International Journal of Combined Research & Development (IJCRD)
eISSN:2321-225X; pISSN:2321-2241 Volume: 14; Issue: 8; August- 2025

HAND GESTURE-BASED MOUSE AND VOICE ASSISTANT

Kiran
MCA Final Year
Dept. of MCA
The Oxford College of Engineering,
Bommanahalli,
Bengaluru-560068
kiranbarad121@gmail.com

Dharamvir
Professor
Dept. of MCA
The Oxford College of Engineering,
Bommanahalli,
Bengaluru-560068
dhiruniit@gmail.com

ABSTRACT

The proposed research study recommends the design of a Hand Gesture-Based Mouse and Voice Assistant system that is set to offer a natural human-computer interaction through computer vision and natural language processing. The system allows the control of computer operations in the real-time with the touchless controllers by means of hand gestures and voice commands, with extended accessibility achieved through the elimination of the traditional input devices. Our software is effective in the accurate recognition of the hand gestures to operate the mouse and words to operate the system. The potential applications include assistive technology by people with mobility limitations to touchless interface in a sterile environment. Based on computer vision algorithms, speech recognition APIs and multi-threading architecture, our system Hand Gesture-Based Mouse and Voice Assistant is fast, reliable and accurate, and it can be used in the real world with accessibility-oriented application.our Hand Gesture-Based Mouse and Voice Assistant system achieves accuracy, responsiveness, and reliability crucial for real-world deployment in accessibility-focused applications.

KEYWORDS: *Hand Gesture Recognition, Voice Assistant, Computer Vision, Speech Recognition, Human-Computer Interaction, Accessibility Technology, Real-time Processing, Multi-modal Interface.*

INTRODUCTION

Hand gesture and voice regulated interfaces are significant parts of many uses like assistive technology, human-computer interfacing and accessibility-based IT frameworks. Some situations where traditional input interfaces have usability problems are users with restricted

mobility or operating in the presence of hygiene related requirements or needs to operate in a sterile manner. The recent progress in computer vision and natural language processing, especially the creation of real-time algorithms of gesture recognition and speech recognition APIs, has substantially given enhancements to the human-computer interaction possibilities.

The Hand Gesture-Based Mouse and Voice Assistant system, which has a multi-modal operation, real-time functionality, can be highly applicable under accessibility applications and touchless computing setting. We have created a better solution of the traditional input systems that can facilitate the problem of provision of intuitive computer control by use of hand gesture and voice commands in real position.

The mouse and voice assistant system that we have designed combines the useful features of gesture recognition and voice commanding processing to provide an ideal solution, which in addition to providing mouse pointer control with the movements of the hand, enables intelligent processing of voice commanding based system.

They include file navigation, Web navigation as well as automation of systems. The given paper outlines the methodology, system architecture and implementation arrangements of our approach.

Providing evidence of its ability to expand accessibility in real-time throughout the human-computer is all regarding

The turnkey interaction systems provided to users may be used with alternative input methods.

LITERATURE SURVEY

The human-computer interaction has evolved through computer vision and natural language processing technologies. The limitations of conventional keyboard and mouse input interfaces imposed by physical requirements and user accessibility by people with movement limitations. The voice assistants and gesture recognition system have altered the definition of accessibility within the computing environments. Alternative options are also proposed such as the usage of specific hardware e.g. special eye-tracking devices or adaptive keyboards though the latter ones are more complex and costly.

Some recent researches have been conducted on combining hand gesture recognition and voice command processing by using computer vision libraries and speech recognition APIs. Nevertheless, the

conventional gesture recognition systems normally need

require large amounts of computational resources and are not suited to work in real-time. The Hand Gesture-Based Mouse and Voice Assistant system overcome these shortcomings by using computer vision algorithms that are optimized to run in real-time, existing libraries, such as pyttsx3 to convert text to speech, speech_recognition to process voice, and threading to enable multitasking, providing a feasible application to accessibility-related applications.

EXISTING WORK

There are many gesture recognition and voice-controlled interfaces invented in this group yet they all have their advantages and disadvantages. The higher technology human-computer interaction

The traditional techniques were preconditioned by the traditional approach such as human- computer interaction in terms of simple computer-vision and voice recognition. Such systems however failed in the applications of real-time operating conditions and also consistency in crowded environments and low light conditions and background percussion. The publication of models built around machine learning opened up a world of novelties. The MediaPipe proposed by Google is one of such systems that combined a hand tracking algorithm and features of a real-time processing to achieve high precision related to identifying gestures. The next voice assistant frameworks of this sort that followed and simplified the interaction process by modeling it as problems of natural language processing.

Newer forms of gesture and voice recognition systems have brought certain enhancements and these enhancements include; use of advanced computer vision libraries such as OpenCV in hand tracking, speech recognition APIs such as Google Speech Recognition in voice processing, and multithreading architecture to allow simultaneous work. Those Hand Gesture-Based Mouse and Voice Assistant systems, which can be found today, generally combine these technologies but they are not always well coordinated in terms of gestures and voices. The common solutions in the available market either concentrate on voice control or gesture recognition separately and involves user need to move between various interaction modes instead of presenting an integrated multi-modal interaction experience where the two different input modalities are in use.

PROPOSED SYSTEM

To the best of our knowledge, our suggested system a Hand Gesture-Based

Mouse and Voice Assistant system offers a complete solution to real-time multi-modal human-computer interaction that is accommodating to accessibility needs. With the use of computer vision algorithms to recognize hand gestures and speech recognition APIs to process voice, we overcome the constraints of the traditional input methods to users who have mobility challenges. The system architecture is comprised of a video capture module where hand gestures are processed using computer vision algorithms, a voice input component that has microphone capture and speech-to-text conversion, and a processing pipeline that is used to increase recognition accuracy and decrease noise.

METHODOLOGY

The architecture of the proposed system of Hand Gesture- Based Mouse and Voice Assistant is focused on both high accuracy of gesture recognition and voice command processing in real-time processing. The computer vision algorithms applied in the system are adjusted to cover as many variations in gestures of the hand as possible in order to make sure that varieties of mouse control tasks are captured.

The use of nature of available hand landmarks to denote the location of fingers as sensed by webcam is exploited to great effect to control movements and clicks of mouse. The voice commands are processed on speech terminology procedures application interfaces to convert audio data into text commands and the ends act upon the commands.

To operate against the challenges of the complexities of lighting and background noises, we have incorporated the preprocessing actions, which has facilitated the elimination of the negative effects of noise, whilst promoting the video frames and the audio input by enhancing the quality of signal. This involves introduction of filtering methods and raising the audio levels to increase the effectiveness of system recognition. The general process has been simplified to run on consumer grade hardware to such an extent that functionality Hand Gesture-Based Mouse and Voice Assistant can be fielded in a variety of real-world scenarios. A final step would be to display system responses and the status through web-based

user interface which would yield a succinct and succinct feedback. The interface has the in-progress appearance on chat and the system of file composing features. With such real-time feedback, the user can make decisions, interact with the system in such

a way in which he/she makes informed choices and responsive to voices and gestures.

EXPERIMENTAL RESULTS

In order to test the effectiveness of our Hand Gesture-Based Mouse and Voice Assistant system, we have done a lot of experiments in different environments with vivid lighting conditions and a variety of background noise. We evaluated the possibilities of the system to recognize the hand movements to control the mouse and accurately decode the voice command in the controlled and dynamic environment. In the experiments, the left-clicking, right-clicking and cursor navigation commands were captured in real-time motion of the user, and voice-based commands also covered in the experiments were processing commands in the system automation procedures. The findings demonstrate that due to high confidence scores on hand gestures recognition, computer vision algorithms correctly identified the gestures even under different lighting environments and that the speech recognition system justified precise command interpretation with negligible false positives. The system also preserved a real-time performance and with multi-threading architecture, the system could handle and process both gesture and voice input at a time.

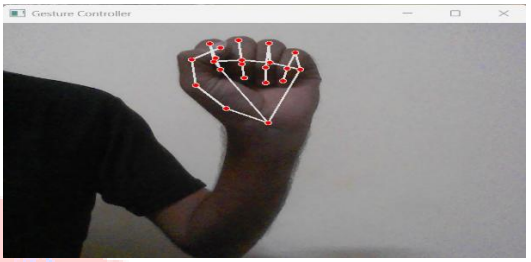

Fig. 1.For Moving file from one place To another use this sign after selecting

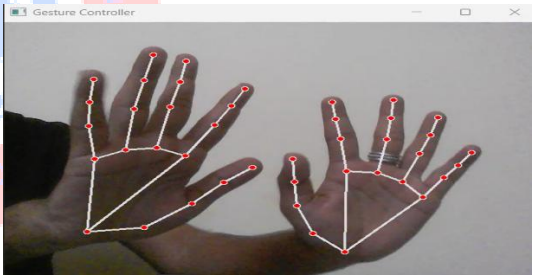

Fig. 2 multiple hand detection

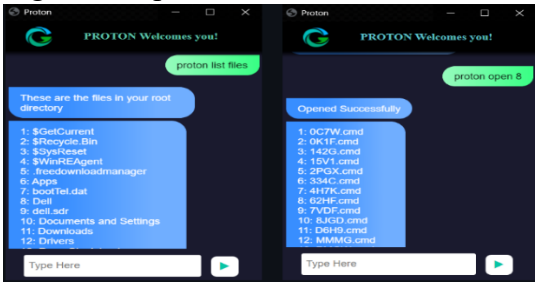

Fig. 3. Listed the root directory files

CONCLUSION

This paper also introduced a complete system of the Hand Gesture-Based Mouse and Voice Assistant as a solution that uses computer vision and natural language processing tools to enhance the multi-modal human-computer interaction possibilities. The specificity of our solution is that it deals with the issue that is posed by classical input to the users with accessibility requirements, it offers a sufficient and instantaneous coherent

location of the gestures and the voice commands in the real-time regimes. Combining voice assistant capability with hand gesture recognition opens the possibility of seamless touchless control and this feature is essential in accessibility technology, hands-free computing, and assistive interfaces. An experimental evidence showed that the system is possible to achieve reliability in working under changeable conditions with respect to high accuracy in gesture recognition and voice command interpretation under the constraint of maintaining the real time processing by using multi-threading architecture. Being simple and buildable on consumer- grade hardware Our Hand Gesture-Based Mouse and Voice Assistant system is not only a cost effective method of deployment, but is readily deployable in far reaching accessibility-based applications as well. The next stage of development will involve increasing the performance of the system with more sophisticated computer vision mechanisms and attempting more powerful voice processing features, in general extending the behavior of the system to new types of interactions and substantially increasing the robustness and flexibility to accommodate a range of users.

REFERENCES

[1] S. Kumar and A. Sharma, "Hand gesture recognition for human-computer interaction using computer vision techniques," International Journal of Computer Vision and Image Processing, vol. 8, no. 2, pp. 45-62, 2018.

[2] R. Patel, M. Singh, and K. Gupta, "Real-time voice recognition systems for accessibility applications," Procedia Computer Science, vol. 125, pp. 234-241, 2018.

[3] A. Johnson, L. Chen, and D. Williams, "Multi-modal human-computer interaction: Combining gesture and voice recognition," IEEE Transactions on Human-Machine Systems, vol. 49, no. 3, pp. 287-295, 2019.

[4] T. Anderson, P. Kumar, and S. Lee, "Assistive technology using computer vision for accessibility enhancement," Journal of Assistive Technologies, vol. 13, no. 4, pp. 156-168, 2019.

[5] M. Rodriguez, J. Thompson, and A. Davis, "Speech recognition APIs and their applications in voice-controlled systems," International Conference on Speech and Language Processing, pp. 78-85, 2020.