International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 05 | May 2022 www.irjet.net p-ISSN: 2395-0072

Gesture Recognition System

Shweta Hiremath[1], Krishnakumar Chavan[2], Pratik Kumar[3], Moksha Bafna [4]

[1,2,3,4] *Student, Dept. of Computer Engineering, Sinhgad College of Engineering, Maharashtra , India*

---***---

Abstract - *Gesture-based hand interaction is more natural and efficient than traditional interaction methods like keyboard, mouse and pen. The identification of hand gestures has become most important aspects of human- computer interaction(HCI). The method of gesture recognition identifies the gesture and uses them to control and provide commands to computer. In this paper, the vision-based, glove-based, and depth-based techniques to hand gesture detection are briefly elaborated. Hand gesture recognition/detection is a fun project for a Computer Vision enthusiast to undertake since it contains an easy-to-follow step-by-step approach that allows you to create more sophisticated things on top of these notions.*

***Key Words*: Computer Vision, Human-Computer Interaction, Artificial Intelligence, Gesture Recognition, Image Processing.**

1. INTRODUCTION

Given the widespread availability of mobile devices with embedded cameras, such as smartphones and notebook, a gesture detection system can be a useful tool for interacting more intuitively with these camera-enabled devices than traditional interfaces. The trend toward embedded, ubiquitous computing in residential settings necessitates human-computer interaction forms that are natural, convenient, and efficient. Hand gestures are indeed a fascinating area of study since they can aid communication and provide a natural mode ofengagement that can be applied to a numerous situations. The information gathered was then processed on a computer that was wired to the glove. Hand gestures for human–computer interaction (HCI) began with the introduction of the data glove sensor, which could be made portable by employing a sensor coupled to a micro- controller. By sensing the exact coordinates, gloves identify gestures.

Curvature sensors, angular displacement sensors, optical fiber transducers, flex sensors, and accelerometer sensors were among the sensors that used the same technique dependant on angle of bending, different physical principles are used. Although the above mentioned procedures have yielded positive results, they have a number of drawbacks that make them inappropriate for the elderly, who may endure discomfort and ailments as a result of cable connection issues.

Owing to the fact that real-time segmentation offoreground objects from a cluttered backdrop is a difficult challenge.The most obvious reason is that when a human looks at an image and a computer look at the same image, there is a semantic gap. Images are essentially 3- dimensional matrices to a computer, yet humans can readily figure out what's in them. As a result, computer vision difficulties continue to be a challenge.

The proposed system architecture can recognize gestures and perform various real-time operations like virtual mouse control, zoom in-out, virtual keyboard control, drag and drop and volume control respectively. The dynamic gestures can interact with each other.

1.1 Aim and Objective

The highest priority of the software is to design the Gesture Recognition module. Gesture Recognition module is the first main step of the software.

Major goal of the software is to keep the memory use in reasonable degree. Since the aim of the software is the software can be implemented into mobile devices, memory use is the most critical points that the software will faced. The upper limit of the memory use will be 20% of the memory.

The other important goal is the speed of the software. The process of the software, which is recognition of the gesture, processing it and providing the action, will be done in at most 0.2 second in order that the software will work in reasonable time.

1.2 Motivation

With the rise of ubiquitous computing, traditional user interaction methods such as keyboard, mouse, and pen are no longer adequate. Because of these devices' limitations, the command set that can be used is similarly limited. Natural interaction can be established by using hands as a mode of input.

2. Literature Review

Hasan [7] implemented multivariate Gaussian distribution to identify hand gestures using non-geometric features. The input hand image is categorized using two different methods; skin-based classification using the HSV color model and composite-based composite methods. Some tasks are performed to capture the shape of the hand to remove the hand element; Direction Analysis analytics algorithm is adopted to determine the relationship between statistical parameters (variance and covariance) from data,which is used to calculate the inclination of an object (hand) and a trend by obtaining direction for hand gestures.

Hasan [7] used a limited familiarity with touch perception based on light element simulation. The inserted image is separated using the thresholding method when the background is dark. Any segmented image is normalized, and the center weight of the image is determined, so that the links are moved to match the centroid of the object at the base of axis X and Y. Since this method depends on the medium size of the object, the images produced are of different sizes, for this reason the standard function of the scales is used to overcome this problem which preserves the image size and time, where each of the four blocks is located, rating with a feature different from other block- chain features. Two methods are used to extract features; firstly through the use of edge mages, and secondly through common features where only light value pixels are calculated and some black pixels are ignored to reduce the vector length of the element. The website consists of 6 different touches, using 10 samples per touch, 5 training samples and 5 test samples. The standard feature recognition standard has achieved better performance than the standard feature method, the 95% recognition standard for the previous method and 84% for the latter.

Wysoski introduced a fixed rotation shape using a boundary histogram. The camera used to capture the input image, the skin color detection filter was applied following the merging process to determine the boundary of each group in the merged image using a standard contour tracking algorithm. The image is subdivided into grids and the parameters are set to normal. The boundary is represented as a series of chord sizes used as histograms, by dividing an image into a number of circuits N in a radial manner, depending on the specific angle. During the separation of Neural Networks, MLP and Dynamic Programming(DP) matching is used. Many experiments started with a different feature format in addition to using a different chord size histogram, the FFT chord size. 26 Standard American Sign Language cases were used in the study, Homogeneous network of hand-to-hand detection. In order to determine the color circuit for color separation based on the skin color filter, the handwriting estimates were obtained using network (SGONG). The three elements were extracted using a fingerprint identification method that determined the number of raised fingers and hand shape features, as well as the Gaussian distribution model used for visualization.

3. Design and Analysis

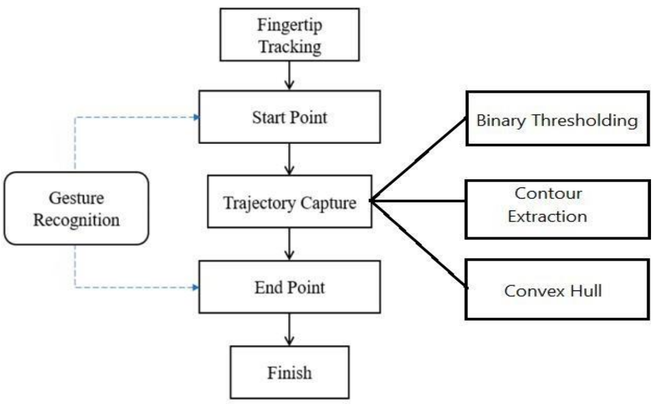

Fig - 1 : System Architecture Diagram

The design of any touch sensitivity system basically involves the following three components: (1) manual detection and tracking; (2) data representation or feature removal and (3) feature enhancement or decision-making. Considering this, a solution that can be applied to any human-computer interaction system is presented in Figure

1. As shown in the diagram, the system first detects and traces the user's hand using a webcam, separates the hand parts from the image and removes the necessary hand features by removing background sound. The features thusobtained are used to identify the user gesture.

4.METHODOLOGY

In this chapter, a method is presented for development of Gesture Recognition System. This also includes explanation various approaches and segmentation proposed.

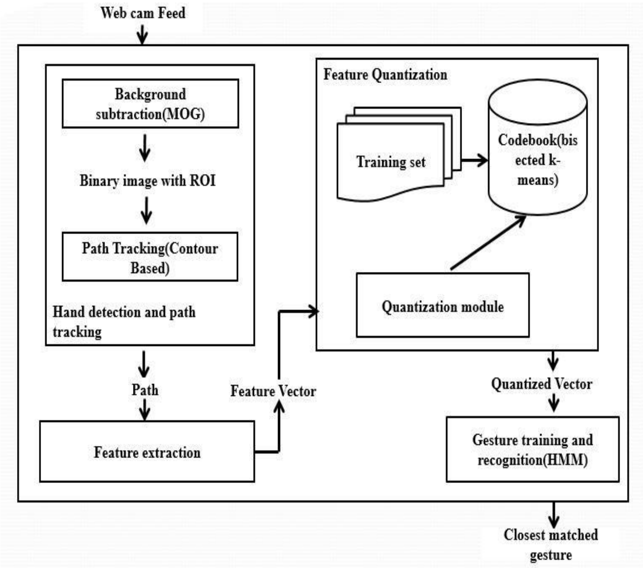

Fig - 2 : System flow Diagram

4.1 Hand Gestures Based on Instrumented Glove Approach

Hand motion and location can be captured using the wearable glove-based sensors. Furthermore, employing sensors linked to the gloves, they can readily offer the exact coordinates of palm and finger placements, orientation, and configurations. However, this strategy requires the user to literally attach hardware to the computer, which prevents the user from interacting with the computer.In addition, these machines are very expensive. The current method based on gloves uses a promising technology which is Industrial-grade haptic technology. Using micro fluidic technology, the glove provides a haptic feedback that allows the user to feel the shape, texture andmovement of the object.

4.2 Hand Gestures Based on Computer Vision Approach

Because this type allows for contact-less human- computers communication, the camera vision-based sensor is a widespread, suitable, and relevant approach. Different camera configurations, such as monocular, fish- eye, TOF, and IR, can be used. Lighting variations, background difficulties, complicated backgrounds, processing time vs. resolution and frame rate are among challenges that this technology faces. The following sections will go over these difficulties.

The commonly used models for the separation of the hands are the discovery of the skin color, which is used in various processes including the separation of objects, the detection of the damaged image, person movement tracking, video observation, HCI applications, facial recognition, hand segmentation, and gesture identification.Two approaches were used to detect the hue of the skin. The first method categories every pixel as skin or not based on its surroundings which is pixel-based skin detection. The second method involves spatially processing skin pixels using information like intensity and texture named area skin detection, which .

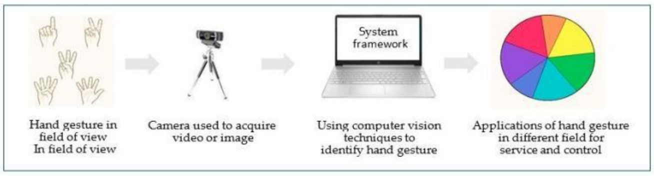

Fig - 3 : Gesture Recognition by Computer Vision approach

4.3 Hand Segmentation

This phase is critical in any procedure to recognize hand movements and to help in the development of the system's functionality by removing extraneous data from the video stream. In general, there are two methods for this purpose.The initial method is based on skin colour, and it is a simple but effective method that is influenced by light conditions in the environment and the type of the background.

The next technique is based on the shape of the hand and uses the notion of convexity to detect the hand. Process of recognizing hand gestures, the position of the palm is particularly significant.

There are other many techniques helpful to detect thehand region from the image may be summarized as:

A. Edge - Detection

B. RGB values as a result of the values of RGB for handcompletely different from the background of the image.

C. Subtraction of background

The colour of the palm is taken from the neighborhood ofthe features mean location once the hand has beendetected. We adopt a Gaussian model to describe handcolour in HSV colour space because of the balancebetween computing expense and accuracy of description. To get a quick hand colour model, only features within thecircle are employed. Our technique outperforms the normalized RGB histogram in terms of segmentation. The histogram approach is based on the presupposition that the user's skin colour in a specific area around the palm is not revealed. If wooden objects or a portion of the user's face pass across this area, the histogram will diverge, and the segmentation results will quickly deteriorate. In thatsituation, our strategy might be more effective.

4.4 Binary Thresholding

Thresholding is an Open-CV method that involves allocating pixel values relative to a threshold value. It is setto zero if the value of pixel is below the limit, otherwise setto the maximum possible value (usually 255). Thresholding is a widely used method of distinguishing between front and rear objects. A threshold(number) is a value consisting of two zones on either side, one below thelimit and one above.

Thresholding is a grayscale image processing technique used in computer vision. Hence, the image must be transformed to grayscale colour space first. Binary Thresholding is an assignment to set the pixel intensity to 0 and 1 based on a certain level of limit so that the only thing we need is taken in the picture.

4.5 Contour Extraction

We can recognize object borders and easily locate them in a picture using contour detection. Many intriguing applications, such as picture foreground extraction,rudimentary image segmentation, detection, and recognition, use it as a starting point. The contour can also be a wave linking spots with comparable colour values, which is useful for form analysis and item recognition. the findContours() function is applied to detect the contours in the image.

```
gray = cv2.cvtColor(src, cv2.COLOR_BGR2GRAY) # convert to grayscale
blur = cv2.blur(gray, (3, 3)) # blur the image
ret, thresh = cv2.threshold(blur, 50, 255, cv2.THRESH_BINARY)
```

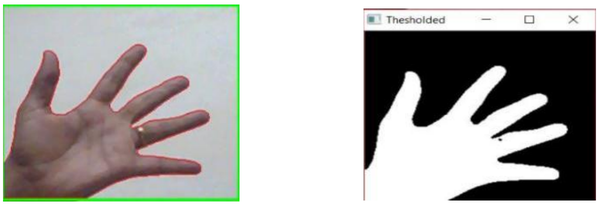

Fig - 4 : Contour Extraction

4.6 Feature Extraction and Recognition

The peaks of the convex hull cluster cover the hand region. We must first understand the idea of the convex Set, which states that all lines connecting any two points within the hull are totally included within it. The precise functioning is carried out after the gesture has been determined. The process of perceiving movement is a dynamic one. Return to the first stage to accept another image to be processed, and so on, after executing the specified command from the gesture.

5. IMPLEMENTATION

This chapter covers the role of various subsystems/modules/classes along with implementation details for the major functionalities.

5.1 Hand Tracking Module

Hand tracking is the process by which a computer uses computer vision to get a hand in embedded image andthen focuses on hand movements and orientation. Hand tracking helps us to develop many programs that use hand movements and postures as their input.

The idea of writing the very same code in different tasks todo manual tracking is part of a program. Creating a manual tracking module solves this issue as the code is written only once. This particular of code is then converted into a module. This module can be imported into various operations.

Here, majorly 2 python libraries are needed :

1. Open-CV - Open-CV is a library used for computer vision applications. Open-CV, is utilized to build an enormous number of applications that work better in real-time. Mainly it is used for image and video processing.

2. Media-pipe - Media-Pipe is a framework used to create audio, video, or timeline data. Using the Media-pipeframework, the most impressive pipelines can be built for a variety of media processing operations.

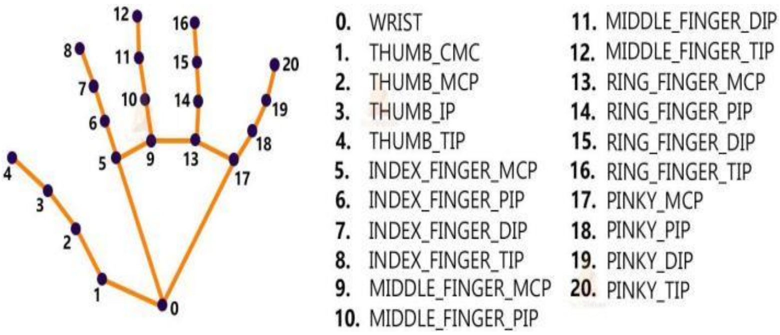

Fig - 5 : Hand Landmark Model

Basically, Media-pipe uses a single-shot palm detection model and once that is done it makes an important point for the 21 3D palm connectors in the obtained hand region. The Media-pipe uses many similar models, the palm- receiving acquisition model that returns the hand-held box from the full image. The cropped image area is provided with a handwritten model defined by a hand detector and returns the most reliable 3D hand points.

5.2 Implementation using Hand Tracking Module

5.2.1 Virtual Mouse

An AI-based method of mouse control was developed using Python and Open-CV with a real-time camera that detects hand gestures, tracking touch patterns instead of real mouse. Here the Machine Learning(ML) algorithm is used. Based on hand gestures, the system can perform the function of a cursor without usage of actual mouse.

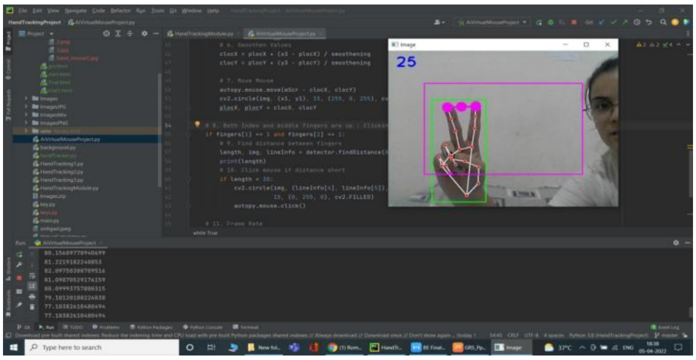

Fig - 5.1 : Hand gesture for Virtual Mouse

5.2.2 Volume Control

We start by looking at the handwriting and then use the local symbols to get the gesture of our hand to change the volume.

Creating a Volume Control with Open-CV can be done in 3easy steps:

1. Find the landmarks of the Hand

2. Calculate the distance between the thumb tip and theindex finger

3. Map distance between thumb tip and index finger tip,then analyze the distance with volume range.

For example, if the distance between the thumb tip and theindex finger tip is within 15 - 220 range then the volume range can be wide - 63.5 - 0.0.

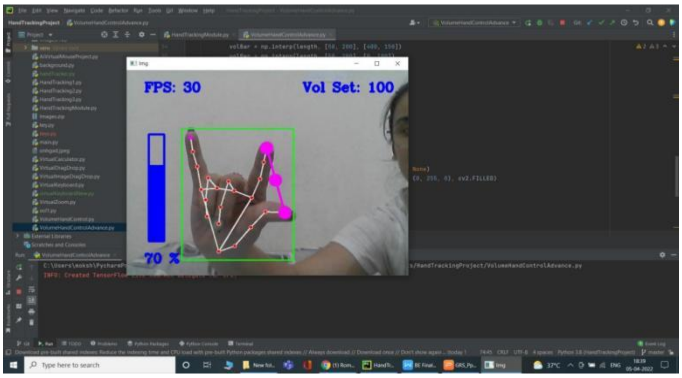

Fig - 5.2 : Hand gesture for Volume Control

5.2.3 Zoom In - Out

Distance between hands utilized for performing zooming operation on GUI. The spread fingers gesture is the inhibitor gesture. Zooming of the object depends upon the distance between two fingers, index finger and . As the distance increases the object zoom out and vice versa. The figure below will depict the operation.

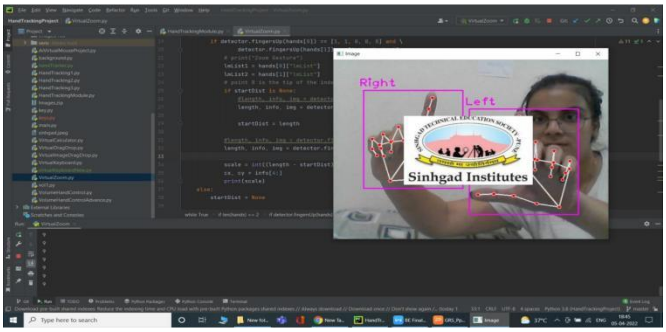

Fig - 5.3 : Hand Landmark Model of Zoom in-out

5.2.4 Drag and Drop

The "drag & drop" operation may be implemented using two hand postures ("grab" and "release") and a motion trajectory necessary for the start and end locations . In this project we are trying to do this operation with the help of single hand. Using two tip index finger we are able to grab and release the objects.

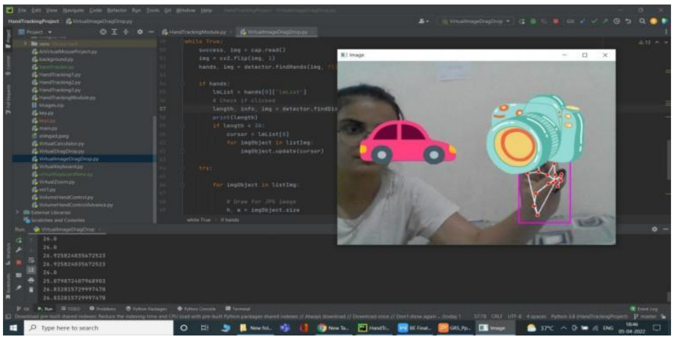

Fig - 5.4 : Hand Landmark Model of Drag and Drop

5.2.5 Virtual Keyboard

We start the hand scanner with the confidence of finding

0.8 and then give it a detector. Then we create more lists depending on the layout of the keyboard and an empty character unit to store typed keys is defined . Launch the keyboard control, then define the function with the word draw () which takes two arguments, an image and a list of buttons and restore the image. Here, within the draw function (), we use the cvzone cornerRect() function to draw rectangular edges at each corner of the key. It will look like the image below.

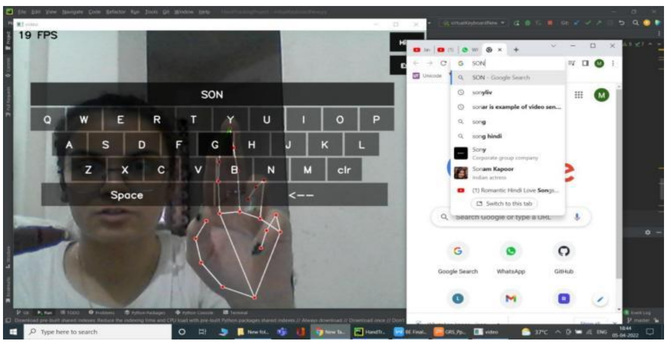

Fig - 5.5 : Hand Landmark Model of Virtual Keyboard

6. APPLICATION

1. Medical Applications – Advanced robotics systems with gesture detection can be placed in hospitals or homes to recognize and treat life threatening conditions like heart attacks or strokes.

2. Entertainment applications – Most video-games today are played either on game consoles, arcade units or PCs, and all require a combination of input devices. Gesture detection can be proved to be very helpful in immersing the players into the gaming world with experience like never before.

3. Automation systems – In homes, offices, transport vehicles and so on, touch recognition can be integrated to maximize usability and reduce the resources needed to create basic or secondary input systems such as remote controls, car entertainment systems with buttons or the likewise.

4. An easier life for the disabled – One of the biggest challenges we face today is to provide different and equitable services to people with disabilities in a different way. Although there are special offers around the world, we still have a lot of room to improve to bring all lives intoequality. Gesture Detection can make life easier for those who are not as lucky as us by eliminating a lot of manual labour.

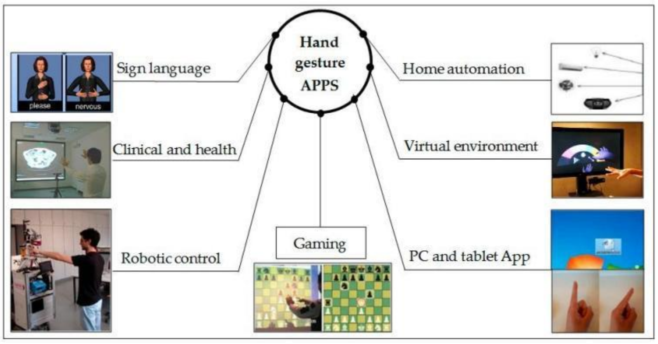

Fig - 6 : Applications of Hand gestures in real-world

7. CONCLUSION

Hand gesture recognition corrects a flaw in computer interface systems. Hand-controlling things is natural, easier, more versatile and cost effective, and problems produced by hardware devices are not needed to be solved because just a camera is needed. The most efficient technology for developing a Gesture Recognition System isComputer Vision. To build an accurate and robust hand gesture detection approach, we combined quick hand tracking, hand segmentation, and multi-scale feature extraction.

It implements adaptive hand segmentation by using colourand motion cues gathered during tracking. We'd like to increase the precision even more in the future, as well as add more gestures to perform more functions.Althoughthis method can identify numerous gestures, it has some cons as well in some cases, such as loss of certain gestures due to variations in the accuracy of the separationalgorithms. Additionally, the process is time consumingthan the initial method due to the same data set in the event since a huge amount of data is used. In addition, the gesture data set cannot be used by other frameworks.

ACKNOWLEDGEMENT

We express our heartiest acknowledgement to all those who supported us and provided valuable guidance whilst completion of this paper. We would like to take this opportunity with great pleasure to express our deep senseof gratitude towards our guide Prof. Jyoti Kulkarni andProf. Mayuri Khade for their valuable guidance, incessant encouragement and co-operation extended towards us during this dissertation work.

REFERENCES

[1] Abhishek, S, K., Qubeley, Fai, L. C., Ho, & Derek. (2016).Glove-based hand gesture recognition sign language translator using capacitive touch sensor. In 2016 IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC) (pp. 334337): IEEE.

[2] Ahuja, M. K., & Singh, A. (2015). Static vision-based Hand Gesture recognition using principal component analysis. Paper presented at the 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE).

[3] Bretzner, L., Laptev, I., & Lindeberg, T. (2002). Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering. Paper presented at the Proceedings of fifth IEEE international conference on automatic face gesture recognition.

[4] Chen, F.-S., Fu, C.-M., & Huang, C.-L. (2003). Hand gesture recognition using a real-time tracking method and hidden Markov models. Image and vision computing, 21(8), 745-758.

[5] Gupta, S., Jaafar, J., & Ahmad, W. F. W. (2012).. Static hand gesture recognition using local gabor filter. Procedia Engineering, 41, 827-832.

[6] Garg, P., Aggarwal, N., & Sofat, S. (2009). Vision based hand gesture recognition. World academy of science, engineering and technology, 49(1), 972-977.

[7] Hasan, M. M., & Mishra, P. K. (2012). Hand gesture modeling and recognition using geometric features: a review. Canadian journal on image processing and computer vision, 3(1), 12-26

[8] G. R. S. Murthy, R. S. Jadon. (2009). "A Review of Vision Based Hand Gestures Recognition, "International Journal of Information Technology and Knowledge Management, vol. 2(2), pp. 405-410.

[9] S. Mitra, and T. Acharya. (2007). "Gesture Recognition: A Survey" IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 37(3), pp. 311- 324,doi:10.1109/TSMCC.2007.893280

[10] Rafiqul Z. Khan, Noor A. Ibraheem, (2012). "Survey on Gesture Recognition for Hand Image Postures", International Journal of Computer And Information Science, Vol. 5(3), Doi:10.5539/cis.v5n3p110

BIOGRAPHIES

Shweta Hiremath, a final year student pursuing Bachelors in Computer Engineering having keen interest in Artificial Intelligence and Neural Networksand wishes to gain deeper knowledge in similar technologies and contribute her skills to real world applications.

Krishnakumar Chavan is currently pursuing Bachelors degree in Computer Engineering. He has sound technical knowledge and is willing to contribute his skills and gain experience.

Pratik Kumar is currently pursuing Bachelors degree in Computer Engineering. He has profound technical knowledge with a great analytic ability and is inclined to contribute his skills and specializations and acquire professional experience.

Moksha Bafna currently pursuing Bachelors degree in Computer Engineering. She has sound knowledge in technical field and is eager to gain more experience by contributing and applying her skills.