a lower-dimensional space using virtuous class reparability so as to decrease computational prices as well as also evade over fitting. The novel linear discriminant was first designated for a two-class issue, in addition it was then afterwards widespread as “Multiple Discriminant Analysis” or “multi-class LDA” through C. R. Rao in the year of 1948. Linear Discriminant Analysis is “controlled” as well as calculates the guidelines (“linear discriminants”) which would probably signify the axes that are applied to make the most of the separation amongst multiple type of classes.

9. References

1. Bowden R, Zisserman A, Kadir T, Brady M. Vision based interpretation of natural sign languages. UK. 2003; 1–2.
2. Aran O. Vision based sign language recognition: modeling and recognizing isolated signs with manual and non-manual components’ (Doctoral dissertation, Bogaziçi University. 2008; 1–169.
3. Cutler R, Turk M. View based interpretation of real time optical flow for gesture recognition’. IEEE International Conference on Automatic Face and Gesture Recognition. 1998. p. 1–6.
4. Hemayed EE, Hassanien AS. Edge-based recognizer for Arabic sign language alphabet (ArS2V-Arabic sign to voice)’. In Computer Engineering Conference (ICENCO). 2010 International IEEE. Egypt. 2010. p. 121–7.
5. Nasser H, Dardas D, Nicolas D, Georganas G. Real-Time Hand Gesture Detection and Recognition using Bag-of-Features and Support Vector Machine Techniques. IEEE Transactions on Instrumentation and Measurement. 2011; 60(11):3592–607.
6. Jayashree R, Pansare P, Sharavan H, Gawande G, Ingle M. Real-Time Static Hand Gesture Recognition for American Sign Language (ASL) in Complex Background. Journal of Signal and Information Processing. 2012; 3(3):364–7.
7. AtiqurRahman MD, Ahsan-Ul-Ambia A, Aktaruzzaman MD. Recognition of Static Hand Gestures of Alphabet in ASL. IJCIT. 2011; 2(1):1–4.
8. Mapari R, Kharat G. Hand Gesture Recognition using Neural Network. International Journal of Computer Science and Network. 2012; 1(6):1–8.
9. Randive AA, Mali HB, Lokhande SD. Hand Gesture Segmentation. International Journal of Computer Technology and Electronics Engineering. 2012; 2(3):1–5.
10. Vaishali S, Kulkarni K, Lokhande SD. Appearance Based Recognition of American Sign Language Using Gesture Segmentation. International Journal of Computer Science and Engineering. 2010; 2(3):560–5.
11. Deval G, Patel P. Point Pattern Matching algorithm for recognition of 36 ASL gestures. International Journal of Science and Modern Engineering. 2013; 1(7):1–5.
12. Singha J, Das K. Indian sign language recognition using eigen value weighted euclidean distance based classification technique. International Journal of Advanced Computer Science and Applications. 2013; 4(2):1–8.
13. Huang DY, Hu WC, Chang SH. Vision based Hand Gesture Recognition Using PCA+Gabor filters and SVM. IEEE Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Taiwan. 2009; 1–4.
14. Karami A, Zanj B, Sarkalesh AK. Persian sign language (PSL) recognition using wavelet Transform and neural Networks. ELSEVIER Journal of Expert Systems with Applications. 2011; 38(3):2661–7.
15. Manigandan MIM, Jackin J. Wireless Vision based Mobile Robot control using Hand Gesture Recognition through Perceptual Color Space. IEEE International Conference on Advances in Computer Engineering, India. 2010; 95–9.
16. Saengsri S, Niennattrakul V, mahatana CAR. TFRS: Thai Finger-spelling Sign Language Recognition System, IEEE, Thailand. 2012; 457–62.
17. Kia A, Sensoy S. Assessment the effective ground motion parameters on seismic performance of r/c buildings using artificial neural network. Indian Journal of Science and Technology. 2014 Jan; 7(12):1–7.
18. Sulthana ESS, Kanmani S. Implementation and Evaluation of SIFT Descriptors based Finger-Knuckle-Print Authentication System. Indian Journal of Science and Technology. 2014 Jan; 7(3):1–9.
19. Hajmohammad MH, Salari M, Hashemi SA, Esfe MMH. Optimization of Stacking Sequence of Composite Laminates for Optimizing Buckling Load by Neural Network and Genetic Algorithm. Indian Journal of Science and Technology. 2013 Aug; 6(8):1–8.
20. Sefat IY, Borgaee AM, Beheshti B, Bakhoda H. Application of Artificial Neural Network (ANN) for Modelling the Economic Efficiency of Broiler Production Units. Indian Journal of Science and Technology. 2014 Jan; 7(11):1–7.
21. Hernawan A, Lee YS. Bayesian Sensing Hidden Markov Model for Hand Gesture Recognition, ACM, ASE BD&SI. 2015 Oct 07 - 09.
22. Hamissi M, Faez K. Real-time hand gesture recognition based on the depth map for human robot interaction. International Journal of Electrical and Computer Engineering (IJECE). 2017, 3(6):770–8.