

# Multi-point Gesture Recognition Using LED Gloves For Interactive HCI

Manisha R.Ghunawat

**Abstract**— The keyboard and mouse are currently the main interfaces between man and computer. In other areas where 3D information is required, such as computer games, robotics and design, other mechanical devices such as roller-balls, joysticks and data-gloves are used. Humans communicate mainly by vision and sound, therefore, a man-machine interface would be more intuitive if it made greater use of vision and audio recognition

This paper describes a gesture-based user interface device, gloves to interact with computer and its integration into application software. Initially there is information about the existing technologies for gesture recognition. Afterwards paper is describing the new robust technology for multi-point gesture recognition and the advantages of this new approach over existing approaches. This system can be used in number of applications ranging from desktop applications to the control of a mobile robot.

**Keywords** -Hue, Saturation, Value color scheme(HSV), JMyron(JM), Red, Green, Blue color scheme. (RGB )

## I. INTRODUCTION

Human society lives through interaction among its entities and their environments. In our daily lives we interact with other people and objects to perform a variety of actions that are important to us. Computers and computerized machines have become a new element of our society. They increasingly influence many aspects of our lives: for example, the way we communicate, the way we perform our actions, and the way we interact with our environment. A new concept of interaction has, thus, emerged: human-computer interaction (HCI). Although the computers themselves have advanced tremendously, the general problem is quite challenging due to a number of issues including the complicated nature of static and dynamic hand gestures, complex backgrounds, and occlusions and the common HCI still relies on simple mechanical devices - keyboards, mice and joysticks - that tremendously reduce the effectiveness and naturalness of such interaction. This limitation has become even more evident with the emergence of a new concept surrounding this interaction - virtual reality. However, new means of HCI have to be available for us to perform interactions in such environments in a more natural way. Ever since the early days of computers we have been attempting to make them understand our speech. But only in the last several years has there been an increased interest in trying to introduce the other means of human-to-human interaction to the field of HCI. These new means include a class of devices based on the spatial motion of the human arm: hand gestures. Human hand

gestures are a means of non-verbal interaction among people. They range from simple actions of pointing at objects and moving them around to the more complex ones that express our feelings or allow us to communicate with others. To exploit the use of gestures in HCI it is necessary to provide the means by which they can be interpreted by computers. The HCI interpretation of gestures requires that dynamic and/or static configurations of the human.

Recently strong efforts have been carried out to develop intelligent and natural interfaces between users and computer systems based on human gestures. Gestures provide an intuitive interface to both human and computer. Thus such gesture based interfaces can not only substitute the common interface devices, but also can be exploited to extend their functionality.

Attacking the problem in its generality requires elaborate algorithms requiring intensive computer resources. The main motto for this work is to make the computer to recognize the 3D hand gestures performed by the human. Due to real-time operational requirements, an efficient algorithm is computed. Early approaches to the hand gesture recognition problem involved the use of markers on the finger tips. An associated algorithm is used to detect the presence and color of the markers, through which one can identify which fingers are active in the gesture. The inconvenience of placing markers on the user's hand makes this an infeasible approach in practice. Device-based techniques use a glove, stylus, or other position tracker, whose movements send signals which the system uses to identify the gesture. For example, sensors on the gloves relay information about the wearer's hand.

In this project, a simple but fast and efficient algorithm is proposed. The user is required to stand in front of the webcam and make hand gestures. The valid frames that are in the field of sight of the cam are captured. Pre-processing is done on the images. This includes blurring or sharpening of the image, obtaining the RGB values, gray scaling and thresholding. Afterwards the finger count is obtained and the vectors are calculated. Depending on the finger count corresponding action is taken for human computer interface. This approach enables users to use any application with a very fast speed and also in real time. This also reduces the ergonomics. The hardware used for it is economical and easily available, which makes the overall cost acceptable to the user. It finds use in number of applications like home security systems, video game controllers and industry robots. [1][2]

## II. EXISTING TECHNOLOGIES

- *Data gloves & electromechanical devices:*

This method employs sensors (mechanical or optical) attached to a glove that transducers finger flexion into electrical signals for determining the hand posture or we can use colored gloves instead. [3]



Fig.1

Disadvantages:

1. These devices limit the speed and naturalness of interaction.
2. This approach forces the user to carry a load of cables which are connected to the computer and hinders the ease and naturalness of the user interaction.
3. Also does the user needs visualization of the gloves impact, because there is no spatial feedback and therefore the user won't know what he is operating.

- *Vision Based Techniques*

Vision based techniques provide a natural way for controlling robots. Visual gesture recognition system for controlling robots by using Fuzzy-C Means Clustering algorithm. The proposed method is applied for recognizing both static and dynamic hand gestures. In dynamic hand gesture recognition, instead of processing all video frames, key frames are extracted by using Hausdorff distance method. After key frame extraction, a sequence of static gesture recognition operations is done for recognizing these key frames. To overcome the limitations of such electro-mechanical devices, vision based techniques are introduced. Vision based techniques do not require wearing of any contact devices on hand, but use a set of video cameras and computer vision techniques for recognizing gestures.

Disadvantages:

1. In the real-world, visual information could be very rich, noisy, and incomplete, due to changing illumination, clutter and dynamic backgrounds, occlusion, etc.



Fig.2

1. The malfunctions or mistakes of Vision-based interaction incur much loss and thus the computer makes more wrong decisions.

- *Hand Gesture Recognition Using Hidden Markov Model*

Hand gesture recognition from visual images has a number of potential applications in human-computer interaction, machine vision, virtual reality, machine control in industry, and so on. Most conventional approaches to hand gesture recognition have employed data gloves, but for a more natural interface, hand gestures must be recognized from visual images without using any external devices. Our research is intended to draw and edit graphic elements by hand gestures. As a gesture is a continuous motion on a sequential time series, the HMM (Hidden Markov Model) must be a prominent recognition tool. The most important thing in hand gesture recognition is what the input features are that best represent the characteristics of the moving hand gesture. [6][7]

Disadvantages:

1. HMMs are computationally expensive and require large amount of training data. Performance of HMM-based systems could be limited by the characteristics of the training dataset.
2. The types of prior distributions that can be placed on hidden states are severely limited.
3. It is not possible to predict the probability of seeing an arbitrary observation.
4. Still slow in comparison to other methods

- *Appearance Based Recognition of American Sign Language Using Gesture Segmentation*

The sign language is the fundamental communication method between the people who suffer from hearing defects. In order for an ordinary person to communicate with deaf people, a translator is usually needed the sign language into natural language and vice versa.

The goal is to develop a system for automatic translation of static gestures of alphabets in American Sign Language. In doing so three feature extraction methods and neural network is used to recognize signs. The system deals with images of bare hands, which allows the user to interact with the system in a natural way. An image is processed and converted to a feature vector that will be compared with the feature vectors of a training set of signs. The system is rotation, scaling of translation variant of the gesture within the image, which makes the system more flexible. The system is implemented and tested using data sets of number of samples of hand images for each signs. Three feature extraction methods are tested and best one is suggested with results obtained from ANN. The system is able to recognize selected ASL signs with the accuracy of 92.22%.

The system is proved robust against changes in gesture. Using Histogram technique we get the misclassified results. Hence Histogram technique is applicable to only small set of ASL alphabets or gestures which are completely different from each other. It does not work well for the large or all 26 number of set of ASL signs. For more set of sign gestures segmentation method is suggested. The main problem with this technique is how good differentiation one can achieve. This is mainly dependent upon the images but it comes down to the algorithm as well. It may be enhanced using other image processing technique like edge detection as done in the presenting paper. We used the well-known edge detector like Canny, Sobel and Prewitt operators to detect the edges with different threshold. We get good results with Canny with 0.25 threshold value. Using edge detection along with segmentation method recognition rate of 92.22% is achieved. Also the system is made background independent. As we have implemented sign to text interpreter reverse also implemented that is text to sign interpreter.

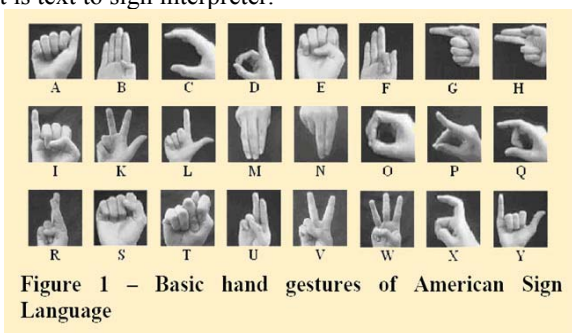


Fig.3

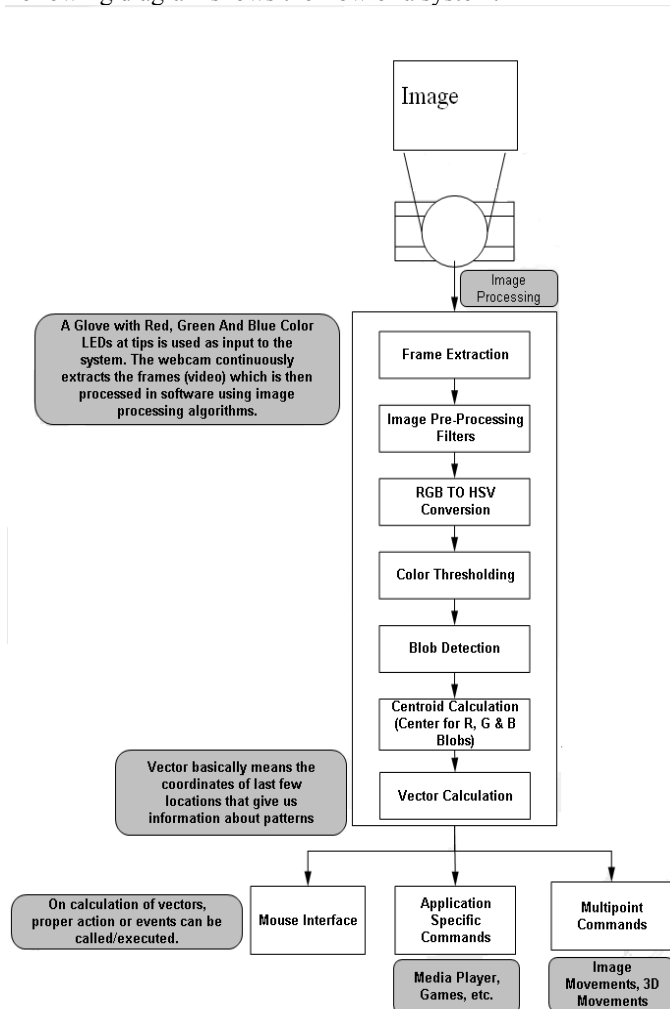
Disadvantages:

1. It does not work well for the large or all 26 number of set of ASL signs.
2. The main problem with this technique is how good differentiation one can achieve.
3. The system deals with images with uniform background.

These are the systems which are already implemented. They have the limitations as described above. So we are developing a system with more advanced features like multi point gesture recognition. We are using multi-color LED's at

### III. NEW APPROACH USING LED GLOVES

In this system we are using a simple gloves with LED's mounted on the tip of the fingers to give a input to the system. Following diagram shows the flow of a system.



#### Description

1. Initially we are capturing the images from the webcam. For that minimum resolution of the image provided by the webcam must be 320 x 240.
2. Input given to the system is in the analog form. Cmos sensor present in the webcam converts analog input into digital. Further processing is done in the digital domain.
3. After that image processing techniques are applied. Following image processing techniques are applied in following sequence

*Image Processing* [8]

- *Gaussian Blur*

Gaussian blur is a widely used effect in graphics software such as Adobe Photoshop, GIMP, Inkscape, and Photofilter. It is typically used to reduce image noise and reduce detail levels. The visual effect of this blurring technique is a smooth blur resembling that of viewing the image through a translucent screen, distinctly different from the bokeh effect produced by an out-of-focus lens or the shadow of an object under usual illumination. Gaussian smoothing is also used as a pre-processing stage in computer vision algorithms in order to enhance image structures at different scales—see scale-space representation and scale-space implementation.

Mathematically speaking, applying a Gaussian blur to an image is the same as convolving the image with a Gaussian or normal distribution. (In contrast, convolving by a circle (i.e., a circular box blur) would more-accurately reproduce the bokeh effect.) Since the Fourier transform of a Gaussian is another Gaussian, applying a Gaussian blur has the effect of low pass filtering the image.

The Gaussian blur is a type of image-blurring filter that uses a normal distribution (also called "Gaussian distribution", thus the name "Gaussian blur") for calculating the transformation to apply to each pixel in the image. The equation of Gaussian distribution is

$$G(r) = \frac{1}{(2\pi(\sigma^2))^{N/2}} e^{-r^2/(2(\sigma^2))}$$

where  $r$  is the blur radius ( $r^2 = u^2 + v^2$ ), and  $\sigma$  is the standard deviation of the Gaussian distribution.



Fig.4

This filter will smoothen the image and remove sharp noise. So Image processing becomes easier.

- *RGB TO HSV*

The RGB color space, used directly by most computer devices, expresses colors as an additive combination

of three additive primary colors of light: red, green, and blue. A commonly used color space that corresponds more naturally to human perception is the HSV color space, whose three components are hue, saturation, and value. The formulas used to convert RGB to HSV depend on which of the RGB components is largest and which is smallest.

The classic RGB color space used in GDI+ is excellent for choosing or defining a specific color as a mixture of primary color and intensity values but what happens if you want to take a particular color and make it a bit lighter or a bit darker or change its saturation. For this you need to be able to use the HSL (Hue, Saturation and Luminance) color space. The conversion of HSL to RGB is a well known algorithm that you can find in numerous places on the web.

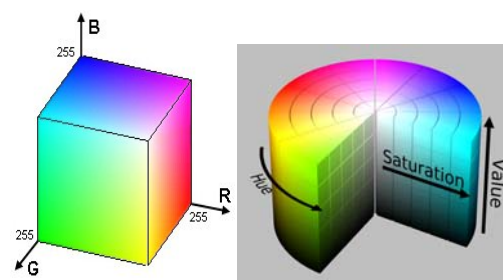


Fig.5

We are performing rgb to hsv conversion because hsv color model is more robust in color detection.

- *Thresholding*

Thresholding is the simplest method of image segmentation. Individual pixels in a grayscale image are marked as "object" pixels if their value is greater than some threshold value (assuming an object to be brighter than the background) and as "background" pixels otherwise. Typically, an object pixel is given a value of "1" while a background pixel is given a value of "0."

*Procedure for Thresholding*

The key parameter in thresholding is obviously the choice of the threshold. Several different methods for choosing a threshold exist. The simplest method would be to choose the mean or median value, the rationale being that if the object pixels are brighter than the background, they should also be brighter than the average. In a noiseless image with uniform background and object values, the mean or median will work beautifully as the threshold, however generally speaking, this will not be the case. A more sophisticated approach might be to create a histogram of the image pixel intensities and use the valley point as the threshold. The histogram approach assumes that there is some average value for the background and object pixels, but that the actual pixel values have some variation around these average values.

However, computationally this is not as simple as we'd like, and many image histograms do not have clearly defined valley points. Ideally we're looking for a method for choosing the threshold which is simple, does not require too much prior knowledge of the image, and works well for noisy images. A good such approach is an iterative method, as follows:

1. An initial threshold (T) is chosen; this can be done randomly or according to any other method desired.
2. The image is segmented into object and background pixels as described above, creating two sets:
  1.  $G_1 = \{f(m,n):f(m,n)>T\}$  (object pixels)
  2.  $G_2 = \{f(m,n):f(m,n)\leq T\}$  (background pixels) (note,  $f(m,n)$  is the value of the pixel located in the  $m^{th}$  column,  $n^{th}$  row)
3. The average of each set is computed.
  1.  $m_1 =$  average value of  $G_1$
  2.  $m_2 =$  average value of  $G_2$
4. A new threshold is created that is the average of  $m_1$  and  $m_2$ 
  1.  $T' = (m_1 + m_2)/2$
5. Go back to step two, now using the new threshold computed in step four, keep repeating until the new threshold matches the one before it.



Original Image                      Image after thresholding

Fig.6

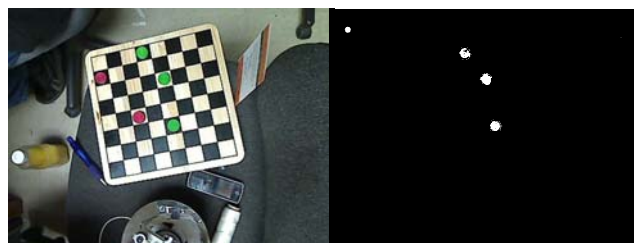
We are performing color thresholding. I.e. only keep colors that belong to gloves. (Convert them to white). Convert everything else to black (background removal).

• *Blob Detection*

In the area of computer vision, 'blob detection' refers to visual modules that are aimed at detecting points and/or regions in the image that are either brighter or darker than the surrounding. There are two main classes of blob detectors (i) differential methods based on derivative expressions and (ii) methods based on local extrema in the intensity landscape. With the more recent terminology used in the field, these operators can also be referred to as interest point operators, or alternatively interest region operators.

*Need for Blob Detection*

There are several motivations for studying and developing blob detectors. One main reason is to provide complementary information about regions, which is not obtained from edge detectors or corner detectors. In early work in the area, blob detection was used to obtain regions of interest for further processing. These regions could signal the presence of objects or parts of objects in the image domain with application to object recognition and/or object tracking. In other domains, such as histogram analysis, blob descriptors can also be used for peak detection with application to segmentation. Another common use of blob descriptors is as main primitives for texture analysis and texture recognition. In more recent work, blob descriptors have found increasingly popular use as interest points for wide baseline stereo matching and to signal the presence of informative image features for appearance-based object recognition based on local image statistics.



Original Image                      Image after blob detection

Fig.7

4. After blob detection we will get x and y co ordinate of each blob. After that we will maintain a queue for each blob. According to points in the queue and we will form the matrix.
5. We will match this matrix with the hardcoded matrix provided in the database. Each hardcoded matrix in the database has some action associated with it. So accordingly action will get performed.

**IV. CONCLUSION**

In today's digitized world, processing speeds have increased dramatically, with computers being advanced to the levels where they can assist humans in complex tasks. Yet, input technologies seem to cause a major bottleneck in performing some of the tasks, under-utilizing the available resources and restricting the expressiveness of application use. Hand Gesture recognition comes to rescue here.

Using multi-colored LED's enables one to specify true multi-point gesture input. Interactive controls can be programmed using multi-point input. This methodology can be extended for more complex applications

## REFERENCES

- [1] Y. B. Lee, S.W. Yoon, C. K. Lee, and M. H. Lee, "Wearable EDA sensor gloves using conducting fabric and embedded system," in *Proc. IEEE Conf. EMBS*, 2006, pp. 6785–6788.
- [2] A. Tognetti, F. Lorussi, M. Tesconi, R. Bartalesi, G. Zupone, and D. De Rossi, "Wearable kinesthetic systems for capturing and classifying body posture and gesture," in *Proc. IEEE EMBS*, 2005, pp. 1012–1015.
- [3] A. Tognetti, N. Carbonaro, G. Zupone, and D. De Rossi, "Characterization of a novel data glove based on textile integrated sensors," in *Proc. IEEE Conf. EMBS*, 2006, pp. 2510–2513.
- [4] F. Axisa, C. Gehin, G. Delhomme, C. Collet, O. Robin, and A. Dittmar, "Wrist ambulatory monitoring system and smart glove for real time emotional, sensorial and physiological analysis," in *Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2004 (IEMBS'04)*, pp. 2161–2164.
- [5] Y. Lee, B. Lee, C. Lee, and M. Lee, "Implementation of wearable sensor glove using pulse-wave sensor, conducting fabric and embedded system," in *Proc. Int. Summer School Med. Devices Biosensors*, 2006, pp. 94–97.
- [6] L.R. Rabiner, "A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proc IEEE*, vol.77, pp.257-285, 1989
- [7] H. Lee and J.H Kim, "An HMM-Based Threshold Model Approach for Gesture Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 961-973, 1999.
- [8] Mark S. Nixon and Alberto S. Aguado. *Feature Extraction and Image Processing*. Academic Press, 2008, p. 88
- [9] F. Iredale, T. Farrington, and M. Jaques, "Global, fine and hidden sports data: Applications of 3D vision analysis and a specialised data glove for an athlete biomechanical analysis system," in *Proc. Annu. Conf. Mechatron. Mach. Vis. Practice*, 1997, pp. 260–264.
- [10] H. H. Asada and M. Barbagelata, "Wireless fingernail sensor for continuous long term health monitoring," MIT Home Automation and Healthcare Consortium, Cambridge, MA, Phase 3, Progr. Rep. 3-1, 2001.
- [11] R. Paradiso and D. De Rossi, "Advances in textile technologies for unobtrusive monitoring of vital parameters and movements," in *Proc. IEEE EMBS*, 2006, pp. 392–395.
- [12] F. Lorussi, E. Scilingo, M. Tesconi, A. Tognetti, and D. De Rossi, "Strain sensing fabric for hand posture and gesture monitoring," *IEEE Trans. Inf. Technol. Biomed.*, vol. 9, no. 3, pp. 372–381, Sep. 2005.
- [13] M. A. Diftler, C. J. Culbert, R. O. Ambrose, R. Platt, Jr., and W. J. Bluethmann, "Evolution of the NASA/DARPA Robonaut control system," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2003, vol. 2, pp. 2543–2548.