Survey on Gesture-Based Virtual Keyboard and Mouse

Prof. S. A. Nagtilak[1], Kranti Bhosale[2], Labhesh Chaudhari[3], Prathmesh Chaudhari[4], Mansi Ingale[5]
Information Technology / Smt. Kashibai Navale College of Engineering / Savirtibai Phule Pune University/India

---***---

Abstract - Envisioning a future where the conventional mouse and keyboard inputs are supplanted by a touchless system that harnesses the power of eye tracking and hand gestures. This visionary transformation in human-computer interaction is not without its challenges, and at the heart of this pursuit lies the goal of achieving unmatched precision and user-friendliness. This innovative system is underpinned by a fusion of technology and creativity. Hand gestures, a symbol of human expressiveness, are seamlessly integrated with Haar Cascade, offering keyboard operation and mouse control. This elegant synergy empowers users to navigate digital landscapes intuitively, minimizing the barriers between humans and machines. However, this journey is marked by the quest for absolute precision and calibration. Insights from past research underscore the significance of precision in interpreting and responding to user gestures, be it the intricate movements of the hand or the subtleties of gaze tracking. Bridging this gap, it aim to make every interaction fluid and error-free, while maintaining a commitment to user-friendliness. Given methodology unfolds as an intricate dance of technology and innovation. It involves the real-time processing of hand gestures and eye tracking data, intricate data fusion in the Gesture Recognition System, and the seamless updating of the user interface. The potential applications of this technology are as diverse as they are impactful. From enhancing accessibility for individuals with disabilities to providing hygienic control of medical equipment in healthcare settings, from delivering immersive gaming experiences to facilitating interactive learning environments in education, the reach of this system extends across various domains. This research is a testament to the belief that technology should be an enabler, not an obstacle. It is a journey that involves overcoming challenges in pursuit of a user-centric experience.

Key Words: AR, CNN, CPU, IMU, VR

1. INTRODUCTION

In an ever-evolving digital landscape, where technology continues to redefine daily interactions, this system emerges as a beacon of innovation and transformation. Embark on a mission to revolutionize the very essence of human-computer interaction envisioning a future where traditional mouse and keyboard inputs are supplanted by a touchless system that harnesses the inherent power of eye tracking and hand gestures. The genesis of the system lies in a profound realization a recognition that the established paradigms of user-computer engagement have room for profound enhancement. The world is replete with technological advancements, from voice recognition to facial recognition. Yet, the simple act of manipulating a cursor or typing on a keyboard persists as an interface bottleneck. The mouse and keyboard, for all their historical relevance, often seem like artifacts of a bygone era in the context of today's technology landscape. With this realization, set forth on a journey that holds at its core the fundamental challenge of ensuring accuracy and user-friendliness in this innovative system. The aim is to transcend the limitations of traditional inputs, making the interaction between humans and machines intuitive, fluid, and devoid of the hindrances that often accompany physical devices.

At the heart of this visionary transformation lie two crucial elements: hand gestures and Haar-Cascade. The former, a universal language of human expression, becomes the conduit for keyboard operation and mouse control. It is the embodiment of natural interaction this method of empowering users to navigate the digital world as they would be the physical ones. The latter, the Haar-Cascade method, adds another layer of intelligence, providing the system with the capability to interpret gestures, furthering the user experience. However, in this relentless pursuit of innovation, recognize that precision and calibration remain essential pillars to overcome. Insights from past research underline the significance of accuracy in interpreting and responding to user gestures, be it the subtleties of hand movement or the precision of gaze tracking. This strives to bridge this gap, ensuring that every user interaction is frictionless and devoid of errors, reinforcing the commitment to a truly user-friendly system.

This methodology is a complex amalgamation of technology, creativity, and real-time processing, and represents the core mechanism by which is intend to fulfil the vision. It involves the intricate choreography of hand gesture recognition and eye tracking data processing, their seamless fusion in the Gesture Recognition System, and the dynamic update of the user interface. This symphony of processes seeks to harmonize technology with the human experience, ultimately culminating in a touchless interface that is not only responsive but also instinctive in its interactions. The applications of this technology are as diverse as they are impactful, reaching across various sectors to bring about transformative changes. From revolutionizing accessibility for individuals with disabilities to ensuring hygienic control of medical equipment in healthcare settings, this system finds relevance in diverse domains. It contributes to immersive gaming experiences and enhances interactive learning environments in education, revealing the myriad of possibilities the touchless system unlocks. This system signifies more than just a technological advancement; it embodies the spirit of a reimagined digital realm. It envisions a world where technology becomes an enabler, rather than an impediment. It advocates the empowerment of individuals, irrespective of their technical proficiency, to interact with technology seamlessly, naturally, and intuitively. This is the vision essence of a touchless input system, where they do not merely address challenges, but celebrate the boundless potential of a more intuitive digital world.

2. Body of Paper

2.1 RELATED WORKS

Samuel Solomon [1] Pattern recognition and image processing have long struggled with the problem of human face detection. The article proposes a new human face detection system based on the underlying Haar-Cascade algorithm supplemented with three additional weak classifiers. The three weak classifiers are based on matching skin hue histograms, detecting mouths, and detecting eyes. First, images of people are processed by a primitive Haar-Cascade classifier, nearly without wrong human face rejection (very low rate of false negative) but with some wrong acceptance (false positive). Secondly, a weak classifier based on face-skin hue histogram matching is employed to eliminate these incorrectly accepted non-human faces, and most non-human faces are eliminated. After that, some remaining non-human faces are identified and disqualified using a weaker classifier that is attached and relies on eye detection. Lastly, the false positive rate is further reduced by using a mouth detection procedure on the remaining non-human faces. Test results on human photos with varying occlusions, illuminations, and degrees of orientation and rotation, in both the training and test sets, using OpenCV as a tool, demonstrate the effectiveness of the suggested approach and its ability to reach state-of-the-art performance. Moreover, its ease of use and straightforwardness of execution make it efficient.

J. Shin and C. M. Kim, [2] claimed that a large number of experiments have been conducted on text input systems that use image-based hand gesture detection. Nevertheless, there are certain issues with hand gesture languages being widely used, such as finger alphabets, sign languages, and aerial handwriting discussed in the earlier studies. Writing and recognition of aerial handwriting take a lot of time. The number of people who can utilise finger alphabets and sign languages is limited because they require a significant amount of practice and education to use. This study suggests a new character input method that can be used to improve human-computer interaction. It is based on hand-tapping movements for both English and Japanese hiragana characters. The hand-tapping gestures are motions for hands to tap keys on virtual keypads in the air. Anyone, including those with hearing impairments, can utilise these gestures as an efficient way to use hand alphabets. When writing in hiragana, the consonant portion of the character and the aerial virtual keypad are determined by the hand that is used to press the key and the number of fingers that are stretched. To enter a character, simply tap the key on the virtual keypad that corresponds to the intended vowel. This employ a key layout that is akin to the English and Japanese flick keyboards seen on smartphones, so anyone can use these hand-tapping movements with just a quick explanation. With this non-touch input technology, which uses the Kinect sensor alone—no keyboard, mouse, or body-worn devices—users may interface with computers efficiently. This Model anticipate that a new avenue for human-computer connection will be opened by this character input method.

L. Cuimei, Q. Zhiliang, J. Nan and W. Jianhua stated that [3] Static or dynamic, hand gestures are a hot topic in research and have several applications in real-time systems for human-computer interaction. Basic methods of interacting with computers are through hand gestures, both static and dynamic. This work proposes an approach to the hand-gesture recognition-based text input mechanism. This portable hand-operated text input system is intended for use with virtual reality (VR) and augmented reality (AR) gadgets. A standard camera takes a picture of the hand to identify and categorise hand gestures. Following background subtraction, the hand is segmented, and the segmented hand gesture is fed into the trained neural network for gesture recognition. Lastly, a convex hull algorithm is used to track and record hand movements. A neural network that has been trained is given the matching written character. After testing the suggested architecture and comparing the experimental findings with those from other approaches, it was found that the suggested approach outperformed conventional approaches and achieved an accuracy of 96.12%, which is an improvement over current approaches overall.

In this paper, C. Lian et al. [4] presented an improved virtual keyboard design using smart rings worn on each hand. The rings have sensors that detect motion during typing. First, this model changed the keyboard layout from a rectangle to an arc shape. This increases the difference in angle between adjacent keys, improving accuracy. Second, our model used data from the sensors - acceleration, gyroscope, and magnetometer - to capture subtle differences in keystrokes. This model evaluated feature importance and correlation to select effective features that contribute highly and have low similarity. Nine features were chosen from angle and magnetometer data for keystroke recognition. Weighing recognition speed, accuracy, and selected features, it increased recognition speed nearly 4x while ensuring 98.53% accuracy. This ring-based virtual keyboard is portable, small, and lower cost than many other human-computer interfaces.

Y. Zhang, W. Yan and A. Narayanan [5] developed a brand-new virtual keyboard that enables users to text on any surface, on any kind of device. Customized and printed on simple paper, the virtual keyboard can be affixed to a wall or mounted on any oblique plane. Then, leveraging the fingertip location and hand skin tone, the device camera is employed for key recognition. The software will recognize a key as an input if the fingertip stays on it for a set period. The results of the tests demonstrate how several personalized virtual keyboards enable users to enter text without noticeable performance compromises, provided they have a physical keyboard in front of them. The total recognition rate of all inputs is 94.62% (true positives divided by all samples). Under natural illumination, keyboard (a) has the best average input recognition rate of 97.7%, whereas under lamplight, keyboard (b) has the worst average input recognition rate of 90.7%.

Pratiksha Kadam, Prof. Minal Junagre, Sakshi Khalate, Vaishnavi Jadhav, and Pragati Shewalep [6] stated that this research project aims to use computer vision to develop an optical mouse and keyboards that can be operated through hand movements. The computer camera will capture images of different hand gestures made by the user, and the mouse pointer or cursor on the computer screen will move accordingly. Different hand gestures can be used to execute right and left clicks. Similarly, the keyboard functions can be performed using different hand actions, such as using a finger to select an alphabet and a four-digit swipe left or right. The virtual mouse and keyboard can be used wirelessly or externally, and the only hardware required for the project is a webcam. The implemented system enables the control of mouse cursor movement by tracking the movement of the user's eyeballs and hand gestures. The system replaces the conventional input devices such as a mouse and keyboard by combining their functionalities. The main aim of this system is to provide a comfortable data entry method that is versatile and portable, especially for small mobile devices. The virtual mouse and keyboard system utilizes gesture recognition, cognition, and image processing to move the mouse cursor in accordance with the eyeball movement and to perform keyboard functions using hand gestures.

Dinh-Son Tran & Ngoc-Huynh Ho & Hyung-Jeong Yang1, Guee Sang Lee [7] declared that one common way to use computers without a mouse device is to use fingertip tracking as a virtual mouse. In this work, they have propose a new approach to virtual mice by utilizing fingertip detection and RGB-D images. Using detailed skeleton-joint information images from a Microsoft Kinect Sensor version 2, the hand region of interest and the palm's centre is first extracted, and the image is subsequently converted into a binary format. Next, a border-tracing algorithm extracts and describes the hand contours. Based on the hand-contour coordinates, the fingertip location is determined using the K cosine algorithm. Lastly, a virtual screen is used to control the mouse cursor by mapping the fingertip location to RGB images. Using a single CPU and Kinect V2, the system tracks fingertips in real-time at 30 frames per second on a desktop computer. The system can function well in real-world settings with a single CPU, according to the experimental results, which demonstrated a high degree of accuracy. The accuracy of fingertip detection was highest (93.25%) for the two-person group. The six-person group had the lowest accuracy rate, at 53.35%. The three, four, and five groups had accuracy rates of 89.78%, 78.03%, and 65.38%, in that order. The findings indicate that when the group size grows, accuracy decreases.

Mishaha MK, Manjusha MS [8] stated that this restriction can be addressed in the suggested virtual mouse and keyboard system by using a webcam or an integrated camera to record hand motions and coloured objects. Convolutional neural networks are the foundation of the hand gesture recognition technique (CNNs).Without a hardware mouse, the computer can be remotely operated to carry out left- and right-click scrolling volume and brightness controls, and other computer cursor functions. A system called Hand Gesture Recognition with Python can identify hand gestures in live videos. The Python-based OpenCV package is utilized to record motions from a computer's webcam or built-in camera. The laptop webcam is utilized in the proposed system to monitor coloured items that aid in its operation using the object tracking approach. Utilizing the keyboard's standard functions, such as space, enter, backspace, and others, and the object tracking system. In well-lit environments, the system performs exceptionally well, and this also solves the issue of the gesture's ability to be recognized against any background. However, because the system does not require a training phase for gesture detection and has a training accuracy of 96.37, it is slightly more responsive than other systems that have been developed previously.

2.2 PROPOSED SYSTEM

Given proposed system’s methodology is an intricate orchestration of technology and creativity, meticulously designed to bring this vision of enhanced computer interaction through eye tracking and hand gestures to life. It represents a holistic approach that traverses multiple phases, each finely tuned to deliver the desired user experience. The journey begins with the precise capture of user inputs, involving the seamless recording of both hand gestures and eye-tracking data in real-time. Hand gesture processing follows, where the intricacies of hand movements are analysed, delving into factors like position, direction, and gesture patterns, all underpinned by advanced computer vision techniques. Concurrently, eye tracking technology comes into play, capturing the user's gaze with precision and detail, and thereby providing insight into their intent.

The heart of methodology lies in the seamless integration of data from hand gesture recognition and eye tracking, allowing for a comprehensive interpretation of user input. This integrated data is then channelled into the Gesture Recognition System, a critical component that distinguishes specific user commands and control signals. Here, machine learning and pattern recognition techniques come into play,

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 01 | January - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

mapping gestures to actions, ultimately defining the user experience. This proposed methodology culminates in the real-time update of the user interface, where the system translates user input into tangible outcomes, be it visual changes, application adjustments, or operating system commands. It is at this juncture that users tangibly experience the responsiveness of the proposed system.

Concurrently, eye-tracking technology comes into play, capturing the user's gaze with precision and detail. This step provides valuable insight into the user's intent, contributing to the overall richness of the interaction. The heart of the proposed methodology lies in the seamless integration of data from hand gesture recognition and eye tracking, allowing for a comprehensive interpretation of user input. This integrated data is then channelled into the Gesture Recognition System, a critical component that distinguishes specific user commands and control signals. Here, machine learning and pattern recognition techniques play a pivotal role, mapping gestures to actions and ultimately defining the user experience. This proposed methodology culminates in the real-time update of the user interface, where the system translates user input into tangible outcomes, such as visual changes, application adjustments, or operating system commands. It is at this juncture that users tangibly experience the responsiveness of the system, particularly in the context of a Gesture-Based Virtual Keyboard and Mouse.

Throughout this methodology, an unwavering commitment to real-time feedback underpins the entire process. Users receive immediate responses to their gestures and gaze tracking, ensuring that interactions are not only fluid but also intuitive. The methodology, deeply rooted in technology and interdisciplinary collaboration, culminates in a responsive, user-centric, and innovative touchless input system. By ensuring accuracy, user-friendliness, and real-time responsiveness at its core, this methodology represents the practical realization of a vision that seeks to elevate the way our model interact with computers.

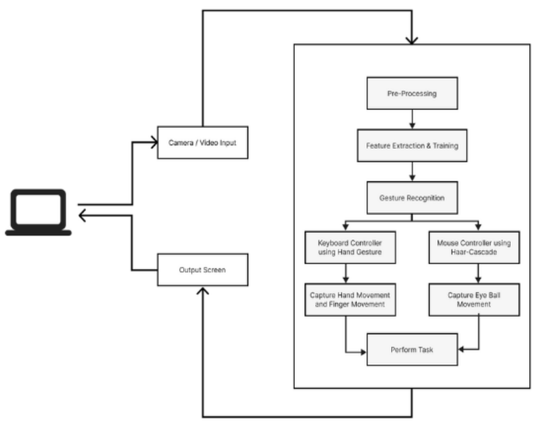

Fig. 2.1 System Architecture

2.2.1 Keyboard

a) The user is given two alternatives after logging in successfully: Virtual Mouse and Keyboard.
b) The user must choose the Virtual Keyboard option to access the virtual keyboard, which causes a keyboard to appear on the desktop screen.
c) The only characters and letters that are pre-programmed into the virtual keyboard are those that have been declared in the system's coding.
d) Using the webcam on their laptop or computer, the user only needs to point at a letter or character with their finger to type it.
e) The letter that has been selected is indicated in a small rectangular field beneath the keyboard as the user points to each letter.
f) Once the desired characters have been selected and displayed in the rectangular field, the user can either copy and paste the text into a document or use it for any other intended purpose.

2.2.2 Mouse

a) On selecting the "Virtual Mouse" option, a tiny window pops up.
b) The first step involves turning on the webcam, which enables real-time video recording of the user's face and surroundings.
c) The webcam then locates and follows the user's eye coordinates, pre-processes the recorded data, and extracts features.
d) The user's eye movement is then translated into the corresponding mouse cursor movement on the screen by using the features that were extracted.
e) By detecting an eye blink, which doubles as a mouse click, the system additionally enables the user to select any location on the screen.

2.3 FUTURE SCOPE

In the future, the privilege of a touchless input system that combines hand gestures and eye tracking has the potential to be revolutionary. Ongoing research will enhance gesture recognition precision. Integration with emerging technologies like AR and VR opens new dimensions. Accessibility will expand with customizable profiles. Healthcare applications, including remote monitoring and telemedicine, will advance. Machine learning and AI will personalize interactions. Security will benefit from biometric authentication. The system's presence in consumer electronics, gaming, education, and enterprise solutions will grow. Ongoing UX design will be user-friendliness. Real-world applications extend to kiosks, smart homes, and robotics. International collaborations will foster global adoption, forging a more interconnected and advanced digital future. Expanding into the gaming industry, the touchless input system has the potential to revolutionize gaming experiences. Its integration into gaming consoles and virtual reality setups can offer an immersive and intuitive way to interact with virtual worlds, making gaming more engaging and responsive. In the automobile industry, the system can find applications in creating touchless interfaces for in-car controls. Drivers and passengers can navigate infotainment systems, adjust climate controls, and even interact with navigation systems without physical contact, enhancing safety and convenience.

3. CONCLUSIONS

In the pursuit of redefining computer interaction, this system has brought to light the potential of eye tracking and hand gestures as touchless input systems. By seamlessly merging these technologies, our model have aimed to bridge the gap in accuracy and user-friendliness. The envisioned touchless human-computer interaction system, seamlessly integrating hand gestures and eye tracking via Haar Cascade without reliance on additional hardware, signifies a remarkable technological advancement. Striking a delicate balance between precision and user-friendliness, the innovative approach aims to eliminate barriers between humans and machines. Meticulous calibration and real-time processing are geared towards creating a seamless and error-free interaction experience. The absence of additional hardware adds to the system's advantages, offering a versatile and accessible solution. The potential applications span diverse domains, from enhancing accessibility to revolutionizing healthcare and education. This research underscores the belief that technology should serve as a facilitator, driving efforts to overcome challenges for a universally user-centric experience.

ACKNOWLEDGEMENT

We are very thankful to all the teachers who have provided us valuable guidance towards the completion of this project work on Gesture Recognition Based Virtual Mouse and Keyboard. We express our sincere gratitude towards cooperative department who has provided us with valuable assistance and requirements for the project work.

We are very grateful and want to express our thanks to Prof. S. A. Nagtilak for guiding us in right manner, correcting our doubts by giving her time whenever we required, and providing her knowledge and experience in making this project work. We are also thankful to the HOD of our Information Technology department Dr. M. L. Bangare for his moral support and motivation which has encouraged us in making this project work. We are also thankful to our Vice Principal Dr. K. R. Borole and our Principal Dr. A.V. Deshpande, who provided his constant support and motivation that made a significant contribution to the success of this project.

REFERENCES

1. J. Shin and C. M. Kim, "Non-Touch Character Input System Based on Hand Tapping Gestures Using Kinect Sensor," in IEEE Access, vol. 5, pp. 10496-10505, 2017, doi: 10.1109/ACCESS.2017.2703783. [IRJASET 2023].
2. L. Cuimei, Q. Zhiliang, J. Nan and W. Jianhua, "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers," 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), Yangzhou, China, 2017, pp. 483-487, doi: 10.1109/ICEMI.2017.8265863.
3. N. Nooruddin, R. Dembani and N. Maitlo, "HGR: Hand-Gesture-Recognition Based Text Input Method for AR/VR Wearable Devices," 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Toronto, ON, Canada, 2020, pp. 744-751, doi: 10.1109/SMC42975.2020.9283348.
4. C. Lian et al., "Towards a Virtual Keyboard Scheme Based on Wearing One Motion Sensor Ring on Each Hand," in IEEE Sensors Journal, vol. 21, no. 3, pp. 3379-3387, 1 Feb.1, 2021, doi: 10.1109/JSEN.2020.3023964.
5. [5]Y. Zhang, W. Yan and A. Narayanan ”A Virtual Keyboard Implementation Based on Finger Recognition” Auckland University of Technology, DOI: 10.1109/IVCNZ.2017.8402452ISSN: 2151-2205, vol(17)[2017 IEEE].
6. Pratiksha Kadam, Prof. Minal Junagre, Sakshi Khalate, Vaishnavi Jadhav, Pragati Shewale, “Gesture Recognition based Virtual Mouse and Keyboard” Department of Computer Engineering, Indira College of Engineering and Management, Pune, India.vol(11), ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538,[2023 IRJASET].
7. Dinh-Son Tran, Ngoc-Huynh Ho, Hyung-Jeong Yang, Guee Sang Lee, “Real-time virtual mouse system using RGB-D images and fingertip detection”, Chonnam National University,
8. Multimedia Tools and Applications 80(7):1-18, DOI:10.1007/s11042-020-10156-5 (Vol 8)[2020 Springer].
9. Mishaha MK, Manjusha MS PG Scholar,” Virtual Mouse And Keyboard For Computer Interaction“, Dept. Computer Science & Engineering, M.Dasan Institute Technology Ulliyeri,
10. Koyilandi, Kozhikoode,673620, DOI:10.46647/ijetms.2022.v06i04.0036 ISSN: 2581-4621 ,vol(6)[2022 IJETMS].
11. Eichelberger, C., Wagner, D., & Schmalstieg, D. "Real-time detection of hand and finger poses in a head-mounted display for 3D interaction.", Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2(4), 168. [IMWUT 2018.].
12. Zhang, X., Sugano, Y., Fritz, M., & Bulling, A., "Gaze estimation using deep learning and its applications in driver monitoring." In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 4117-4126.[2015]
13. Rahim, M. S. I., & Molla, M. M. M., "A survey of hand pose recognition with Microsoft Kinect.", Computer Vision and Image Understanding, 131, 18-39.,[2015].
14. Zhang, W., & Liu, P., "A survey of vision-based hand gesture recognition.", In Proceedings of the 2017 IEEE International

Conference on Robotics and Automation (ICRA), pp. 5135-5142.[2017].
15. López, S. M., Lozano, M., Sappa, A. D., & Sánchez, A., "Evaluation of a Vision-Based Interface by Tracking Eyes and Hands in an Indoor Environment.", Sensors, 19(15), 3290.[2019].
16. Castelló, A. B., Alvarez, A. L., Moreno, A. L., & Naya, F. "Hand gesture recognition using a Myo armband and a Thalmic Labs software library.", In Proceedings of the International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS), pp. 139-150[2015].