Gesture Based Os Navigation And Control

Dr. Doddegowda B J[1], Monika C L[2], Manoj S[3], Manoj K[4], Lakshmikanta[5]

[1]Professor,DepartmentofComputerScience&Engineering,AMCEngineeringCollege,Bengaluru,India

[2-5]Students,DepartmentofComputerScience&Engineering,AMCEngineeringCollege,Bengaluru,India

ABSTRACT

This paper proposes a system for Gesture-based OS navigation using OpenCV, an open-source computer vision library. The system uses a webcam to capture video of the user's hand movements and apply computer vision algorithms at real time to recognize specific gestures and commands. The system is designed to provide an intuitive and natural way for users to interact with computers and other devices.

The Gesture-based OS navigation system allows users to navigate through menus and launch applications using simple hand gestures. The system recognizes gestures such as swipes, taps, and pinches, and translates them into corresponding commands. Notes maker allows users to draw or write in mid-air using their fingers as brushes. The system tracks the movement of the user's hand and draws a corresponding line or shape on the screen.

The proposed system is implemented using OpenCV, Python programming language, and a standard webcam. The system's performance is evaluated using a dataset of hand gesture images, and the results demonstrate the system's high accuracy and real-time processing capability. The proposed system has potential applications in various fields, including gaming, education, and entertainment. The system offers a promising platform for exploring new possibilities in human-computer interaction, and for developing innovative applications that take advantage of natural hand gestures and movements.

Keyword: - *OpenCV, OS Navigation, Gesture based OS Navigation, Notes Maker.*

1. INTRODUCTION

Gesture based operating system navigation and control allow users to interact with their device using hand movements. Gesture recognition technology uses sensors such as cameras and gyroscopes to detect and interpret hand movements and translate them into commands that the device can understand.

Gesture-based OS navigation and control can be found on a variety of devices, including smart phones, tablets, laptops, and smart home devices. Gesture-based OS navigation and control is a technology that allows users to interact with their computer using gestures rather than traditional input methods, such as a keyboard and mouse. Gesture-based OS navigation and control can be used to perform a variety of tasks, such as launching applications, navigating menus and settings, and controlling media playback. It can also be used to input text or data, or to control the cursor on a screen.

Gesture-based OS navigation and control can be an alternative to traditional input methods, especially in situations where it is difficult or impractical to use them, such as when the user is standing or in a confined space. It can also be a more natural and intuitive way to interact with a computer, as it can mimic the way we interact with the physical world. It can be used to move cursors on a screen. We can do right click, left click, double click as mouse performs.

It also contains one more feature called Notes Maker in this project. NotesMaker is a software application that allows users to create and collaborate on digital whiteboards using hand gestures and touchless controls. The software uses a combination of gesture recognition technology, machine learning, and computer vision to detect and interpret hand movements, allowing users to draw, write, and manipulate digital content in a hands-free, touchless way.

1.1 OBJECTIVE

- The objective of gesture-based operating system-based navigation and control are to provide great flexibility for user to interact with their devices. By using hand movements instead of using mouse, gesture-based navigation aims to make it easier and more efficient to users to perform tasks and access features of their devices.

- Gesture-based navigation also has the potential to improve user productivity and reduce the risk of repetitive stress injuries associated with traditional input methods like mouse. By allowing users to perform tasks and navigate their devices using hand gestures, they can work more efficiently and with less strain on their hands and wrists.

1.2 DRAWBACKSOFTHEEXISTINGSYSTEM

As of now we are using Mouse to perform tasks on the screen. Mouse-based navigation has been a standard input method for decades. Using a mouse for long periods of time can cause physical strain on the hand, wrist, and arm, leading to repetitive strain injuries like carpal tunnel syndrome. To overcome that physical pain we are using gesture-based operating system navigation and control to use hand movements for control volume button and cursor moves on the computer screen. We can not write or draw something on the screen in existing system. We can perform only right click, left click, double click, scrolling and select all as mouse performs.

2. LITERATURESURVEY

i. J. Cheng, L. Li, & X. Li [1] Proposed definition for gesture-based operating system navigation and control. He defined gesture-based OS navigation is a system that enables users to interact with their computer's operating system through hand gestures, without the need for a physical input device like a mouse or keyboard.The system typically uses computer vision algorithms to recognize and interpret these gestures, and then maps them to various functions and commands within the operating system.

ii. R. C. Jain & S. S. Saurabh [2] Proposed definition for "Gesture Recognition for Operating System Navigation". He defined Gesture-based OS navigation systems often use machine learning algorithms to recognize and classify hand gestures, which can improve the accuracy and responsiveness of the system.The survey also highlighted some of the challenges and limitations of gesture recognition systems, including the need for high-quality sensors and cameras to capture accurate data, the need for large datasets of annotated gestures to train machine learning models, and the challenge of dealing with variations in lighting, background noise, and other environmental factors that can impact the accuracy of the system.

iii. L. Guo, C. Huang, and Z. Zhao [3] conducted literature survey on A Multi-user Collaborative System for Creative Drawing. The authors described the technical components of the system, including the computer vision algorithms used for gesture recognition, the communication protocol for transmitting data between devices, and the user interface design. They also evaluated the system's performance in terms of accuracy, latency, and user satisfaction through a user study.

3. METHODOLOGY

The objective is to develop a gesture-based operating system navigation and control and Notes Maker are provide better experience for the user to use their devices using their hand movements. It enables more future for the user to use their devices. Using hand movements user interact with the computer and can draw something on the computer screen. Using mouse make human hand wrist so painful. This gesture-based operating system navigation and control alternative for mouse.

The system needs an algorithm that can accurately interpret hand gestures and translate them into commands that the OS can understand.The user interface for the OS must be designed to support touch and gesture-based inputs. This could involve creating a custom UI that enables users to navigate and control the OS using gestures or adapting an existing UI to support gestures. To support the Notes Maker feature, the system must include software that allows users to draw and interact with virtual objects in the air. This could involve developing a custom application or integrating existing air canvas software into the OS.
Following are the Packages need to be installed for the gesture-based operating system navigation and control.

1.CV2:

CV2 is a popular computer vision library for Python. It is the Python interface for OpenCV (Open-Source Computer Vision Library), a powerful and comprehensive library for image and video processing. It provides a wide range of functions and tools for tasks such as image manipulation, object detection, facial recognition, feature detection, and more. It supports various image and video file formats, including BMP, JPEG, PNG, and MPEG.

Some of the key features of cv2 include image and video capture, image processing and filtering, object detection and tracking, and feature detection and extraction. It also includes functions for creating graphical user interfaces for computer vision applications.cv2 is widely used in the computer vision and image processing fields and is popular among developers and researchers. Its powerful functionality and ease of use make it a valuable tool for a wide range of applications, from simple image editing tasks to complex computer vision research projects.

2.MediaPipe:

MediaPipe is a cross-platform framework for building multimodal applied ML pipelines. It is developed by Google and provides a collection of building blocks for creating complex computer vision and machine learning pipelines. MediaPipe offers pre-built components for tasks such as object detection, hand tracking, face detection and recognition, pose estimation, and many others. MediaPipe provides a flexible and easy-to-use pipeline for developers, allowing them to build custom models and pipelines without having to implement everything from scratch. The framework supports a variety of platforms, including desktop, mobile, and the web.

MediaPipe is an open-source project and can be used under the Apache 2.0 license. It is built using the TensorFlow framework and provides pre-trained models for many tasks, which can be easily integrated into your own projects.MediaPipe is a powerful tool for developers who want to add computer vision and machine learning capabilities to their applications, without having to spend a lot of time on building and training models from scratch.

3.PyAutoGUI:

PyAutoGUI is a Python package for automating mouse and keyboard operations on a computer. It allows users to control their computer programmatically by simulating mouse clicks, keyboard typing, and other interactions.PyAutoGUI can be used for a variety of tasks, such as automating repetitive tasks, testing GUI applications, or creating macros. It provides a simple and easy-to-use interface for controlling the mouse and keyboard and can also take screenshots and perform image recognition tasks.

PyAutoGUI is compatible with Windows, macOS, and Linux, and supports both Python 2 and Python 3. It can be installed using pip, the Python package manager, and the package also comes with a number of example scripts to help users get started.While PyAutoGUI can be a powerful tool for automating tasks, it is important to use it responsibly and with caution. It is recommended to always test your scripts thoroughly before running them on important systems, and to be aware of potential risks such as unintentional clicks or keystrokes.

4.Math:

The Python standard library includes a built-in math module that provides a wide range of mathematical functions for performing operations such as trigonometry, logarithms, and more.The math module is part of the Python standard library, which means that it is included with every Python installation and does not require any additional installation or configuration. To use the math module in your Python code, simply import it at the beginning of your script using the import math statement.

5.Enum:

The Enum package in Python provides a way to define enumerations as a distinct type, rather than just a set of named values. Enumerations created with this package are more powerful than simple constants or named tuples because they have a richer set of behaviours, such as iteration, comparison, and support for custom methods. The Enum package provides a more powerful and expressive way to define enumerations in Python, making it easier to work with symbolic names and sets of named values in your code.Enums are often used in programming to define sets of related constants that are easier to read and maintain than raw integer values.

6.Ctypes:

The ctypes package is a Python library that provides a way to access functions in shared libraries or dynamic-link libraries (DLLs) written in C or other low-level languages from within Python. It is particularly useful for working with C libraries that expose a C API, allowing you to call functions from those libraries directly from Python code.To use ctypes, you first need to define the interface to the C library you want to use. This involves defining the functions you want to call, the data types they expect as arguments and the data types they return. Once you have defined the interface, you can load the library into your Python program using ctypes.cdll or ctypes.windll (for Windows DLLs).

7.Comtypes:

The comtypes package is a Python library that allows you to access and interact with Component Object Model (COM) components from within Python. COM is a binary interface standard for software components introduced by Microsoft in the 1990s, and is used for inter-process communication and integration between software components on the Windows platform. Comtypes allows you to create and use COM objects, invoke methods on those objects, and interact with their properties. It can be used to automate Windows applications, access Microsoft Office applications, and work with other COM-based software components.comtypes also provides a number of other functions and classes for working with COM components, including comtypes.gen, which generates Python code from COM type libraries, and comtypes.server, which allows you to create your own COM components in Python.

8.Pycaw:

The pycaw package is a Python library that provides a simple way to interact with the Windows Audio Session API (WASAPI) to control audio playback and recording devices on Windows. WASAPI is a low-level audio interface introduced in Windows Vista that allows applications to interact with audio devices at a low level.With pycaw, you can get information about the current audio devices and sessions, and manipulate audio playback and recording volume and mute settings. You can also monitor audio events, such as when a new audio device is added or removed.We then use Activate to activate the default device, and AudioEndpointVolume to get the volume control for the device.

9.Google.protobuf.json_format:

The google.protobuf.json_format package is a Python library that provides utilities for converting between Protocol Buffer messages and JSON format. Protocol Buffers is a language- and platform-neutral data serialization format developed by Google, and is often used for message passing between distributed systems. With google.protobuf.json_format, you can convert Protocol Buffer messages to JSON format and vice versa, and customize the serialization and deserialization behavior using options and custom converters.

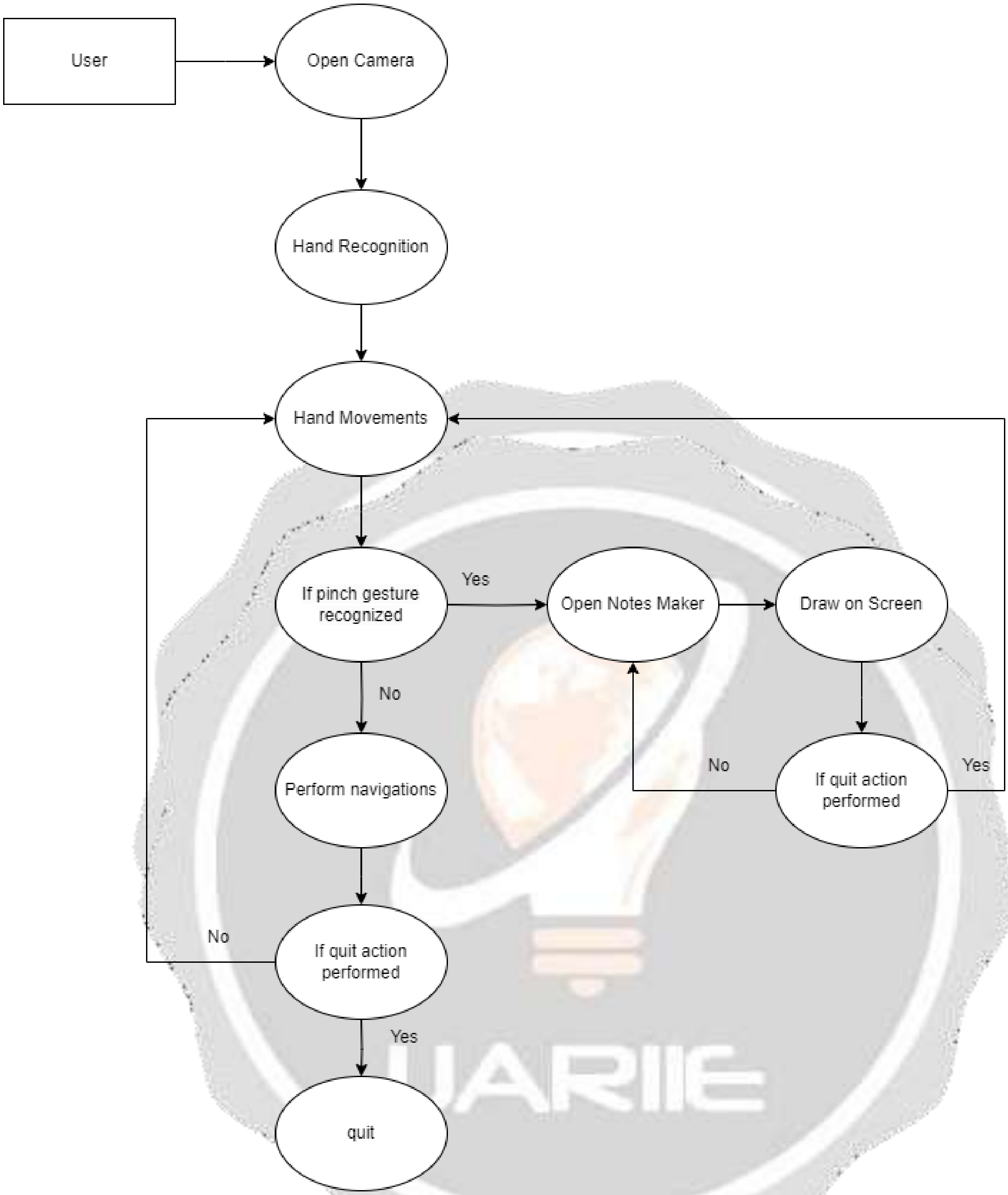

Figure-1 Use case diagram of gesture based OS navigation.

Following are some gesture movements to perform actions on the screen.

1.Move Cursor:
We can move cursor on the screen using by moving index finger and middle finger simultaneously. We can halt or stop the movements by showing all 5 fingers at the same time.The speed of the cursor movement is proportional to the speed of the hand.

2.Left Click:
We can perform left click action by half folding index finger and straitening the middle finger.

3.Right Click:
We can perform right click action by half folding middle finger and straitening the index finger.

4.Double Click:
We can perform double click action by half folding both middle finger and index finger simultaneously.

5.Scrolling:
Dynamic Gestures for horizontal and vertical scroll. The speed of the scroll is proportional to the distance moved by the pinch gesture from the start point. We can perform Scrolling by half folding middle finger and index finger at the same time and moving wrist up and down.

6.Drag and Drop:
It can be performed by folding all fingers and making a fist to drag any files or windows and by straightening all the fingers showing a palm drops the file or window.

7.Volume Control:
It can be performed by joining index finger and thumb to make a pinch gesture and moving it up or down controls the volume.

8.Multiple Item Selection:
It can be performed by folding all fingers and making a fist to drag and select items.

9.Notes Maker Board:
It can be performed by joining index finger and thumb to make a pinch gesture and moving it to left or right opens the Notes Maker index finger can be used to draw and write on the screen.

3.1 SYSTEMARCHITECTURE

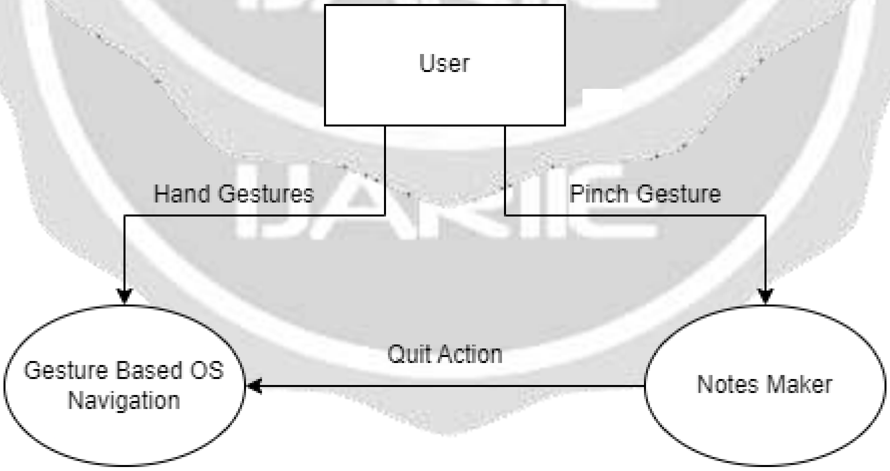

Fig-2 System architecture

Figure-2 shows the how gesture-based operating system navigation and control and Notes Maker works. It describes the how user can perform gesture actions on the screen. It also describe the how user can draw something on the Notes Maker board using pinch gesture.
.

4.CONCLUSION

This project was developed with an aim to provide an alternative way for users to interact with their computer, using gestures rather than traditional input methods such as a mouse. This technology uses sensors, such as cameras or infrared sensors, to detect and interpret specific movements of the hand and fingers. Translates these gestures into commands or actions that can be understood by the computer. Additionally, gesture-based systems and Notes Maker technology have the potential to improve accessibility for people with disabilities, as well as to enhance healthcare and industrial applications.

As these technologies continue to evolve and become more advanced, we can expect to see them integrated into more aspects of our daily lives, from smart homes and wearables to retail and entertainment experiences. However, there are also challenges to consider, such as the need to ensure these technologies are secure and reliable, as well as the need for proper training and education to ensure users can fully utilize their capabilities. Nonetheless, the future for gesture-based navigation and control and Notes Maker technology is promising, and we can expect continued innovation in these areas.

5.ACKNOWLEDGEMENT

We would like to thank Dr. Doodegowda B J for sharing his insightful observations throughout the qualitative research.

6. REFERENCES

[1] “Notes Maker: Interaction Techniques for Precision Graphics on Large Displays using Handheld Augmented Reality” by Peter Rukavina and Mark Hancock (2014).
DOI: 10.1145/2669485.2669522.

[2] “Gesture-Based Interaction for Touchless Automotive HMI” by Xinxin Zhu, Andreas Riener, and Doris Hausen (2018).
DOI: 10.1109/ACCESS.2018.2834242.

[3] “Design and Implementation of a Gesture-Based Interface for Controlling an Intelligent Home Environment” by José A. Gallud, et al. (2019).
DOI: 10.3390/s19092113.