Gesticulation of PowerPoint Presentation

Sandhiya M   
Department of Information Technology Madras Institute of TechnologyChennai,Tamil Nadu  
sandhiyam1302@gmail.comSivaranjani S   
Department of Information Technology Madras Institute of TechnologyChennai,Tamil Nadu  
s.sivaranjani0213@gmail.comPoomala M   
Department of Information Technology Madras Institute of TechnologyChennai,Tamil Nadu  
poomala2003@gmail.com

*Abstract*—Actions speak more than words. In context to the above statement the importance of gestures and using them to control a system have become popular. Many pre-trained models are available to recognize the gestures of human such as Media Pipe and so on. These models can be used to identify the gestures more efficiently and respective action can be assigned based on the identified gestures. Thus, the control system of PowerPoint presentations can be made easy and convenient using these gestures-based control. Thus, the proposed model improves the interaction between the human and the computer.

Keywords—gesture-based control, Media Pipe, Human Computer Interaction (HCI)

# Introduction

Gesture-based control has become increasingly important in day-to-day life, as it offers a more intuitive and convenient way to interact with technology. From smartphones to home appliances, gesture-based control allows users to perform tasks and navigate interfaces without physically touching the device or using traditional input methods such as a mouse or keyboard. One of the significant benefits of gesture-based control is its potential to enhance accessibility for individuals with disabilities. For instance, people with mobility impairments can use gesture-based control to operate devices without the need for complex or specialized assistive technology. Additionally, gesture-based control can improve the user experience for everyone by simplifying interactions and reducing the cognitive load required to use technology. gesture-based control has the potential to transform the way we interact with virtual and augmented reality environments. By using hand and body gestures, users can manipulate virtual objects and navigate digital spaces in a more natural and intuitive way.

# LITERATURE SURVEY

A.N.Paithane , Monika Marathe , Apurva Shelar , Pratiksha Ambhore [1] authors recommended the use of gesture devices that are integrated with sensors. Authors have used the sensor to record the movements of the hand in a few limited directions that could hold control over the slides. The communication between the gesture device and slides happens with RF transmission and reception. The division of the process is the gesture transmitter and computer receiver. The authors used a wearable glove controller, microcontroller, and RF trans receiver for gesture transmission. Accelerometer and Flex sensors were used for gesture recognition. They detect the flexing and movements of the hands and detect the movements by axis tracking the hands. Slide switches are traced by the accelerometer. There are many actions happening hence different threshold levels are set for every action. Sensors send analog signals as their output. The ADC of the microcontroller converts the received signals to digital signals. Microcontrollers use the signals generated and processes to send them to RF transmitter. The transmission of signals takes place from RF transmitter to RF receiver that gives it to the microcontroller. As the visual basic application runs by, the microcontroller gives data to it. The operations are controlled using visual basic application based on the data received.

Stephen M. Chang [2] the design and implementation of a speech and gesture recognition system used to control a PowerPoint presentation using the Microsoft Kinect. This system focuses on the identification of natural gestures that occur during a PowerPoint presentation, making the user experience as fluid as possible. The system uses an HMM to classify the performed gestures in conjunction with an SVM to perform real-time segmentation of gestures. The fusion of these two models allows the system to classify gestures in real time as they are being performed instead of waiting until completion. The incorporation of speech commands gives the user an additional level of precision and control over the system. The torso is used for identifying the general location of the body with respect to the Kinect. The two arms are analyzed separately allowing the gestures to be performed with either arm independently of the other. The primary goal of the system is to free the user from these restraints and automatically react to the naturally spoken words and gestures throughout the presentation.

Muhammad Idrees et.al [3] proposed a model to control different operations of the PowerPoint slideshow through gestures. This research has used Machine Learning to detect gestures with subtle differences and tried to map them with some fundamental PowerPoint slideshow controlling functions using Python. Vision-based gesture detection can also be improved using deep learning. The author has used deep learning in system to make it more useable, natural, and reliable. The only vision-based tactic is superb for static gestures and gesture detection by counting contours but not very efficient for dynamic gestures. The glove-based strategy and use of Kinect are expensive as we must buy sensors or Microsoft Kinect for these approaches. A massive dataset of mini clips of different gestures, distinct backgrounds, lighting conditions, and skin tones, is required to accomplish the objective. The dataset contained videos at a specific distance hence the detection can be achieved at that distance only. The detection was good with the use of GPU, but the detection accuracy was very low on the CPU.

Ms.Triveni Paradkar Prof. Anil Wanare [4] proposed a model with the methodology that uses a camera , projector, laptop and colour markers to make system that will project information to be displayed to the surface like wall. Matlab contains software routines of camera initialization, rgb2hsv conversion, morphological operations, and centroid locating, virtual mouse driver and projector interface. Image obtained by camera goes under series of operations mentioned in algorithm below it, virtual mouse driver will enable cursor getting controlled by Matlab.

Bhairavi Pustode et.al [5] proposed an intelligent presentation system employing hand gestures gives a simple method to update or control the slides. There are several pauses during presentations to operate the presentation using the keyboard. This system's purpose is to enable users to use hand gestures to control and explore the slideshow. The technique employed machine learning to identify various hand gestures for many tasks. A recognition technique offered an interface for human system communication. This model limited in the number of actions performed over the slideshow.

On surveying all these research papers, the existing models have a common limitation of the distance of the instructor who is presenting the PPT. But this limitation is negligible one because there should be technique to identify the instructor gesture which can be done using the distance they stand. The above discussed models also did not include more features to be handled using the gesture. They all have limited number of gesture-based features. All the above-mentioned models have gone through tough development process of using gesture dataset and training them to recognize the hand and then identify the gesture and performing the required action. Thus, the proposed model includes various other features of slide show to be handled using the gesture such as saving the slide with annotations, changing the color of the pointers, undoing the last annotations etc. The proposed system also uses light-weight models to recognize and identify the gestures performed in front of the camera using MediaPipe and its models such HandTrackingModel.

# PROPOSED SYSTEM

The proposed system uses the pre-trained models to identify the gestures performed in front of the camera and the respective action are assigned when the gesture is performed. Initially the PowerPoint is converted into the image folder where each slide of the PowerPoint (PPT) is converted into the image format as either JPG or JPEG or PNG. The MS PPT as the provision to directly convert a PPT into image folder which can be further used for our gesture-based control. Open-Source Computer Vision Library (OpenCV) is an open-source computer vision and machine learning software library designed to help developers create real-time computer vision applications. It is written in C++, but has interfaces for several programming languages including Python. OpenCV is a powerful tool for image and video processing, offering a wide range of functions for tasks such as image and video capture, image and video manipulation, feature detection and matching, object recognition, and machine learning. Next, the step is to open camera for recognising the gesture. OpenCV supports system to open camera and capture the frames. The captured frame is show as a small-sized window in the image frame of the PPT show in the widow using OpenCV.

Python hand tracking models are used to detect and track hand movements in real-time using machine learning techniques. One popular hand tracking model for Python is the Mediapipe Hand Tracking library developed by Google. This open-source library uses deep learning to detect and track the landmarks of the human hand, such as the fingertips, knuckles, and wrist, in real-time video streams. The Mediapipe Hand Tracking library provides an easy-to-use interface for developers to integrate hand tracking functionality into their Python applications. It is highly accurate and efficient, making it suitable for a variety of applications, including gesture recognition, sign language interpretation, and robotics. The library can be installed using pip and integrated with popular computer vision libraries like OpenCV for real-time hand tracking. Now, the HandTrackingModel is used to recognise both the hands in the frame generated by opening the camera. Once the hands are recognized using the HandTrackingModel the fingers that are lifted is noted. Based on the fingers lift up the actions are taken place.

A User Interface is also developed with the help of tkinter framework in python. The UI takes the path of the image folder and it class the procedure to handle the file with the gestures recognized from the camera generated frame by OpenCV. The list of action with their respective gesture is listed below Table 1.

**Table 1. List of gestures and their actions**

|  |  |  |  |
| --- | --- | --- | --- |
| **S. No** | **Gesture** | **Fingers Involved** | **Action** |
| 1. |  | Thumb pointing to left | Move to previous slide |
| 2. |  | Little finger up | Move to next slide |
| 3. |  | Index and middle finger up | Shows up the pointer on slide |
| 4. |  | Index finger up and movement | Draws in the slide |
| 5. |  | Index and little finger up | Changes color of the stokes and points. |
| 6. |  | All fingers up | Saves the current slide with the changes. |
| 7. |  | Index, middle, ring and little fingers up. | Undo operation. |

# FLOW OF THE MODEL

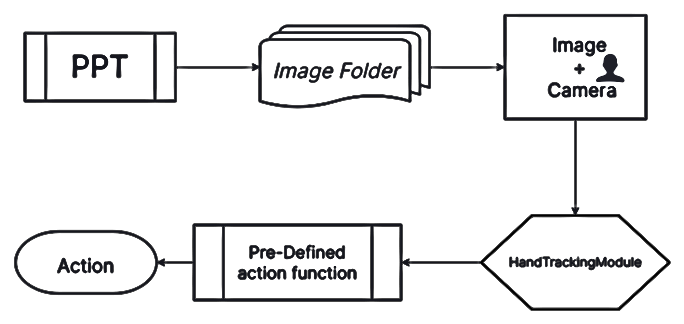


Figure 1. Flow of the proposed system.

The architecture of the proposed system deeply explains the flow of the proposed system. As discussed earlier the PPT is converted into the image folder in which each of the slide is converted into image format. This folder is being used to move the slides when the respective gesture is performed in front of the camera. The outcome of model is shown below figures.

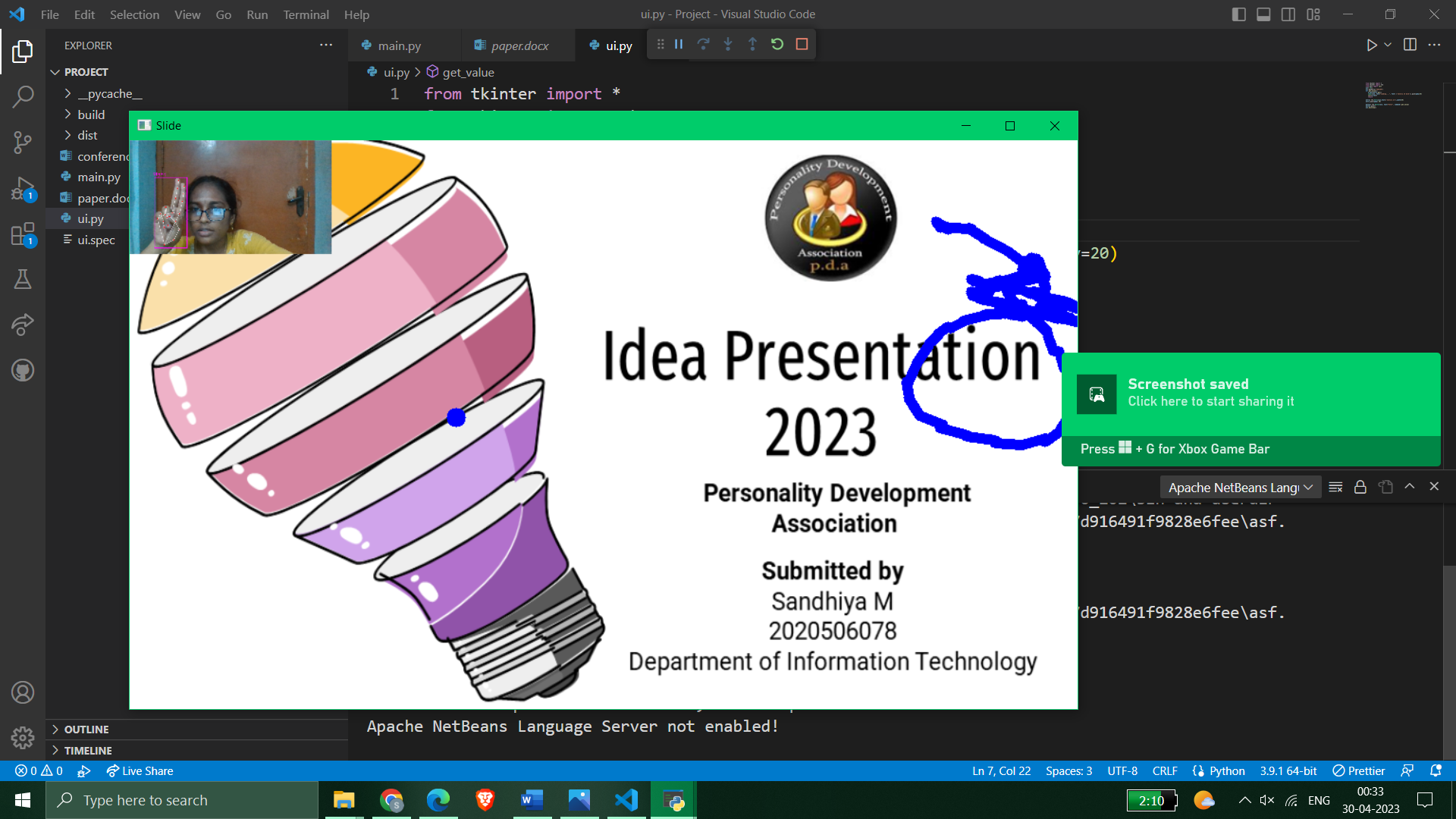


Figure 4.1

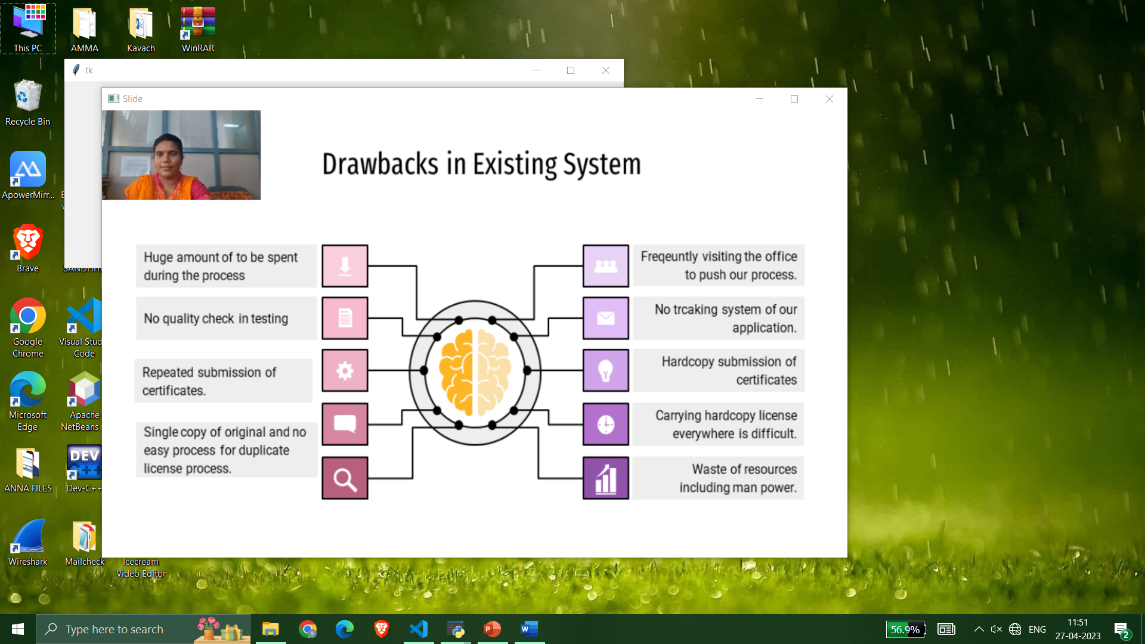


Figure 4.2

# ALGORITHM

The proposed algorithm basically follows the flow represented in the figure 1. The algorithm for gesture-based control of PowerPoint Presentation is below.

**Algorithm 1 Gesture-based control**

*procedure* Gesture (Source address)

Read the Images from Source

*images* ← image𝑠 [𝑆𝑜𝑢𝑟𝑐𝑒]

*camera* ← System [Camera]

*handTracker* ← HandTrackingModel

*frame*← frames[camera]

*images*← images + frame

**if**(handTracker(*frame*)) then

*temp\_image*←images

fingers ← handTracker(frame)

**if** (fingers [0]) then

*images* ← prev\_image

**else if** (fingers [4]) then

*images* ← next\_image

**else if** (fingers [1] and fingers [2]) then

*images*← images + pointer

**else if** (fingers [1]) then

*images* ← images + draw the location of fingers [1]

**else if** (fingers [0] and fingers [1] and fingers [4]) then *images* ← images + color\_change(pointer)

**else if** (fingers [0] and fingers [1] and fingers [2] and fingers [3] and fingers [4]) then

Source[*images*] ← images

**else if** (fingers [1] and fingers [2] and fingers [3] and fingers [4]) then

*images* ← temp\_image

show*(images)*

**endif**

**end**

**end** procedure

# ARCHITECURE

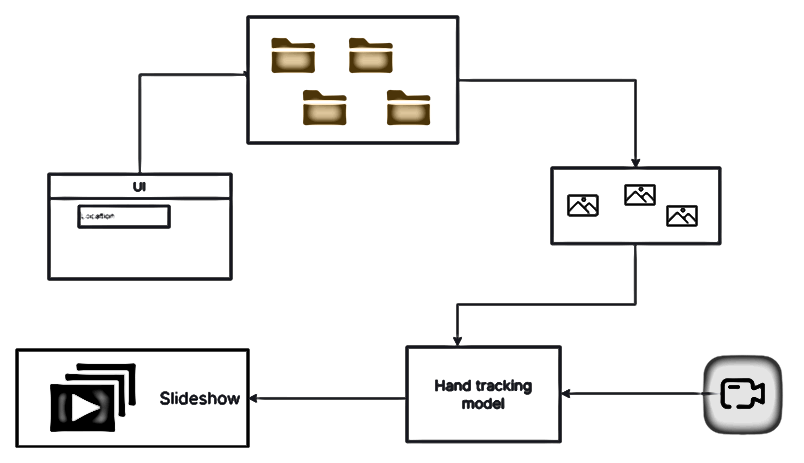


Figure 2. Architecture of the proposed model

# FUTURE WORK

The proposed model can be further developed to a customizable gesture-based application where the users can customize the gesture and map with the action. Instead of converting the slides into image an extension of this system can be added to the PPT to do the same operation on within the application without depending on a third-party application. This method can also be made available to all other devices. The proposed system works with single and there is issue in right-left orientation this can be improved in the future models.

# LIMITATIONS AND FUTURE SCOPE

The main issue found in the model is that the distance of the tutor from the camera. The HandTrackingModel does not work fine that is it could not detect the hands when the tutor is far away from the model. But still, this is not a big limitation since we need a fixed distance tracking of such that only the presenter could manipulate the PPT.

Gestures are widely used and important in many computers’ science-related domains. Gestures represent the future of Human-Machine Interactions because, as previously noted, there is a need for a more natural manner of interacting with computers and machines. This project aids a presenter in his presentations and assists students with Human-Computer Interaction (HCI) when gestures are used. Gestures can be used to control almost everything in a computer system, and when combined with speech recognition, we can create a system in which one does not need to learn a set of commands or a procedure to work with a machine, but it can all be done through natural communication.

# CONCLUSION

In a nutshell, the proposed system could able to recognize the gesture performed from the frames generated by OpenCV library and the respective actions are taken place on the images of the slides. Thus, the model helps in day-to-day life to the tutors and presenters in controlling the presentation from the place they stand with gestures. This system also improves the interaction between the Human and Computer which is termed as Human Computer Interaction (HCI). The goal of this project is to make human-machine interfaces simple and easy to use. The virtual connection between the laptop and the user without the use of hardware such as a mouse or keyboard makes the entire system more user friendly and simpler. The addition of sound response to the system makes this system useful to visually impaired people. Clearly, it has the potential to become the ultimate "transparent" user interface for accessing information.

# REFERENCES

1. A.N.Paithane , Monika Marathe , Apurva Shelar, Pratiksha Ambhore, “Power Point Presentation Control in Classroom Teaching by Hand Gesture” Vol.6 Issue No.5, International Journal Software Engineering and Computer Science.
2. Stephen M. Chang, “Using Gesture Recognition to Control PowerPoint Using the Microsoft Kinect”, Massachusetts Institute of Technology (2013).
3. Idrees, Muhammad & Ahmad, Ashfaq & Butt, Muhammad Arif & Danish, Hafiz. (2021). Controlling PowerPoint Using Hand Gestures in Python. Webology. 18. 1372-1388.
4. Triveni Paradkar and Prof. Anil Wanare,   
   “Advance Processing Unit With Virtual Human Machine Interface”, Issue 4 volume 4, July-August 2014, International Journal of Advanced Scientific and Technical Research.

<https://rspublication.com/ijst/2014/AUG14/33.pdf>

1. Bhairavi Pustode,Vedant Pawar, Varun Pawar, Tejas Pawar, Samiksha Pokale, “Smart Presentation System Using Hand Gestures”, Research Square.

1. Ahmed Kadem Hamed AlSaedi, Abbas H. Hassin Al Asadi, A New Hand Gestures Recognition System, Indonesian Journal of Electrical Engineering and Computer Science, Vol 18, (2020).
2. Xingfeng Wang, Kaihuai Qin , “A Six-degree-of-freedom Virtual Mouse based on Hand Gestures”, 2010 International Conference on Electrical and Control Engineering, pp. 257-260.
3. Apoorva Bharambe1, Divya Chanekar2, Divya Naik3, Prof. A. B. Vitekar4, “ Automatic Hand Gesture Based Remote Control for Home Appliances,” International Journal of Advanced Research in Computer Science and Software Engineering 5(2), February - 2015, pp. 567-571.
4. Al Saedi, A. K. H., & Al Asadi, A. H. H. (2020). A new hand gestures recognition system. Indonesian Journal of Electrical Engineering and Computer Science, 18(1),49–55. <https://doi.org/10.11591/IJEECS.V18.I1.PP49-55>.
5. D. Jadhav, Prof. L.M.R.J. Lobo, Hand Gesture Recognition System to Control Slide Show Navigation IJAIEM, Vol. 3, No. 4 (2014).