

# VIRTUALLY CONTROLLING COMPUTER USING HAND GESTURES AND VOICE COMMANDS

<sup>1</sup>S.Akhil, <sup>2</sup>G.Kalyan, <sup>3</sup>M.Sourav, <sup>4</sup>M.Sridhar, <sup>5</sup>Dr. R. Jegedeesan, <sup>6</sup>Dr. M Sujatha, <sup>7</sup>G.Srikanth

IV Year CSE Students<sup>1,2,3,4</sup> Associate Professor<sup>5,6,7</sup>

Jyothishmathi Institute of Technology And Science, Karimnagar, 505 481, Telangana.

## ABSTRACT

Human-computer interaction (HCI) is all over in our daily lives and is usually achieved using a physical controller such as a mouse, keyboard or touch screen, this hinders the natural user experience as it creates a barrier between the user and the computer as well as it is very expensive and distracting. A lot of space on the desk, there are few hand tracking systems available in the market, but they are complex and expensive, we present here the development of a manual tracking system without tags and gesture recognition using low cost devices. The proposed system is a simple yet effective method that allows quick manual tracking despite the background. Complex, as this system removes motion blur and can detect gestures such as clicking and hand tracking, and the detected gestures are converted into specific functional inputs such as clicking and mouse movement to control other applications. A voice module is integrated to receive and carry out voice commands from the user, allowing an intuitive HCI interface and interactive gaming.

Index terms: Human-computer interaction (HCI), RGB, YCrCb, light-emitting diodes.

## INTRODUCTION

Computer technologies have grown in large scale over the past 50 years, and as technologies have advanced further, current HCI technologies have become narrowing. The most common mode of HCI is keyboard and mouse. These devices have become familiar, but less natural and intuitive, besides, with the presence of normal controllers, there is a barrier between the user and the computer, and therefore a new type of HCI is needed in the market for users to have a better experience while interacting with the computer. The system presented also uses a simple camera that recognizes hand and voice gestures and translates them into system inputs. This is a tag less system and therefore more users friendly. The audio portion converts the users' spoken commands into text and executes them on the computer.

Voice assistants are programs that help you facilitate your daily tasks, such as viewing a weather report, creating reminders, creating shopping lists etc. They can take commands via text messages (online chatbots) or by voice. Voice-based intelligent assistants need a recall word or an alert word to activate the listener, followed by the command. We have a lot of virtual assistants, like Siri from Apple.

## LITERATURE REVIEW

Three modules provided are as follows, camera module, detection module and interface module. The camera module captures and processes the image to facilitate manual

detection, and the detection module detects hands in the image and recognizes the gesture. The interface unit converts these gestures into computer inputs.

**Camera Module:** Here a simple USB webcam is used for image acquisition. The received frames are then scrolled further to subtract the background. The image is converted from RGB to YCrCb and then thresholding is applied to this for a given range to remove the background. Then, Haar-like features are applied to the image to detect the face. The discovered face is then removed because the face color is similar to that of our hand and can be misunderstood as the hand in further steps. Sometimes the face may not be detected because it may appear to be away from the camera and only one connected contour will be detected.

**Detection module:** Here first unwanted small lines are removed and then contour extraction and polygon approximation are applied to extract the desired contour and to make this contour more suitable for shape analysis. Then, the center of the palm can be found by finding the point farthest from the circumference of the contour.

**Interface module:** Using a timer, the system can detect various dynamic gestures such as pinch to zoom. These recognized gestures are passed as input to the system using APIs. The goal is to present the creation of a natural human-computer interaction. Interactive computer games improve whether the computer can understand the player's hand gestures.

## **EXISTING SYSTEM**

It is well known that there are different types of physical computer mouse in modern technology, below discusses the types and differences about physical mouse.

The mechanical mouse is known as the trackball mouse that was popular in the 1990s, and the ball inside the mouse is supported by two rotating rollers in order to detect the movement of the ball itself. One cylinder detects forward/backward movement while the other detects left/right movement. The balls inside the mouse are made of steel and covered with a layer of hard rubber, so that the detection is more accurate.

Optical and laser mouse is a commonly used mouse these days, optical mouse movements rely on light-emitting diodes (LEDs) to detect movements related to the underlying surface, while a laser mouse is an optical mouse that uses coherent laser lights. Compared to its predecessor, the mechanical mouse, the optical mouse no longer relies on pulleys to determine its movement, but instead uses an imaging array of photodiodes. The purpose of implementing this is to eliminate the degradation limitations that plagued the current predecessor, giving it more durability while providing better accuracy and precision. However, there are still some downsides, although the optical mouse works on most diffuse opaque surfaces, it is unable to detect movements on the polished surface.

Google Assistant was an extension of Google Now, and it was designed to be personal while extending existing Google "OK Google" voice controls. Originally, Google Now intelligently pulled relevant information for you. He knows where you work, your meetings and travel plans, which sports teams you like, and what interests you so he can provide the information that matters to you. Google killed Google Now a long time ago, but the Assistant lives in the same space, integrating these personal items with a wide range of voice control. Google Assistant supports both text or voice input and will follow the conversation whichever input method you use.

## **PROPOSED SYSTEM**

Existing systems use portable devices such as a mouse, keyboard, or joystick to interact with the computer system. This hinders the normal interaction between the user and

the computer. Existing systems also include hand gesture recognition but they are only used to recognize simple app-specific hand gestures which are not very popular and rarely used.

The system developed here uses a simple camera available on any laptop or may require a simple, cheap and readily available USB webcam. This system easily recognizes the hand in the image and cuts it without using a depth detection camera. This system recognizes hand gestures, such as touching and moving the hand in a specific direction that you use to control your virtual keyboard and mouse.

This voice assistant can perform operations like google search, find location on google maps, navigate files, current date and time, copy and paste and turn on gesture recognition. Virtual mouse can perform all operations like original mouse and control brightness and volume options..

Algorithm:

- 1) Deep learning object detection: A model is used to give the exact bounding box of the hand in the image.
- 2) Skin Color Threshold: This converts the skin color portion of the bounding box containing the hand to white and the remainder to black.
- 3) Contour extraction and polygon approximation: Helps in find the maximum contour area which is the hand.
- 4) Extraction of convex and convex hull defects: It helps to find the fingertips in the image.
- 5) Simple distance and angle formulas are used to find the click event and the direction of the hand in which it is moving.
- 6) The PyAutoGUI library is used to give system commands.
- 7) The speech recognition library is used to convert voice to speech.

Block Diagram:

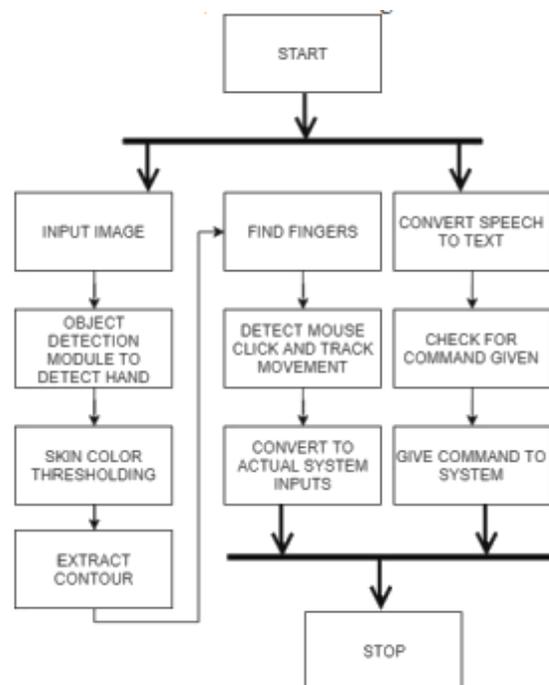


Fig 1:Block Diagram

Figure 1 shows the block diagram of the proposed system. Object detection module: A pre-trained TensorFlow object detection model on a manual dataset is used here. This model takes the picture from the webcam and detects the hand in the picture. The form returns the

coordinates of the frame where the hand is located. A black image is formed in which the frame containing the hand is colored. Skin tone threshold: Skin tone determination is performed as the hand is separated from the remaining background.

Contour extraction: The largest contour of the image is found which is the area of the hand. The Centroid formula is used to get the centroid of the hand.

Finger search: Convex structure defects and convex defects are used to find fingertips and joints in the picture Event Detection: Simple distance and angle formulas are then applied to find the click event and the direction in which the hand moves to control the movement of the mouse. A virtual keyboard that is controlled by the click event is also displayed.

Speech-to-text conversion: In the audio portion, a speech-to-text library is used to convert the user's spoken commands into text. Command Check: The received text is then parsed to perform a specific task.

## CONCLUSION

In conclusion, it is not surprising that a physical mouse has been replaced by a non-physical virtual mouse in human-computer interactions (HCI), where all mouse movements can be quickly performed from your fingers everywhere and at any time without any environmental constraints. A system developed for gesture recognition and hand tracking without HCI tags. Users can interact with computer applications with hand gestures instead of relying on a mouse and keyboard. Therefore, it is more natural and intuitive. The solution presented here is intuitive and simple. It can be easily calibrated to adapt to any skin tone and work with almost any type of background. You can track your hand in real time and can recognize simple hand gestures like a tap. For a more intuitive and simple experience, a voice pane has also been integrated that accepts voice commands and converts them into system commands.

## FUTURE SCOPE

This system can be further improved by using the r-cnn mask model that completely splits the hand from the image without using a color threshold, but is more expensive than computation. This system can be combined with a gesture detection model to detect complex gestures that will aid in controlling games. The NLP model can be combined with the voice part to make it a fully functional conversational robot that can also help control the entire system. This system can be extended to android \ IOS devices where it can be used to control games. There are many features and improvements required to make the software more user-friendly, accurate and flexible in different environments.

## REFERENCES

1. Rafiqul Zaman Khan and Noor Adnan Ibraheem, "Recognition of hand gestures: a review of the literature", International Journal of Artificial Intelligence and Applications (IJAIA), Vol. 3, No.4, July 2012.
2. Ankit Chaudhary, J. L. Raheja , Karen Das , Sonia Raheja, " Intelligent Approaches to interact with Machines using Hand Gesture Recognition in Natural way". International Journal of Computer Science & Engineering Survey (IJCSES) Vol.2, No.1, Feb 2011.
3. Dr.A.Nirmal Kumar, Dr.R.Jegadeesan, Dr.D.Baswaraj, J.Greeda 2019 "Improved Migration Performance In Virtualized Cloud Datacenters" International Journal of Scientific & Technology Research. Volume 8, Issue 09,page no.1515-1518 September 2019. (Scopus indexed)
4. Nolkner C., Ritter H., Visual Recognition of Continuous Hand Postures, IEEE Transactions on neural networks Vol 13, No. 4, July 2002, pp. 983- 994.
5. Nguyen D.D., Pham T.C., Jeon J.W., Fingertip Detection with Morphology and Geometric Calculation, IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis ,USA, Oct 11- 15, 2009, pp. 1460-1465.
6. Dr R Jegadeesan, Dr.C.N.Ravi, Dr.A.Nirmal Kumar 2020 "Automatic Rice Quality Detection Using Morphological and Edge Detection Techniques" ICCCE 2020 3rd International Conference on Communications and cyber Physical Engineering, Metadata of the chapter that will be visualized in Springer Link. Volume, issue, May.2020 Page No.233-242, Springer conference.
7. Lee D. and Park Y., Vision-Based Remote Control System by Motion Detection and Open Finger Counting, IEEE Transactions on Consumer Electronics, Vol. 55, issue 4, Nov 2009, pp. 2308-2313.
8. Joseph J. LaViola Jr., "A survey on techniques and technologies for the recognition of hand gestures and postures".
9. Annadi Jahnvi, Hanumandla Bhavana, Dr. R. Jegadeesan "An Implementation of Detecting Password Pattern In Dictionary Attack" International Journal of Advanced Science and technology. ISSN 2005-42386, Page no: 84-92 July,2019(Indexed by Scopus, Elsevier).
10. S. S. Rautaray and A. Agrawal, Vision Based Hand Gesture Recognition for Human Computer Interaction: A survey, Springer Transaction on Artificial Intelligence Review, pp. 1-54, (2012).
11. Dr.A.Nirmal Kumar, Dr.R.Jegadeesan, Dr.C.N.Ravi, J.Greeda2019 "A Secure Transaction Authentication Scheme using Blockchain based on IOT" International Journal of Scientific & Technology Research. VOLUME 8, ISSUE 10,Page no:2217-2221 OCTOBER 2019. ISSN 2277-8616 (Scopus indexed)
12. M. M. Hasan and P. K. Mishra, Hand Gesture Modeling and Recognition using Geometric Features: A Review, Canadian Journal on Image Processing and Computer Vision, vol. 3(1), pp. 12-26, (2012).
13. Banerjee, A., Ghosh, A., Bharadwaj, K., & Saikia, H. (2014). Mouse control using a web camera based on colour detection. arXiv preprint arXiv:1403.4722.
14. Murthy and Jadon, Hand Gesture Recognition using Neural Networks, In 2nd IEEE International Advance Computing Conference (IACC), pp. 134-138, (2010).