Department of Computer Science and Engineering
St. Thomas College of Engineering and Technology
Mattannur

Virtual Hands: Real Time Keyboard, Desktop & Application Navigation using Gestures

Author:
ABHISHEK UK (STM21CS006)

Supervisor:
Dr. SHINU MATHEW JOHN

Academic Year
2024-25
October 15, 2024

OUTLINE

INTRODUCTION

- ⇛ Computer systems rely on input and output devices, with the keyboard being the most basic and essential peripheral for user interaction.
- ⇛ The frequent use of keyboards and mice, especially among IT professionals in offices and on-site locations, has raised concerns about the transmission of infections through surface contact.
- ⇛ The pandemic increased awareness of these risks, which makes it essential to explore alternatives to physical interaction with computer accessories to ensure safety in both professional and public settings.
- ⇛ With growing concerns about health and safety in both public and professional environments, there is a growing demand for innovative contactless input solutions to minimize direct interaction with potentially contaminated surfaces.

INTRODUCTION

- ⇛ This seminar introduces a contactless virtual keyboard and Windows navigation system using webcam and computer vision technology to mitigate the risk of contamination.
- ⇛ This system enables hands-free operation, making it ideal in situations where hygiene is critical, such as when handling dirty objects or working with mechanical parts.
- ⇛ By leveraging Mediapipe and OpenCV, the system accurately detects hand movements and performs actions through gestures, ensuring seamless interaction with the computer without physical contact.
- ⇛ In addition to addressing hygiene concerns, this method can be applied in various industries where hands-on interaction with keyboards and devices could compromise cleanliness or safety.

- ⇛ The system uses facial gestures, such as mouth opening and eye winking, to control mouse functions such as cursor movement, left-click, right-click.
- ⇛ Using Mediapipe, it detects facial landmarks with high accuracy, enabling precise mouse control.
- ⇛ Designed for people with disabilities, this system allows hands-free computer interaction, eliminating the need for a physical mouse.
- ⇛ The system requires only a standard webcam, making it more accessible and affordable than solutions that require additional hardware, such as Kinect or sensors.

Pros

- ⇛ Requires only a webcam
- ⇛ Precise control of mouse functions.

Cons

- ⇛ May generate false positives, leading to unintended actions.
- ⇛ Prolonged use can cause eye strain and discomfort from fixed gaze.
- ⇛ More complex operations like drag-and-drop are not yet fully implemented.

- ⇒ Uses a camera and mini projector to detect finger interactions.
- ⇒ Detects touch based on changes in the shape of projected buttons when a finger interacts.
- ⇒ A fast, robust algorithm processes button distortions to judge touch actions.

Pros

- ⇒ Requires only one camera and a mini projector.
- ⇒ Processes button distortions quickly, delivering near real-time feedback with a 96.92% touch detection rate.
- ⇒ Edge detection algorithm resists finger shadows and environmental noise, enhancing reliability.

Cons

- ⇒ Requires a flat surface for projection and touch detection, limiting use in non-planar or dynamic environments.
- ⇒ Touch detection relies on visible button distortion, which can be affected by finger shadows and changing lighting conditions.
- ⇒ Limited hover detection can lead to false positives, misidentifying hovering fingers as touch actions.

PROPOSED SYSTEM

- ⇛ The proposed system introduces a fully contactless virtual keyboard and Windows navigation solution that uses a webcam and computer vision technology.
- ⇛ The system detects hand gestures using OpenCV and MediaPipe modules, allowing users to input commands and navigate their desktops without touching physical devices.
- ⇛ This technology addresses hygiene concerns, particularly in environments where hands may be dirty or contaminated.

OpenCV

- ⇛ An open-source computer vision library designed for real-time computer vision and image processing.
- ⇛ Widely used for tasks like object detection, face recognition, and machine learning integration.

Keyboard Layout

- ⇛ The keyboard layout is drawn on each frame captured from the camera, ensuring the interface stays updated in real-time as the user interacts.
- ⇛ The *cv2.rectangle()* function draws a virtual keyboard layout directly on the camera feed, creating an interactive interface.
- ⇛ **Syntax**: cv2.rectangle(frame, start_point, end_point, color, thickness)

Figure: Keyboard layout drawn using cv2

PROPOSED SYSTEM

MediaPipe

- ⇛ MediaPipe is an open-source framework developed by Google that enables real-time machine-learning solutions for cross-platform applications.

MediaPipe Solutions

- ⇛ Prebuilt libraries for tasks such as object detection, face tracking, and hand tracking.
- ⇛ Designed for easy integration into applications with minimal setup.

MediaPipe Hands library

- ⇛ Detects 21 landmarks on the hand for real-time hand and finger tracking.
- ⇛ Detects hand gestures accurately even in challenging environments.
- ⇛ The Mediapipe detection mechanism works in two phases:
  - ⇛ BlazePalm Detector
  - ⇛ Hand Landmark Model

PROPOSED SYSTEM

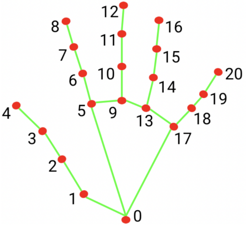
BlazePalm Detector

⇛ This first phase detects the palm of the hand using a machine learning model, which is crucial for hand tracking.

Hand Landmark Model

⇛ After detecting the palm, the hand landmark model predicts 21 hand landmarks, including their x, y coordinates and the relative depth (z).

⇛ This model precisely tracks hand movements, finger joints, and tips.



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

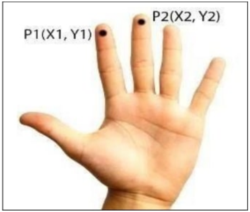
Figure: Hand Landmarks after Process

Pointing keys in the keyboard

- ⇛ Output of landmark from mediapipe is in between 0.0 to 1.0
- ⇛ Normalize the landmark to the keyboard image by,

$$X = landmark.x \times width$$

$$Y = landmark.y \times height$$



Key Press Activation

- ⇛ When distance between INDEX_FINGER_TIP and MIDDLE_FINGER_TIP is less than 30 pixels, the corresponding key is clicked.

$$Distance = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$$

Registering the key

- ⇛ Using PyAutoGUI to register the key to os
- ⇛ **Syntax:** PyAutoGUI.press(key)

PROPOSED SYSTEM

Virtual Desktop Navigation

- ⇒ Switch between created virtual desktops.
- ⇒ Squeeze hand to navigate between virtual desktops.
- ⇒ Feedback: Displays message like "Switched to Desktop 3" after switching.

Gesture-Based Navigation

- ⇒ Squeeze gesture detected when index finger distances are below threshold:
  - ⇒ 8 & 12 points: < 30 pixels
  - ⇒ 16 & 20 points: < 50 pixels

PyVDA Module for Virtual Desktops

- ⇒ Used for handling virtual desktop operations.
- ⇒ Supports app pinning, switching, and managing multiple desktops.

PROPOSED SYSTEM

Gesture-Based Application Navigation

- ⇛ Navigating between active open windows.
- ⇛ Switch between active applications using hand gestures.
- ⇛ Squeeze hand to navigate between open applications.
- ⇛ Displays the name of the application after switching.
- ⇛ Allows smooth switching between different applications.

Underlying Technology

- ⇛ Win32API: Manages window switching through process control.
- ⇛ Win32GUI: Interacts with the active window using 'hwnd' ID.
- ⇛ PyVDA: Provides smooth management of applications on the current desktop.

Figure: Key highlights as finger is placed

Figure: Window on all desktops triggered by PyVda module

RESULT

- ⇒ The system achieved 95% accuracy in detecting key presses and gestures during testing, resulting in a minimal 5% error rate mainly due to missed actions or incorrect key presses.
- ⇒ Efficient hand tracking and key registration occur in real time, with minimal latency of about 0.2 seconds between actions, ensuring a smooth user experience.
- ⇒ The system successfully switches between virtual desktops using hand gestures, providing precise control over window management.
- ⇒ The system performed effectively in cleanliness-critical environments, such as working with mechanical parts or dirty hands, demonstrating its practicality beyond pandemic-related concerns.

PERFORMANCE COMPARISON

- ⇛ The virtual keyboard provides better hygiene and safety by avoiding surface contact, although physical keyboards still excel in speed and user familiarity.
- ⇛ Unlike previous systems that require additional hardware like projectors or external devices, this system uses only a standard webcam, making it more cost-effective and easier to implement.
- ⇛ The MediaPipe BlazePalm detector and hand landmark model enable more accurate hand tracking than older methods like shadow gestures or printed keyboard surfaces, which often faced issues with positioning and fragility.
- ⇛ This system is more versatile than similar systems that require specific surface or finger positioning, as it functions effectively even when hands are not perfectly aligned with the camera.

ADVANTAGES & DISADVANTAGES

pros

- ⇛ Completely contactless, making it ideal for environments where hygiene is crucial by reducing contamination risks.
- ⇛ Requires only a webcam, which is typically built into most computers, eliminating the need for extra hardware.
- ⇛ Works across multiple platforms (Windows, Linux) using open-source tools, providing flexibility for various settings like offices, workshops, and kitchens.

cons

- ⇛ Users familiar with physical keyboards may initially find the gesture-based system slower and less intuitive.
- ⇛ The system's performance may decline in low-light conditions, requiring improvements in fingertip detection accuracy.
- ⇛ The system performs essential tasks but lacks advanced features found in physical keyboards, such as complex shortcuts and multi-key combinations.

CONCLUSION

- ⇛ The proposed virtual keyboard and desktop navigation system provides an effective way to interact with computers without touching physical devices, addressing hygiene concerns in various environments.
- ⇛ This contactless system is valuable in various scenarios where cleanliness is essential or physical interaction is undesirable.
- ⇛ The system shows strong potential but can be enhanced by improving low-light performance and expanding its capabilities for multiuser interactions.
- ⇛ With 95% accuracy and no extra hardware needed, the system provides a practical and cost-effective hands-free solution, meeting the demand for touchless technology in workplaces and public spaces.

REFERENCES

Harshit Sharma, Richa Saxena, Satendra Kumar, *“Virtual Hands: Real Time Keyboard, Desktop & Application Navigation using Gestures,”* 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP), vol. 5, pp. 262–267, Nov. 2022, doi: 10.1109/icfirtp56122.2022.10059450.

Akula Kumar Raja, Chidakash Sugandhi, Gorantla Nymish, *"Face Gesture Based Virtual Mouse Using Mediapipe,"* 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, 2023, pp. 1-6, doi: 10.1109/I2CT57861.2023.10126453.

Jun Hu, Guolin Li, Xiang Xie, Zhong Lv, and Zhihua Wang, *"Bare-fingers Touch Detection by the Button’s Distortion in a Projector–Camera System,"* in IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 4, pp. 566-575, April 2014, doi: 10.1109/TCSVT.2013.2280088.

Thank You for Your Attention!