IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

A Smart System using Hand Gestures and Voice

Prof Girish B G[1], Arvind P R[2], Aditya M Gowda[3], Bhoomick R[4], Sushanth U V[5]
Project Guide and Assistant Professor, Department of Computer Science and Engineering[1]
Students, Department of Computer Science and Engineering[2,3,4,5]
Sri Jagadguru Chandrashekaranatha Swamiji Institute of Technology, Chikkaballapur, Karnataka, India

Abstract: *Human-Computer Interaction (HCI) grew enormously over the years. Hand gesture recognition technology brought a brand-new era to the artificial intelligence branch of human-computer interaction. Hand Gesture-based communication is one of the most effortless and natural methods. Identifying and interpreting hand gestures from a continuous sequential stream of input data is called gesture recognition. Amateur or aged people find it hard to identify and press the exact alphabet that they need. Overcoming this difficulty stands as the ultimate goal for our proposed system. Typing has taken many forms, first using a keyboard, gradually changing to touch screens, and now too much easier finger motion tracking systems. The gesture is made by the user is detected by the machine via the image processing techniques and the operation unique to the machine is carried out, thereby minimizing the requirement of any hardware input device. This paper proposes a virtually controlled system that uses hand gestures and voice to perform operations. The system will allow the user to manage cursor, keyboard, and writing functions using hand gestures, and also provides the user with an additional feature to convert the user's voice to text. The proposed system utilizes a webcam as an input device. This application will add a brand-new era of HCI implementation for educational purposes.*

Keywords: Artificial Intelligence, Gesture Recognition, Human-Computer Interaction, Webcam, Efficient Interaction, Intuitive communication

I. INTRODUCTION

This Human society has been using hand gestures for communication since generations. Different hand gestures such as shaking of hands, Thumbs up and Thumbs down have been ever existing in the environment. It is believed that gestures are the easiest way of interacting with anyone. So then why not utilize it to the machines that we are using. In this work, we are demonstrating, real- gestures. The initial setup includes a low-cost USB web camera that can be used for providing input to the system. The complete process is divided into 4 steps which are Video-capturing, frame processing, region extraction, and feature-matching. To the extreme, it can also be called hardware because it uses a webcam for tracking hands

The prospects of this latest technology can serve as a solution to the problem considered. Representing an alphabet using hand movements might not be too difficult in comparison to typing with keyboards for the elderly or amateur people, provided they know the language. Considering a scenario where the person needs to type a small document or even an e-mail and if typing using a keyboard or touch was the only option, they would have to stand behind others to help them. In an effort to take human-computer interaction to further heights, the latest form that typing has adopted is seen to be the much easier "Gesture Recognition Systems". Recognizing gestures is a terminology that deals with the interpretation of human gestures with the help of several mathematical algorithms. Gesture recognition helps users to interact with computer systems naturally without the use of any Physical or hardware devices. Gesture recognition systems overthrew the usage of touch pads, bringing in more efficient communication with devices.

In the world of digital technology, the traditional art of writing is being replaced by digital art. Digital art includes many ways of writing by using a Keyboard, Digital pen, touch-screen devices, using electronic hand gloves, controllers, etc. But in this system, we are using hand gesture recognition with the use of OpenCV and by using python programming, which creates natural interaction between man and machine.

As the technologies are evolving day to day the devices are becoming compact in size. Some devices have become wireless, while some of them have gone latent. The technique of establishing an interconnection between humans and computers has evolved since computer technology's invention. The mouse is an exceptional invention in Human-

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

Computer Interaction (HCI)technology. Though wireless methods such as Bluetooth mouse is developed still, that technology is not completely hardware-free. This paper proposes a system that could make a progress in the field of Human-computer Interaction (HCI) by making device-free technologies.

Speech recognition technology is the ability of a machine or program to identify words and phrases in spoken language and convert them to a machine-readable format. Many speech recognition applications, such as voice dialing, simple text entry, and speech-to-text are in existence today. It is an alternative to traditional methods of interacting with a computer, such as textual input through a keyboard. An effective system can reduce, or replace the reliability of, standard keyboard input. However, human language has numerous exceptions to his own rules. The way words and phrases are pronounced can be vastly altered by accents, dialects, and mannerisms.

II. LITERATURE SURVEY

The proposed system is an application that will detect the numbers written by hand gestures. This will work as a virtual board where numbers and letters can be written by using gesture. The purpose of this project is to create a virtual board which can be used to make online confrontation by moving fingers as the hand gesture. The gesture is made by the user who is detected by the machine through the image processing and the operation unique to the machine is carried out, thereby eliminating the requirement of any hardware input device. The system utilizes nothing but a camera with good quality and can follow the hand of the user in 2D and identify up to four mouse-defined hand motions. The input picture from the camera is transformed first into the colour space of HSV, which detects the skin and removes the backdrop. The application is developed using Open CV and Pytorch.[1]

To provide an immersive freehand painting experience, the system is proposed with a flexible airbrush model making use of the hands tracking capability of Leap Motion Controller. When the user moves hands over the screen, the brush model continually captures his/her hands movement data and extracts multiple control signals which describes multiple gestures. The virtual airbrush moves along with the user's hands movement and its properties change with gesture's change. When the virtual airbrush interacts with the screen, it continually exerts paints over the screen. User test shows that the user can easily create multifarious brush stroke effects by directly operating on the screen. The proposed system is a flexible and customizable virtual 3D airbrush based on the hands tracking technologies. With a common PC and a Leap Motion controller that are affordable for most people, the brush enables user to directly paint on the screen by hands movement without any extra tools or wearable devices.[2]

In the recent years writing in air has been one of the most fascinating and challenging research areas in field of image processing and pattern recognition. It can improve the interface between man and machine in numerous applications. Object tracking is considered as a important task within the field of Computer Vision. The invention of faster computers, good quality video cameras, availability of inexpensive and demands of automated video analysis has given popularity to object tracking techniques. The system describes about the use computer vision to trace the path of the finger. The generated text can also be used for various purposes, such as sending messages, emails, etc. The project focuses on developing a motion-to-text converter that can potentially serve as a software for intelligent wearable devices for writing from the air. The project takes an advantage of this gap and focuses on developing a motion-to-text converter that can potentially serve as a software for intelligent wearable devices for writing from the air.[3]

In human-computer interaction, virtual mouse is implemented with fingertip recognition and hand gesture tracking based on image in a live video is one of the studies. Image processing, a part of signal processing, can consists of an image or a video as input and output as an image or various parameters of it. The main objective is to find the solution for the finger tracking in the real world and the cursor control of a computer is still performed physically. The proposed system describes virtual mouse control using fingertip identification and hand gesture recognition. This study consists of two methods for tracking the fingers, one is by using coloured caps and other is by hand gesture detection. This includes three main steps that are finger detection using colour identification, hand gesture tracking and implementation on on-screen cursor. Hand gesture tracking is generated through the detection of contour and formation of a convex hull around it.[4]

Finger Motion Tracking System intends to identify the English character written in air using our finger. The character is identified and converted into text and displayed on the screen. A web cam is used to capture the LED fitted finger movements and the patterns are identified with one of the characters available in the database. Once the characters are identified, they are printed onto the display screen. A redcoloured LED pointed light source is attached to the finger of

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

the user to make the process easier for finger movement tracking, as it is easier to track down the red colour than to track the colour of the moving finger. It is assumed that the alphabet drawn is a valid English alphabet and there is no red-coloured object other than the LED light within the focus of the web camera. The ultimate objective of the implemented system is to provide an economic solution that can convert the finger movement of an English alphabet to text in a text editor and serve as an efficient means of Human-Computer Interaction.[5]

Hand motion acceptance is critical for human PC connection. The proposed framework is vision-based and uses AI methods and contributions from a PC webcam. The hand motions acknowledgment framework is to supplant the essential pointing gadgets utilized in PC to portray hand motions. Vision-based signal acknowledgment following and motion acknowledgment in the project structure, the hand part is separated from the foundation with the foundation subtraction technique. At this point, fingers are portioned in order to identify and perceive the fingers. A standard classifier is applied to anticipate the names of hand motions. The examinations on the informational index of 1300 pictures show that this strategy performs well and is exceptionally productive. Besides, the technique shows preferred execution over a condition of workmanship strategy on another informational collection of hand motions.[6]

Computer vision has reached its pinnacle, where a computer can identify its owner using a simple program of image processing. In this stage of development, people are using this vision in many aspects of day-to-day life, like Face Recognition, Colour detection. In this framework, computer vision is used in creating an Optical mouse and keyboard using hand gestures. The camera of the computer will capture the image of different gestures performed by a person's hand and according to the movement of the gestures the Mouse or the cursor of the computer will move and also perform right and left clicks using different gestures. The keyboard functions are used with some different gestures, like using one finger gesture for alphabet selection and a four-figure gesture to swipe left and right. The only hardware aspect made used in the project is a webcam and the coding is done in python using the Anaconda platform.[7]

Speech acts as a bridge to communication between two individuals and helps them in expressing their feelings, thoughts, emotions, and ideologies among each other. The process of establishing a communicational interaction between the machine and human is known as Natural Language processing. Speech recognition aims in translating the spoken language into text. The framework will come up with a Speech Recognition model that converts the speech data given by the user as an input into the text format in his desired language. The system is developed by adding Multilingual features to the existent Google Speech Recognition model based on some of the natural language processing principles. The goal is to build a speech recognition model that even facilitates an illiterate person to easily communicate with the computer system in his regional language.[8]

The aim is to control mouse functions using only a simple webcam instead of a traditional or regular hardware device. The Virtual Mouse works as a bridge between the user and the machine and only uses a webcam. It helps the user to interact with the system without any physical devices and control the mouse functions. The concept is implemented using Python programming language and Computer Vision-based library OpenCV. This system has the potential to replace the hardware mouse. The lighting barrier acts as the barrier to this system. That's why the system still can't be enough to replace the traditional mouse as most of the systems are used in poor lighting conditions.[9]

Air-writing is the latest form of user interface that allows us to write through finger movement in the air. Due to the different writing patterns and styles, it became more complicated and difficult than gesture recognition. The written characters are displayed on an imaginary plane without haptic feedback. The user uses a virtual coordinate to canvas the character. The main problem is when it is written as a continuous motion track or trajectory, where each and every point is identified. The system uses a large character dataset. A huge amount of data is required for training the deep learning model. Dataset consists of 30,000 characters, among which, 26,000 are used for training and validation. The rest of the data is used for testing and verifying the model. A simple and robust CNN model is used for creating the Air writing system with an accuracy of 97.5%.[10]

III. PROPOSED SYSTEM

The aim of this system is to develop a cursor control system virtually using hand gestures that perform operations such as left-click, right-click and cursor movement. The project framework focuses on developing a motion-to-text converter that can potentially serve as a software for writing on air. The system will use computer vision to trace the finger's path, and in that way, one can write on air. The generated text can be used for numerous purposes, such as sending messages,

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

emails, etc. A virtual Keyboard is developed which can be used by the user to create documents and also the user can access all the keyboard functions using hand gestures, thus minimizing the use of hardware components. The System will have a voice-to-text converter that interprets the voice message and converts it into a relevant text format. This proposed system overcomes the limitations of the existing system by avoiding the use of the leap motion controller which is an expensive hardware device used for hand detection, instead, we use the in-built python libraries and functions which makes the system less expensive and effective.

IV. METHODOLOGY

The whole system is divided into different modules

Hand Tracking Module: The webcam is used to capture the hand gestures made by the user. The image is captured frame by frame so as to not miss anything. The webcam continuously provides a sequence of images at a particular speed of FPS (Frames per second). The system analyzes each frame and then converts the frame from one color to another. Using Mediapipe, a python library initial hand location is detected. The webcam is used to capture the hand gestures made by the user. The image is captured frame by frame so as to not miss anything. The webcam continuously provides a sequence of images at a particular speed of FPS (Frames per second). The system analyzes each frame and then converts the frame from one color to another. Using Mediapipe, a python library initial hand location is detected. Mediapipe library consists of two different models namely the Palm Detection Model and Hand detection Model in which a full image of a hand is identified and draws a box around the hand, and operates on this boxed image formed by Palm Detector and provides high fidelity 2D hand key point coordinates.

Cursor control Module: This module is designed to control the mouse of the computer and perform basic mouse functions such as movement, scrolling, and clicking using just our two fingers. To enable cursor operations, we use "autopy" library, which is an in-built library in Python language. This library is basically used to enable mouse operations in an easy and efficient way.

Keyboard Module: With this module, we can control the keyboard operations of the computer. The hand tracking module is imported to detect the hand. To enable the keyboard buttons and operations, we use "pynput" library, which is an in-built library in python language. This package pynput.keyboard contains classes for controlling and monitoring the keyboard functions.

Virtual Painter Module: The function of this module is to enable the user to draw on the screen using their fingers. We first import the Hand tracking module which tracks the 21 points on our hand. The fingertip can be used for canvasing and three fingers together can be used for erasing the canvased content on the screen.

Speech Recognition Module: This module is basically used to convert the user's voice to text. To take the input, that is the user's voice we make use of the Pyaudio which allows the user to record voice using the microphone. Then we make use of a speech recognition package in python which converts the spoken words to text.

V. RESULTS AND DISCUSSION

We did a thorough evaluation to check and verify the working of the main phases of the system.

Hand Tracking Module: We can observe that the webcam captures the video and is shown in a discrete window. As we can see that the hand above the wrist is assigned with 21 hand knuckle points. The points are highlighted with a different color and the line that links the point is depicted using a different color. As shown in the figure the whole hand is initially detected and is boxed. Then, the points are localized inside the boxed image.

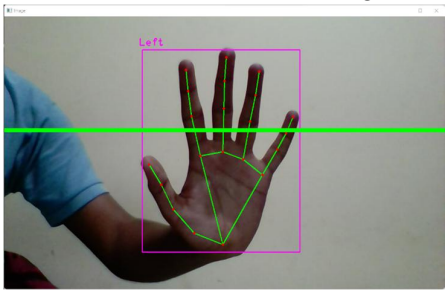

Figure 1. Hand Tracking Module

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

Cursor control module: We can observe that the hand tracking module is used to assign 21 hand knuckle points using which we can control the mouse. The working of the cursor and display is similar to the hand tracking window. As seen in the below figure, the index finger is used to control the scrolling operation whereas the middle and index finger together is used for the clicking operation. The red-colored dot represents the clicking operation.

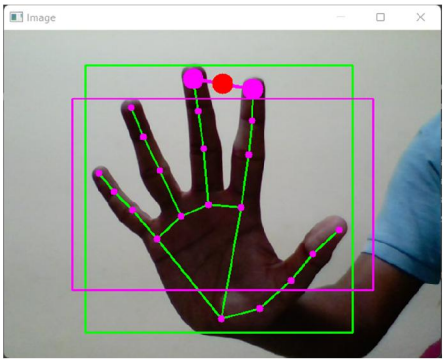

Figure 2. Clicking option using Index and Middle Finger

Keyboard Module: The hand tracking module is used to detect the hand using which we can control the keyboard virtually. As we can observe the index finger is used to scroll over the keyboard and using the index and thumb fingers together, we can select any key on the keyboard.

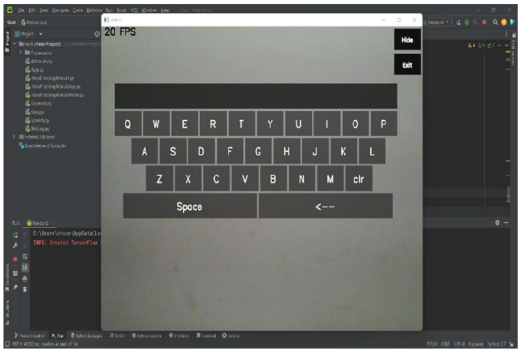

Figure 3. Virtual Keyboard with functions.

Virtual Painter Module: This module also makes use of the hand tracking module to detect the hand. The index finger is used to write, the index and middle fingers together are used to control the cursor functions and the index, middle, and ring fingers together are used to erase the written content.

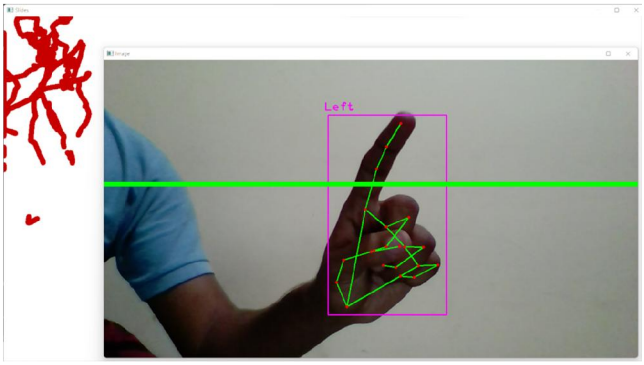

Figure 4. Writing Using the Index Finger.

Speech Recognition Module: We can see that in this module the user's voice is converted to relevant text. The user is allowed to record his/her voice through a microphone which is then recognized by the system and is shown as text to the user. The user is given an amount of time to say or give the input through the microphone. The system then analyses the audio given by the user and then converts it to text and displays it to the user.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 6, June 2022
Impact Factor: 6.252

VI. CONCLUSION

The main purpose of this paper is to implement the idea of a Virtually controlled system. The main tasks performed are cursor operations, Keyboard operations, writing, and voice-to-text converter. From all the above discussion it can be concluded that a Hand gesture-based system plays a significant role in interacting with the computer as a virtual machine and also reduces the hardware cost by eliminating the use of a mouse and keyboard. This system is developed in such a way that a simple Webcam is sufficient enough to identify hand gestures and perform operations. The developed system can be used for educational purposes. Smart classrooms can be implemented using hand gestures which reduces time and also minimizes the use of hardware. The system also acts as a user-friendly device for physically challenged and aged people by providing them with an interface for communication. This system can be convenient for presentations and useful for reducing the workspace and burden of extra hardware devices. Since it replaces the hardware devices, it brings the user and their workspace much closer to them than before.

REFERENCES

[1]. F. Soroni, S. a. Sajid, M. N. H. Bhuiyan, J. Iqbal and M. M. Khan, "Hand Gesture Based Virtual Blackboard Using Webcam," 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2021, pp. 0134-0140, doi: 10.1109/IEMCON53756.2021.9623181.

[2]. R. Lyu et al., "A flexible finger-mounted airbrush model for immersive freehand painting," 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), 2017, pp. 395-400, doi: 10.1109/ICIS.2017.7960025.

[3]. Prof. S.U. Saoji, Nishtha Dua, Akash Kumar Choudhary, Bharat Phogat "Air Canvas application using OpenCV and NumPy in python"International Research Journal of Engineering and Technology, Volume: 08, Issue 08 Aug 2021

[4]. K. V. V. Reddy, T. Dhyanchand, G. V. Krishna and S. Maheshwaram, "Virtual Mouse Control Using Colored Finger Tips and Hand Gesture Recognition ", 2020, doi:10.1109/HYDCON48903.2020.9242677.

[5]. P. Ramasamy, G. Prabhu and R. Srinivasan, "An economical air writing system converting finger movements to text using web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), 2016, pp. 1-6, doi: 10.1109/ICRTIT.2016.7569563.

[6]. C. D. Sai Nikhil, C. U. Someswara Rao, E. Brumancia, K. Indira, T. Anandhi and P. Ajitha, "Finger Recognition and Gesture based Virtual Keyboard," 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020, pp. 1321-1324, doi: 10.1109/ICCES48766.2020.9137889.

[7]. S. Bano, P. Jithendra, G. L. Niharika and Y. Sikhi, "Speech to Text Translation enabling Multilingualism," 2020 IEEE International Conference for Innovation in Technology (INOCON), 2020, pp. 1-4, doi: 10.1109/INOCON50539.2020.9298280.

[8]. S. R. Chowdhury, S. Pathak and M. D. A. Praveena, "Gesture Recognition Based Virtual Mouse and Keyboard," 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), 2020, doi: 10.1109/ICOEI48184.2020.9143016.

[9]. K. H. Shibly, S. Kumar Dey, M. A. Islam and S. Iftekhar Showrav, "Design and Development of Hand Gesture Based Virtual Mouse," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1-5, doi: 10.1109/ICASERT.2019.8934612.

[10]. M. S. Alam, K. -C. Kwon and N. Kim, "Trajectory-Based Air-Writing Character Recognition Using Convolutional Neural Network," 2019 4th International Conference on Control, Robotics and Cybernetics (CRC), 2019, pp. 86- 90, doi: 10.1109/CRC.2019.00026.

[11]. S.S. Abhilash, L. Thomas, N. Wilson and C. Chaithanya, "Virtual Mouse Using Hand Gesture", International Research Journal of Engineering and Technology (IRJET), vol. 5, no. 4, pp. 3903-3906, 2018

[12]. S. Song, D. Yan and Y. Xie, "Design of control system based on hand gesture recognition," 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 2018, pp. 1-4, doi: 10.1109/ICNSC.2018.8361351.