Smart Presentation System using Hand Gesture

Amrendra Shekhar
Computer Science Engineering
Arya Institute of Engineering and Technology
Jaipur, India
amrendrashekhar890@gmail.com

Divyanshu Mishra
Computer Science Engineering
Arya Institute of Engineering and Technology
Jaipur, India
divyanshum174@gmail.com

Priyanshu Mishra
Computer Science Engineering
Arya Institute of Engineering and Technology
jaipur, India
priyanshumishra2822@gmail.com

Harsh Jha
Computer Science Engineering
Arya Institute of Engineering and Technology
jaipur, India
jhah63448@gmail.com

***Abstract*—In the current context, there is contact between humans and computers at all levels, so we have diverse approaches.One of them is Gesture Recognition which play a quit important role while interacting with Computers. It include advancement of technology to support learning System.Human-Machine Interaction (HMI) enables used to control presentation using natural Hand gesture or hand Movement.This paper propose a simple system that uses machine learning algorithm to identify the hand structure and perform various action based on the command given by the hand movement.Our suggested system uses Mediapipe 2D skeletal data to detect the hand landmark and recognise the gesture.Because there is a substantial barrier between the user and the computer, using a physical device like mouse or keyboard for human Comuter interaction limits natural interface. We have developed a powerful marker-less hand gesture detection system that can track both static and dynamic hand gesture in this study.Our system converts the detected motion into action like moving the slides forward and backward, writing something and erasing which written. In presentation, the dynamic gesture used to shuffle the slides.[5] According to the test, the hand gesture application can hadle in real-time with an accuracy of more than 90%.**

***Index Terms*—Mediapipe, Gesture, machine learning, Human-Machine Interaction(HMI).**

I. Introduction

Every day in our modern culture, a massive volume of diverse information is bought into the commercial and educational environment. Because they are often constrained to using a devices such as a mouse, keyboard, or wireless presenter tool to interact with presentation content, current presentation system provide a barrier to sharing this quantity of knowledge.Presenters would find it more convenient to make their presentation utilising only hand motion, gesture, finger counts. This dilemma drives our research to create a smart presentation system that recognise human behaviour and provide a high level of naturalness and flexibility when presenting [7]. Managing the show using the smart presentation system is as simple as utilizing hand gesture instruction. As a result, the presenter not only gives the presentation, but also make it more appealing,understandable and communicative. Hand Gesture, hand movement and finger are first two recognition system proposed in the smart Presentation System. Recognition systems can use keyboard and mouse actions as a result of their work. Users can operate and engage with any application using a combination of recognition systems and mouse-keyboard movements, such as Microsoft Power Point or Google slide presentation.[10]

Identify applicable funding agency here. If none, delete this.

II. Literature Review

Hoang-A. Le and colleagues.[1] Developed a system that allows presenters to control their presentations with natural gestures and vocal directions. The system combines three interaction modules: gesture recognition using Kinect 3D skeletal data, key idea detection using HMM approach from natural speech, and small-scaled hand gesture recognition using haptic data from a smartphone sensor, all of which are analysed in real time. To enhance the thrilling experience for audiences, the system uses events provided by the three modules to trigger pre-defined scenarios in presentation. The necessity to train the word that you wish to recognise, as well as the lack of an Indonesian pre-defined corpus, are both issues when utilising HMM speech recognition.

J. Han-Shao et al .[2] This study provides an in-depth look at recent Kinect-based computer vision methods and applications. The studied options are categorised according to the types of vision problems that the Kinect sensor can address or improve. Preprocessing, object tracking and recognition, human activity analysis, hand gesture analysis, and indoor 3-D mapping are among the topics covered. Outlined the primary algorithmic contributions for each type of approaches and summarised their advantages/differences from their RGB counterparts. Finally, give an outline of the field's issues and potential research directions. For Kinect-based computer

vision researchers, this document is expected to serve as a tutorial and a source of references.

This article by Ren Zhou and colleagues[3] focuses on developing a comprehensive hand gesture detection system using the Kinect sensor. To deal with the chaotic hand shape collected from the Kinect sensor, we suggest Finger-Earth Mover's Distance, a new distance metric for measuring hand dissimilarity (FEMD). Because it simply matches fingers rather than the entire hand form, it can better discern subtle changes in hand gestures.

Sritrusta et al.[4] have a similar kind of project which uses kinect as a device that can perform different action like gesture recognition, finger tracking, then draw while the hands are detected that there is sketch on the display hand movement earlier results, and research uses Mircosoft Visual studio Application Developer.

F. Morbini et al .[5] This research examines the suitability of a number of publicly available automated speech recognizers (ASRs) for usage in various types of dialogue systems. This study focuses on cloud-based ASRs that have recently become available to the public, covering ASR system capabilities as well as desiderata and requirements for various dialogue systems, taking into account the system's genre, user type, and other factors. Then, for six different dialogue systems, display the speech recognition results. As a result, different ASR systems perform better on different data sets.

Vivien A. Et al.[6] demonstrate how they used CMU Sphinx to create a multimedia player that can be operated with an Indonesian voice recognition system.

III. Current Opportunity For Hand Gesture

It's about locating a location where gesture commands can be collected as commands or where users can receive feedback. As advanced significantly in recent years, there is a particular interest in the location, or interface, of gesture commands. Electronic analogues of pencil and paper are gestural interfaces. Gestural interfaces feature a variety of possible benefits as well as a few potential drawbacks. large screen, PC/Laptop based, LED light, audio-visual, mobile handheld, and other interfaces are among the researches in the table. The study reveals the possibilities for implementing technology in several fields. The following is a list that is not exhaustive.[13]

A. Entertainment

For whatever type of user, gesture technology can deliver more entertainment options. GestureTek offers a variety of gesture-based entertainment options, including interactive advertisements, signage, movies, and screens. Wii, Sony Eyetoy, and Microsoft's X-box have all displayed many entertainment options such as music playing, personalised games, and so on.[15]

B. Artificial Intelligence

People, technologies, and computing will become increasingly intertwined and will soon become a part of our daily lives. In this intelligent life, gesture-based technologies will play an essential role. Any portion of the body can be used to deliver communication commands or even to operate the window curtain. As part of the Human Robot Interaction, the robotic industry is also adopting gesture technologies to manage and control the robot's actions. Many researches, like Select-and-Point, are well received by consumers and can greatly improve users' interaction with numerous gadgets in a ubiquitous computing environment. Users can connect various devices using networking technologies and hand gestures.[9]

C. Simulation

Body gestures are used to simulate human body activity on the screen. In a variety of ways, physical simulation can improve the realism of the final gestural animation. GestureTek creates a fun virtual reality therapeutic world in which patients are guided through interactive rehabilitation exercises, games, and activities that can target specific body parts as therapists prescribe. [8] The patient's progress is monitored and evaluated.

D. Training and Education

It is possible to create a digital solution for training and education. People in rehabilitation or fitness centres can be automatically trained based on their profile and body structure. The capacity to take natural input from body movements is the most important advantage over a mouse or keyboard. [3]

E. Assistive living

Multi-agent systems, secure communications, hypermedia interfaces, rich environments, increased home appliance intelligence, and collaborative virtual environments are all convergent and represent an important enabling factor for the design and development of virtual elderly support community environments. TeleCARE's goal is to build and construct a customizable framework for virtual communities that support older people. During the research, we tried to pay closer attention to gesture-controlled devices for the elderly and disabled. We must continue to strive more for them as their numbers grow. New technology has the potential to help elderly and disabled persons live independently and with assistance. They can do so by restoring power and independence, both of which are psychologically desirable. [4] Telehealth, telecare, telemedicine, and personal safety systems all exhibit this trend. They use a multimodal strategy due to external conditions or personal preferences. Several factors influence interface design considerations when building for elderly users. The removal of age-restricted users' isolation has little impact on interface design.

IV. Problem Statement

A "Smart Presentation System based on Hand Gesture using Camera" determined by visual processing. There has been a lot of study in gesture recognition combining Kinect Sensor and a HD camera in the last year, however cameras and kinect sensor are more expensive. This research used a simple web camera to minimise the cost and improve the robustness of the suggested system.

V. Proposed Work

The majority of gesture recognition system have three major steps. The detection of objects is the initial stage. This stage's goal is to identify hand items in digital images or videos. At this step, many environment and images problem must be resolved in order to ensure that hand pattern or region can be retrieved precisely in order to improve recognition accuracy. Unstable brightness and contrast are all common images issues. These issues can be efficiently by a better atmosphere and camera devices.[11]

When the gesture recognition system is used in the real world or becomes a product, however, it is difficult to regulate. As a result, the image processing method provides a better way to handle these picture difficulties and build a robust and adaptive gesture detection system. Object recognition is the second stage. To identify the gestures, the observed hand objects are recognised. Differentiated characteristics and successful classifier selection are important issues in most studies at this point. The final stage involves analysing sequential motions to determine what users are instructing or doing.[2]

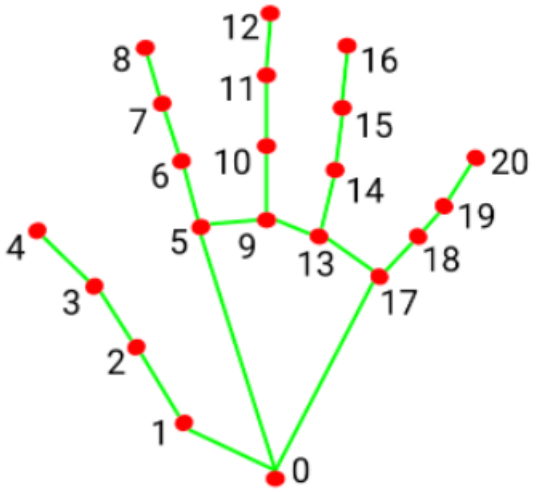

Fig. 1. LandMarks on Hand.

The demand which is more participatory and engaging with interacting experience has increased, and interactive presentation technology has give a solution. For example, if you can flip papers while giving a presentation with a gesture, or if equipment can change the background of a live event using gesture control, the presentations will be more interactive and engaging to the audience.This is a gesture application which detect hand gesture and sends wireless signals over a great distance that perform some specific functionality. People can also use an interactive projection system to create more appealing creative exhibits, which include the interaction with floors and walls. This product is about the creation of interactive projection technology that recognises the user's gesture in real time and gives a more active and fascinating watching experience.[7] By walking into the audience, the presenter can give a performance. A few interactive applications have been created using this technology.

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

Fig. 2. Names Given to Landmarks.

Hand gestures can also be used to operate electronic gadgets with the system. As a result, this system will function as a remote control for all consumer electronic items in a home. Hand gestures are detected and recognised and utilised as command signals to control gadgets. The item can also be used to show text or an image. Disabled people might utilise this to display their needs. All that is required of the user is that they wear the product and make a pre-determined motion. This single product aids in the fields of presentation, home automation, and accessibility for the disabled.[13] It's a significant technological advancement.

A. Separation and Detection Process

Whenever we talk about detection, whether it's of an object, a person, an animal, or, in our case, hands, the first step is to establish the model with valid parameters. This is true regardless of whether we're using Mediapipe or Yolo, but initialising the model is crucial.[15] On both desktop and mobile devices, Mediapipe provides models with high accuracy and low latency. It gives a 3D Hand Landmark model that uses machine learning techniques to predict 21 points from a single picture and can be used on desktops, mobile devices, and browsers, among other platforms. Here is a list of all the points provided by mediapipe for a single hand.

Hand tracking is the method by which a computer detects a hand from an input image using computer vision and maintains focus on the hand's movement and orientation. Hand tracking enables the creation of a variety of applications that employ hand movement and orientation as input.

In order to execute hand tracking as part of our programme, we frequently reuse code from previous projects. Because we just write the code once, creating a hand tracking module overcomes this problem. After that, we turn this code into a module. We can use this module to perform hand tracking in any Python project that we are working on.You now know how to construct a hand tracking software and have all of the necessary skills. You also possess the necessary expertise to turn the code into a module.[15]

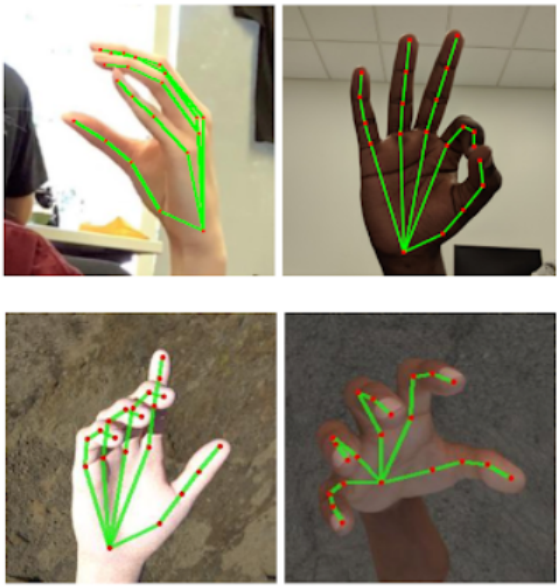

Fig. 3. Hand Detection using Mediapipe.

Fig. 4. Hand Tracked As per the Landmarks

B. Tracking Process

There are two stages to hand tracking with MediaPipe:

- Palm detection - MediaPipe analyses the entire input image and generates a hand-cropped image.
- Hand landmarks recognition - MediaPipe recognises 21 hand landmarks on a cropped image of the hand.

We use a simple algorithm to derive the gestures from the expected hand skeleton. To begin, the cumulative angles of joints establish the status of each finger, such as whether it is bent or straight. The set of finger states is then mapped to a set of pre-defined movements. We can estimate basic static gestures with reasonable accuracy using this simple yet effective technique. Counting motions from many cultures, such as American, European, and Chinese, as well as numerous hand signs such as "Thumb up," "closed fist," "OK," "Rock," and "Spiderman," are supported by the existing pipeline. The hand is tracked in order to analyse the landmark and determine how it might be allocated to a specific function. Here are some factors to consider for efficient hand Tracking:

- Static image mode:- Indicates whether the incoming images should be regarded as separate and unrelated (True) or as a video stream (False) (False). We'll set the value to False, which means that after successfully detecting hands in a video frame, the algorithm will localise the landmarks and simply monitor the landmarks in subsequent frames without invoking another detection until it loses track of any of the hands.
- Max num hands: The number of hands that can be identified in total. Despite the fact that this parameter's default value is 2, we'll explicitly set it to the same value to demonstrate its use.
- Min detection confidence: The minimum level of confidence (between 0 and 1) that the hand detection is effective. We'll put it at 0.7.
- Min tracking confidence: Minimum level of confidence (between 0 and 1) that the hand landmarks have been correctly tracked. We'll put it at 0.7.

C. Presentation Control

Thanks to gesture detection, you may control your creative presentation without touching any physical appliances which include remote or even your touch Screen.Given that some slide can be used create Slides and demos, using gesture graphics and symbols in our presentations to demonstrate how a programme is used may be beneficial. If you can back and forth pages with a gesture during a presentation, for example, the presentation will be more interactive and appealing in any presentation. hence this will also be beneficial for upcoming digital Era that can comprehend gestures, such as the Samsung Smart TV. Google Slides is an essential component of our professional and academic lives. We confront a number of issues while giving a presentation or seminar on a particular topic, such as the necessity for an operator to run the slides, which can cause issues if the speaker and operator are not well prepared.[3]

The Python code for the presentation controller is set up in such a way that the user can do the following with only a single finger gesture:

- To display certain stored images
- Forwarding the Slides to next.
- Moving slides to previous one.
- Draw and Erase makes more interactive.
- Using gesture will always be a fun and interactive way to keep up your digital life.

D. Advantages of Proposed System

The proposed technique is intended to eliminate all of the current method's drawbacks. It will have some new functions

that aren't available in the existing system. The proposed system is a low-cost replacement for the current system's manual processes.

- A projection system that is interactive
- There will be no physical contact with the computer.
- Real-time recognition of gestures
- Using a gesture, flip pages.
- It's simple to set up and utilise on any platform.
- Wireless transmission for smart interaction.

We conceived and designed a system for vision-based hand gesture identification that included several stages that we explained using an algorithm. The gesture recognition system's functioning flowchart is illustrated below:

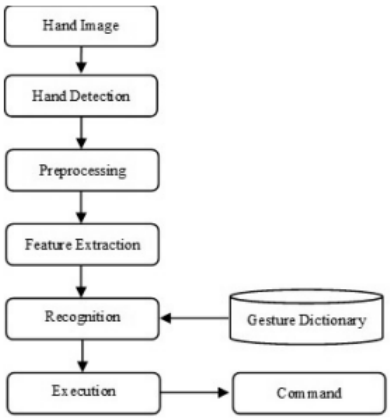

Fig. 5. Flow Chart of designed System

VI. Result and Discussion

In today's world, there are a variety of facilities and modalities for delivering input to any programme. Unfortunately, despite the ever-increasing smart environments and matching input technologies, there are currently few applications that can be operated utilising the samrt digital method for delivering the data by hand gesture. The best significant benefit of utilizing this kind of hand gesture based system is that user can interact with machine without using any physical devices. The present study proposes and implements the use of hand gestures to manipulate objects in a virtual environment, resulting in an efficient and user-friendly human computer interface. The user can interact with virtual objects using hand gestures instead of other physical input devices with the aid of this programme. Because the application allows users, particularly physically challenged users, to design gestures based on their feasibility and convenience of use, the gestures can be defined in a variety of ways. For several objectives, we proposed an approach based on a gesture recognition system. Experiments have shown that continual user training leads to improved skills and, as a result, better performance. This technology will allow more applications to be controlled with hand gestures in the future. Paint, slide displays, multi-player games, and many other applications are among them.[16]

After multiple repetitions of the practical testing, it was discovered that the system performs exceptionally well in favourable lighting circumstances and in a plain background environment devoid of any skin-like items. The method isn't very robust because it has a hard time detecting the hand against a complicated background. The system's performance is also influenced by the threshold value used to compute the radius through a series of practical approximations. Despite this, the system is a little more responsive than other systems that have been developed previously because it does not require a training phase for gesture recognition.[11]

VII. Conclusion and Future Work

There are various facilities accessible in today's world, when technology is at its best, for providing input to any programmes running on computer systems; some of the inputs can be provided through physical touch, while others cannot (like speech, hand gestures, head gestures etc.). Many users may control applications without touching them by using hand gestures. However, there are numerous applications that cannot be operated with hand gestures. Physically challenged folks can benefit greatly from this technique since they can tailor the gesture to their specific needs. Although the current system we have created appears to be more user-friendly than modern device or command-based systems, it is less robust in terms of detection and identification, as we saw in the previous phase. We need to improve our system and strive to create a more robust algorithm for both recognition and detection, even when the background is busy and the illumination is normal. We'll also need to expand the system to include additional gesture classes, as we've only developed it for six so far. This system, on the other hand, can be used to control applications such as power point presentations, games, media players, and Windows Picture Manager, among others.[13]

VIII. References

References

[1] S.Rautaray and A. Agrawal, Vision Based Hand Gesture Recognition for Human Computer Interaction: A survey, Springer Transaction on Artificial Intelligence Review, pp. 1–54, (2012).

[2] Siddharth S. Rautaray Anupam Agrawal," Vision based hand gesture recognition for human computer interaction: a survey", ArtificialIntelligence Review, DOI 10.1007/s10462-012-9356-9, Publisher Springer Netherlands Issues 06 Nov 2012.

[3] Steve Furber, ARM System-On-Chip Architecture (2nd Edition), Pearson Education Rajkamal, Embedded Systems: Architecture, Programming and Design, Tata McGraw Hills Publications.

[4] N. A. Ibraheem, R. Z. Khan and M. M. Hasan, Comparative Study of Skin Color Based Segmentation Techniques, International Journal of Applied Information Systems (IJAIS), vol. 5(10), pp. 24–38, (2013)

[5] Ray Lockton, Balliol College, Oxford University, Hand Gesture Recognition using Computer Vision.

[6] Sritrusta Sukaridhoto, Mohammad Hasbi Assidiqi and Novenda Nur Alfath Salam, "Simple Interactive Projector Based on Hand Gesture Movement", IES, 2014.

[7] Vivien Arief Wardhany, Sritrusta Sukaridhoto and Amang Sudarsono, "Indonesian Automatic Speech Recognition for Command Speech Controller Multimedia Player", EMITTER International Journal of Engineering Technology, vol. 2, no. 2, December 2014.

[8] Simple-OpenNI OpenNI library for Processing, [online] Available: https://code.google.com/p/simple-openni/.

[9] Processing library that performs real-time finger tracking from depth images [Software]. Available from, [online] Available: https://github.com/atduskgreg/FingerTracker.
[10] A. Zhang, Speech Recognition (Version 3.1), 2015.
[11] Li Lantian and Thomas Fang Zheng, "Gender-dependent feature extraction for speaker recognition", Signal and Information Processing (ChinaSIP) 2015 IEEE China Summit and International Conference on, pp. 509-513, 2015.
[12] Zheng Thomas Fang, Qin Jin, Lantian Li, Jun Wang and Fanhu Bie, "An overview of robustness related issues in speaker recognition", Asia-Pacific Signal and Information Processing Association 2014 Annual Summit and Conference (APSIPA), pp. 1-10, 2014.
[13] [4] Sritrusta Sukaridhoto, Mohammad Hasbi Assidiqi, Novenda Nur Alfath Salam, "Simple Interactive Projector Based on Hand Gesture Move-ment", IES 2014.
[14] [4] Sritrusta Sukaridhoto, Mohammad Hasbi Assidiqi, Novenda Nur Alfath Salam, "Simple Interactive Projector Based on Hand Gesture Move-ment", IES 2014.
[15] Logistic Regression is Still Alive and Effective:The 3rd YouTube 8M Challenge Solution of the IVUL-KAUST team Merey Ramazanova, Chen Zhao, Humam Alwassel, Mengmeng Frost Xu, Sara Rojas Martinez, Bernard Ghanem, Fabian Caba (IVUL-KAUST, ranked at 11)
[16] Steve Furber, ARM System-On-Chip Architecture (2nd Edition), Pearson Education Rajkamal, Embedded Systems: Architecture, Programming and Design, Tata McGraw Hills Publications