International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Gesture Control System

Undergraduate of Computer Science and Engineering, Acropolis Institute of Technology and Research, Indore, India

Shambhavi Narulkar
(Tel: +91-9009941560; Email: shambhavinarulkarcs19@acropolis.in)
Vanshika Jain
(Tel: +91-7828047459; Email:vanshikajaincs19@acropolis.in)
Vinay Sharma
(Tel: +91-9399198477 ; Email:vinaycs19@acropolis.in)
Yash Kharche
(Tel: +91-9009171494; Email: yashkharchecs19@acropolis.in)

Abstract - For a user, a few gestures could be defined for frequently used tasks- save, exit, print, next slide, previous slide, etc. With the advancement in technology, the next frontier of Human-Computer interaction is gesture-controlled computers. The problem we are tackling here is Hand-gesture recognition using just the computer's webcam i.e., without any external hardware and sensors to perform functions like save, print, and exit operations in Word and PowerPoint. This hand gesture recognition technology can be extended to various fields like virtual reality, healthcare, consumer electronics, etc.

Key words: TensorFlow, Python, Media Pipe, OpenCV, Machine Learning, Gesture recognition, Automation.

Introduction

Gesture recognition is a technology that recognizes hand movements and gestures made, to read and interpret hand movements as commands. This has minimized the need for text interfaces and GUIs (Graphical User Interface). Gestures are useful for computer interaction since they are the primary and expressive forms of human communication. The project deals with recognizing Hand-gestures to execute the most frequently performed action such as save, print, open, close, volume control, etc. The project needs to be able to recognize and identify a hand gesture accurately. To execute a command which is linked to the performed hand gesture correctly. The project can be useful for touchless technology, as people can control a system sitting from a valid distance from the computer. It can also be used by visually impaired people. Gesture recognition is a technology that recognizes hand movements and gestures made, to read and interpret hand movements as commands. This has minimized the need for text interfaces and GUIs (Graphical User Interface). Gestures are useful for computer interaction since they are the primary and expressive forms of human communication. The project deals with recognizing Hand gestures to execute the most frequently performed action such as save, print, open, close, volume control, etc. The project can be useful for touchless technology, as people can control a system sitting from a valid distance from the computer. It can also be used by visually impaired people with some training.

Literature Review

Object Detection and Recognition has been studied for more than four decades now and significant efforts have been paid to develop representation schemes and algorithms aiming at recognizing generic objects in images taken under different imaging conditions. Within a limited scope of distinct objects, such as handwritten digits, fingerprints, faces, and road signs, substantial success has been achieved. In addition, significant progress towards object categorization from images has been made in the recent years. This enables hand gesture recognition.

Analysis of Current System

While there is no fully developed software for hand gesture recognition to control the personal computer, there exist some software for professional use which are mentioned ahead.

No.	Existing software	Features	Disadvantages	Limitation /Gaps
1.	**Motion Gestures**	Eliminates building own gesture recognition software	No plans available for personal use	Requires internet connection to run the software.
2.	**Mano Motion**	Eliminates need to operate machines manually.	Used only by industries Not for individua l use.	Not for personal use.
3.	**GetSure**	Eliminates the need to bridge the sterile field and risk surgical site infection.	Some external hardware required. And only for Surgeons	External hardware required.

Solution Proposed

The solution is to create a system that can recognize hand gestures and perform desired actions. The steps involved are:

1. Hand Detection

2. Gesture Recognition

3. Corresponding Command execution

The project's expected outcome is to enable the system to be able to recognize a hand gesture and hence perform a desired action.

We'll first use MediaPipe to recognize the hand and the hand key points. MediaPipe returns a total of 21 key points for each detected hand. These key points will be fed into a pre-trained gesture recognizer network to recognize the hand pose.

The program begins with capturing and reading each frame from the webcam using OpenCV. MediaPipe works with RGB images but OpenCV reads images in BGR format. So, using 'cv2.cvtCOLOR()' function we convert the frame to RGB format. The 'process' function takes an RGB frame and returns a result class. Then we check if any hand is detected or not, using 'result.multi_hand_landmarks' method.

After that, we loop through each detection and store the coordinate on a list called landmarks. Here image height (y) and image width(x) are multiplied with the result because the model returns a normalized result. This means each value in the result is between 0 and 1.

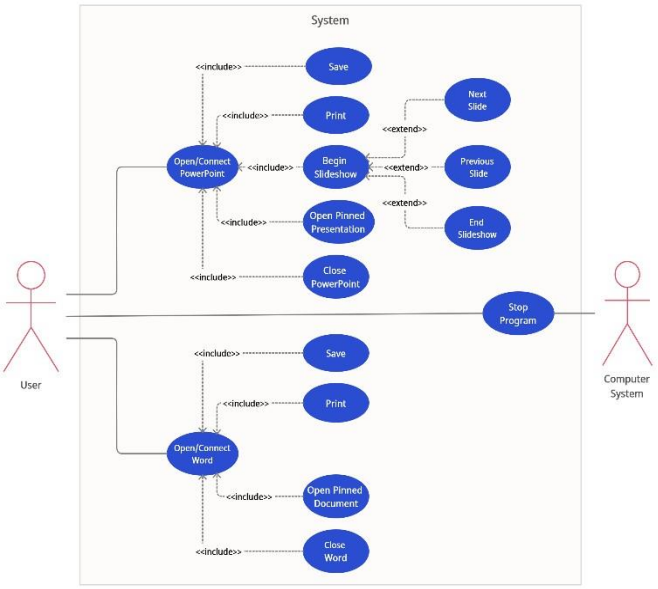
And finally using 'mpDraw.draw_landmarks()' function we draw all the landmarks in the frame. The 'model.predict()' function takes a list of landmarks and returns an array contains 9 prediction classes for each landmark.

The output looks like this-

[[2.0691623e-18, 1.9585415e-27, 9.9990010e-01, 9.7559416e-05, 1.6617223e-06, 1.0814080e-18, 1.1070732e-27, 4.4744065e-16, 6.6466129e-07, 4.9615162e-21]] 9]

'Np.argmax()' returns the index of the maximum value in the list. After getting the index we can simply take the class name from the classNames list which changes according to the stage at which the program is in, i.e., if it is in the main menu, ppt menu or the doc menu. Then using the 'cv2.putText' function we show the detected gesture into the frame. Finally, we execute the required command using 'PyWinAuto', which is a set of python modules to automate the Microsoft Windows GUI.

Use Case diagram



Technology Used

Python

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation. Python is dynamicallytyped and garbage-collected. It supports multiple programming paradigms, including structured, object-oriented and functional programming.

Tensorflow

TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks.

OpenCV

OpenCV is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel, it was later supported by Willow Garage then Itseez. The library is cross-platform and free for use under the open-source Apache 2 License.

PyWinAuto (Python Window Automation)

PyWinAuto is a set of python modules to automate the Microsoft Windows GUI. At its simplest it allows you to send mouse and keyboard actions to windows dialogs and controls.

MediaPipe

Mediapipe is a cross-platform library developed by Google that provides amazing ready-to-use ML solutions for computer vision tasks. MediaPipe is able to achieve its speed thanks to the use of GPU acceleration and multi-threading. Such development techniques are generally difficult, but MediaPipe makes it easier. MediaPipe Hands is a high-fidelity hand and finger tracking solution. It employs machine learning (ML) to infer 21 3D landmarks of a hand from just a single frame. Whereas current state-of-the-art approaches rely primarily on powerful desktop environments for inference, our method achieves real-time performance on a mobile phone, and even scales to multiple hands.

Limitation of Work

The performance of a system highly depends on the quality of the training data. A system trained with good quality data is trained well. The model for the project is trained with good quality data. While the model is able to recognize all the gestures with high accuracy, it may give inaccurate results if the captured camera frame is too cluttered or the lighting is improper. The project also cannot operate if two Word files or two PowerPoint files are open at the same time.

Conclusion

In summary, Gesture recognition technology is the turning point in the world of Interactive Touchless technology. It can allow seamless non-touchable control of computerized devices to create a highly interactive, yet fully immersive and flexible hybrid reality. The inclusion of this technology in multiple applications across various sectors is further revolutionizing human-computer communication.

Acknowledgement

We thank our Institution **Acropolis Institute of Technology and Research** for giving us an opportunity to work under such profound minds.

There are number of people without whom this projects work would not have been feasible. Their high academic standards and personal integrity provided me with continuous guidance and support.

We owe a debt of sincere gratitude, deep sense of reverence and respect to our guide and mentor **Prof. Preeti Shukla,** Professor, AITR, Indore for his motivation, sagacious guidance, constant encouragement, vigilant supervision and valuable critical appreciation throughout this project work, which helped us to successfully complete the project on time.

We express profound gratitude and heartfelt thanks to **Prof. Kamal Kumar Sethi**, HOD-CSE, AITR Indore for his support, suggestion and inspiration for carrying out this project. We are very much thankful to other faculty and staff members of CSE Department, AITR Indore for providing me all support, help and advice during the project. We would be failing in our duty if do not acknowledge the support and guidance received from **Dr S.C. Sharma**, Director, AITR, Indore whenever needed. We take opportunity to convey my regards to the management of Acropolis Institute, Indore for extending academic and administrative support and providing me all necessary facilities for project to achieve our objectives.

We are grateful to **our parent** and **family members** who have always loved and supported us unconditionally. To all of them, we want to say “Thank you”, for being the best family that one could ever have and without whom none of this would have been possible.

Reference

[1] https://www.tensorflow.org/

[2] https://www.kaggle.com/

[3] https://mediapipe.dev/

[4] https://google.github.io/mediapipe/solutions/hands.html

[5] https://opencv.org/

[6] https://pywinauto.readthedocs.io/en/latest/

[7] https://www.motiongestures.com/

[8] https://www.manomotion.com/