Smart Presentation using Hand gesture Recognition

P. Sathvika 201IT141
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025
Email: psathvika.201it141@nitk.edu.in

H. Jayachandra 201IT123
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025
Email:hanumanthujayachandra.201it123@nitk.edu.in

Mayur Jinde 201IT135
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025
Email:jindemayur.201it135@nitk.edu.in

***Abstract*——In this project we are developing hand gesture recognition in real-time, where we can change the slide without touching mouse or keyboard. This system makes possible the control of Power Point presentation through distance. The user does not have to be concerned about operating the power point and changing slides or controlling the presentation with a mouse, keyboard, or laser pointer, or any physical device. The proposed system takes input data from the portable web camera consisting the gestures to process the input and map the gesture to its corresponding functionality which is defined in the model so that the system can perform the appropriate function to control the power point presentation.**

I. Introduction

In today's environment, gesture recognition is a critical component of human-computer interaction and plays a major role in interaction with computers. Today as the interaction scenario with computers is changing and evolving people wish not want to use mechanical things to interact with their system so they are shifting to the gesture-based interaction where one can give the command to the system with a single wave of hand. In case of power point presentation also people don't want to use mechanical device to control it as the hand of the presenter will be occupied during presentation, which exhibit the use of both hand in the presentation will hamper the impact of the presentation so to address this issue, we suggest a methodology for controlling the presentation via hand gestures at a distance, which allows the presenter to use both his hands and complete movement which will increase the impact of the presentation and also give good experience of interaction with the system. To design such a model one of the methods to detect hand gesture is by calculating the centroid of the object region with the help of distance transformed method and then determine the finger count [2]. Another method is with hog feature extraction, in which an image is segmented, categorized, and made fit for comparison with a gesture dataset [1] and also can be done using circular profiling for determining the finger count [4].

In our proposed model we are using MediaPipe to detect the hand landmark which is developed by Google® on the basis of detected hand landmarks we are defining the gesture to control the power point. After selecting the gesture all the gestures were mapped with corresponding functionality to control the power point. In the proposed model the drawing feature is also integrated inside the power point only which can be used to avoid the digital pad for writing on the power point to demonstrate anything this will also increase the interaction time of the user with the system. The proposed model will give better results on high frame per second cameras. In the 30 frame per second or less webcam the latency will be more which can reduce the user experience.

II. Literature Survey

This segment gives a quick overview of the various tutorials and projects, we referred to the existing Hand Gesture Models to operate the Power Point Presentation. We referred to a base paper where they used the images as input, from which they trained a model which will detect the Image gesture. Firstly, they used a skin detection algorithm to see the hand and then resized the image obtained to a fixed size after the skin

detection algorithm. Then they eliminate the background from the image, allowing them to precisely identify the hand in the image. Then from the image, they extracted the features. The image was then classified into one of four categories using a clustering method. They then applied the appropriate action to the image after categorizing it. They got 80% accuracy in classifying images correctly. The other IEEE papers we referred to used different classification algorithms instead of the K-NN algorithm. So almost all existing systems used classification algorithms and got 75 - 80 % accuracy.

III. Problem Statement

Developing a system that controls Power-Point Presentation operations using different Hand Gesture through distance.
How we can change the slide without touching mouse and keyboard and any other physical device. So, it would become simpler and easier to use.

A. Objectives

- We want to create a gesture recognition system that can control a PowerPoint presentation using natural gestures.
- To enhance the experience of presentation for presenter as well as listener
- To reduce the user's physical strain of using mouse and other physical device
- To make the presentation more impactful and seamless

IV. Methodology

In the proposed model of PowerPoint presentation using hand gesture recognition, we divided the methodology in three separate modules in order to thoroughly understand the requirements analysis and gather all the requirements for the development of a complete model.

The different modules are:

- Gesture Selection
- System Design
- GUI Development

*A. **Gesture Selection***

As our project is about control presentation using gestures, it is most important that we have selected gestures on our model and mapped with the associated function to trigger the appropriate action on controlling the ppt.
So, we are using media pipe framework which is developed by Google®which offers several functionalities such as the perception system in Nest Cam, real time object detection using webcam, augmented reality ads, cloud vision API etc. The Media-Pipe perception pipeline is called a Graph. Let's consider the example of the first solution, Hands. We feed a stream of images to the webcam as input which comes out with hand landmarks as shown in the Fig 1. [11].

The flowchart below fig 2 represents the MP (Abbr. Media-Pipe) hand solution graph [11].

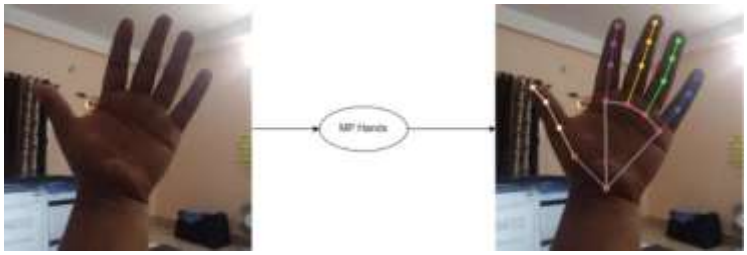


Fig. 1. Flow chart of how hand land marks will be displayed

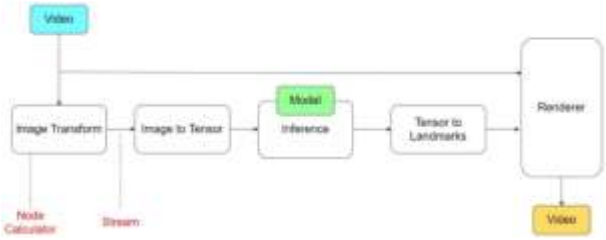


Fig. 2. Workflow of how Media-Pipe extracts Hand Landmarks

A graph is a collection of nodes connected by edges in computer science. The nodes of the Media-Pipe Graph are known as Calculators, while the edges are known as Streams. A sequence of packets with escalating time stamps is carried by each stream. So, this principal media-pipe works. The gesture which we chose is shown in Fig 3 to perform the task in PowerPoint presentation. The actions mapped with each gesture are previous slide, next slide, drawing, pointer, undo respectively as these are the basic operations which are necessary to operate the presentation.

*B. **System Design***

Fig 4 shows the required steps for executing the complete system. In this system the input is real time which captured using Computer's WEB-CAMERA of resolution 1280×720 . The captured image is then passed into the Media-Pipe module which give the landmark on the hands as shown in fig 1.Then using the landmark make the gesture and map gesture with corresponding functionality to control the PowerPoint presentation, in our case the chosen gesture is shown in fig 3. These are the gesture which we are using in our module to control the PowerPoint.

*C. **GUI Development***

To present the PowerPoint with the integrated system which we have built. We provide a graphical user interface in the form of a website. So, the website is built with the help of hypertext markup language and cascade style sheets and bootstrap. These frameworks are used to build the website's front end. The Flask API is used in the back-end which uses

Gesture	Action
	Previous Slide
	Next Slide
	Drawing
	Pointer
	Undo

Fig. 3. Action with respect to gesture

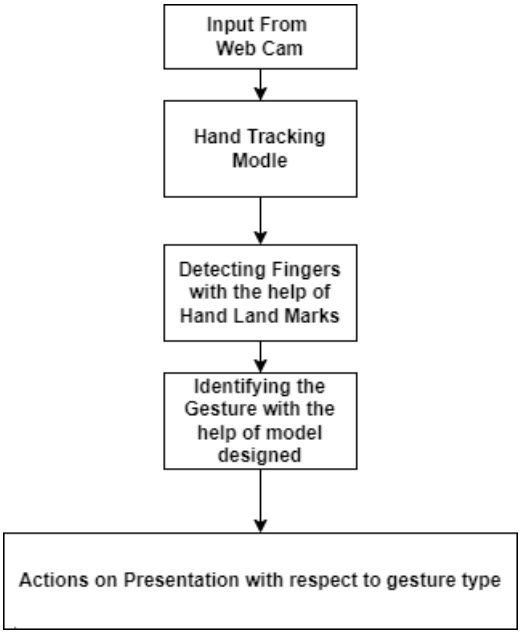


Fig. 4. Flow Chart of recognizing Gesture

python language, as our system is created on python so it is easily integra table with the system. In the web application as shown in the fig 5 the user needs to open the website there all the information regarding the use case such as which gesture is used for which operation and which gesture is used to terminate the GUI is given. In the same page a button is given which direct the user into next page where user need to upload the ppt and press the display button which will render the presentation with “power point using hand gesture recognition” feature enabled.

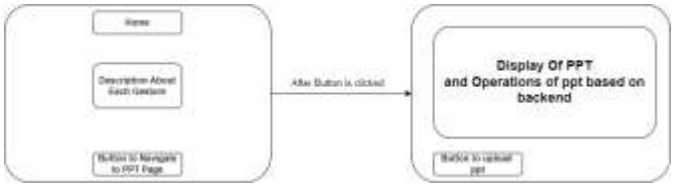


Fig. 5. Front End Design

V. Experiments Results and Analysis

For Experimental Analysis we tested our product with 30 people and calculated the response time of each gesture when the 30 people were using it. Also, we took feedback from them about the performance of the product.

A. *Likert Test*

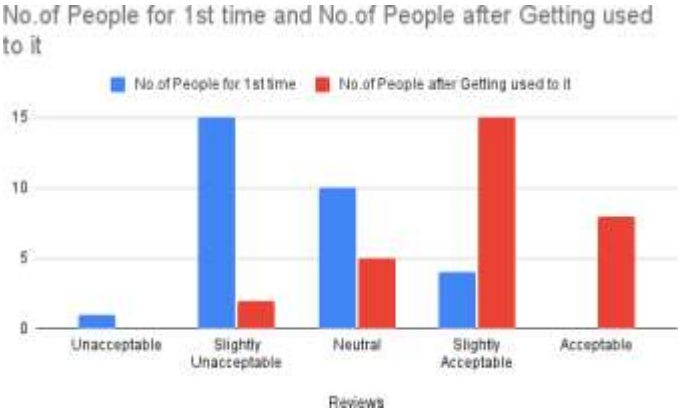
The Survey contains two parts. We took the feedback from the users after using it for the very first time and also took the feedback after they got used to it. We got different feedback during two times of survey. The feedback of 2 surveys is kept in the below table.

Review	No.of People selected
Unacceptable	1
Slightly Unacceptable	15
Neutral	10
Slightly Acceptable	4
Acceptable	0

Fig. 6. 1st Survey

Review	No.of People selected
Unacceptable	0
Slightly Unacceptable	2
Neutral	5
Slightly Acceptable	15
Acceptable	8

Fig. 7. 2nd Survey (After Getting used to Product)



From the Bar Graph we can see that Many people who are using the product for the very first time found it difficult to use with Hand Gestures. But after they got used to the product those who said the product was not satisfied, they said it was satisfied.

So, from the survey we can conclude that as for people, using hand gestures is a new thing (Which they didn't use before). So, it takes time to get used to it.

B. *Product Performance Analysis based on Response Time*

Response time analysis means calculating the time taking by the using to perform the particular operation from a steady position.

We calculated Response time for:

- 15 Trails using Hand Gestures
- 15 Trails using Mouse

When we say a person to do operation, we start the stop watch and end it when the operation is done.

From this way we are calculating the response time both by using Hand Gesture and Mouse

We collected the data of all 30 trails i.e. 15 trails using Hand Gesture and 15 trails using Mouse and plotted the graph.

We finally calculated average response times for both.

Average response time of person using Hand Gesture = 1.614666667 sec

Average response time using Mouse = 1.471333333 sec

So as the average response times by using both gesture and mouse are almost same. So, our product performance is almost as equal as using mouse.

C. *Statistical Analysis*

We performed One-way ANOVA test for statistical analysis. Survey from 2 groups of users is taken assuming with each

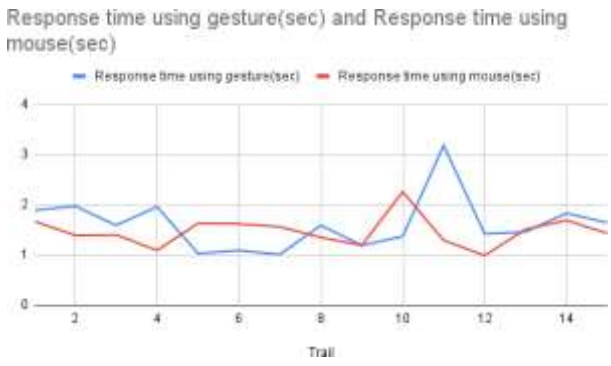


Fig. 9. Response Time of a person in 15 trails

user of different area of domain, such as people familiar to laptop mouse, little familiar to laptop, users who don't use laptop with a different age group other than the previous two groups. The test involves collecting the response time from 2 groups, while they are performing operation of ppt using our product. Now each user's response time is averaged. Figure 10. shows the average time taken by user of each group.

S No	Group 1	Group 2
1	1614	6012
2	2152.8	6732
3	1874	4533
4	3403.2	7841.3
5	3788	8921
6	1543	5453
7	1899.3	7022.5
8	2576	7564.5
9	3112.1	8015
10	3473.5	6003.5
11	1933	5187.5
12	4123.6	6134.5
13	2358	8279
14	3102.5	4214
15	1719	7899

Fig. 10. Average Response times of people in both groups

One-way ANOVA test is performed to verify if there is statistical evidence that these users are actually in different groups. Tool available in internet [11] is used to do the ANOVA test the calculated F-test-statistic value is 89.57447, F critical value in F table with

$$alpha = 0.10$$

is 2.89385. With this it can be said that F-test-statistic value is greater than F critical value with the given value of alpha, which proves that there is a significant difference between the two groups. The same details of the values involved in ANOVA test are included shown in Figure 11,12.

ANOVA Summary

Source	Degrees of Freedom Df	Sum of Squares SS	Mean Square MS	F-Stat	P-Value
Between Groups	1	124602506.8393	124602506.8393	89.5745	0
Within Groups	28	38948379.9825	1391049.2851		
Total:	29	163551886.8219			

Fig. 11. Anova Test Summary

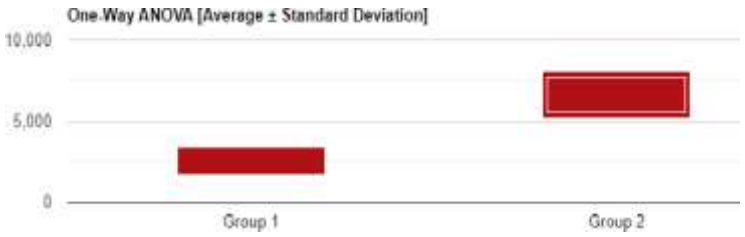


Fig. 12. Anova Test Summary

.

D. *Snapshot of results*

Below given snapshots are the final result of our mode in Fig 13 user is accessing the pointer through gesture and in the Fig 14r user is writing in the power point through drawing feature which is added in the model.

VI. CONCLUSION

The proposed system which is presented in this paper allows humans to control the PowerPoint through distance without using any hardware and medium. It will only take use of the webcam. The system is designed by using a MediaPipe module which provides hand landmarks. With extracted hand landmarks the all 5 gestures which are used in this project

Fig. 13. Using Pointer with Hand Gesture

Fig. 14. Drawing with Hand Gesture

to control the various functionality of the PowerPoint are designed and used to control the power point. To demonstrate the practical use efficiency of the designed system we performed response time analysis on controlling the power point using our system and hardware (mouse) through analysis we found that the average response time of mouse is 1.47133 sec and for our system it is 1.61466 sec which is very near to that of mouse so from this test we can say that outcomes are nearly same so it is better to control the power point using gesture.

REFERENCES

[1] https://www.semanticscholar.org/paper/Power-point-control-using-hand- gesture-recognition-Salunke-Bharkad/7599de204059b35e41207eb06076fea24305ffdf

[2] https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7282130&tag=1

[3] https://www.sciencedirect.com/science/article/pii/S1877050917319130

[4] https://www.semanticscholar.org/paper/Finger-Pointing-Gestur-e-Analysis-for-Slide-Harika-Prihatmanto/541d61cfa21af385a7649c2d43b8d 04c75aee371