

# A Review of Various Gesture Recognition Techniques

Vaibhavi S. Gandhi<sup>1</sup>, Akshay A. Khond<sup>2</sup>, Sanket N. Raut<sup>3</sup>, Vaishali A. Thakur<sup>4</sup>, Shabnam S. Shaikh<sup>5</sup>

<sup>1</sup>Student (UG) ,Department Of Computer Engineering,  
AISSMS COE Pune, Maharashtra, India.

[vaibhavi.gandhi23@yahoo.co.in](mailto:vaibhavi.gandhi23@yahoo.co.in)

<sup>2</sup> Student (UG) ,Department Of Computer Engineering,  
AISSMS COE Pune, Maharashtra, India.

[akshaykhond@gmail.com](mailto:akshaykhond@gmail.com)

<sup>3</sup>Student (UG) ,Department Of Computer Engineering,  
AISSMS COE Pune, Maharashtra, India.

[sanketraut19@gmail.com](mailto:sanketraut19@gmail.com)

<sup>4</sup>Student (UG) ,Department Of Computer Engineering,  
AISSMS COE Pune, Maharashtra, India.

[thakurrani81@gmail.com](mailto:thakurrani81@gmail.com)

<sup>5</sup>Assistant Professor, Department of Computer Engineering,  
AISSMS COE, Pune, Maharashtra, India.

[shabnamfsayyad@gmail.com](mailto:shabnamfsayyad@gmail.com)

**Abstract:** *Gesture recognition is a subfield of Human Computer Interaction (HCI). Human Computer Interaction has become very attractive field in recent years. Various conventional devices such as mouse, keyboard, joysticks now can be replaced by touch free technologies. To achieve touch-free environment for interacting with computer systems, various algorithms and methodologies have been proposed. In this paper we are surveying methodologies that are proposed previously for hand gesture recognition. We are comparing these techniques on the basis of various parameters such as high accuracy, robustness and lower complexity. The evolution of hand gesture technology from glove based sensing to most recent model based sensing is explained in accordance with their advantages and limitations. Vision based sensing has an advantage that there is no hardware component required and it also gives a benefit of directly using natural motion of hand. For vision based sensing, latest approaches are listed on the basis of parameters of correctness and user friendliness.*

**Keywords:** Human Computer Interaction, Augmented Reality, Gesture Recognition, RGB, HSV, YC<sub>b</sub>C<sub>r</sub>

## 1. Introduction

Interaction of humans with computers using mouse or a touch-pad is now obsolete. Instead of using these devices, a touch free technology can be experienced. Today is the world of using gestures for various technological applications. Gestures are basically non-verbal communication which may include moving a hand with or without any object. Gestures in the field of computers provide a better and efficient Human Computer Interaction (HCI) [1]. It has applications in various fields like Augmented Reality (AR) [2], virtual environment and number of electronic devices. Gestures are used for manipulation, selection and navigation tasks.

Two types of gestures present so far are static and dynamic gestures. Static gestures are those which do not involve any

kind of motion, while the dynamic gestures involve the movement of body parts. Right from data gloves to 3D model based gestures, the technology has evolved. Sensor based technology used external hardware that hindered the natural motion of hand and was restrictive to use. This disadvantage of sensor based technology was overcome by vision based technology which made use of images of gestures to interact with computer.

The paper is organized as follows. Section I gives an introduction and basic terminologies which might be required for reference. Section II gives a description about the gesture recognition approaches which is followed by gesture recognition model described in Section III. Section IV gives a detailed description on dynamic gesture recognition and its constraints. Section V concludes the paper.

## 1.1 Basic Terminologies:

### A. Human Computer Interaction:

It consists of the analysis, design and uses of the interaction which involves the users and computers. The HCI (Human Computer Interaction) is considered to be an association of computer science along with behavioural sciences, designing and several other fields of study.

### B. Augmented Reality:

The technology which is used to superimpose a computer-generated image or a structure on the user's view of the real world, which thus provides a composite view, is called as Augmented Reality (AR).

### C. Gesture Recognition:

The Gesture Recognition [3] is a method in which computer science and language technology are combined together to interpret human gestures via mathematical algorithms.

### D. RGB Color Model:

The RGB color model consists of the primary colors red, green, and blue combined together in various ways to obtain an array of colors [4]. The name of the model is derived from the initials of the three primary colors which form the model.

### E. HSV Color Model:

The HSV color model is the auxiliary representation of the points in the RGB color model. The HSV color model has a cylindrical structure. In HSV color model, H represents hue, S represents saturation and V represents value. In this 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).

### F. $YCbCr$ Color Model:

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

## 2. Gesture Recognition Approaches

### A. Sensor Based Approach:

Sensor based approach makes use of various sensors [5] [6][7] to collect the data of gesture performed. This data is then analysed and conclusions are drawn in accordance with the recognition model. In case of hand gesture recognition various sensors are placed on hand and when the hand performs any gesture, the data is recorded and is further analysed. Data gloves is an example of sensor based technique. The invention of the first data glove was done in 1977[8]. Sensor based approach uses external hardware which hinders the natural motion of hand. Also complex gestures cannot be performed using this method which is major disadvantage of this method.

### B. Vision Based Approach:

Vision based approach [9] makes use of image data of gesture. This method focuses on captured image of gesture and extracts the feature and recognizes it. Color markers is one of the methods of vision based approaches [10]. There are restrictions on use of color bands or color markers. Therefore gestures performed using bare hands are preferred over color markers.

In bare hand gesture recognition system for detection of hand, various skin detection algorithms [11][12][13] are used. The problems faced by these types of system are of cluttered and dynamic background conditions. These problems can be resolved using proper background subtraction methods.

Vision based techniques have evolved with time. These techniques are further divided into model based and appearance based technique. In appearance based technique by observing the appearance of captured gesture the recognition model is created but in case of model based technique by analyzing the captured image of gesture the model of gesture is created and in accordance with that gesture recognition is done[14].

## 3. Gesture Recognition Model

Gesture recognition using sensor based model does not involve image processing activities. It uses data oriented approach. Current technological scenario shows use of vision based technology over sensor based technology to a greater extent. In vision based model various gesture recognition models make use of different algorithms but their approach for gesture recognition is similar. Most of the algorithms follow the flow, first detection of hand, second feature extraction, and third recognition of gesture.

These methodologies make use of different supporting algorithms such as background subtraction, to improve accuracy and robustness of system. Algorithms which are effective as well as are having moderate complexity are selected because very much complex algorithms make the system slower and complex.

Every type of gesture recognition model comprises of following key steps:

### A. Hand Tracking and Segmentation:

For dynamic gestures, the video of hand gestures is captured and it is then divided into frames for further processing. The objective is tracking hand and its segmentation. To do this various approaches and techniques are divided into two categories i.e. sensor based and vision based. In vision based techniques algorithms to track hand, skin detection are used. The results may vary due to different light illumination conditions. HSV and  $YCbCr$  color models are very efficient in skin detection. So, skin is detected by converting the image into HSV color model because RGB color model is very sensitive to light illumination conditions.

Other methods for detecting hand is wearing data gloves or using color markers. But it hinders the natural motion of hand and it is very restrictive. Thus the motion of bare hands is preferred. One of the methods for hand tracking was using Wiimote. Dr. Johnny Chung Lee worked on a project "Tracking your fingers with Wiimote [15]" using the motion tracking method. But later on improved Wiimote having image manipulations came up[16]. Since the Wiimote's camera is sensitive to bright IR light source, array of about 100 LED's is used. Both this LED and Wiimote are placed in front of the

user performing gestures. LED will emit the light on the users hand and the reflected gesture is captured by the Wiimote camera.

In some cases the cluttered background is removed using background subtraction method. HAAR or Viola and Jones methods [17] could be used to detect the face and remove it from background. Also sometimes the unwanted region of whole hand like the wrist is discarded either from the point where the straight skin colour region is encountered [18] or by calculating the wrist point[19]. For tracking the motion of hand, centroid of hand is calculated from each consecutive frame and these points are then joined to form the gesture path [20].

#### B. Feature Extraction:

Once the hand is detected, we have to first extract the feature before gesture recognition. The accuracy and efficiency of the gesture recognition system depends on quality of extracted features. There are three main factors namely orientation, velocity and location [20]. The orientation is the key factor and in some algorithms it is calculated and according to that the feature is classified. Camshift algorithm[21] is used to extract the feature points. In model based approach for creating the model feature points such as finger position, orientation, are extracted from the image[14].

#### C. Gesture Recognition:

One of the method to recognize the gestures is creating the model of hand by using the number of defects present in captured image of hand and comparing this model with our prototype [22]. Another method is using bag of features. In this the captured gesture image is reduced to very small size and only the hand region is present in that image. The key points extracted from the image of hand are provided to SIFT algorithm [17]. The SIFT algorithm i.e. Scale Invariance Feature Transform, is used for detecting same object with different orientations and views. These key points are then matched with the key points in Bag of Features algorithm using K-means clustering [17]. In many approaches contour of hand region is considered for gesture recognition [23].

HMM i.e. Hidden Markov Model [20] is one of the important approaches for gesture recognition as it calculates the probability which is more accurate. Along with HMM, different supporting algorithms like Viterbi algorithm is used for determining the optimal gesture path [20].

## 4. Dynamic Gesture Recognition

The steps mentioned above are applicable to static gestures as well as dynamic gestures. But in case of dynamic gestures the hand is not still, but in motion. This motion of hand is recorded into video with the help of webcam. This video is then analyzed for gesture recognition. For analyzing this gesture first the video is divided into frames and these frames are analyzed. By considering each individual frame at a time as well as relation between consecutive frames, the dynamic gestures are recognized.

### 4.1 Constraints in dynamic gesture recognition:

A. Background-Cluttered Background, Dynamic Background:

The problem of dynamic background could be resolved by using skin detection technique. This will detect the hand and motion of only skin colored object is considered for recognition. In other case if there are two moving objects having skin color then the object having biggest blob [25] could be considered or else the object having more extent of motion is treated as object of interest. Many various methods can be applied to improve efficiency of gesture algorithm by using background subtraction techniques.

#### B. Light Illumination Conditions:

Many vision based approaches gives different results in different light illumination conditions. This leads to failure of system. To overcome this problem a more robust color model against light illumination conditions such as HSV or YCbCr can be used instead of using RGB model images. RGB is very sensitive to illumination conditions compared to other two mentioned models.

Therefore the captured image or frame which is in RGB form is first converted into HSV or YCbCr and then it is processed further. This do not completely overcomes the drawback but the performance is increased.

#### C. Scaling and Orientation Defects:

Scaling and orientation defects occur due to different view and orientation of same object. This defect can be successfully removed by making use of SIFT algorithm [17].

#### D. Processing Speed:

Due to restrictions on processor speed and increased complexity of the system the computational time increases to too much extent. This issue could be solved by using algorithms which are having low complexity and moderate efficiency.

## 5. Conclusion

Gesture recognition approaches discussed in the paper are anomalous in their own way with each of them having their pros and cons. Sensor based approach being more complex in terms of hardware and constraints on natural hand motion whereas vision based approach is more content and comprehensible. Vision based approach is further categorized into appearance based and model based approach with the recent approach being 3D model based approach. Almost all hand gesture recognition models comprises of similar layout of techniques for recognition but their algorithms vary. Gesture recognition models discussed in the paper are analyzed on the basis of their accuracy, robustness and complexity.

Further, Dynamic Gesture Recognition and its methodologies on how the process takes place is explained. Various constraints on dynamic gesture recognition such as dynamic background, light illumination conditions and processing speed are discussed and possible solutions are given. The accuracy using these methods is intended to reach 100% mark, but this has not been the case due to several interdependent entities such as more complex algorithm increases processing speed whereas the opposite takes place if less complex algorithms are used. Therefore, all the entities should be optimized according to their requirement in the algorithm.

## References

- [1] Hewett, Baecker; Card; Carey; Gasen; Mantei; Perlman; Strong; Verplank. "ACM SIGCHI Curricula for Human-Computer-Interaction". ACM SIGCHI. Retrived 15 July 2014.
- [2] Graham, M., Zook, M., and Boulton, A. "Augmented reality in urban places: contested content and the duplicity of code." Transactions of the Institute of British Geographers.
- [3] Mitra S. Acharya, T. , "Gesture Recognition: A Survey" in Systems, Man and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on (Volume:37, Issue:3)
- [4] Charles A. Poynton. Digital Video and HDTV: Algorithms and Interfaces. Morgan Kaufmann. ISBN 1-55860-792-7.
- [5] Zhou Ren, Junsong Yuan, JingjingMeng, Zhengyou Zhang. "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor". IEEE Transactions on Multimedia, Vol.15, No.5, August 2013.
- [6] Thomas Sholmer, Benjamin Poppinga, NielsHenze, Susanne Bol. "Gesture Recognition with a Wii controller". In TEI'08 Proceedings of the 2<sup>nd</sup> international conference on Tangible and embedded interaction Pages 11-14.
- [7] Popa, M. "Hand Gesture recognition based on Accelerometer Sensors". IEEE Networked Computing and Advanced Information Management(NCM),2011.
- [8] Sturman, D.J.,cZeltzer, D. (January 1994). "A survey of glove-based input". IEEE Computer Graphics and Applications 14(1): 30-39.
- [9] Ying Wu and Thomas S. Huang, " Vision-Based Gesture Recognition: A Review" in Gesture-Based Communication in Human-Computer Interaction Lecture Notes in Computer Science Volume 1739,1999, pp 103-115
- [10] P.Mistry, P.Maes. "Six Sense-A Wearable Gestural Interface" .in Proceedings of SIGGRAPH Asis 2009.
- [11] Guoliang Yang, Huan Li, Li Zhang, Yue Cao. " Research on a Skin Colour Detection Algorithm Based on Self-Adaptive Skin Colour Model". IEEE Communications and Intelligence Information Security(ICCIIS),2010.
- [12] Lei Yang, Hui Li, Xiaoyu Wu, Dewei Zhao, Jun Zhai. " An algorithm of skin detection based on texture". IEEE Image and Signal Processing(CSIP),2011.
- [13] Ghouzali, S. ,Hemami, S. , Rziza, M. , Aboutajdine, D. , Mouaddib, E.M. "A skin detection algorithm based on discrete Cosine Transform and generalized Gaussian Density". IEEE image Processing ,2008.ICIP 2008.
- [14] Ali Erol, George Bebis, MirceaNicolescu, Richard D. Boyle, XanderTwombly, Vision-based hand pose estimation: A review in Computer Vision and Image Understanding 108 (2007) 52–73.
- [15] Johnny Chung Lee, Hacking the Nintendo Wii Remote, IEEE Pervasive Computing 7 (3) (2008) 39–45.
- [16] Jiaqing Lin, Hiroaki Nishino, Tsuneo Kagawa, KouichiUtsumiya,A method of two-handed gesture interactions with applications based on commodity devices in Computers and Mathematics with Applications 63 (2012) 448–457.
- [17] Nasser H. Dardas and Nicolas D. Georganas, Fellow, IEEE, Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and SupportVector Machine Techniques in IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 60, NO. 11, NOVEMBER 2011.
- [18] Prashan Premaratne, Sabooh Ajaz, Malin Premaratne, Hand gesture tracking and recognition system using Lucas–Kanade algorithms for control of consumer electronics in Neurocomputing 116 (2013) 242–249.
- [19] Zhi-hua Chen,Jung-Tae Kim,Jianning Liang,Jing Zhang and Yu-Bo Yuan, Real-Time Hand Gesture Recognition Using Finger Segmentation in Hindawi Publishing Corporation The Scientific World JournalVolume 2014, Article ID 267872.
- [20] Chang-Yi Koa\* and Chin-ShyurngFahn, A Human-Machine Interaction Technique: Hand Gesture Recognition Based on Hidden Markov Models withTrajectory of Hand Motion in Procedia Engineering 15 (2011) 3739 – 3743.9.
- [21] D.U. Guanglong, Ping Zhang, Human–manipulator interface using hybrid sensors with Kalman filters and adaptive multi-space transformation.
- [22] Siddharth S. Rautaray, AnupamAgrawal, Real Time Gesture Recognition System for Interaction in Dynamic Environment in Procedia Technology 4 ( 2012 ) 595 – 599
- [23] Peixoto, P.;Goncalves, J.;Araujo, H. "Real time Gesture Recognition System Based on Contour Signatures".in IEEE Pattern Recognition 16<sup>th</sup> International Conference.
- [24] Bellarbi, A. ,Belghit, H.;Benbelkacem, S.;Zenati, N.; Belhocine, M. "Hand Gesture Recognition using contour based method for TableTop surfaces". IEEE Networking,Sensing and Control(ICNSC),2013.
- [25] JoyeetaSingha, Karen Das. "Hand Gesture Recognition Based on Karhunen-Loeve Transform".

## Author Profile



**Vaibhavi S. Gandhi** is pursuing Bachelor's degree in Computer Engineering from Savitribai Phule Pune University from A.I.S.S.M.S College of Engineering, Pune, Maharashtra, India.



**Akshay A. Khond** is pursuing Bachelor's degree in Computer Engineering from Savitribai Phule Pune University from A.I.S.S.M.S College of Engineering, Pune, Maharashtra, India.



**Sanket N. Raut** is pursuing Bachelor's degree in Computer Engineering from Savitribai Phule Pune University from A.I.S.S.M.S College of Engineering, Pune, Maharashtra, India.



**Vaishali A. Thakur** is pursuing Bachelor's degree in Computer Engineering from Savitribai Phule Pune University from A.I.S.S.M.S College of Engineering, Pune, Maharashtra, India.



**Shabnam S. Shaikh** is an Assistant Professor, Department of Computer Engineering, Savitribai Phule Pune University in AISSMS College Of Engineering, Pune, Maharashtra, India.