Human Hand Tracking System(“H2TS=Hatuus”) Development Using Python

Ananya Sharma (20BCS3049), Magan Jyot Kaur(20BCS3041), Lokesh Choudhury(20BCS3033)
Department of CSE, Chandigarh University
Gharuan(140413), Mohali, Punjab, India
20BCS3049@cuchd.in
20BCS3041@cuchd.in
20BCS3033@cuchd.in

Abstract— **Hand tracking system allows user to interact without any physical input to the device by detecting the position and orientation of hands and the configuration of fingers. In this way user’s hand act like an external device which inputs the data to execute a specific algorithm. In this work, we present a real-time method for performing different features using hand tracking incorporated into one program. H2TS system is based on a futuristic technological approach which will make human and software interaction much more virtual and convenient without using hardware. The advanced feature which can be incorporated are voice command, touchless operations {which can be done through web cam using hand landmarks’ detection (total hand landmarks detection=21)} which will help- in cost saving (Keyboard), less maintenance cost of hardware, enough distance from hardware equipment ultimately better eye care and many more technological advantages. In previous systems, data gloves or markers were used to input data into the system. We have no such limitations when it comes to using the system. So, here SAHVI (Software and Human Virtual Interaction) a virtual assistant will help user for features' usages to do each task one by one through voice command like virtual volume controller, virtual mouse controller, virtual game controller, virtual calculator and virtual painter. It is not constrained by the number of tasks we can add more tasks to it.**

Keywords— **Gesture Recognition, Virtual Assistant, Python, Human-computer interaction, Machine learning.**

I. INTRODUCTION

Human computer interaction (HCI) is an interdisciplinary field of study which focuses on computer technology and majorly on communication between two processors that is computer and human being where humans are the users who feed the required data into computer for it to function. While in the past HCI was majorly concerned with the interaction between humans and computers, HCI has since branched out to cover a large number of information technology design. For this project we are focusing on a specific branch of human-computer interaction I.e., gesture recognition or human hand tracking system. Gesture recognition is a technology which helps the computer to interpret data(image) uploaded or the movement of hands (dynamic hand movements taken using webcam) to produce an output. It is considered as one of the most viable and desired solution for improving human computer interaction. Using gesture recognition, a computer reads human body’s motions via webcam or any external camera connected to the device. To perform the actual hand detection and tracking on our input image we made use of media pipe and to perform operations which are associated with computer vision we made use of OpenCV.

SECTION - A

Hand Landmarks Identification: After the palm detection process is completed over the whole input image hand landmark performs precise key point localization. It identifies 21 3-dimensional hand-knuckle coordinated inside the detected hand regions via lapsing, that so direct coordinates can be predicted and then considered to produce desired output.

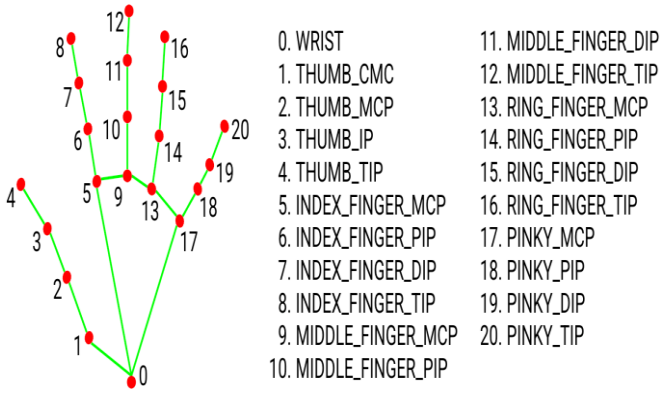


Fig 1: Hand Landmarks Identification(Count)

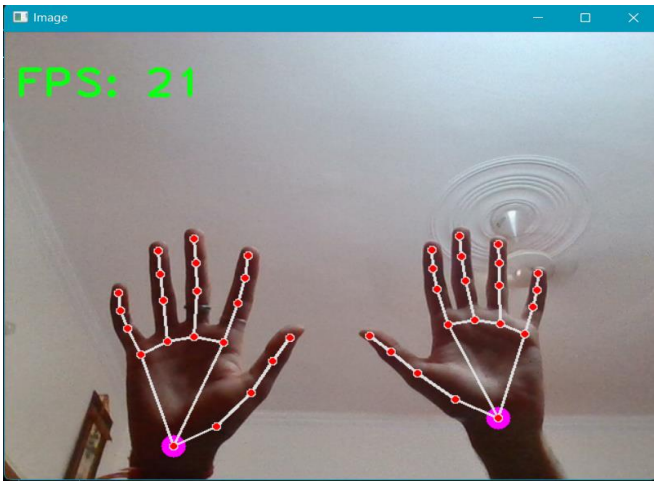


Fig 2: Hand Landmarks Identification (In real time)

Palm Detection: Media pipe works on the complete image (image which was uploaded dynamically using webcam) and returns a positioned or oriented hand bounding box. It precisely crops hand image to the hand landmark model significantly decreasing the amount of data argumentation required (rotation, translation and scale) allowing the network to devote most of its capacity towards coordinating prediction accuracy.

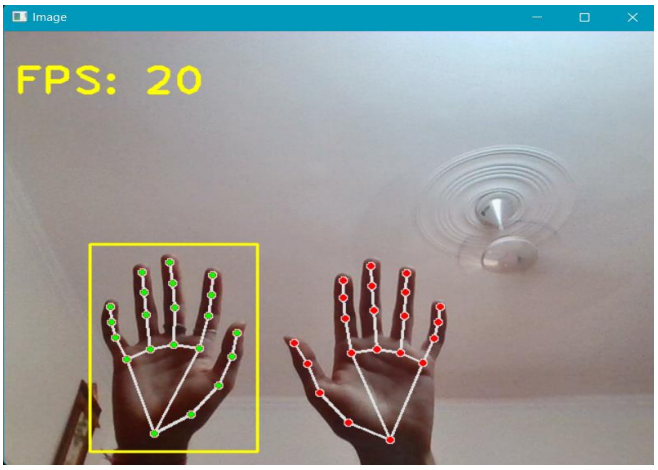


Fig 3: Palm Detection (In real time)

SECTION - B

Research: H2TS project started from background study about this technology where we started to gather the knowledge about the developments already done in this field by other software developers. We got the really good guidance from internet to an extent to move ahead with this project in python.

Web Cam operationalization: Web cam operations executed through coding in python.

Hand Tracking Module Creation: In this stage the main task was to track the hand landmarks which are 20 specifically in numbers by creating a hand tracking file.

Now the stage comes where the previous file ("HTMF = Hand Tracking Module File") would be used to create for further specific features of H2TS system which are as follows: -

1) *Virtual Volume Controller (VVC):* With the use of previous file (HTMF) a virtual volume controller feature would be created through coding in python.

2) *Virtual Mouse Controller (VMC):* With the use of previous file (HTMF) a virtual mouse controller feature would be created through coding in python.

3) *Virtual Game Controller (VGC):* With the use of previous file (HTMF) a virtual handle drive controller feature would be created through coding in python.

4) *Virtual Calculator(VC):* With the use of previous file (HTMF) a virtual calculator feature would be created through coding in python.

5) *Virtual Painter(VP):* With the use of previous file (HTMF) a virtual painter feature would be created through coding in python.

6) *Virtual Assistant ("SAHVI = Software and Human Virtual Interaction"):* SAHVI is a virtual assistant for this program which will assist the user for features' usages through voice command.

II. FEATURES AND CHARACTERISTICS

The features and characteristics of a project are defined as that the project must be specific they must be measurable, achievable, relevant, efficient

and the objectives must be achieved. The proposed project provides certain features and characteristics which makes this project distinct then other projects. This project allows users to interact with the system without using any physical input device like keyboard, pen or mouse. The proposed project detects the position and orientation of hands and the configuration of fingers to perform certain tasks which are available in in the project. In this project the user's hand act as an external device which inputs the data by performing a certain movement of user's hands and then the webcam present in the system captures it frame by frame on real time and performs image processing algorithm. The H2TS system is based on a futuristic technological approach which will make human and software interaction much more virtual and convenient without using hardware. The advanced features which are incorporated in this project are voice command which can be used to open different tasks present in the project and touchless operations which will help in reducing hardware cost thus saving less maintenance cost of hardware. So the proposed system uses SAHVI (Software and Human Virtual Interaction) which is a virtual assistant that will help user for features usages to perform each task one by one through voice command.
Currently it will assist five different tasks: -

1) *Virtual Volume Controller (VVC):* Under this feature we develop an interface which will capture human hand gesture dynamically and will control volume level by thumb and index figure positioning. It detects our hand with 21-point landmarks so as it can see the distance between our thumb fingertip and index fingertip. We can manipulate devices volume by increasing and decreasing distance between out thumb fingertip and index fingertip. When we increase the distance between thumb fingertip and index fingertip the volume will increase and when we decrease the distance between thumb fingertip and index fingertip the volume will decrease.

2) *Virtual Mouse Controller (VMC):* Under this feature we develop an interface or a mouse simulation system which will capture human hand gesture dynamically and will perform all the functions performed by mouse corresponding to hand movements and gestures. It retrieves necessary data and executes in absolute notation that is acceptable by predefined notations which are stored in the computer. The fingertip of the finger which we're using as a replacement of cursor is detected by the program and then all the work is done by that suspected finger. Selection can be done by clicking index fingertip and middle fingertip.

3) *Virtual game Controller (car) (VGC):* Video gaming is a form of entertainment which is at its peak. A major part of population uses gaming as a way to release their stress or a way to escape their problems. For higher involvement in the game, we use human computer interaction to attract audience. Human computer interaction enhances their experience. The camera detects landmarks points of index fingertip, finger joint point and wrist. The angle between these landmarks is calculated and based on the angle between these landmarks, the game can be played by the user.

4) *Virtual Calculator:* Under this feature we develop a system for virtual calculator that uses finger movements or gestures to operate. The camera detects the landmark points of index fingertip and middle fingertip (fingers which are used to operate virtual calculator). Calculator is visible on the top right side of the screen where the user can move around their fingers any perform any calculation they want. By clicking middle fingertip and index fingertip user can select an operand or operator to perform the desired operation.

5) *Virtual Painter (VP):* We develop virtual paint application which empower user to virtually paint in the air by using their fingers which are detected by the webcam. It can be used to draw on our system screen based on the position of our index finger movements (landmarks are detected beforehand). For drawing mode or active mode, user needs to show their index finger. To select different color or the eraser the user must select it by clicking their index fingertip and middle fingertip together at top of the icon of the function user wants to use. To erase a drawing user can just slide the eraser on top of the drawn picture after selecting the eraser function.

6) *Virtual Assistance (SAHVI):* SAHVI stands for software and human virtual interaction. It is a virtual assistance for this program which will assist the user select features through voice command. It gives a preface about the features and helps the sure to transvers through the project. User can ask their assistance to direct them to a specific feature.

III. DISCUSSION

Human-computer interaction (HCI) research focuses on the interfaces between users and systems in the design and use of computer technology. It can alternatively be thought as as two processors attempting to interact with each other, with one unit (computer) relying on data analysis provided by the other unit (people). There has already been a lot of study done in this sector, and as technology advances, fresh research and ideas are being developed to assist this field thrive. Human-computer interaction (HCI) is the study of how people interact with computers and how computers might be made to be more user-friendly, practical, and intuitive.

In the subject of HCI, the Human Hand Tracking System has been a significant and active area of research. The Human Hand Tracking System may convert observed hands or gestures into various functional inputs and interfaces with other programmes using a variety of techniques. It allows users to communicate with computers in a more interactive way. The Human Hand Tracking System feeds real-time data to the computer, allowing it to carry out the user's requests. In terms of gesture control, a touchless user interface is a new form of technology. The method of commanding a computer using body motion and gestures rather to a keyboard, mouse, or screen is known as touchless user interface (TUI). Machine learning algorithms that recognise hand and finger postures are learned using depth pictures of hands.

The human hand tracking system is mostly composed of three levels:

• *Detection:* A gadget detects hand or finger motions with the use of a camera, and a machine learning algorithm segments the real-time images to find hand edges and positions. This method uses machine learning (ML) to infer 21 3D critical points of a hand from a single picture, resulting in high-fidelity hand and finger tracking. There may be two steps in this section:

- *Palm Detection:* A palm detection model is implemented or used to detect initial hand placements. The Blaze Palm detector model works on the entire image and delivers an orientated hand bounding box.

- *Hand Landmarks Identification:* Following palm detection throughout the entire image, the hand landmark model uses regression to conduct exact key point localization of 21 3D hand-knuckle coordinates within the discovered hand areas. Even with partially visible hands and self-occlusions, the model develops a consistent internal hand posture representation.

• *Tracking:* A device captures every movement and provides reliable input for data processing by monitoring motions frame by frame. Media Pipe Box Tracking can be combined with machine

learning to create useful and efficient pipelines. For example, to develop an item recognition and tracking pipeline, box tracking can be combined with ML-based object detection. This pipeline has various advantages to executing detection each frame when tracking:

- It supports instance-based tracking, which means that the object ID is preserved across frames.
- Detection does not have to be performed at every frame. This allows for more accurate detection models to be run while keeping the pipeline light and real-time on mobile devices.
- With the use of tracking, object localisation is more temporally consistent, resulting in reduced jitter across frames.

• *Pattern Recognition:* The system looks for patterns in the data. The system conducts the action connected with the gesture when it detects a match and understands it. On the cropped image of the hand, there are 21 hand landmarks. Hand gestures can be understood by evaluating the placement of landmarks, and then a specific job connected with that hand motion can be done.

With the advancement of technology, HCI has become a popular field to research on. As of now many researches have been done on the Human hand tracking related system. After reading thoroughly into the HCI field we came to find out that there is no research available of an integrated project. To make the experience more pleasant we have come up with a solution named SAHVI (Software and Human Virtual Interaction) a virtual assistant. SAHVI will help user for features' usages to do each task one by one through voice command. Currently it will assist five different tasks: -

- Virtual Volume Controller
- Virtual Mouse Controller
- Virtual Game Controller
- Virtual Calculator
- Virtual Painter

It is not constrained by the number of tasks we can add more tasks to it. As hand tracking system is a good option to eliminate interaction with hardware and with the use of virtual assistant the interaction with the software will be very smooth.

Activity diagrams are graphical representations of workflows that include step-by-step activities and actions, as well as support chevalier, iteration, and concurrency.
Depicting flow of control using activity diagram:

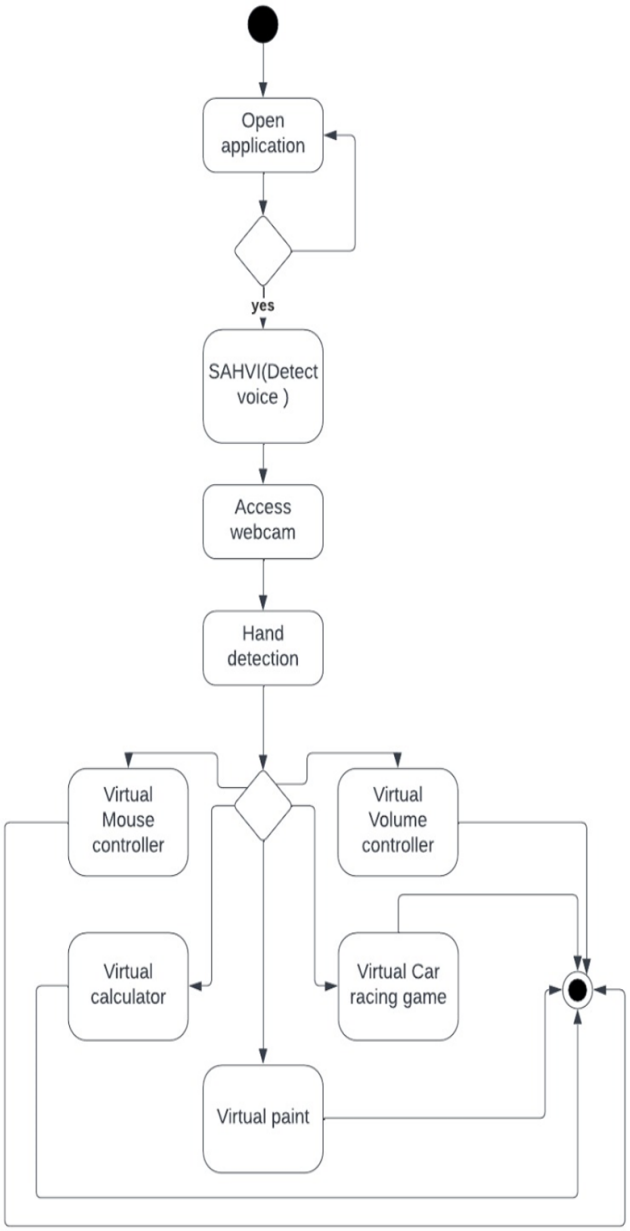


Fig 4: Activity Diagram

IV. Conclusions

The paper proposes a virtual assistant based human hand tracking system that can recognize hand movements and gestures throughout the different tasks of the software. The proposed system uses a webcam and hand detection algorithm to perform different tasks which are used in everyday life and makes it easy for young generation to adapt this technology. The system makes use of the OpenCV and Media Pipe which makes it very efficient in executing different tasks of the project as the features of this libraries are fast and easily calculable. Our system is focused more on making it user friendly using voice based command and also making it faster as compared to other hand detection recognition rate. The system is also focused more on reducing cost and improve robustness of the system as compared to other proposed system. Further improvements will be focused more on including more number of tasks and making it more efficient and lag free system.

V. Result Analysis

Our technology is a real-time gesture categorization system that can automatically recognize gestures in natural lighting. So, here's a virtual computer environment that can perform a variety of touchless tasks utilizing a webcam and hand landmark detection, all with the help of a virtual voice assistant dubbed SAHVI (Software and human virtual interaction). The interface is also efficient, user friendly and customizable that it can include more and more number of tasks in the future. Overall quality of the project is optimized then other proposed projects and the cost of the system has been reduced.

Acknowledgment

We would like to express our special thanks of gratitude to all the people who have played a crucial role in the research for this project, without their active co-operation the preparation of this project could not have been completed within the specified time limit.

I am very thankful to our respected Co-Supervisor, Ms. Neha Sharma for motivating us to complete this project with complete focus and attention.

And special thanks to our project guide lecturer, Mr. Gyan Chand Yadav who supported us as a Supervisor throughout this project with utmost co-operation and patience and for helping us in doing this project completed successfully.

Finally, we would like to thanks "CHANDIGARH UNIVERSITY" for providing us an opportunity to showcase our talent through this project.

References

[1] J.Jenkinwinston (96207106036), M.Maria Gnanam (96207106056), (96207106306), Anna University of Technology, Tirunelveli: Recognitionsystem Using Haar Wavelet.

[2] Kay M. Stanney HANDBOOK OF VIRTUAL ENVIRONMENTS Design, Implementation, and Applications, Gesture Recognition Chapter #10 by Matthew Turk.

[3] Daniel Thalman, Gesture Recognition Motion Capture, Motion Retargeting, and Action Recognition.

[4] J. Heinzm ann and A. Zelinsky Robust Real - Time Face Tracking and Gesture Recognition.

[5] Vladimir Vezhnevets, Vassili Sazonov, Alla Andreeva, Moscow State University A Survey on Pixel-Based Skin Color Detection Techniques.

[6] Ari Y. Benbasat and Joseph A. Paradiso in MIT Media Laboratory, Cambridge An Inertial Measurement Framework for Gesture Recognition and Applications.

[7] Ray Lockton, Balliol College, Oxford University, Hand Gesture Recognition Using Computer Vision.

Plagiarism Report

Scan Properties

Number of Words : 995
Results Found : 0

To or From
Binary Translator

To or From
PDF Converter

0% Plagiarism
100% Unique

Start New Search

To check plagiarism in photos click here

Reverse Image Search

Scan Properties

Number of Words : 873
Results Found : 0

To or From
Binary Translator

To or From
PDF Converter

0% Plagiarism
100% Unique

Start New Search

To check plagiarism in photos click here

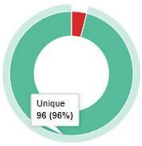
Reverse Image Search

Scan Properties

Number of Words : 957
Results Found : 2

To or From
Binary Translator

To or From
PDF Converter



4% Plagiarism
96% Unique

Make it Unique
Start New Search

To check plagiarism in photos click here

Reverse Image Search