Project Phase II Report

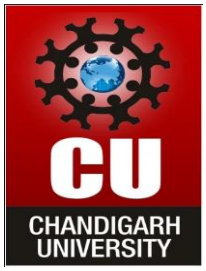
On

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

Submitted for the requirement of Project course

BACHELOR OF ENGINEERING

COMPUTER SCIENCE & ENGINEERING

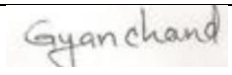
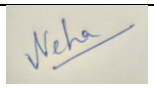


Submitted to: Mr. Gyan Chand Yadav
Project Teacher (Supervisor)
(Name – Mr. Gyan Chand Yadav & ID - E12247)

Co - Supervisor
(Name – Ms. Neha Sharma & ID - E12270)

Submitted by: Ananya Sharma
Student Group - 3
Name - Ananya Sharma
Magan Jyot Kaur
Lokesh Choudhury

UID - 20BCS3049
20BCS3041
20BCS3033

Supervisor Signature - 	**Co - Supervisor Signature -** 

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CHANDIGARH UNIVERSITY, GHARUAN

June 2022

Ananya Sharma – 20BCS3049
Magan Jyot Kaur – 20BCS3041
Lokesh Choudhury – 20BCS3033

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.

Introduction

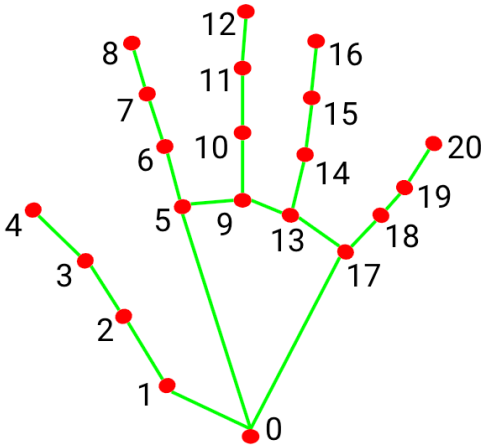
Human-computer interaction (HCI) is a multidisciplinary field of study focusing on the design of computer technology and, in particular, the interaction between humans (the users) and computers. While initially concerned with computers, HCI has since expanded to cover almost all forms of information technology design. For this project we are focusing on a specific branch of Human-computer interaction I.e., gesture recognition or human hand tracking. Gesture recognition helps computers to understand body language. It is considered as one of the most viable and popular solution for improving human computer interaction. Using gesture recognition, the human body's motion is read by the computer camera. The ability to perceive the shape and motion of hands can be a vital component in improving the user experience across a variety of technological domains and platforms.

We made use of OpenCV to perform operations associated with computer vision and Media Pipe to perform the actual hand detection and tracking on our input image.

Ananya Sharma – 20BCS3049
Magan Jyot Kaur – 20BCS3041
Lokesh Choudhury – 20BCS3033

Palm detection: Media Pipe works on the complete input image and returns an oriented hand bounding box. It provides a precise cropped hand image to the hand landmark model significantly reducing the need for data argumentation (rotation, translation and scale) allowing the network to dedicate most of its capacity towards coordinating prediction accuracy.

Hand landmarks identification: After the palm detection over the whole image hand landmark model performs precise key point localization of 21 3D hand-knuckle coordinates inside the detected hand regions via regression, that so direct coordinate prediction.



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

1)Virtual Volume Controller [VVC]

Under this feature we develop an interface which will capture human hand gesture dynamically and will control the volume level. The camera in our device is used for this. It detects our hand with points in it (21 landmarks) so as it can see the distance between our thumb fingertip and index fingertip. The distance between the tips of these two fingers is directly proportional to the volume of device, in other words, via increasing and decreasing distance between thumb fingertip and index fingertip we can increase and decrease device's volume respectively.

Ananya Sharma – 20BCS3049
Magan Jyot Kaur – 20BCS3041
Lokesh Choudhury – 20BCS3033

2)Virtual Mouse Controller [VMC]

Under this feature we develop a mouse simulation system which performs all the functions performed by mouse corresponding to hand movements and gestures. It retrieves necessary data and implements it to the mouse interface of the computer according to predefined notions. The finger tip of finger which is suspected to work as replacement of cursor is detected. Desired item can be selected by clicking index fingertip and middle fingertip.

3)Virtual Game Controller(car) [VGC]

Video gaming is a form of entertainment which is advancing at a rapid rate leading industry to pick up its pace. For higher involvement in the game, we use human computer interaction as to increase the audience immersion. Landmarks of index fingertip, index finger joint point and wrist are detected. The angle between these landmarks is calculated and based on the angle the game can be played by the user.

4)Virtual Calculator [VC]

Under this feature we develop an interface for virtual calculator that uses finger movements to operate. The camera detects the landmarks of index fingertip and middle fingertip. Calculator is visible on the right side of screen where the user can move around their fingers and click their index fingertip and middle fingertip to select an operand or operator to perform the desired operation.

5)Virtual Painter [VP]

Virtual Painter is an application that enables one to virtually paint in the air using their fingers. It can be used to draw on your system screen based on your index finger movement. As soon as the presence of hand is detected, application draws a bounding box around the hand. If user shows

Ananya Sharma – 20BCS3049
Magan Jyot Kaur – 20BCS3041
Lokesh Choudhury – 20BCS3033

only index finger than the user is in drawing mode. To select different color or eraser the user must select it by clicking his index and middle finger together at the top of icon. If the user wants to erase the drawn painting, then user needs to select the eraser and just slide the eraser on top of the drawn image.

6)Virtual Assistant (SAHVI)

SAHVI (Software and Human Virtual Interaction) is a virtual assistance for this program which will assist the user for features usages through voice command. SAHVI gives a preface about the features and helps us to traverse through the project. User can ask their assistant to direct them to a specific feature.

Literature review

Human-computer interaction is research in the design and the use of computer technology, which focuses on the interfaces between user and systems. It can also be viewed as two processors trying to communicate with each other where one of the units (computer) depends on the data analysis which were input by the other unit I.e., humans. Plenty of research has already been done on this field and as the technology is advancing new researches and ideas are helping this field to flourish. Human-Computer interaction studies how people interact with computing technology and how a computer is designed more easily, more practically and more intuitively.

Human Hand Tracking System has been an important and active area of research in the field of HCI. Human Hand Tracking System is able to translate the detected hands or gestures into different functional inputs and interfaces with other applications via several methods. It enables the user interactive Human Computer Interaction. Human Hand Tracking System provides real time data to computer to make it fulfill the user's commands. Touchless user interface is an emerging type of technology in relation to gesture control. Touchless user interface (TUI) is the process of

Ananya Sharma – 20BCS3049
Magan Jyot Kaur – 20BCS3041
Lokesh Choudhury – 20BCS3033

commanding the computer via body motion and gestures without touching a keyboard, mouse, or screen. Machine learning algorithms trained based on the depth images of hands allowing them to recognize hand and finger positions. Working of Human hand tracking system is mainly based on three levels:

- Detection: With the help of a camera a device detects hand or finger movements and a machine learning algorithm segments the real time images to find hand edges and positions. This approach provides high-fidelity hand and finger tracking by employing machine learning (ML) to infer 21 3D key points of a hand from just a single frame. Under this part there can be two stages:
    - Palm Detection: To detect initial hand locations, palm detection model is implemented or used. Palm detector model or Blaze Palm operates on the full image and returns an oriented hand bounding box.
    - Hand Landmarks Identification: After the palm detection over the whole image subsequent hand landmark model performs precise key point localization of 21 3D hand-knuckle coordinates inside the detected hand regions via regression, that is direct coordinate prediction. The model learns a consistent internal hand pose representation and is robust even to partially visible hands and self-occlusions.

- Tracking: A device monitors movements frame by frame to capture every movement and provide accurate input for data analysis. Media Pipe Box Tracking can be paired with ML inference, resulting in valuable and efficient pipelines. For instance, box tracking can be paired with ML-based object detection to create an object detection and tracking pipeline. With tracking this pipeline offers several advantages over running detection per frame:
    - It provides instance-based tracking, i.e., the object ID is maintained across frames.
    - Detection does not have to run every frame. This enables running heavier detection models that are more accurate while keeping the pipeline lightweight and real-time on mobile devices.
    - Object localization is temporally consistent with the help of tracking, meaning less jitter is observable across frames.

Ananya Sharma – 20BCS3049
Magan Jyot Kaur – 20BCS3041
Lokesh Choudhury – 20BCS3033

- Recognition: The system tries to find patterns based on the gathered data. When the system finds a match and interprets a gesture it performs the action associated with this gesture. There are 21 hand landmarks on the cropped image of the hand. By analyzing the location of landmarks, hand gesture can be interpreted and then a specific task associated with that hand gesture is performed.

So, after doing all the research about it, we came to know that there are many systems of hand tracking related to human computer interaction, but because all of them are not available at one place, 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 like -

- 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.

Problem definition

Biometric technologies make use of various physical and behavioral characteristics of human such as fingerprints, expression, face, hand gestures and movement. Among the biometric sector hand gesture recognition are gaining more and more attention because of their demand regarding security for law enforcement agency as well as in private sectors such as surveillance systems. Hand gestures are important to intelligent human and computer interaction to build fully automated

Ananya Sharma – 20BCS3049
Magan Jyot Kaur – 20BCS3041
Lokesh Choudhury – 20BCS3033

systems that analyze information contained in images, fast and efficient hand gesture recognition algorithms are required.

Objectives

- First objective of this project is to allow the communication between human and computer by the use of gestures and hand movements to be more intuitive.
- Second objective of this project is to create a complete system to detect, recognize and interpret the hand gestures through computer vision.
- Third objective of this project is to do more and more work on the computer with the help of hand gestures such as to control volume of the system, mouse control, game control and many more.

References

Research paper - https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3702844

AnalyticsVidhya - https://www.analyticsvidhya.com/

Geeksforgeeks - https://www.geeksforgeeks.org/

GitHub - https://github.com/