International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682
Vol.4, No.5, May 2015 DOI:10.15693/ijaist/2015.v4i5.25-31

Resilient Real-time Hand Gesture Recognition System

Mohd. Abbas Murtaza
Department of ECE
JUIT, Solan, India

Ruchita Gupta
Department of ECE
JUIT, Solan, India

Meenakshi Sood
Department of ECE
JUIT, Solan, India

Abstract — **Gesture based Human Computer Interaction (HCI) is one of the most natural and intuitive modes of communication between individuals and machines. Intelligent gesture recognition systems pave the way for a new era where the focus lies on the task itself, as it should be, and not on the interaction modality. Gestures of the hand itself can be used to communicate effectively as compared to computer peripherals, thereby empowering the visual and speech impaired population. In this paper, the authors propose a technique that employs vision based approach together with dynamic gesture recognition techniques. The prototype architecture of the application comprises of a 2D webcam which continuously records the trajectory of the hand extracting information using a suitable colour model. This work focuses on developing a prototype of a robust, real-time hand gesture recognition system which works well under different degrees of background complexity and illumination conditions for dynamic gestures. The system is cost effective and works efficiently while interacting with objects in virtual environment using hand gestures. As an application, features of media player were controlled with hand gestures and an overall success rate of 97.5 per cent was achieved. While conducting the experiments, the authors have taken into account the challenges faced due to complex background, presence of non-gesture hand motions and different illumination environments in order to develop a system which can be used for innumerous further applications.**

Index terms - Hand Gesture Recognition, Colour Models, Hand Movement Tracking

I. INTRODUCTION

Human gestures constitute motions expressed by the body, face, and/or hands. Gestures have been extensively used as an alternative form of communication between humans and machines in an easy way. Among a variety of gestures, hand gesture is the most expressive and the most frequently used. The human-machine interface with gestures only allows a user to control a wide variety of devices. Most of the recent work related to hand gesture interface techniques has been categorized as sensor-based method and vision-based method [1]. Sensor based approach makes use of various sensors to collect the data of gestures. Vision based approach involves capturing image of gesture and extracting the feature to recognize it. Colour markers is one of the methods of vision based approaches. These techniques are further divided into model based and appearance based technique. In appearance based technique, the appearance of captured gesture is observed to create a recognition model whereas in model based technique, the captured image of gesture is analysed and the model of gesture is created for gesture recognition. While working on hand gesture system, gestures are distinguished on the basis of movement: static and dynamic [2]. A static gesture is a particular hand configuration represented by a single image. A dynamic gesture is a moving gesture, represented by a sequence of images. In this paper, we focus our attention to vision-based recognition of hand gestures and our motive is to recognize dynamic gestures on complex background.

A low cost computer vision system that can be executed in a common personal computer equipped with USB web cam is one of the main highlights of our approach. The system should be able to work under different degrees of background complexity and illumination conditions. The present paper is organized as follows: Section II throws light on the various techniques used by previous researchers. Section III briefly describes the methodology proposed for this work. Section IV presents various colour models which could be used and demonstrates the proposed recognition colour method. The results of the proposed system are discussed in Section V and the conclusion and future work are given in Section VI.

II. RELATED WORK

Over the past two decades, a lot of research has been carried out in the area of vision based hand gesture recognition with the objective of helping speech and language impaired people communicate effectively in their day to day life. Previous researchers have used orientation histograms as a feature vector for gesture classification and interpolation [3]. Employing special hardware or offline learning, several researchers have developed successful systems to recognize general hand gestures [4, 7].

Hidden Markov Models have been used to recognize continuous gestures on stationary background [5]. S. Nagarajan et al proposed a static hand gesture recognition system for sign language alphabets using edge orientation histogram and multi class SVM [6]. Background Subtraction has also been extensively used to solve the issue of background complexity [8].

A technique utilizing Features from Accelerated Segment Test (FAST) corner detection and Adaptive

Resonance Theory (ART) neural network has also been proposed [9]. Jayashree R. Pansare et al. proposed a system to recognize 26 static hand gestures for ASL alphabets from a complex background [10] using the Euclidean distance measure. The neural network is used to classify the hand gestures for alphabets in [11] after segmenting the hand from the input image using the Lab color space and the extraction of peaks and valleys as the features. [12] Presented a hand gesture recognition system for American Sign Language using back propagation neural network. Raheja used PCA as a tool for real-time robot control. PCA is assumed to be a faster method for classification as it does not necessarily require a training database [13].

III. METHODOLOGY

The previous real-time hand gesture recognition systems required special hardware or lengthy training analysis. Our proposed technique uses only 2D video input and a single colour band. The removal of extensive wiring and hardware not only simplifies the system but also provides an easy mode of real time communication between humans and machines.

The proposed technique is divided in four stages and depicted in Figure 1:

1. *Acquisition* of continuous video frames from the webcam.
2. D*etection* of the red coloured marker from the video image and forming a new binary image containing two regions; the marker region and the background.
3. *Tracking* the center of the marker region and extracting the total number of pixels indicating it.
4. *Pattern Recognition* based on the movement of the center of hand and the number of marker pixels.

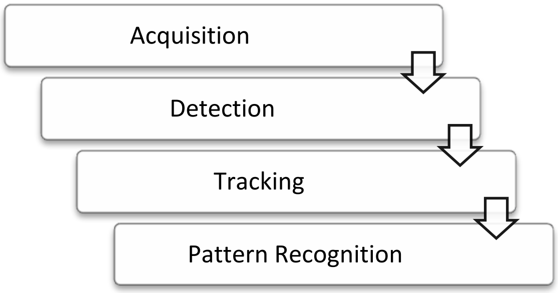


Figure 1. Sequential order of the steps involved in the proposed technique

Image acquisition involves the action of retrieving an image from a source that can be processed. Real-time image acquisition usually involves retrieving images from a source that is automatically capturing images. It creates a stream of files, called frames that can be automatically processed, queued for later work, or stitched into a single media format. Three-dimensional (3D) image acquisition is one of the advanced methods of image acquisition which require use of two or more cameras that have been aligned at precise points around a target, forming a sequence of images to create a 3D or stereoscopic scene. Cutting down on the expenses of implementation, our proposed technique uses only a simple 2D camera. Figure 2 depicts the steps involved in image acquisition for acquiring the images in the required form.

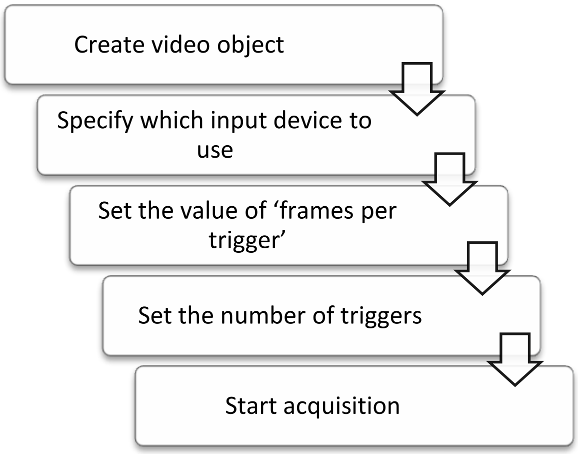


Figure 2. Sequential order of the steps involved in the Image Acquisition phase of the proposed technique

The next step in gesture recognition systems is the detection of hands and the segmentation of the corresponding image regions. This segmentation is crucial because it isolates the task-relevant data from the image background, before passing them to the subsequent tracking and recognition stages. We have utilized the YCbCr colour space in order to extract information of our colour marker from the acquired image which has been detailed in section III. The tracking of the marker is based on the total number of marker pixels and the center of the marker, based on which different functions are controlled. The movement of the marker is tracked, the gesture is recognized and accordingly the respective action is implemented.

IV. COLOUR MODELS

Skin colour segmentation has been utilized by several researchers for hand detection. A major decision is the selection of the colour space to be employed. Several colour spaces have been proposed including RGB, normalized RGB, HSV, YCbCr, LAB, etc. Colour spaces that efficiently separate the chrominance components from the luminance components of colour are typically considered preferable. This is due to the fact that by employing chrominance-dependent components of colour only, some degree of robustness to illumination changes can be achieved.

A. RGB Colour Model

The three primary colours (red, green, and blue) and their combination is the visible light spectrum. With different weights, their combination can indicate different colours.

B.GRAYSCALE Colour Model

A gray colour is one in which the red, green and blue components all have equal intensity in RGB space, and so it is only necessary to specify a single intensity value for each pixel, as opposed to the three intensities needed to specify each pixel in a full colour image. Often, the gray scale intensity is stored as an 8-bit integer giving 256 possible different shades of gray from black to white. If the levels are evenly spaced then the difference between successive gray levels is significantly better than the gray level resolving power of the human eye.

C. HSV Colour Model

In HSV colour model, H represents hue, S represents saturation and V represents value. In its cylindrical coordinate representation , the angle which is around the central vertical axis of the cylinder represents hue(H), the distance from the central axis represents saturation(S) whereas the height of the cylinder represents value(V).

D. CMY Colour Model

This stands for cyan-magenta-yellow and is used for hardcopy devices. In contrast to colour on the monitor, the colour in printing acts subtractive and not additive. A printed colour that looks red absorbs the other two components G and B and reflects R. Thus its (internal) colour is G+B=CYAN. Similarly R+B=MAGENTA and R+G=YELLOW. Thus the C-M-Y coordinates are just the complements of the R-G-B coordinates:-

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

E. LAB Colour Model

LAB colour model was proposed by Commission International del'Eclairage (CIE) as the international standard of colour survey in 1931. A colour can be defined by a lightness component (L) and two colour components (A and B) where A shows the degree from green to red and B indicates the degree from blue to yellow.

F. YCbCr Colour Model

The YCbCr is the collection of colour models which are implemented as part of colour image pipeline. The Y represents the luma component, Cb represents the blue colour difference and Cr is the red colour difference. YCbCr is a method of encoding the RGB colour information.

$$Y = 0.299R+0.587G+0.114B$$
$$Cr = R-Y$$
$$Cb = B-Y$$

Figure 3 illustrates the comparison of the various colour models discussed above. After careful evaluation, it was observed that the colour marker could be separated from the hand effectively using the YCbCr clolour model.

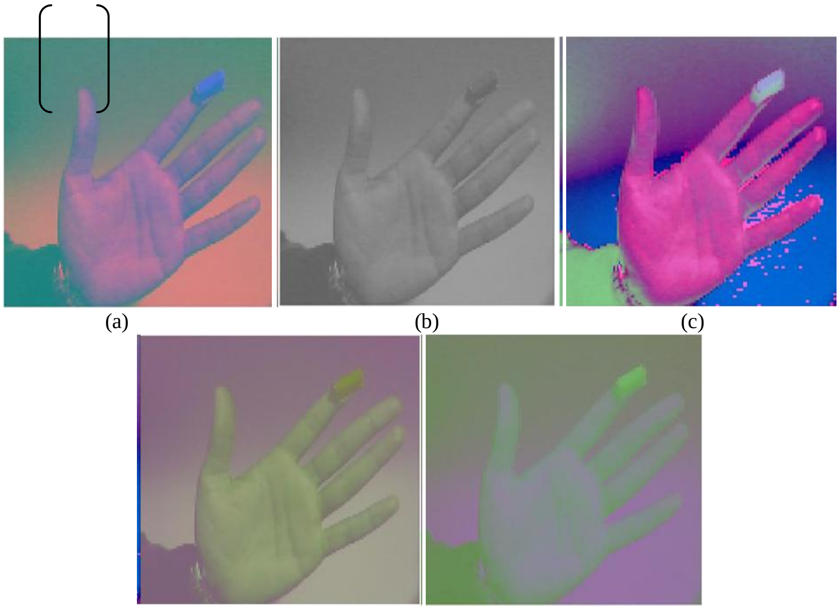

(a) (b) (c)
(d) (e)

Figure 3. Image of Hand Using (a) RGB (b) GRAYSCALE (c) HSV (d) LAB and (e) YCbCr Colour Model

V. RESULTS AND DISCUSSION

Colour information is an efficient tool for identifying skin areas and can be adapted for different lighting environments. Unlike RGB, it has separate luminance and chrominance components which make this colour space attractive for skin colour segmentation. This fact leads to avoid the use of RGB, because the red, green and blue components are highly correlated and dependent on lighting conditions. YCbCr colour space has been defined in response to increasing demands for digital algorithms in handling video information and has become a widely used model in a digital video. In contrast to RGB, the YCbCr colour space is luminance independent that results in better performance. In YCbCr model different threshold values are defined to differentiate between different colours. After exhaustive trials, the threshold values used in our algorithm for the red colour band are taken as:-

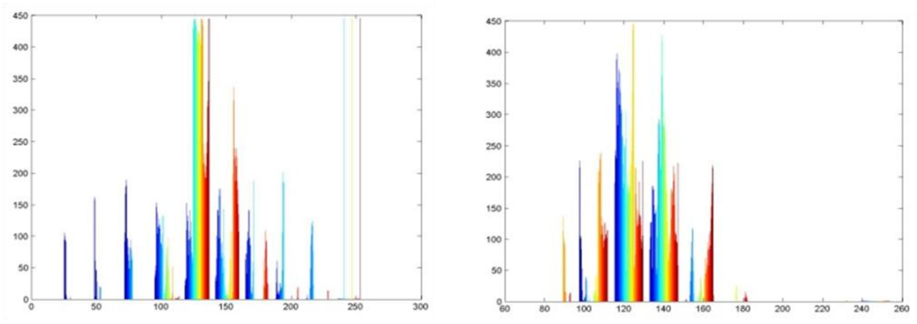
$$160 < Cb < 175$$
$$105 < Cr < 140$$

Our proposed method is based on YCbCr colour model for skin colour segmentation as it has separate luminance and chrominance components. Figure 4(a) depicts the histogram of a RGB image showing a high of colour intensities present in it viz. 0-225 whereas Figure 4(b) portrays the histogram of the of the same image in the YCbCr colour space in which only the desired intensities lying inside our threshold region viz. (90-160) are present which makes it convenient for us to detect the red colour marker thereby tracking it efficiently and enhancing our implementation. Figure 5(a) shows the histogram of the Y component of the image, Figure 5(b) and 5(c) shows the histograms of the Cb and Cr components of the image respectively.

Based on Y, Cb and Cr components, red colour is extracted from the image and a new binary image is formed where red colour is denoted by 1 and background is denoted by 0.

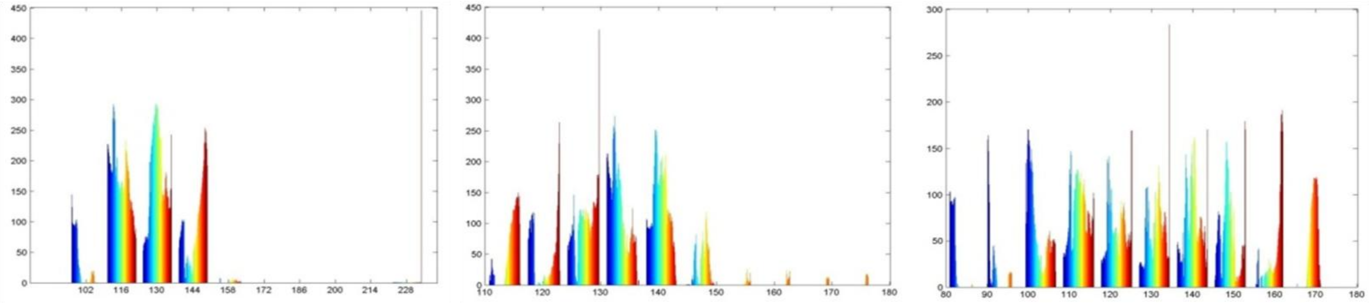
Total numbers of pixels having intensity level as 1 are found along with the center of this region. Functions are implemented based on these parameters as follows:-

1. When the finger is brought closer to the laptop's webcam, the total number of red pixels as compared to background pixels increases.
2. When the finger is taken away from the laptop's webcam, the total number of red pixels as compared to background pixels decreases.
3. When the finger is moving towards the left, the center of the region containing the red marker moves correspondingly.
4. When the finger is moving towards the right, the center of the region containing the red marker moves correspondingly.



(a) *(b)*

Figure 4. Histogram of (a) RGB Image of hand and (b) YCbCr Image of hand



(a) *(b)* *(c)*

Figure 5. Histogram of (a) Y Component, (b) Cb Component and (c) Cr Component of the YCbCr Image of hand

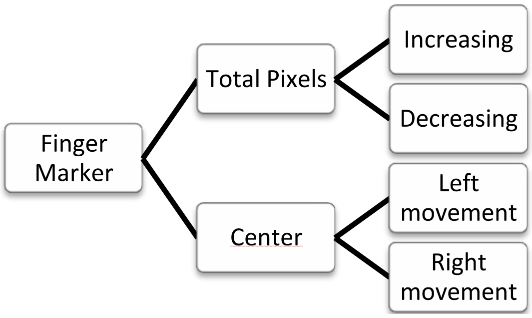


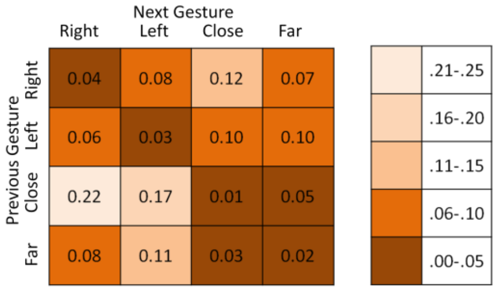
Figure 6. Tracking and implementation of marker

We have implemented a real-time version of hand gesture recognition system, using an ordinary workstation with no special hardware beyond a video camera input and a red colour band. The technique works well under different degrees of background complexity and illumination conditions. The proposed method is implemented by using an optimized Matlab code. The experimental results are illustrated in Table 1. The vertical column shows the next gesture whereas the horizontal column depicts the previous gesture. The table shows the accuracy of transition from one gesture in the vertical column to the other in the horizontal column. We have tested four different moving gestures from different angles and under different illumination conditions. The transition from one gesture to another was tested 100 times for each gesture. The recognition rate for this system was found to be 97.5 per cent. The error matrix obtained for transitions between different gesture movements is tabulated in Table 2. The darkest region indicates least error (less than 5 per cent) and the brightest one represents the maximum error which is only in one case.

Table 1. Recognition Results for Different Gestures

		Next Gesture			
		RIGHT	LEFT	CLOSE	FAR
Present Gesture	RIGHT	96	92	88	93
	LEFT	94	97	90	90
	CLOSE	78	83	99	95
	FAR	92	89	87	98

Table 2. Error Matrix for different gestures

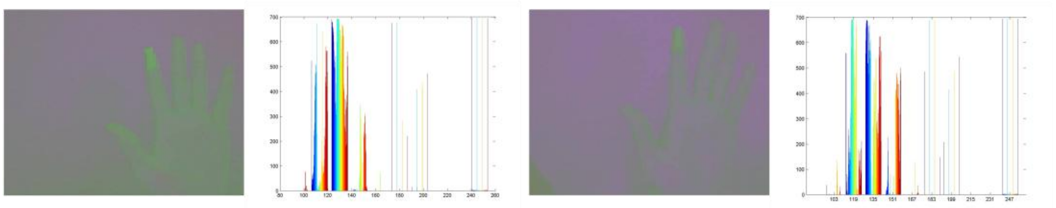


		Next Gesture			
		Right	Left	Close	Far
Previous Gesture	Right	0.04	0.08	0.12	0.07
	Left	0.06	0.03	0.10	0.10
	Close	0.22	0.17	0.01	0.05
	Far	0.08	0.11	0.03	0.02

.21-.25
.16-.20
.11-.15
.06-.10
.00-.05

The proposed approach is applied to control the various functions of a music player (Figure 9). The efficiency achieved is 97.5 per cent. Also, orientations of hand at different angles were tested to ensure that system is robust and is able to detect the gesture with different orientations. Four different angles each at different distance from the webcam were taken to check the adaptability of the system. The results showed that the system was able to recognize the gesture with no deterioration in the actual efficiency. Along with the testing of different orientations of the hands, we also tested our system under different illumination conditions (Figure7). Images having different degrees of brightness, contrast and saturation were captured. It was found that the efficacy of the system was sustained even after changing the illumination conditions and there was no significant change on the recognition rate of the system.

Furthermore, we tried to combat one of the main challenges of gesture recognition system viz. the background complexity. The system was put to trial under different background complexities. The results portray that the system was able to recognize the performed gesture efficiently even with different backgrounds. Figure 8 shows images and histograms of hand under different background conditions.



Brightness	0	Brightness	46
Contrast	32	Contrast	37
Saturation	64	Saturation	42

Figure 7. Histograms of YCbCr Images under Different Illumination Conditions

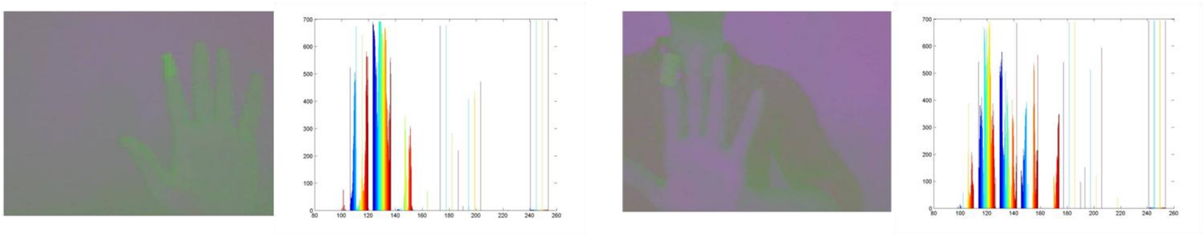


Figure 8. Histograms of YCbCr Images with Different Backgrounds

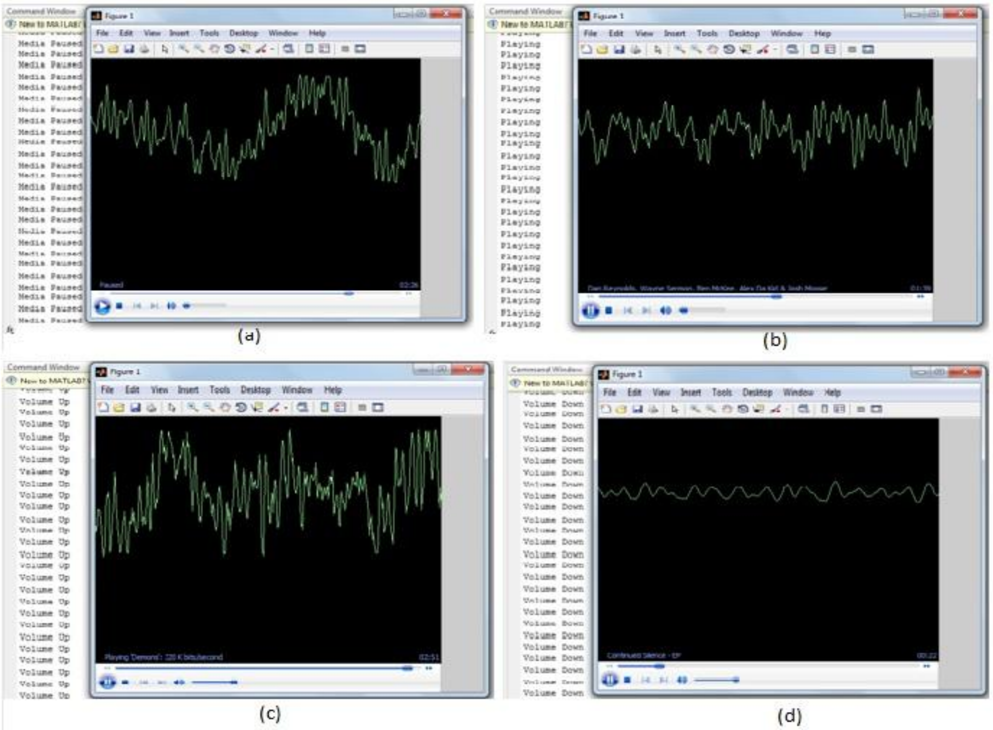


Figure 9. Controls of music player (a) pause (b) play (c) volume up (d) volume down

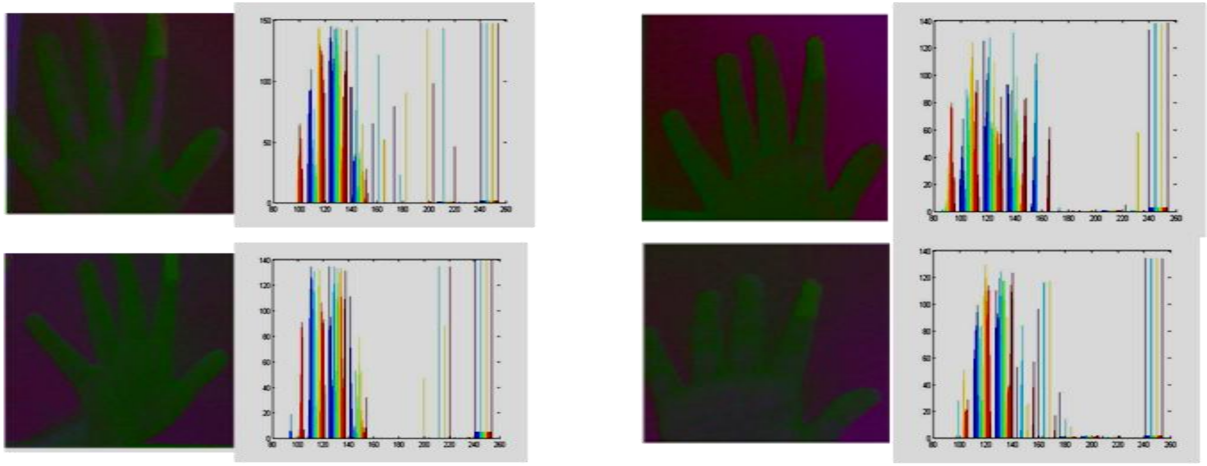


Figure 10. Histograms of YCbCr Images with Different Orientations of the Hand

VI. CONCLUSION AND FUTURE SCOPE

In our research we have explored the area of hand gesture recognition to enhance the interface between human computer interactions. A system was developed which used vision based approach to work with the dynamic gestures in real time. The lags were dealt to a sufficient degree of success. The prototype architecture of the application comprises of a webcam which tracks the movement of the hand; uses a suitable colour model so that the desired information can be extracted; and then on the basis of number of pixels and center analyzes the variations in the hand locations, and finally recognizes the gesture. The experimental results show that the technique works well under different degrees of background complexity and illumination conditions with an overall success rate of 97.5 per cent. A lot of future work is possible in order to increase the adaptability of a gesture recognition system. The gesture recognition accuracy can be further improved by collection of additional gesture information. Moreover, multi stage gestures can also be introduced along with two handed gestures which would further increase the efficacy of the proposed hand gesture recognition systems.

REFERENCES

[1] Feng-Sheng Chen, Chih-Ming Fu, Chung-Lin Huang, "Hand gesture recognition using a real-time tracking method and hidden Markov models," Elsevier -Image and vision computing, 2003, pp. 745–758.
[2] William T. Freeman, Michal Roth, Orientation histograms for hand gesture recognition, Mitsubishi Electric Research Laboratories TR94-03.
[3] Hervé Lahamy and Derek Litchi, "Real-time hand gesture recognition using range cameras," International Journal of Innovative Research & Development, 2013.
[4] Mohamed Alsheakhali, Ahmed Skaik, Mohammed Aldahdouh, Mahmoud Alhelou, "Hand gesture recognition system," 9th International Conference on Information, Communication and Signal Processing (ICICS), 2011.
[5] S. Nagarajan et al "Static hand gesture recognition for sign language alphabets using edge orientation histogram and multi class SVM," International Journal of Computer Applications, 2013, pp. 28-35.
[6] Hassan Ahmed, Todd C. Alexander, Georgios C. Anagnostopoulos "Real time, static and dynamic hand gesture recognition for human-computer interaction".
[7] Radhika Bhatt, Nikita Frenandes and Archana Dhage, "Vision based hand gesture recognition for human-computer interaction," International Journal of Engineering Science and Innovative Technology, 2013.
[8] Sauvik Das Gupta, Souvik Kundu, Rick Pandey, Rahul Ghosh, Rajesh Bag, Abhishek Mallik, "Hand gesture recognition and classification by discriminant and principal component analysis using machine learning techniques," International Journal of Advanced Research in Artificial Intelligence (IJARAI), 2012, pp. 46-51.
[9] Jayashree R.Pansare, Sharavan 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, pp 364-367.
[10] Rajesh Mapari, Dr.Govind Kharat, "Hand Gesture Recognition using Neural Network", International Journal of Computer Science and Network, 2012.
[11] Md. Atiqur Rahman, Ahsan-Ul-Ambia, Md. Aktaruzzaman, "Recognition of Static Hand Gestures of Alphabet in ASL", IJCIT, 2011.
[12] J.L. Raheja, R. Shyam, U. Kumar, P.B. Prasad, “Real-Time Robotic Hand Control using Hand Gesture”, 2nd international conference on Machine Learning and Computing, Bangalore, India, 2010, pp. 12-16.
[13] B. Dorner, "Hand shape identification and tracking for sign language interpretation", Looking at people workshop, Chambery, France, 1993, IJ-CAI.

Authors Profile

Ruchita Gupta is currently pursuing **B.Tech. i**n electronics and communication engineering from Jaypee University of Information Technology, Solan, India. Her research interest includes image processing, gesture recognition.

Mohd. Abbas Murtaza is currently pursuing **B.Tech.** in electronics and communication engineering from Jaypee University of Information Technology, Solan, India. He has published research papers in the field of management, specializing in banking and finance. His research interest also includes gesture recognition.

Meenakshi Sood is Assistant Professor in ECE department of Jaypee University Of Information Technology Solan, Himachal Pradesh, India. She has more than 15 years of teaching experience. Her research interest includes biomedical signal processing, antenna design, meta-materials, image processing and Neural Networks and Fuzzy logic**.**