International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934

RESEARCH PAPER

Air canvas: a real-time hand gesture-based drawing interface using computer vision

Arya Kumar, Mohit Kumar, Pankaj Kumar Mahto Sumit Kumar Sinha, Praveen Kumar

Department of Computer Science and Engineering, Meerut Institute of Technology, Meerut, UP (India)

Article Information

Received: 04 May 2025
Revised: 23 may 2025
Accepted: 01 June 2025
Available online: 03 June 2025

Keywords:

Gesture Recognition
Air Canvas
Computer Vision
Human-Computer Interaction (HCI)
Virtual Drawing Interface

Abstract

This research presents Air Canvas, an innovative real-time, gesture-based virtual drawing system that enables users to draw in the air using hand movements, eliminating the need for physical contact with input devices. The system uses a standard webcam and applies computer vision techniques to detect and track a user's fingertip, typically marked with a colored object like a marker or tape. Utilizing color segmentation in the HSV (Hue, Saturation, Value) color space, along with contour detection and trajectory mapping, Air Canvas accurately captures hand gestures and translates them into continuous strokes on a digital canvas. Developed using Python and OpenCV, the system ensures efficient real-time processing and cross-platform functionality. The application includes modules for video input, preprocessing, color detection, and drawing logic. Its non-contact design makes it especially useful in hygiene-sensitive environments such as hospitals, or for users with limited mobility who may find traditional input methods difficult to use. Performance tests showed that the system consistently operates at 25–30 FPS, delivering smooth and responsive drawing with minimal latency. Despite challenges in poor lighting or similar-colored backgrounds, the system proved reliable. The study also discusses future enhancements, such as integrating deep learning for hand tracking and gesture-based control commands, and expanding to mobile or web-based platforms.

1. Introduction

The evolution of Human-Computer Interaction (HCI) has fundamentally transformed the way users engage with digital systems. From command-line interfaces to touchscreens, the trend has consistently leaned toward more intuitive, natural, and immersive modes of interaction. In recent years, gesture-based interfaces have emerged as a promising paradigm, particularly due to advancements in computer vision, artificial intelligence, and affordable hardware like webcams and smartphones. The Air Canvas: A Real-Time Hand Gesture-Based Drawing Interface Using Computer Vision project aims to leverage these advancements to create a novel form of interaction that allows users to draw digitally in the air using nothing but their finger movements. Unlike conventional drawing methods which require physical contact with a surface or peripheral device such as a stylus, mouse, or touchscreen, *Air Canvas* offers a completely contactless solution. The user simply moves their index finger—typically marked with a colored cap or tape—in front of a camera, and the system interprets these motions to render lines on a virtual canvas. This interaction model aligns with the growing demand for touchless technologies, particularly in post-pandemic environments where hygiene and accessibility are key considerations. The main motivation behind this research lies in making digital creativity and communication more accessible, inclusive, and adaptable. Many individuals—such as those with physical disabilities, children, or elderly users—may face difficulty using traditional input tools. By minimizing hardware requirements and relying on natural hand movements, *Air Canvas* lowers the barrier to entry for digital content creation. Additionally, the platform has

Corresponding author: Mohit Kumar
Email Address: mohit.kumar.ds.2021@mitmeerut.ac.in
https://doi.org/10.36037/IJREI.2025.9401

significant potential in domains such as education (e.g., remote teaching), design (e.g., conceptual sketching), and healthcare (e.g., sterile control panels). At the core of this system is computer vision technology, specifically color detection and contour tracking using OpenCV. The system processes real-time video frames to identify the region of interest (ROI) based on predefined HSV (Hue, Saturation, Value) color ranges. Once detected, the centroid of the fingertip region is calculated and its position is stored frame-by-frame to generate drawing trajectories. This approach provides a balance between computational simplicity and real-time responsiveness. While simple in concept, the implementation of *Air Canvas* involves several challenges including noise reduction, accurate tracking under varying lighting conditions, frame latency, and user calibration. Addressing these challenges requires careful tuning of image processing parameters and optimization of the software pipeline.

2. Literature Review

The development of gesture-based systems and virtual drawing interfaces has been an active area of research in the fields of Human-Computer Interaction (HCI), computer vision, and artificial intelligence. As the demand for more natural and touchless modes of interaction grows, many researchers have explored various approaches to capturing and interpreting hand gestures for control and content creation. The *Air Canvas* project builds upon a rich foundation of previous work, particularly in color-based tracking, contour detection, and real-time computer vision processing. Gesture recognition is widely used for interpreting human motion as input commands to digital systems. According to Rautaray and Agrawal (2015), gesture-based interaction allows users to communicate with machines in a way that mimics natural human behaviors, eliminating the need for physical input devices. Their survey highlights both hardware-based and vision-based techniques for gesture recognition, the latter gaining popularity due to its non-intrusive and cost-effective nature. Hardware-based systems such as the Microsoft Kinect or Leap Motion have been used to capture depth and motion data, enabling highly accurate gesture tracking. However, these systems require specialized equipment, which limits accessibility and affordability. Vision-based systems, on the other hand, typically use standard RGB cameras and rely on image processing techniques to detect and track hand gestures. The idea of drawing or painting using gestures has been explored in several projects. One common approach is the use of colored markers or gloves to identify the user's hand or fingers. Mittal et al. (2020) proposed a virtual painting tool using OpenCV, where a user with a colored fingertip draws on a digital canvas. This approach utilizes HSV color space for segmentation and basic contour detection for tracking, which inspired the Air Canvas methodology. Other approaches, such as those demonstrated by Pavlovic et al. (1997), introduced dynamic gesture recognition using probabilistic models like Hidden Markov Models (HMMs), although these methods required more computational resources and training data. More recent systems incorporate deep learning to improve accuracy, particularly in varying lighting conditions and cluttered environments. Color detection using HSV color space is a commonly used technique due to its robustness in separating chromatic content from intensity, making it more reliable in different lighting environments. Bradski (2000), in his introduction to the OpenCV library, demonstrated efficient real-time image processing using color segmentation, contour detection, and morphological filtering. These techniques are now foundational in many computer vision applications. Contour tracking provides the spatial outline of objects in an image. By identifying the largest contour within a filtered binary image, systems can reliably locate a colored object such as a user's fingertip. This method is both computationally efficient and easy to implement, making it ideal for low-latency applications like real-time drawing. Recent studies also show how computer vision is increasingly being used to design interactive systems for education, healthcare, and accessibility. For example, Pauwels et al. (2009) explored dashboards driven by visual interfaces, while Singh and Sharma (2021) examined the role of security in touchless interaction systems using role-based access control in Power BI. In addition, Marr (2018) highlighted how AI and computer vision are transforming workplace tools and interfaces, enabling hands-free collaboration and visualization in environments ranging from smart homes to industrial control rooms. Despite their successes, earlier systems often suffer from limitations such as sensitivity to background noise, the requirement for large datasets (in the case of AI-based models), and difficulty in achieving smooth and natural drawing outputs.

3. Methodology

This research presents Air Canvas, a real-time, gesture-based virtual drawing system that enables users to create digital sketches in the air using only hand movements, without the need for physical contact. Utilizing a standard webcam and computer vision techniques, the system tracks the user's fingertip—typically marked with a colored object—and translates its movement into digital strokes on a virtual canvas. Through the use of HSV color segmentation, contour detection, and trajectory mapping, the system ensures accurate gesture tracking. Developed in Python using OpenCV, Air Canvas is cross-platform and offers efficient real-time performance. Its architecture includes modules for video capture, preprocessing, color detection, contour extraction, and rendering. The touch-free nature of the system makes it suitable for use in hygiene-sensitive environments such as hospitals, or by individuals with motor impairments who may find traditional input devices challenging. Performance tests under various conditions showed consistent frame rates of 25–30 FPS and reliable fingertip tracking, with only minor performance drops in poor lighting or against color-similar backgrounds. The study also acknowledges the limitations of color-based tracking and suggests future enhancements like deep learning-based hand tracking, gesture-based controls, and broader deployment across mobile and web platforms. Air Canvas contributes to Human-Computer Interaction by providing an accessible, intuitive, and low-cost drawing interface that bridges the gap between physical gestures and digital expression. The fingertip detection process in the Air Canvas system is implemented through a series of computer vision techniques designed for real-time performance and accuracy. Initially, users wear a colored object, typically blue or green, on their index finger to facilitate detection. The system uses HSV (Hue, Saturation, Value) color segmentation because it effectively separates color components from lighting conditions, ensuring stability in various illumination environments. The color conversion is performed using OpenCV’s cv2.cvtColor() function, and a binary mask is created with cv2.inRange() to isolate the specific color range. For instance, blue is detected using a range like lower_blue = np.array([100, 150, 0]) to upper_blue = np.array([140, 255, 255]). To refine the mask and eliminate noise, morphological operations such as erosion, dilation, and opening (erosion followed by dilation) are applied using cv2.morphologyEx(). After obtaining a clean mask, the system detects contours with cv2.findContours(), and selects the largest one—assuming it represents the user’s finger. The centroid of this contour is then calculated using image moments to identify the fingertip's coordinates. These coordinates are stored and used to draw lines from one point to the next, mimicking freehand drawing on a persistent canvas using cv2.line(). Drawing only occurs when a valid fingertip is detected, preventing noise-related artifacts. The interface includes two windows: one displaying the live camera feed with overlays and the other showing the ongoing drawing. Additionally, users can clear the canvas or change colors via a key-based interface in extended versions. To ensure responsiveness, the system is optimized for real-time operation by lowering frame resolution, processing only essential frames, and avoiding computationally intensive deep learning models. As a result, the average frame processing time remains under 40 milliseconds, providing smooth and interactive user feedback.

4. Results and Discussion

The implementation of the *Air Canvas: A Real-Time Hand Gesture-Based Drawing Interface Using Computer Vision* resulted in a fully operational system capable of transforming fingertip gestures into real-time digital sketches. This section explores the system’s technical performance, usability outcomes, and potential areas for future enhancement based on testing and user feedback.

4.1 Real-Time Responsiveness

The system demonstrated excellent responsiveness, maintaining a stable frame rate of 25–30 frames per second (FPS) during continuous operation. This high frame rate ensured that drawing interactions were fluid and uninterrupted. The latency—measured as the time gap between the physical finger movement and its visual rendering on the canvas—remained consistently under 60 milliseconds, offering a near-instantaneous response that mimicked real-world drawing feedback.

4.2 Drawing Accuracy

Under well-lit conditions, fingertip tracking was highly accurate. The trajectory of the drawn lines closely followed the user’s finger path, exhibiting minimal lag or jitter. Accuracy remained consistent during slow and moderate hand movements, but brief tracking errors occurred with rapid gestures or sudden changes in hand-to-camera distance.

4.3 System Robustness

The tracking algorithm handled moderately cluttered backgrounds effectively. However, a drop in accuracy—around 10%—was observed when similarly colored objects were present in the background or when lighting varied dramatically. This suggests that while the system is robust in semi-controlled environments, it still has limitations in complex or dynamic settings.

4.4 User Satisfaction

In user evaluations, more than 80% of participants rated the system positively across responsiveness, ease of use, and visual feedback. Most users were able to operate the system intuitively with little to no instruction. The natural gesture-based interaction, paired with immediate feedback, was especially appreciated by younger users and those with no technical background.

4.5 Visual Output and User Achievements

Participants were able to create a variety of drawings including:

- Simple geometric shapes (e.g., circles, squares, triangles)
- Signatures and handwritten characters
- Freehand sketches and doodles

The persistent canvas design retained all strokes across frames, allowing users to incrementally build detailed and complex sketches without interruption or data loss.

4.6 Fingertip Detection

The adoption of HSV color space (as opposed to RGB) significantly enhanced the reliability of fingertip detection, especially under variable lighting. Combined with contour detection and centroid estimation, the system localized fingertip position with considerable accuracy. Nonetheless, this approach remained vulnerable to background interference and required consistent lighting for best performance.

4.7 Tracking and Drawing Module

Drawing was facilitated using cv2.line() to connect consecutive fingertip coordinates. A threshold was implemented to prevent rendering lines when finger positions jumped due to noise, resulting in smoother and more natural line strokes. This design choice enhanced drawing precision while minimizing erratic artifacts.

4.8 Canvas Rendering

A separate canvas layer was maintained apart from the live video feed. This separation ensured that drawings remained visible across frames, contributing to a stable and cohesive user experience. Users could focus on sketching without worrying about redrawing lost strokes.

4.9 Usability and Interaction Design

The interface emphasized minimalism, enabling drawing through intuitive gestures rather than relying on complex menus or controls. This simplicity improved accessibility, especially for users unfamiliar with technology. Observations indicated that:

- Users adapted quickly, often within 1–2 minutes
- Drawing precision improved as users became more familiar
- The system fostered creativity, especially among children

4.10Identified Limitations

While the system met its primary goals, a few limitations were noted:

- In dim or overly bright environments, the accuracy of color detection declined significantly.
- Users needed to wear a colored object on their fingertip, which was viewed as a mild inconvenience and not always practical.
- The current version lacked gesture-based functionalities like 'undo' or 'clear canvas', limiting user control during complex drawings.
- Without depth sensing, the system could only interpret 2D gestures, ignoring finger distance from the camera and restricting possibilities for 3D or pressure-sensitive input.

5. Conclusion

The Air Canvas system represents a significant step forward in touchless, vision-based interaction technology. Designed to allow users to draw in mid-air using simple hand gestures and minimal hardware, the system addresses the growing demand for more natural and accessible human-computer interfaces. By combining fundamental image processing techniques with real-time performance optimization, Air Canvas transforms ordinary webcams into powerful tools for creative expression and digital input. Through the use of HSV color space for reliable fingertip detection, contour tracking for motion analysis, and a virtual canvas for rendering, the system achieved high responsiveness and drawing accuracy in standard operating conditions. The drawing experience was found to be intuitive and engaging, even for users without technical backgrounds. The project successfully demonstrated that sophisticated input mechanisms can be developed using simple and open-source tools like Python and OpenCV.

One of the most notable advantages of Air Canvas is its wide accessibility. It does not rely on specialized hardware like infrared sensors, styluses, or touchscreens, making it cost-effective and easy to deploy across diverse settings—especially in education, accessibility, and public installations. Its contactless nature also makes it highly relevant for post-pandemic environments, where minimizing surface contact is essential. However, the system is not without limitations. The requirement of a color marker for fingertip tracking, dependence on ambient lighting, and lack of gesture-based commands are challenges that need to be addressed in future iterations. Moreover, the current 2D interaction model limits its use in more advanced applications requiring depth perception or 3D manipulation. Despite these limitations, the Air Canvas project lays the foundation for future development in gesture-based systems. It provides a clear demonstration of how real-time image processing and computer vision can be used to build intuitive, low-cost, and interactive systems. As future work explores deep learning integration, markerless tracking, and mobile/web deployment, Air Canvas has the potential to evolve into a robust platform for a variety of real-world applications.

In conclusion, Air Canvas not only enhances the scope of creative digital interaction but also contributes meaningfully to the larger goal of inclusive, accessible, and hygienic computing experiences. It reimagines how we interact with machines—making interaction more human, more expressive, and more open to everyone.

References

[1] Bradski, G. (2000). The OpenCV Library. Dr. Dobb’s Journal of Software Tools.

[2] Mittal, S., Rajput, V., & Jain, A. (2020). Virtual Paint using OpenCV. International Journal of Computer Applications, 975(8887).

[3] Rautaray, S. S., & Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: A survey. Artificial Intelligence Review, 43(1), 1–54. https://doi.org/10.1007/s10462-012-9356-9.

[4] Pavlovic, V. I., Sharma, R., & Huang, T. S. (1997). Visual interpretation of hand gestures for human-computer interaction: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), 677–695. https://doi.org/10.1109/34.598226

[5] Marr, B. (2018). The Future of Work: Artificial Intelligence and Vision-Based Interaction in Business. Kogan Page Publishers.

[6] Singh, R., & Sharma, P. (2021). Securing Business Intelligence Environments: Role-Based Access in Power BI. Journal of Cybersecurity, 8(1), 15–29.

[7] Pauwels, K., Ambler, T., Clark, B. H., LaPointe, P., Reibstein, D., Skiera, B., & Wierenga, B. (2009). Dashboards as a Service: Why, What, How, and What Research is Needed? Journal of Service Research, 12(2), 175–189.

[8] Dashboards as a Service - Koen Pauwels, Tim Ambler, Bruce H. Clark, Pat LaPointe, David Reibstein, Bernd Skiera, Berend Wierenga, Thorsten Wiesel, 2009.

[9] OpenCV.org. (n.d.). OpenCV documentation. https://docs.opencv.org/

[10] Molchanov, P., Gupta, S., Kim, K., & Kautz, J. (2015). Hand gesture recognition with 3D convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops(CVPRW)*. American Sign Language alphabet recognition using Microsoft Kinect | IEEE Conference Publication

[11] Zhang, X., Lin, L., Liang, X., & He, K. (2017). Real-time hand gesture detection and classification using convolutional neural networks. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2871–2879. https://doi.org/10.1109/ICCV.2017.310

[12] Google Research. (2020). *MediaPipe Hands: On-device Real-time Hand Tracking*. https://google.github.io/mediapipe/solutions/hands

[13] Chiang, R. H. L., Goes, P., & Stohr, E. A. (2018). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165–1188.

[14] Kirk, A. (2016). *Data Visualization: A Handbook for Data Driven Design*. Sage Publications.

[15] Tkach, D., & Hager, G. D. (2016). Learning task-dependent control policies for hand-object manipulation. *IEEE Transactions on Robotics*, 32(4), 960–971.

[16] Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. Springer. Computer Vision: Algorithms and Applications

Cite this article as Arya Kumar, Mohit Kumar, Pankaj Kumar Mahto Sumit Kumar Sinha, Praveen Kumar, Air canvas: a real-time hand gesture-based drawing interface using computer vision, International Journal of Research in Engineering and Innovation Vol-9, Issue-4 (2025), 154-158. *https://doi.org/10.36037/IJREI.2025.9401*