**JOURNAL OF INFORMATION, KNOWLEDGE AND RESEARCH IN ELECTRONICS AND COMMUNICATION ENGINEERING**
**ISSN: 0975 – 6779| NOV 16 TO OCT 17 | VOLUME – 04, ISSUE – 02**

# "Hand gesture recognition using Hu moment"

D. U. Dalavi Deshpande Damayanti D[1] . Malekar Shivani D[2]. Taru Pallavi A[3].

Department Of Electronics and Telecommunication,

*Shri Chhatrapati Shivajiraje college of engineering,Dhangawadi*

*Pune,India*

[1]damayantid22@gmail.com

[2]4shivani5@gmail.com

[3]7pallavi9@gmail.com

[4]dalvidaya@gmail.com

## ABSTRACT

**The task of hand gesture recognition is highly challenging due to complex background, presence of non-gesture hand motions, and different illumination environments. Hand gesture recognition system can be used for interfacing between computer and human using gesture. This work presents a technique for a human computer interface through hand gesture recognition that is able to recognize gestures from the American Sign Language. The paper uses skin detection in YCbCr color space for hand regional image and process a hand gesture recognition method based on Hu moments which have invariance property of translation, rotation and size. By extracting Hu moments of the hand contour, this paper applies $K$ - Nearest Neighbors ( $K$ -NN) method to achieve the hand gesture recognition. Experiments prove that the proposed method has very good precision and stability in virtual reality interaction.**

**Keywords**— Image segmentation, Hu moment, $K$ - Nearest Neighbors

## 1. INTRODUCTION

Computer is used by many people either at their work or in their spare-time. Special input and output devices have been designed over the years with the purpose of easing the communication between computers and humans. With the need to find new means of interactions with computers, hand gestures recognition has become a challenging topic of research [1]. Vision-based hand gesture recognition techniques have many advantages, compared with devices such as mice, keyboards or electronic gloves. It allows using directly the hand to interact or communicate with a computer, thus it provides more intuitive means of interaction. However recognizing the shape (posture) and the movement (gesture) of the hand in images is a complex task. The hand is indeed a deformable object with a large number of degrees of freedom.

Hand gesture is the means of communication among the deaf and mute community. Sign Language emerges and evolves naturally within hearing impaired community. Sign Language

communication involves manual and non-manual signals where manual signs involve fingers, hands, arms and non-manual signs involve face, head, eyes and body. Sign Language is a well-structured language with a phonology, morphology, syntax and grammar. Sign language is a complete natural language that uses different ways of expression for communication in everyday life. Sign Language recognition system transfers the communication from human-human to human-computer interaction. The aim of the sign language recognition system is to present an efficient and accurate mechanism to describe text or speech, thus the "dialog communication" between the deaf and hearing person will be smooth. There is no standardized sign language for all deaf people across the world. However, sign languages are not universal, as with spoken languages, these differ from region to region. There are two main approaches used in the sign language recognition that is Glove/Device based and Vision based. In the glove based method the user has to wear a device which carries a load of cables so as to connect the device to a computer. Such devices are expensive and reduce the naturalness of the sign language communication. In contrast, the Vision based method requires only a camera and directly deals with image gestures. It is a two-step process: sign capturing and sign analysis .Vision based methods provide a natural environment to the user and reduces the complications as in the glove based method. Every country has its own sign language with a high level of grammatical variations. The sign language exists in India is commony known as Indian Sign Language (ISL). It has been argued that the same sign language is used in Nepal, Sri Lanka, Bangladesh, and border regions of Pakistan. Examples of other sign languages are the American Sign Language (ASL) [2, 10], British Sign Language (BSL), Korean Sign Language (KSL), and so on.

## 2 BLOCK DIAGRAM

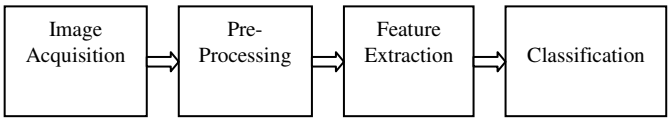


Fig 1: Block diagram of hand gesture recognition

### 2.1IMAGE ACQUISITION

In image acquisition, image is captured by using web camera or camera which is interface with the computer. This capture image then given as an input to the preprocessing step.

### 2.2 PREPROCESSING

Preprocessing is very much required task to be done in hand gesture recognition system. Preprocessing is applied to images before extract features from hand images. Preprocessing consists of two steps:

#### 2.2.1 Skin color Segmentation [4]

Skin color Segmentation is done by converting input RGB image to YCbCr. Values of Cb and Cr are used to set the threshold, which are given in Eq. (2)

$$Cb \in [77\ 127] \quad \text{and} \quad Cr \in [133\ 173] \qquad (2)$$

#### 2.2.2 Image Binirization

Otsu algorithm is used for segmentation purpose and gray scale images are converted into binary image consisting hand and background.

#### 2.2.3 Morphological filtering [4, 5]

Morphology is the set of operation that process images based on shapes. Morphological operations apply a structuring element to an input image, creating an output image of the same size.

## 3 FUNCTIONAL DESCRIPTION

### 3.1 HU MOMENT

Moment invariants have been widely applied to image pattern recognition in a variety of applications due to its invariant features on image translation, scaling and rotation [7, 8, and 9]. The moments are strictly invariant for the continuous function. However, in practical applications images are discrete. Consequently, the moment invariants may change over image geometric transformation. To address this research problem, an analysis with respect to the variation of moment invariants on image geometric transformation is presented, so as to analyze the effect of image's scaling and rotation.

In 1962 Hu derived the equations for the invariant moments. His proposed method is based on algebraic moment invariants. He introduced seven nonlinear functions which are invariants under objects translation. These moments are used as features for image processing, remote sensing, shape recognition etc.

It is possible to calculate moments which are invariant under translation, changes in scale and rotation.

### 3.2 $k$ -NN CLASSIFIER

The $k$ -nearest neighbor ($k$ -NN) [11], decision rule has been a ubiquitous classification tool with good scalability. Past experience has shown that the optimal choice of $k$ depends upon the data, making it laborious to tune the parameter for different applications. One of the various classifier, $k$ -NN classifier' is a case based learning algorithm which is based on a distance or Similarity functions for various pairs of observation such as the Euclidean distance function. It is tried for many applications because of its effectiveness, non-parametric & easy to implementation properties. However, under this method, the classification time is very long & it is difficult to find optimal value of $k$ .Generally, the best alternative of $k$ to be chosen depends on the data. Also, the effect of noise on the classification is reduced by the larger values of k but make boundaries between classes less distinct. By using various heuristic techniques, a good '$k$' can be selected. In order to overcome the above said drawback, modify traditional $k$ -NN with different $k$ values for different classes rather than fixed value for all classes [15].

## 4 DESCRIPTION IN DETAILS

### 4.1 PREPROSESSING

Preprocessing is very much required task to be done in hand gesture recognition system. The database taken here [1] which is standard database in gesture recognition i.e. ASL. Total 15 signs each sign with 32 images are taken. Preprocessing is applied to images before extraction of features from hand images. Preprocessing consist of two steps.

Step 1. Segmentation

Step 2. Morphological filtering

The image captured by camera is the input to the preprocessing step. Captured image is in RGB color space.

### 4.2 COLOR SPACE

A color space is a 4method by which color can specify, create and visualize. As humans, color is define by its attributes of brightness, hue and colorfulness. A computer may describe a color using the amounts of red, green and blue phosphor emission required to match a color [11]. A printing press may produce a specific color in terms of the reflectance and absorbance of cyan, magenta, yellow and black inks on the printing paper [8].A color is thus usually specified using three co-ordinates, or parameters.

#### 4.2.1 Types of Color Space

a. RGB (Red Green Blue)

This is an additive color system based on tri-chromatic theory. Often found in systems that use a CRT to display images. RGB is easy to implement but non–linear with visual perception. It is device dependent and specification of colors is semi–intuitive.

b. HSL (Hue Saturation and Lightness)
This represents a wealth of similar color spaces, alternative names include HSI (intensity), HSV (value), HCI (Chroma / colorfulness), HVC, TSD (hue saturation and darkness) etc. Most of these color spaces are linear transforms from RGB and are therefore device dependent and non–linear. Their advantage lies in the extremely intuitive manner of specifying color. It is very easy to select a desired hue and to then modify it slightly by adjustment of its saturation and intensity.

c.YIQ, YUV, YCBCR, YCC (Luminance - Chrominance)
These are the television transmission color spaces, sometimes known as transmission primaries. YIQ and YUV are analogue spaces for NTSC and PAL systems respectively while YCbCr is a digital standard. This color spaces separate RGB into luminance and chrominance information and are useful in compression applications (both digital and analogue). These spaces are device dependent but are intended for use under strictly defined conditions within closed systems.

4.2 SEGMENTION

The division of an image into meaningful structures, image segmentation, is often an essential step in image analysis, object representation, visualization, and many other image processing tasks. A great variety of segmentation methods has been proposed in the past decades, and some categorization is necessary to present the methods properly here. A distinct categorization does not seem to be possible though, because even two very different segmentation approaches may share properties that defy singular categorization. The categorization presented in this section is therefore rather a categorization regarding the emphasis of an approach than a strict division.

a.Threshold based segmentation
Histogram thresholding and slicing technique are used to segment the image. They may be applied directly to an image, but can also be combined with pre- and post-processing techniques.
b. Edge based segmentation
With this technique, detected edges in an image are assumed to represent object boundaries, and used to identify these objects
c. Region based segmentation
Edge based technique may attempt to find the object boundaries and then locate the object itself by filling them in, a region based technique takes the opposite approach, by (*e.g.*) starting in the middle of an object and then "growing" outward until it meets the object boundaries.
d. Clustering techniques
Although clustering is sometimes used as a synonym for(agglomerative) segmentation techniques, it is used here to denote techniques that are primarily used in exploratory data analysis of high-dimensional measurement patterns. In this context, clustering methods attempt to group together patterns that are similar in some sense. This goal is very similar to what we are attempting to do when we segment an image, and indeed some clustering techniques can readily be applied for image segmentation.
e. Matching
When object is known, then wish to identify in an image (approximately) looks like, this knowledge is used to locate the object in an image. This approach to segmentation is called matching

4.3 FEATURE EXTRCTION

Feature extraction and object recognition are large research areas in the field of image processing and computer vision. Recognition is largely based on the matching of descriptions of shapes. Numerous shape description techniques have been developed, such as analysis of scalar features (dimensions, area, number of corners etc.), Fourier descriptors, moment invariants and boundary chain coding. These techniques are well understood when applied to images and have been developed to describe shapes irrespective of position, orientation and scale. There are different types of feature extraction method. Feature extraction method based on edge, contour, texture, Fourier coefficient moment invariants, scalar features these are the few examples of feature extraction methods. Feature extraction is very important in terms of giving input to a classifier .Here prime feature is Hus seven invariant moment In feature extraction first find edge and region of interest of the segmented and morphological filtered image. By using the algebraic equations of invariant moments, find the central moment of hand in image. With respect to the central moment find out features for the trainee images.

4.4 CLASSIFICATION

In general, process start with a set of data, each data point of which is in a known class. Then want to be able to predict the classification of a new data point based on the known classifications of the observations in the database. For this reason, the database is known as training set, since it trains us what objects of the different classes look like. The process of choosing the classification of the new observation is known as the classification problem, and there are several ways to tackle it. Here consider choosing the classification of the new observation based on the classifications of the observations in the database which it is most similar. However, deciding whether two observations are similar or not is quite an open question. For instance, deciding whether two colors are similar is a completely different process to deciding whether two paragraphs of text are similar. Then before deciding whether two observations are similar, need to find some way of comparing objects. The principle trouble with this is that data could be of many different types - it could be a number, it could be a color, it could be a geographical location, it could

be a true or false (Boolean) answer to a question, etc. This would all require different ways of measuring similarity. It seems that this first problem is one of preprocessing the data in the database in such a way as to ensure that it can compare observations. One common way of doing this is to try to convert all characteristics into a numerical value, such as converting colors to RGB values, converting locations to latitude and longitude, or converting Boolean values into ones and zeros. Once every value is in numbers then imagine a space in which each of characteristics is represented by a different dimension, and the value of each observation for each characteristic is its coordinate in that dimension. Then observations become points in space then the distance between them as their similarity can interpret (using some appropriate metric). Even once decided on some way of determining how similar two observations are, still have the problem of deciding which observations from the database are similar enough to new observation for us to take their classification into account when classifying the new observation new sample point, or by taking only a certain number of the nearest points.

## 5 DATABASE DISCRIPTION

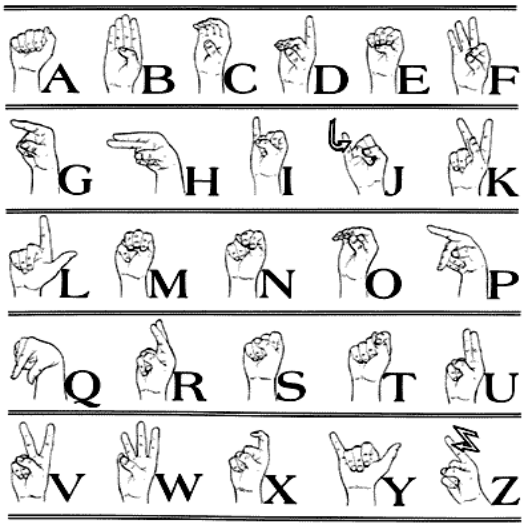


Fig 2: American Sign Language [10]

In this paper all operations are performed on gray scale image. The database consists of 15 hand gesture of International sign language [3]. The letter ae and z has been discarding for their dynamic content. New image for z is introduced. The system works offline recognition i.e. Give test image as input to the system and system tells us which gesture image given as input. The system is purely data dependent. Take RGB image here for the segmentation problem. A uniform background is placed behind the performer to cover all of the workspace.

A good quality of camera is used to capture the hand gesture performed by performer. It produces 24-bit depth level image. Each of the gestures is performed in front of a white background and the users perform the different gesture form ASL, hence easy segmentation of the hand is possible. Each gesture is performed at various scales, translations, and a rotation in the plane parallel to the image-planes. There are total 480 images, 32 images per gesture. The odd images from the database of gestures are used for training and even images for tasting. Here the images are taken with constant color background.

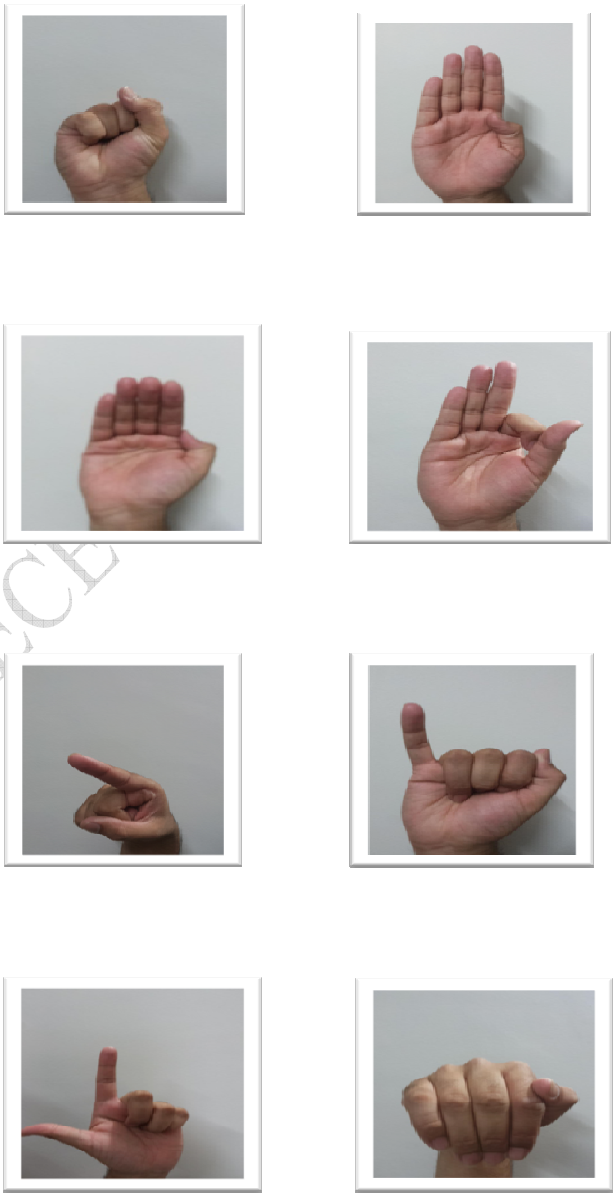

Fig 3: Gesture from Database

## 6. APPLICATION AND ADVANTAGES

### 6.1 APPLICATIONS

#### 6.1.1. Robotics

In robotics, hand gesture recognition used to control the action of robot. According to the gesture detection, robot can perform the action in desired direction.

10.2.2. Secure Communication of Information
Gesture recognition system, transmitter and receiver have the information about gesture shape. According to Gesture shape, transmitter sends the information.

10.2.3. Military Application
Military can use hand gesture recognition system for the surveillance of border line or in Naksli area to detect the mine.

6.2 ADVANTAGES

6.2.1 Sign Language Converter
The hand gesture recognition system used to convert the sign language which is used by deaf and dump people into meaningful signal. So this people can easily communicate with normal world.

6.2.2 High Speed
For this system speed of operation is very fast. The given command will be executed within microsecond as the various operations are performed simultaneously.

6.2.3 High Efficiency
Recognition rate of the hand gesture recognition system is above 90%.

## 7. CONCLUSION

The proposed method, hand Gesture recognition system is developed for dumb and deaf people. The Hu moment is very important feature and used to differentiate between two gestures. In proposed method, feature extraction is done by using Hu moments and classification is done by using $k$ –NN classifier. The proposed method gives 94.58% recognition rate.

## REFERENCES

[1] Ho-Sub Yoon, J. Soh and Y.J.Bae, "Hand gesture recognition using combined features of location, angle and velocity," Patcog 2000.

[2] J. Clerk Chan Wah Ng, S. Ranganath, "Real time gesture recognition system and application," Image and vision computing 2002.

[3] X Shen, G. Hua, Lance William's and Ying Wu, "Dynamic hand gesture recognition: An exemplar-based approach from motion divergence field," Image and vision computing 2011.

[4] S. Conseil, S. Bourennane, L. Martin, "Comparison of FD and Hu moment for hand posture recognition," EUSIPCO 2007.

[5] Pujan Ziaie , Alois Knoll , "An invariant based approach to static Hand Gesture Recognition," 18th ICART 2008

[6] B. Vanjakshi et al," Real object recognition using moment invariants," IJCSC 2012.

[7] Muharrem Mercimek, "Classification of boundary and region shapes using Hu invariant," Sadhana Vol. 30, Part 6, December 2005, pp. 765–775.

[8] Laura Keyes, " Using moment invariants for classifying shapes on large-scale maps," Enterprise Ireland's Strategic Research Grants scheme,1998.

[9] C.Lakshmi Deepika, C. Vimal," Invariant Feature Extraction from Fingerprint Biometric Using Pseudo Zernike Moments," ( IJCCIS)– Vol2. No1. ISSN: 0976–1349,July–Dec2010.

[10] Tiara V. Malloy," Sign Language Use for Deaf, Hard of Hearing, and Hearing Babies," American Society for Deaf Children, July 2003.

[11] drian Ford," Color Space Conversions," August 11, 1998(b).

[12] Yang Song," Informative K-Nearest Neighbor Pattern Classification," The Pennsylvania State University, University Park, PA 16802, U.S.A.

[13] Hussein Rady," Face Recognition using Principle Component Analysis with Different Distance Classifiers," IJCSNS International Journal of Computer Science and Network Security, VOL.11 No.10, October 2011.

[14] Oliver Sutton," Introduction to k Nearest Neighbor Classication and Condensed Nearest Neighbor Data Reduction," February, 2012

[15] Per-Erik Danielsson," Euclidean Distance Mapping," Computer Graphics and Image Processing 14, 227-248,1980.