# Hand Gesture Recognition Technique

**Malav Raj Solanki[1], Shailendra Mishra[2]**

[1]ME Student, Parul Institute of Engineering & Technology,
Waghodia, Vadodara, Gujarat, India
*malavrajsolanki@yahoo.co.in*

[2]Assistant Professor, Parul Institute of Engineering & Technology,
Waghodia, Vadodara, Gujarat, India
*shailendrabemtech@gmail.com*

**Abstract**: ***In this paper, we proposed a robust real-time hand gesture recognition method. In our method, firstly, a specific gesture is required to trigger the hand detection followed by tracking; then hand is segmented using motion and color cues; finally, in order to break the limitation of aspect ratio encountered in most of learning based hand gesture methods, the scale-space feature detection is integrated into gesture recognition. In order to enable a more natural communication with virtual reality systems, automatic hand gesture recognition appears as a suitable means. Hand gesture recognition making use of digital images has been a research topic for many years. However, the use of range cameras for automatic gesture recognition is in its infancy. Hand gesture recognition system can be used for interfacing between computer and human using hand gesture. This work presents a technique for a human computer interface through hand gesture recognition that is able to recognize 25 static gestures from the American Sign Language hand alphabet. The objective of this thesis is to develop an algorithm for recognition of hand gestures with reasonable accuracy.***

**Keywords**: Human Computer Interface (HCI), Hand tracking and Segmentation, Hand Gesture Recognition.

## 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, the two most known are the keyboard and mouse. Every new device can be seen as an attempt to make the computer more intelligent and making humans able to perform more complicated communication with the computer. This has been possible due to the result oriented efforts made by computer professionals for creating successful human computer interfaces. As the complexities of human needs have turned into many folds and continues to grow so, the need for Complex programming ability and intuitiveness are critical attributes of computer programmers to survive in a competitive environment. The computer programmers have been incredibly successful in easing the communication between computers and human. With the emergence of every new product in the market; it attempts to ease the complexity of jobs performed. For instance, it has helped in facilitating tele operating, robotic use, better human control over complex work systems like cars, planes and monitoring systems. Earlier, Computer programmers were avoiding such kind of complex programs as the focus was more on speed than other modifiable features. However, a shift towards a user friendly environment has driven them to revisit the focus area .The idea is to make computers understand human language and develop a user friendly human computer interfaces (HCI). Making a computer understand speech, facial expressions and human gestures are some steps towards it. Gestures are the non-verbally exchanged information. A person can perform innumerable gestures at a time. Since human gestures are perceived through vision, it is a subject of great interest for computer vision researchers. The project aims to determine human gestures by creating an HCI. Coding of these gestures into machine language demands a complex programming algorithm. An overview of gesture recognition system is given to gain knowledge. With the development of ubiquitous computing, current user interaction approaches with keyboard, mouse and pen are not natural enough for them. On PC platform, there are applications such as interactive entertainments and augmented reality requiring more natural and intuitive interface. For mobile or hand held devices, their relatively small size leads to limited input space and encumbered experience with tiny keyboard or touch screen. Hand gesture is frequently used in people's daily life. It's also an important component of body languages in linguistics. So a natural interaction between humans and computing devices can be achieved if hand gestures can be used for communication between human and computing de-vices. Vision based hand gesture interface has been attracting more attentions due to no extra hardware requirement except camera, which is very suitable for ubiquitous computing and emerging applications. Methods for vision based hand gesture recognition fall into two categories: 3D model based methods and appearance model based methods. The performance of vision based gesture interaction is prone to be influenced by illumination changes, complicated backgrounds, camera movement and specific user variance. Many researchers have made effective efforts to deal with these problems. For a real-time application, the expectation is to obtain the best possible images of the hand gesture within the lowest possible time. Some experiments have been conducted with the purpose of defining the best configuration for imaging the hand. This configuration includes, among others, the relative position of the hand and the camera, the influence of the integration time of the camera, the amplitude threshold, and the lighting conditions of the environment, the surrounding objects and the skin color. Some of these parameters will be discussed in this paper which is organized in seven parts as follows. The range camera is presented as well as some of the parameters for defining the best configuration for imaging the user's hand.

## 2. What are Gestures?

It is hard to settle on a specific useful definition of gestures due to its wide variety of applications and a statement can only specify a particular domain of gestures. Many researchers had tried to define gestures but their actual meaning is still arbitrary. Bobick and Wilson have defined gestures as the motion of the body that is intended to communicate with other agents. For a successful communication, a sender and a receiver must have the same set of information for a particular gesture .As per the context of the project; gesture is defined as an expressive movement of body parts which has a particular message, to be communicated precisely between a sender and a receiver. A gesture is scientifically categorized into two distinctive categories: dynamic and static. A dynamic gesture is intended to change over a period of time whereas a static gesture is observed at the spurt of time. A waving hand means goodbye is an example of dynamic gesture and the stop sign is an example of static gesture. To understand a full message, it is necessary to interpret all the static and dynamic gestures over a period of time. This complex process is called gesture recognition. Gesture recognition is the process of recognizing and interpreting a stream continuous sequential gesture from the given set of input data.

**Gesture based Applications:**

Gesture based applications are broadly classified into two groups on the basis of their purpose: multidirectional control and a symbolic language.

**3D Design**: CAD (computer aided design) is an HCI which provides a platform for interpretation and manipulation of 3-Dimensional inputs which can be the gestures. Manipulating 3D inputs with a mouse is a time consuming task as the task involves a complicated process of decomposing a six degree freedom task into at least three sequential two degree tasks. Massachusetts institute of technology has come up with the 3DRAW technology that uses a pen embedded in polhemus device to track the pen position and orientation in 3D.A 3space sensor is embedded in a flat palette, representing the plane in which the objects rest .The CAD model is moved synchronously with the users gesture movements and objects can thus be rotated and translated in order to view them from all sides as they are being created and altered.

**Tele presence:** There may raise the need of manual operations in some cases such as system failure or emergency hostile conditions or inaccessible remote areas. Often it is impossible for human operators to be physically present near the machines. Tele presence is that area of technical intelligence which aims to provide physical operation support that maps the operator arm to the robotic arm to carry out the necessary task, for instance the real time ROBOGEST system constructed at University of California, San Diego presents a natural way of controlling an outdoor autonomous vehicle by use of a language of hand gestures. The prospects of tele presence includes space, undersea mission, medicine manufacturing and in maintenance of nuclear power reactors.

**Virtual reality**: Virtual reality is applied to computer-simulated environments that can simulate physical presence in places in the real world, as well as in imaginary worlds. Most current virtual reality environments are primarily visual experiences, displayed either on a computer screen or through special stereoscopic displays. There are also some simulations include additional sensory information, such as sound through speakers or headphones. Some advanced, haptic systems now include tactile information, generally known as force feedback, in medical and gaming applications.

**Sign Language**: Sign languages are the most raw and natural form of languages could be dated back to as early as the advent of the human civilization, when the first theories of sign languages appeared in history. It has started even before the emergence of spoken languages. Since then the sign language has evolved and been adopted as an integral part of our day to day communication process. Now, sign languages are being used extensively in international sign use of deaf and dumb, in the world of sports, for religious practices and also at work places. Gestures are one of the first forms of communication when a child learns to express its need for food, warmth and comfort. It enhances the emphasis of spoken language and helps in expressing thoughts and feelings effectively .A simple gesture with one hand has the same meaning all over the world and means either 'hi' or 'goodbye'. Many people travel to foreign countries without knowing the official language of the visited country and still manage to perform communication using gestures and sign language. These examples show that gestures can be considered international and used almost all over the world. In a number of jobs around the world gestures are means of communication .In airports, a predefined set of gestures makes people on the ground able to communicate with the pilots and thereby give directions to the pilots of how to get off and on the run-way and the referee in almost any sport uses gestures to communicate his decisions. In the world of sports gestures are common. The pitcher in baseball receives a series of gestures from the coach to help him in deciding the type of throw he is about to give. Hearing impaired people have over the years developed a gestural language where all defined gestures have an assigned meaning. The language allows them to communicate with each other and the world they live in.

**Example of Sign Language**

AMERICAN SIGN LANGUAGE Below figure 1 shows the basic sign language

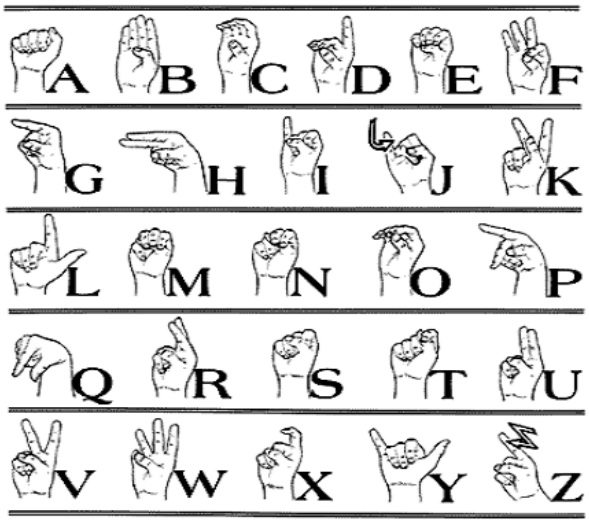


**Figure 1:** American Sign Language

The recognition of gestures representing words and sentences as they do in American and Danish sign language undoubtedly represents the most difficult recognition problem of those applications mentioned before. A functioning sign language recognition system could provide an opportunity for the deaf to communicate with non-signing people without the need for an interpreter. It could be used to generate speech or text making the deaf more independent. Unfortunately there has not been any system with these capabilities so far. In this project our aim is to develop a system which can classify sign language accurately.

## 3. What is Hand Gesture Recognition System?

Vision based analysis, is based on the way human beings perceive information about their surroundings, yet it is probably the most difficult to implement in a satisfactory way. Several different approaches have been tested so far.

• One is to build a three-dimensional model of the human hand. The model is matched to images of the hand by one or more cameras, and parameters corresponding to palm orientation and joint angles are estimated. These parameters are then used to perform gesture classification.

• Second one to capture the image using a camera then extract some feature and those features are used as input in a classification algorithm for classification.

In this paper we have used second method for modeling the system. In hand gesture recognition system we have taken database from standard hand gesture database, prima database. Segmentation and morphological filtering techniques are applied on images in preprocessing phase then using contour detection we will obtain our prime feature that is Local Contour Sequence (LCS). This feature is then fed to different classifiers. We have used three classifiers to classify hand gesture images. Linear classifier is our first classifier and then we have used support vector machine (SVM) and least square support vector machine (LSSVM).

**Block diagram of the system**

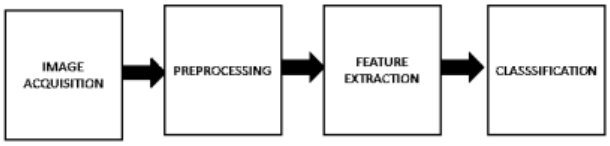


**Figure 2:** Block Diagram Of Hand Gesture Recoginition System .

## 4. Database Description

In this paper all operations are performed on gray scale image .We have taken hand gesture database .The database consist of 25 hand gesture of International sign language. The letter j, z and have been discard for their dynamic content. Gesture is produced as it is a static gesture .The system works offline recognition i.e. we give test image as input to the system and system tells us which gesture image we have given as input. The system is purely data dependent. We take gray scale image here for ease of segmentation problem. A uniform black background is placed behind the performer to cover all of the workspace. The user is required to wear a black bandage around the arm reaching from the wrist to the shoulder. By covering the arm in a color similar to the background the segmentation process is fairly straight forward. A low-cost black and white camera is used to capture the hand gesture performed by performer .it produces 8-bit gray level image. The resolution of grabbed image is 256 * 248. Each of the gestures/signs is performed in front of a dark background and the user's arm is covered with a similar black piece of cloth, 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-plane .There are total 1000 images, 40 images per gesture.

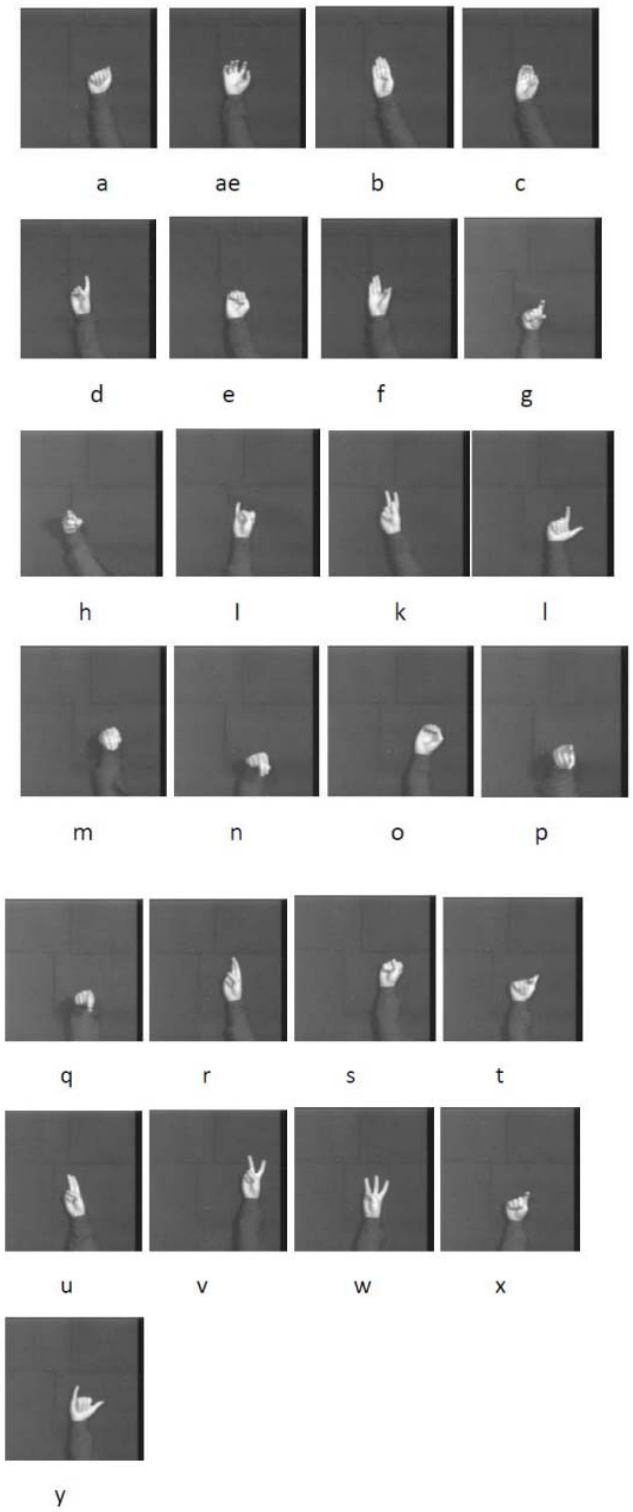


**Figure 3:** Samples of Images rom Database

## 5. Proposed System

### 5.1. Hand Detection

Most of hand detection methods are sensitive to complicated background. Skin color based hand detection is unreliable for the difficulty to be distinguished from other skin-colored objects and sensitivity to lighting conditions. Approaches using shape models require sufficient contrast between object and background. There has been some effort to detect hand in grey image like Ada boost in. It's similar to method in face detection and is adopted in our hand detection. Hand detection in our method is an initial step of interaction. It's important for a gesture interface as it functions as a switch to turn on the interface. The detection uses extended Ada boost method which adopts a new type of feature four box feature like image (d) and (e) in Fig.4 It's a feature type similar to the diagonal rectangle feature in an extension of Viola and Jones's work in . It allows for almost arbitrary area comparisons since the rectangles' locations and sizes are less constrained; even overlapping areas are permitted.


**Figure 4:** (a-c) traditional feature type. (d-e) extended feature type.

### 5.2. Hand Segmentation

Once the hand is detected, the color of hand is collected from the neighborhood of features mean position. Considering trade-off between computational cost and accuracy of description, we use a single Gaussian model to describe hand color in HSV color space. In Fig.5 (a), the bigger white dot denotes features' mean position. Only features within the circle are used to get instant hand color model. Compared with normalized RGB histogram, our method can get better segmentation results. Histogram method is based on the assumption that no other exposed skin color part of user in the certain area around the hand. If there are wooden objects or part of user's face passing by such area, the histogram will deviate and segmentation results will be rapidly degraded. [8]In that case our method can get better results like Fig.5 (b)


**Figure 5:** Hand Segmentation method and results.

### 5.3. Gesture Recognition

In general, recognizing various hand configurations is a difficult and largely unsolved problem. Ong and Bowden [10]gestures as shown in Fig. 6 for navigation interface which is indispensable in interactive interface: These gestures execute operations like shifting focus. LEFT, RIGHT, UP and DOWN are for shifting focus in four directions; OPEN and CLOSE are used to open and close previewof selected image. CLOSE is also interpreted as STOP when shifting focus.


**Figure 6:** Gesture Recognition

### 5.4. The Range Camera and its Sensitivity for Imaging the Hand

The range camera used in this research is the Swiss Ranger SR4000. It is manufactured byMESA Imaging AG, a Zurich-based company of Switzerland. The SR4000 (Figure 7) is a time-of-flight camera that takes both range and intensity images at the same time using an integrated sensor. It has a low resolution of 176 × 144 pixels. Once the image is acquired, the range information is used for generating the x, y, z coordinates in meter for each pixel. The range camera produces images at a rate of up to 54 frames per second.[7]

**Figure 7:** SR 4000

## 6. Experiments

From the[6] experiments, it can be concluded that the gesture recognition algorithm is independentof the hand type (left or right) and the distance between the hand and the camera (Figure 8).

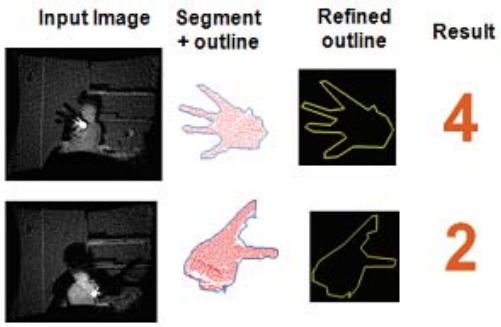


**Figure 8:** Confusion Matrix of First Application

### Moving Object Application

An interface has been designed to visualize simultaneously the range image, the segmented hand and the moving object (Figure 9). Though this research is intended for manipulation of oil and gas reservoirs in a virtual environment, it has been considered in this paper as moving object, a 3D cube with different colours on its six faces.

**Figure 9:** Translation of a moving object

The simultaneous translation and rotation determined from the hand blob and applied to the moving object appear realistic (Figures 9 and 10)

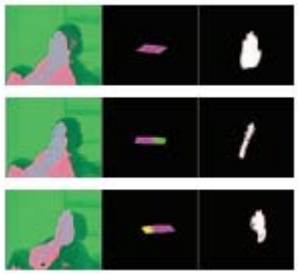

**Figure 10:** Rotation of a moving object

## 7. Conclusions

In recent years a lot of research has been conducted in gesture recognition. The aim of this Paper was to develop an offline Gesture recognition system. We have shown in this paper that offline gesture recognition system can be designed using SVM. It is determined that contour is very important feature and can be used for discrimination between two gesture. The processing steps to classify a gesture included gesture acquisition, segmentation, morphological filtering, contour representation and classification using different technique. The work was accomplished by training a set of feature set which is local contour sequence.

• Otsu algorithm is used for segmentation purpose and gray scale images is converted into binary image consisting hand or background .Morphological filtering techniques are used to remove noises from images so that we can get a smooth contour.

• We have used Local contour sequence as our prime feature. Canny edge detection technique is used to detect the border of hand in image. A contour tracking is applied to find the contour and pixel in contour is numbered sequentially. Local contour sequence for any arbitrary pixel is calculated as perpendicular distance from the chord connecting end points of window size w.

• The main advantage of LCS is that it is invariant to rotation ,translation and scaling so it is a good feature to train the learning machine as we have done with SVM and LSSVM .We have achieved 98.6% accuracy with SVM and 99.2% accuracy with LSSVM.[3]

Altogether our method combines fast hand tracking, hand segmentation and multi-scale feature extraction to develop an accurate and robust hand gesture recognition method. It takes advantage of color and motion cues acquired during tracking to implement adaptive hand segmentation. On the basis of segmentation, multi-scale feature extraction is executed and gestures are recognized with palm-finger decomposition.

Extensive experiments show this method has promising performance with various hand gesture posture aspect ratios and under complicated backgrounds.

## 8. Future Work

- Area of Hand gesture based computer human interaction is very vast. This paper recognizes hand gesture off-line so work can be done to do it for real time purpose. Hand recognition system can be useful in many fields like robotics, computer human interaction and so make this offline system for real time will be future work to do.

- Local contour sequence (LCS) is our prime feature in this project. LCS can be used with used with other features so that it can be optimizing to achieve higher recognition accuracy.

- Support Vector Machine can be modified for reduction of complexity. Reduction of complexity leads us to a less computation time. Reduced complexity provides us less computation time so we can make system to work real time.

## References


[1] 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

[2] Shigeo Abe, “Support Vector Machines for Pattern Classification, second edition”, Kobe University ,Graduate School of Engineering, 2nd edition, Springer-Verlag London Limited 2005, 2010

[3] Simon Haykin, “Neural Networks: A Comprehensive Foundation”, 2nd edition Prentice Hall PTR,1998

[4] O. L. Mangasarian, David R. Musicant, “Lagrangian Support Vector Machines”, Journal of Machine Learning Research, 161-177 (2001)

[5] A. Smola and B. Schölkopf, “Sparse greedy matrix approximation for machine learning,” in Proc. 17th Int. Conf. Mach. Learn., San Francisco, CA, pp. 911–918, 2000.

[6] J. Triesch and C. von der Malsburg, “Robust classification of hand posture against complex background,” in Proceedings of Int. Conf. on Face and Gesture Recognition. Killington, Ver-mont, Apr. 1996, pp. 170–175.

[7] Lars Bretzner, Ivan Laptev, and Tony Lindeberg, “Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering,” in Proceedings of Int. Conf. on Automatic Face and Gesture Recognition. Washington D.C., May 2002, pp. 423–428.

[8] Mathias Kolsch, “Vision based hand gesture interfaces for wearable computing and virtual environments, “PHD Dissertation, UCSB, 2005.

[9] Mathias Kolsch and Matthew Turk, “Robust hand detection,” in Proceedings of Int. Conf. on Automatic Face and Gesture Recognition. Seoul, Korea, May 2004, pp. 614 – 619.

[10] Eng-Jon Ong and Richard Bowden, “A boosted classifier tree for hand shape detection,” in Proceedings of Int. Conf. on Automatic Face and Gesture Recognition. Seoul, Korea, May 2004, pp. 889 – 894.

[11] M. Jones and P. Viola, “Fast multi-view face detection,” Technical Report TR2003-96, MERL, July 2003.

[12] Hyeon-Kyu Lee and Jin H. Kim,” An HMM-Based Threshold Model Approach for Gesture Recognition” IEEE transactions on pattern analysis and machine intelligence, vol. 21, no. 10, October 1999.

[13] Rick Kjeldsen and John Kender, “Finding skin in colour images”, In Proc. IEEE Int. Conf. on autom. Face and Gesture Recognition, pages 3 12-3 17, 1996.

[14] Etsuko Ueda, Yoshio Matsumoto, Masakazu Imai, Tsukasa Ogasawara. “Hand Pose Estimation for Vision-based Human Interface”, IEEE Transactions on Industrial Electronics, Vol. 50, No. 4, pp. 676–684,2003.

[15] Chan Wah Ng, Surendra Ranganath, “Real-time gesture recognition system and application”, Image Vision

[16] Comput, 20(13-14): 993-1007, 2002.


## Author Profile

**Malavraj Solanki** received the B.E degree .in Information Technology From Dharmsinh Desai Institute Of Technology ,Gujarat, India 2009 and M.E. Student in Computer Engineering from Parul Institute of Engineering & Technology, Whaghodia , Vadodara ,India and respectively.