ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering &Technology (IJARCET)
Volume 2, Issue 3, March 2013

Real Time Static Hand Gesture Recognition System in Complex Background that uses Number system of Indian Sign Language

Jayshree R.Pansare, Hrushikesh Dhumal, Sanket Babar,Kiran Sonawale, Ajit Sarode

***Abstract*— Hand gestures are powerful means of communication among humans and sign language is the most natural and expressive way of communication for deaf and mute people. Communication between computers (or robot) and humans, just as we humans interact with one another has been the prime objective of human computer interaction (HCI) research. This paper describes a real-time system for human computer interaction through gesture recognition for Indian Sign Language (ISL). ISL number system includes nine characters which are considered for this paper. The work is divided into four stages- image pre-processing, region extraction, feature extraction, feature matching. First stage converts captured RGB colour space image YIQ space images, then using grey threshold to detect the skin colour. Second stage extracts hand region using blob for getting region of interest(ROI). In third stage, we extract the features of ROI; the system compares four feature extraction techniques. Finally comparison is done between the current and database images in fourth stage using Euclidian distance.**

***Index Terms*— Complex background,Pattern matching techniques, Real time system, Static hand gestures.**

I. Introduction

Recognition of sign language is one of the major concerns for mute and deaf people. Sign language recognition is a research area involving pattern recognition, computer vision, natural language processing. Sign language recognition is a comprehensive problem because of the complexity of the visual analysis of hand gesture and the highly structured nature of sign language. As well as it is considered as a very important function in many practical communication applications, such as sign language understanding, entertainment, and HCI(human computer interaction). Among natural human gestures occurring during non-verbal communication, pointing gesture can be easily recognized and included in more natural new HCI.

The main approaches for analyzing and classifying hand gestures for gesture recognition applications include glove-based techniques and vision-based techniques. The glove-based techniques use sensors to measure the positions of the fingers and the position of the hand in real-time. However, gloves tend to be quite expensive and the weight of the glove as well as the cables of the associated measuring equipment hinders free movement of the hand. The vision-based techniques are usually glove-free and can be divided into the three-dimensional (3-D) and the two-dimensional (2-D) approaches. In the 3-D approach, gesture classification is based upon the parameters of a 3-D model of the human hand. Gesture classification is based upon the parameters of an image of the gesture in the 2-D approach. Because 3-D hand models are quite complex, the classification of gestures from parameters derived from 3-D models is computationally extensive making real-time classification difficult. The 2-D models are relatively less complex than the 3-D models. However, 2-D models do not carry the finger movement and finger position information required for the classification of complex dynamic gestures. Therefore, the 2-D approach is restricted to the less complex problem of classifying well-defined static gestures as suggested in [2].

The images are frequently influenced by the background changes such as illumination changes and changes due to adding or removing parts of the back-ground. Therefore, the quality of the foreground and the segmented image of hand gesture severely drop. The real time system is user invariant so, so we propose a novel system which is based on 1-D obtained from color model YIQ ("I" stand for "inphase" and the "Q" for "quadrature," which isthe modulation method used to transmit the colour information) under condition of uniform scene illumination. We considered the YIQ color space in our work since it is effective in modeling human skin-color. By considering the chrominance components only, the feature space is reduced from 3-D to 1-D, thus reducing the computational complexity of the classification algorithm as proposed in [8] where 3-D to 2-D conversion is used.

The YIQ color system that is known to be less sensitive to lighting than the RGB is adopted. In the YIQ, "Y" means intensity, while "I" and "Q" represent color information [9]. "I" and "Q" called chroma, jointly describe the hue and

Manuscript received Feb, 2013.
***JayshreeR.Pansare**, Dept. of Computer Engineering, M. E. Society's College of Engineering, Pune, India.Mobile No.919850762329.*
***Hrushikesh Dhumal**, Dept. of information technology, AISSMS IOIT,Pune-01, IndiaMobile No. 917709524730.*
***Sanket Babar**, Dept. of information technology, AISSMS IOIT,Pune-01, IndiaMobile No. 919028424737.*
***Kiran Sonawale**, Dept. of information technology, AISSMS IOIT,Pune-01, IndiaMobile No.919028133286,*
***Ajit Sarode**, Dept. of information technology, AISSMS IOIT,Pune-01, IndiaMobile No. 918087994571.*

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering &Technology (IJARCET)
Volume 2, Issue 3, March 2013

saturation attribute of an image. The "I" value is used to build up a skin-color model and to extract hand areas from input images by calculating the similarity between a model and an input image. It aims at maintaining a more robust one that could be adaptive to changing environmental conditions.

II. FLOW OF HAND GESTURE RECOGNITION SYSTEM

The flow of hand gesture recognition system algorithm is presented as shown in Fig 1 as suggested by Prof J.R. Pansare in [12]. The hand image with a resolution of 160 × 120 is first captured using Web camera. A hand region is then extracted from image using skin detector. Feature extracted from filtered binary image and feature matching compares running and training images. Hand recognition uses closer matching hand gesture. Fig 2 shows ISL symbols and their corresponding gestures stored in the system database. The database consists of binary images. For each gesture ten ideal images are stored in the database.

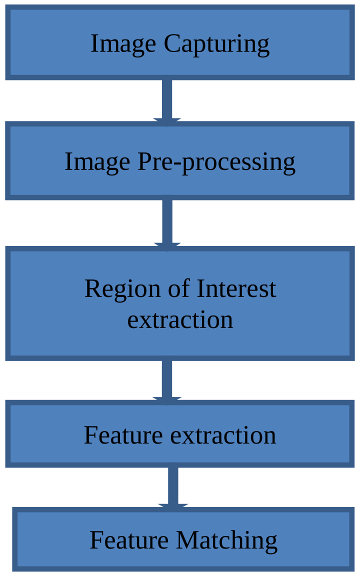

Fig: 1 Flow Of System

III. PROPOSED HAND GESTURE SYSTEM

A. Image Capturing

Color is a robust feature. It is vulnerable to changing lighting conditions and it differs among people. In this proposed method, first of all, the snapshot of RGB image is captured by the 16 megapixel web camera. In this step, the image size is 160× 120. If this size of image is used for the processing, the computation time will be very less. Also the image is stored using ".tiff" (Tagged Image File Format) for lossless compression.

B. Image Pre-processing

The RGB image is converted into YIQ images in this step. The "I" part from YIQ colour space am extracted. A global threshold (level) is calculated that is used to convert an intensity image to a binary image as in Fig 3. Based on this value I from YIQ is extracted to generate a black and white image. Otsu segmentation algorithm is used for this purpose. The *Otsu segmentation algorithm* is a grey level thresholding algorithm based on discriminant analysis. The algorithm treats the segmentation of image into a binary image as a classification problem in which the two classes (in this case, hand and background) are generated from the set of pixels within the image [1]. Morphological operations such as 3-by-3 neighborhood method the majority pixel value is assigned to the selected pixel to reduce inconsistency. Also "thinning" as mentioned in [15] is applied on the image.

C. Region of Interest Extraction

The *biggest continuous blob* is selected from the binary image by using the bilinear interpolation as in [3].The image is once again resized and stored inresolution of 80×60. Median filter is applied as suggested in [10]. Once again morphological operation removes interior pixels and exposes the edge [15].

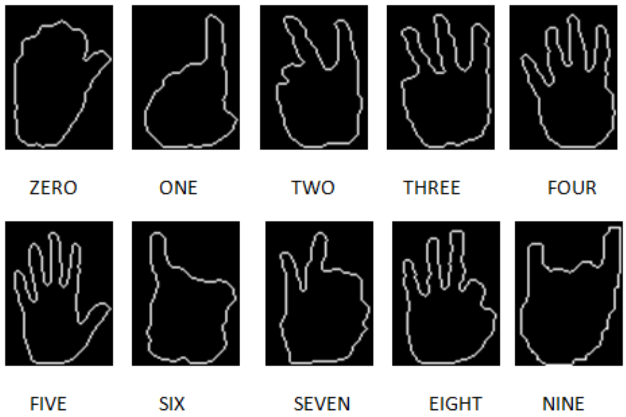

Fig: 2 Hand gesture in Indian Sign Language

D. Feature Extraction:

Centroid matching

Feature vector is formed using centroid given by (Cx, Cy) as in [12].

. $$(C_x, C_y) = \frac{1}{n}\left(\sum_{t=1}^{n} x_t, \sum_{t=1}^{n} y_t\right) \qquad (1)$$

Discrete Cosine Transform

This is the formula to calculate Discrete Cosine Transform of 2-D image. The definition of the two-dimensional DCT for an input image A and output image B in is [16]

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos\frac{\pi(2m+1)p}{2M} \cos\frac{\pi(2n+1)q}{2N}, \quad \begin{matrix} 0 \le p \le M-1 \\ 0 \le q \le N-1 \end{matrix}$$

Where

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, p = 0 \\ \sqrt{\frac{2}{M}}, 1 \le p \le M-1 \end{cases}$$

And

$$\alpha_q = \begin{cases} \frac{1}{\sqrt{N}}, q = 0 \\ \sqrt{\frac{2}{N}}, 1 \le q \le N-1 \end{cases} \quad (2)$$

Fourier transform

A fast Fourier transform (FFT) is an algorithm to compute the discrete Fourier transform (DFT) and it's inverse as in [15].

$$X(k) = \sum_{j=1}^{N} x(j)\omega_N^{(j-1)(k-1)}$$

$$x(j) = (1/N)\sum_{k=1}^{N} X(k)\omega_N^{-(j-1)(k-1)}$$

where

$$\omega_N = e^{(-2\pi i)/N}$$

is an Nth root of unity. (3)

Edge Oriented Histogram (EOH)

Counts the number of values in vector x that fall between theelements in the edges vector (which must contain monotonically non decreasing values).

Feature vectors of training images are stored in .mat files of MATLAB and feature vector of input hand gesture image are calculated at run time.

E. Feature Matching

Feature matching is done using Euclidian distance as suggested in [10]. The geometrical distance between two points in N dimensional space is known as the Euclidean distance. Several samples are involved in training this classification system. For Euclidean distance, features of each training sample are used in distance calculation. Euclidian distance between feature vector of input real- time image and feature vector of each training image is calculated using following formula. In the Euclidean plane, if p = (p1, p2) and q = (q1, q2) then the distance is given by:

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} \quad (4)$$

Least Euclidian distance is used for recognition of perfect matching hand gesture.

IV. EXPERIMENTAL RESULT

In this section, the performance of real time static hand gesture recognition system in complex background is evaluated for each of the 10 hand gestures 2 images for each number considered in various complex backgrounds. Fig3 showsthe various test cases considered. The images are acquired using 16 megapixel web camera, 1.3 megapixel web camera and MATLAB R2010a.

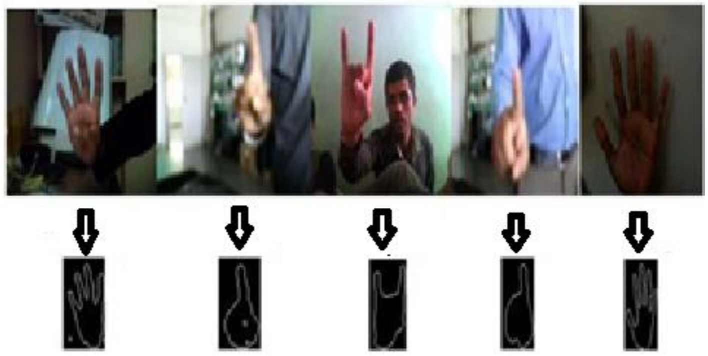

Fig: 3 Test cases

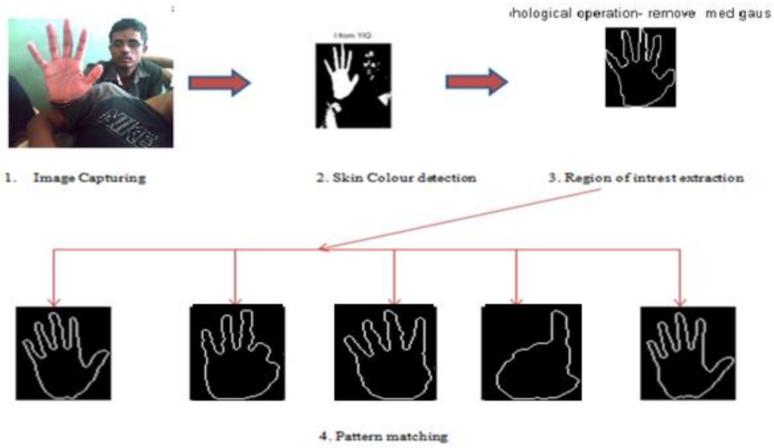

Fig:4 System implementation

As we can see in Fig 4 image is captured from camera then undergoes preprocessing to extract the region of interest which is then compared with the database images. Each of the test images is tested for different feature extraction techniques to find out which techniques are more robust and efficient to support the real time system. The Table 1and Fig 5 shows the obtained test results.

Gesture	Centroid	DCT	Fourier	EOH
0	2	2	2	1
1	2	2	2	1
2	2	2	2	2
3	2	2	1	2
4	2	2	2	2
5	1	2	2	1
6	2	1	1	1
7	1	2	1	1
8	1	1	1	2
9	2	2	2	2
Total	17	18	16	15

Table: 1 Result table

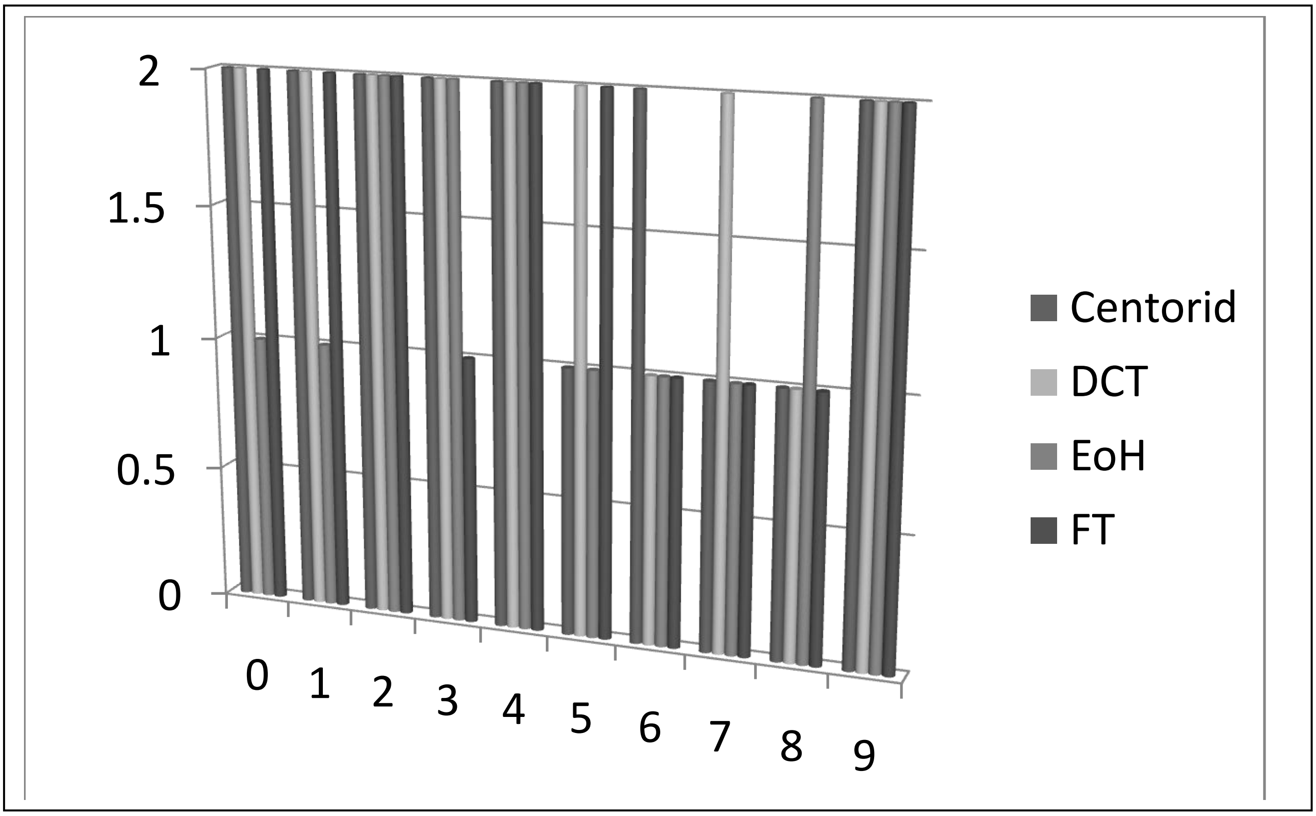

Fig: 5 Graphical representation of result

V. CONCLUSION

The single hand gesture recognition system works successfully for real-time static hand gesture recognition in different lightning conditions such as dark or from different light sources, however for system to work uniform illumination is necessary. From experiments it is observed that the results are more accurate with Discrete cosine transform as compared to edge oriented histogram, centroid, Fourier transform. The results from low-cost 16 mega pixel are better as com-pared to 1.3 megapixel camera. Fixed position of Web camera mounted on the top of monitor or attached to laptop produces more results for imagescaptured from 0.5 meter distance from camera in complex background. Low-level feature extraction from extracted region reduces computation and works efficiently while matching feature vectors of real- time hand gesture images with feature vectors of training dataset. Samples in different lightning conditions with different hand shapes and sizes produce more accuracy by using Euclidian distance for feature matching.

The system recognises skin-colour, so it fails to extract the hand region if other body part is in a plane closer to camera. The system is sensitive to red colour, so gestures should be made without red colour in background.

REFERENCES

[1] Wei Du Hua Li CAD Lab, Institute of Computing Technology, Chinese Academy of Sciences, 100080, Beijing, China “Vision based gesture recognition system with single camera Supported by National 973 Project of China (No. G1998030600) and National Key Project of Chinamo.96-920-02-02-03”, 0~7803-5747-7/0 2000 IEEE.

[2] Lalit Gupta and Suwei Ma “Gesture-Based Interaction and Communication: Automated Classification of Hand Gesture Contours”, IEEE Transactions on Systems, Man and Cybernetics-part C: Applications and Reviews, Vol. 31, No. 1, February 2001.

[3] Lars Bretzner, Ivan Laptev, Tony Lindeberg, Computational Vision and Active Perception Laboratory (CVAP) Centre for User Oriented IT Design (CID) Dept of Numerical Analysis and Computing Science, KTH, 100 44 Stockholm, Sweden “Hand Gesture Recognition using Multi-Scale Colour Features, Hierarchical Models and Particle Filtering”, Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR’02) 0-7695-1602-5/02 2002 IEEE.

[4] JongShill Lee, YoungJoo Lee, EungHyuk Lee, SeungHong Hong Department of Electronic Engineering, Inha University, Korea “Hand region extraction and Gesture recognition from video stream with complex background through entropy analysis”, Proceedings of the 26th Annual International Conference of the IEEE EMBS San Francisco, CA, USA. September 1-5, 2004.

[5] JiaJia, Jianmin Jiang, Dong Wang EIMC, School of Informatics, University of Bradford, West Yorkshire BD7 1DP, UK “RECOGNITION OF HAND GESTURE BASED ON GAUSSIAN MIXTURE MODEL” , 978-1-4244-2044-5/08 2008 IEEE.

[6] JuanfangXu, Shouqian Sun, Yuanwu Shi ZhanxunDongI “Implementation of Digital Chime-bell Interaction System Driven by Hand Gesture”, 'College of Computer Science, Zhejiang University, P.R. China, 978-1-4244-2020-9/08 2008 IEEE.

[7] NasibehAlijanpour, HosseinEbrahimnezhad, AfshinEbrahimi Computer Vision Research Lab, Electrical Engineering Department, Sahand University of Technology, Tabriz, Iran "Inner Distance Based Hand Gesture Recognition for Devices Control", 978-1-4244-3397-1/08 2008 IEEE.

[8] SurachaiOngkittikul Stewart Worrall AhmetKondoz School of Electronics and Physical Sciences, University of Surrey, Guildford, Surrey GU2 7XH, UK "Two Hand Tracking using Colour Statistical Model with the K-means Embedded Particle Filter for Hand Gesture Recognition", 978-0-7695-3184-7/08 2008 IEEE.

[9] Ho-Sub Yoon, Byung-Woo Min, Jung Soh, Young-laeBae and Hyun Seung Yang Department of Computer Science / Korea Advanced Institute of Science and Technology 373-1 Kusung-Dong Yusung-Ku Taejon, Korea "Human Computer Interface for Gesture-Based Editing System", 2009 IEEE.

[10] M.K. Bhuyan, D. Ghosh and P.K. Bora Department of Electronics and Communication Engineering Indian Institute of Technology Guwahati, India "Designing of Human Computer Interactive Platform for Robotic Applications", 2009 IEEE.

[11] Baoyun Zhang, Ruwei Yun Digital Entertainment Research Center, Nanjing Normal University, Nanjing, China "Gesture Recognition System Based on Distance Distribution Feature and Slin-Color Segmentation", 978-1-4244-5858-5/10 2010 IEEE.

[12] Jayashree R. Pansare, Shravan H. Gawande, Maya Ingle "Real-Time Static Hand Gesture Recognition for American Sign Language (ASL) in Complex Background", Journal of Signal and Information Processing, 2012, 3, 364-367 Published Online August 2012.

[13] Archana S. Ghotkar&Gajanan K. Kharate "Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction", International Journal of Human Computer Interaction (IJHCI), Volume (3) Issue (1) : 2012.

[14] Lam, L., Seong-Whan Lee, and Ching Y. Suen, "Thinning Methodologies-A Comprehensive Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 14, No. 9, September 1992, page 879, bottom of first column through top of second column.

[15] Pratt, William K., *Digital Image Processing*, John Wiley & Sons, Inc., 2007.

[16] Jain, A.K. *Fundamentals of Digital Image Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1989.