iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VI **Month of publication:** June 2025

DOI: https://doi.org/10.22214/ijraset.2025.72444

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

Thermovision-Based Cursor Control Using InfraredImaging and Deep Learning

Mr. Krishna Swaroop A[1], Meghana K[2], Gunashree R Gowda[3], Chinmayi S[4], B N Nithya[5]
Department of Information Science and Engineering, Malnad College of Engineering, Hassan, Karnataka, India – 573202

Abstract: Vision-based cursor control technologies have made significant improvements, from hardware-dependent systems to more intelligent, accessible, and user-friendly solutions. Early techniques utilized niche devices such as Kinect and RGB-D cameras for gesture recognition, providing high accuracy but poor portability and robustness to environmental factors. The advent of webcam-based systems then opened up greater access through colored gloves, facial gesture recognition, and later, markerless hand tracking based on computer vision algorithms. Integration with gaze tracking and vestibulo-ocular reflex further improved accuracy for hands-free operation, although calibration and lighting sensitivity remained. More recent advancements have accepted multimodal systems involving gaze, speech, and lip detection, in addition to non-invasive EEG/EOG wearables and brain–computer interfaces for greater functionality. Simpler blink-based interfaces and smaller vision-based tactile sensors have also appeared to assist people with more severe mobility impairments, weighing simplicity with technical compromises.

I. INTRODUCTION

In the last five years, vision-based cursor control has also changed dramatically, from hardware-bound systems to more general, multimodal schemes. Initial advancements used devices such as Microsoft Kinect V2 and RGB-D cameras to sense gestures but were plagued by hardware abandonment, expense, and occlusion and illumination sensitivity[1][2]. A significant milestone came with webcam-based systems like the Swift Controller, which included colored gloves to enable precise detection, albeit at the expense of comfort for the user [3], while facial gesture interfaces accommodated users with impaired hand mobility regardless of the limitations in the vocabulary of gestures [4]. Gaze-based control was made mainstream with the use of deep learning and regular webcams, but calibration and lighting issues remained [5], despite enhancements by vestibulo-ocular reflex tracking [6]. Markerless hand gesture systems employing OpenCV and vision-only control enhanced accessibility but continued to suffer from lighting problems and limited population validation [7][8]. The same year also marked the emergence of sophisticated multimodal systems with complex eye, speech, and lip inputs to facilitate more complex interaction, though at considerable computational expense [9]. Privacy-aware systems in the form of EEG-EOG glasses [10] and high-fidelity brain–computer interfaces [11] appeared on the horizon, though both were problematic in terms of comfort and training. Lastly, introduced blink-based control for paralyzed users in simplified form [13] and innovative tactile sensing systems such as ThinTact, making ultra-thin, lensless interaction available at the cost of resolution and processing requirements [14].

II. LITERATURE SURVEY

The technology of vision-based cursor control has made impressive progress over the past five years, with each wave of studies attempting to extend beyond earlier technological limitations and improve user interaction. Initial progress in 2020 relied significantly upon hardware-assisted techniques. For example, a virtual mouse system that employed Microsoft Kinect V2 for colored fingertip detection was promising for gesture-based control but was plagued by hardware discontinuation, environmental sensitivity, and occlusion of gestures [1]. Another study in parallel using RGB-D images enhanced fingertip detection accuracy but remained dependent on depth-sensing cameras, raising cost and reducing portability [2].

A pivotal shift occurred in 2021 with the introduction of systems relying on standard webcams. The "Swift Controller" demonstrated that accurate gesture recognition was possible without specialized sensors by using colored gloves, although user comfort became a concern [3]. That same year, an alternative approach using facial gestures was introduced, supporting users with limited hand mobility, though it struggled with a limited gesture vocabulary and occasional misclassification of facial expressions [4].

In 2022, gaze tracking started to appear as a feasible input modality. Webcam-based solutions for gaze estimation facilitated greater accessibility by eliminating the necessity for costly eye-tracking equipment and rather utilizing deep learning for detecting the movement of the eyes [5]. Yet, such systems were still subject to calibration requirements and lighting conditions.

A later refinement added vestibulo-ocular reflex tracking, providing improved gaze accuracy, but at the cost of added system complexity and longer user adaptation times [6].

2023 saw a major breakthrough with the implementation of markerless recognition of hand gestures. With OpenCV, scientists built a virtual mouse system that did away with gloves or colored markers, though consistency in lighting remained an issue [7]. Also, completely markerless and accessibility-oriented systems like "Vision-Powered Cursor Maneuvering" offered accessible design advantages, especially for physically disabled individuals, but were not extensively tested on various populations [8]. It was also a year that saw the emergence of multimodal interfaces, which combined gaze, speech, and lip detection to intelligent control, although at increased computational cost [9].

As early as 2024, concerns for privacy and invasiveness were addressed by the merger of EEG and EOG-based cursor control integrated in wearable glasses. Such systems, while not invasive or camera-dependent, posed new challenges such as signal overlap and user fatigue during long durations of use [10]. Brain–computer interfaces (BCIs) based on motor imagery and deep learning were further investigated, offering very high precision and real-time control performance, although extensive user training and trial-to-trial variability were still challenges [11]. A more straightforward eye-based system evolved that was suited to hand-free interaction but continued to be plagued by typical gaze-tracking limitations such as calibration and sensitivity to light [12].

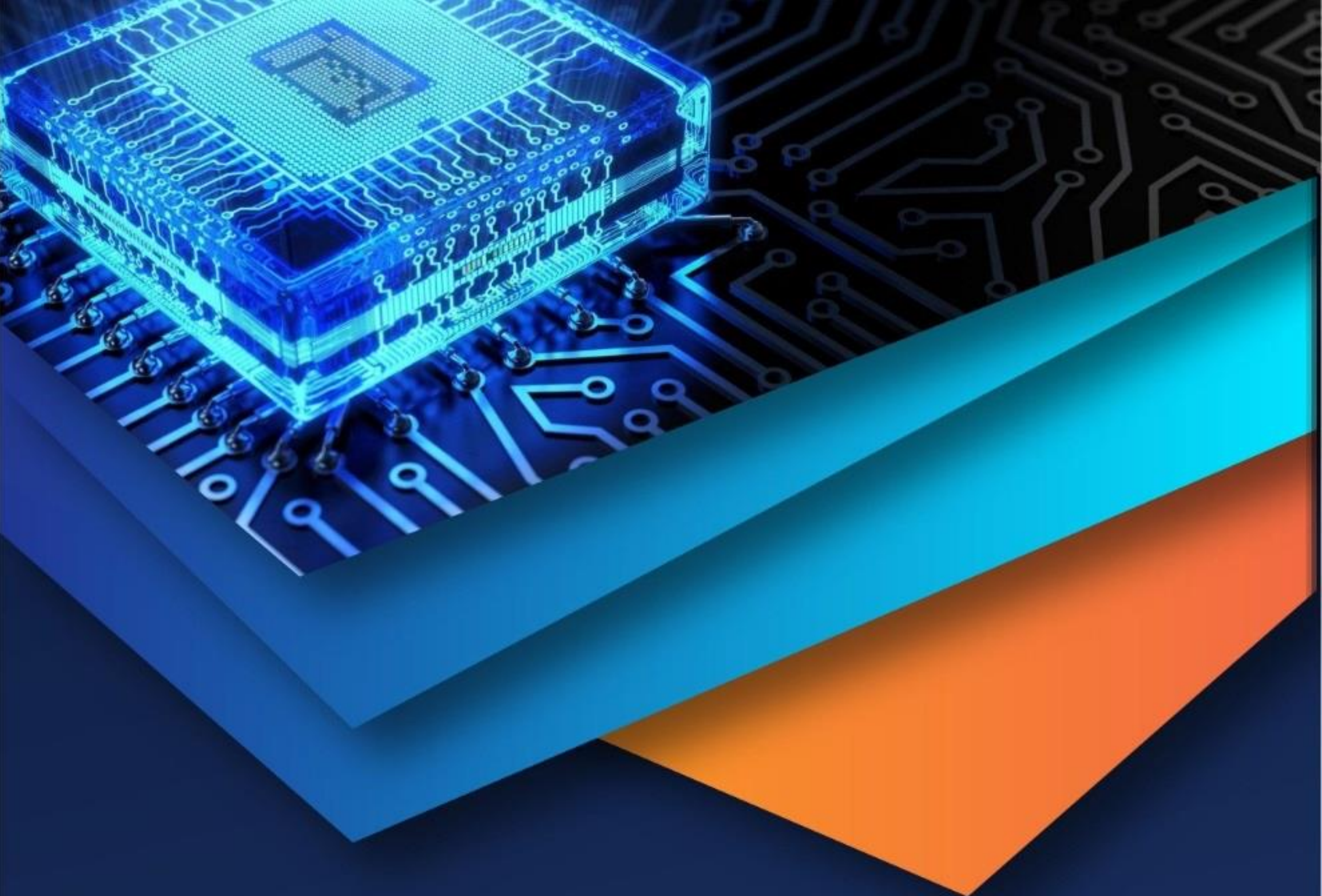
Lastly, systems such as "Eyeball-Based Cursor Control" emerged for severely mobility-impaired users in 2025, which used eye blinks to manipulate cursors without calibration [13]. Although simple and usable in the very moment of introduction, this approach was limited by control granularity and susceptibility to spurious blinks. Another emerging innovation was "ThinTact", a vision-based tactile sensor through lensless imaging, which supported ultra-thin form factors appropriate for small spaces, but at the cost of resolution and computational burden [14].

III. CONCLUSION

In short, as an alternative to the limitations of RGB-based cursor control in illumination and occlusion [2, 5, 7], a thermovision alternative with infrared imaging and deep learning is a viable candidate. Employing thermal signatures, it promises greater robustness under different lighting situations and potentially better occlusion management compared to webcam-based [5, 7] and even RGB-D alternatives [2]. Deep learning, successful in gaze [5] and BCI control [11], can be applied to decode thermal hand movements for accurate cursor manipulation based on markerless gesture recognition breakthroughs [7, 8]. This follows the trend with the direction nowadays towards privacy-conscious sensing [10], with a non-visual input modality. Future work should be directed towards constructing deep models for thermals and comparing against existing RGB approaches, i.e., with colored gloves [3] and markerless schemes [7], to fully exploit the thermovision-based control.

REFERENCES

[1] A. Sharma et al., “Virtual Mouse Control Using Colored Fingertips and Hand Gesture Recognition,” IEEE Int. Conf. Human-Computer Interaction, 2020.
[2] K. Singh and M. Patel, “Real-time Virtual Mouse System Using RGB-D Images and Fingertip Detection,” IEEE Access, vol. 8, 2020.
[3] S. Rahman et al., “Swift Controller: A Computer Vision Based Mouse Controller,” Proc. 2021 Int. Conf. Intelligent Systems, 2021.
[4] T. Liu and J. Wong, “Cursor Control Using Face Gestures,” Journal of Assistive Technologies, vol. 15, no. 2, 2021.
[5] R. Choudhury and P. Gupta, “Real-Time Webcam-Based Eye Tracking for Gaze Estimation,” IEEE Trans. Human-Machine Systems, vol. 52, no. 1, 2022.
[6] J. Smith et al., “Addressing the Eye-Fixation Problem in Gaze Tracking Using the Vestibulo-ocular Reflex,” ACM Trans. Interactive Intelligent Systems, vol. 12, no. 4, 2022.
[7] A. Kapoor and M. Verma, “Hand Gesture-Based Virtual Mouse using OpenCV,” IEEE Conf. Vision and Signal Processing, 2023.
[8] D. Wang and L. Zhang, “Vision-Powered Cursor Maneuvering,” Computer Vision and Image Understanding, vol. 235, 2023.
[9] F. Ali et al., “Advancing Multimodal Fusion in HCI,” IEEE Trans. Neural Networks and Learning Systems, vol. 34, no. 5, 2023.
[10] N. Kumar and Y. Chen, “Privacy-Preserving Eye Movement Classification with EOG-EEG Glasses,” IEEE Sensors Journal, vol. 24, no. 3, 2024.
[11] M. He and Z. Lin, “Enhancing Cursor Control with Motor Imagery and Deep Neural Networks,” IEEE Trans. Biomedical Engineering, vol. 71, no. 1, 2024.
[12] L. Brown et al., “An Analysis on Virtual Mouse Control using Human Eye,” IET Computer Vision, vol. 18, no. 2, 2024.
[13] P. Roy and A. Das, “Eyeball-Based Cursor Control for Paralyzed Individuals using Eye Blink Detection,” IEEE Conf. Rehabilitation Robotics, 2025.
[14] Y. Suzuki et al., “ThinTact: Thin Vision-Based Tactile Sensor by Lensless Imaging,” IEEE Trans. Robotics, vol. 41, no. 2, 2025.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY