The Use of Hand Gestures for Drawing and Control in the Context of Presentations

Abderrafii Abdou and Arthur Oliveira de Rosso

January 2025

1 Introduction

In modern education and professional settings, presentations have become indispensable tools for communication and learning. However, many educators and presenters face limitations in effectively engaging their audiences when traditional tools like whiteboards or chalkboards are unavailable. This absence often hinders the ability to annotate slides, draw diagrams, or emphasize specific points in real time, reducing the interactive and dynamic nature of presentations. Addressing this gap calls for innovative solutions that combine convenience and interactivity.

Imagine a university professor entering a classroom, prepared to discuss key concepts before an important test, only to find no chalk or whiteboard—just a screen projector. In such scenarios, the inability to spontaneously write or draw could compromise the delivery of critical information. To overcome this, we propose a gesture-based solution that enables users to write or draw directly on slides using hand gestures, such as those facilitated by a wearable glove, and controlling slides using simple gestures (pointing toward the wanted direction). This approach eliminates reliance on traditional tools while maintaining the spontaneity of real-time interaction.

Our proposed system offers transformative potential by introducing a new mode of interaction. It allows presenters to control slides easily and annotate them dynamically, fostering engagement and enabling students to focus on specific points as they are being discussed. Moreover, this system provides a digital record of annotations, making it easier for students to revisit key insights without the need for extensive manual note-taking.

By removing constraints such as physical boards, markers, or chalk, and integrating gesture-based controls seamlessly into presentations, this solution not only enhances flexibility, but also bridges the gap between traditional and modern teaching methods. It redefines how presentations can support real-time interaction and long-term learning.

2 Design

2.1 State of the art

Hand gestures have been extensively explored as a natural interface for human-computer interaction. In the context of presentations, the integration of gestures for drawing can enhance communication by allowing presenters to express ideas dynamically. Researchers have explored various technologies and methodologies to integrate gesture-based drawing into presentations, with mixed results. The following is a review of the literature, key approaches, systems, and their effectiveness.

2.1.1 Approaches for Gesture-Based Drawing

Several approaches have been explored to enable gesture-based drawing during presentations:

Computer Vision-Based Recognition Real-time hand gesture recognition leverages cameras and advanced techniques like deep learning-based pose estimation to identify hand positions and interpret gestures. Systems using tools such as OpenCV detect hand contours for interpreting drawing gestures as shown in the paper by Orovwode et al. [1], while libraries like Mediapipe [2] accurately track hand landmarks. This approach is non-intrusive, requiring no additional hardware, and its scalability benefits from advancements in machine vision technology. However, it faces challenges such as dependency on consistent lighting and backgrounds, along with the computational demands of real-time processing.

Sensor-Based Recognition Wearable devices, such as sensor-equipped gloves or controllers, are used to track hand movements and gestures with high precision. Examples include devices like Leap Motion [3] (which was used in this project), the Myo armband, and smart gloves embedded with accelerometers and gyroscopes. These systems are highly effective in diverse environments without relying on camera input. However, they come with challenges, including the cost and potential inconvenience of additional hardware, as well as the need for calibration and user adaptation.

Hybrid Approaches Hybrid systems combine vision-based technologies with sensors to enhance the accuracy and robustness of hand gesture recognition. For example, setups integrating depth sensors like Microsoft Kinect with hand-tracking libraries improve 3D gesture recognition. These systems leverage the strengths of both approaches and adapt well to varying environmental conditions, but they also involve higher implementation complexity and increased costs. Despite these challenges, we adopted this hybrid approach for our solution, as it provided the precision and adaptability required to deliver an optimal user experience.

2.1.2 Implemented Systems

Several systems have been designed and tested for gesture-based drawing:

Interactive Whiteboards with Gesture Recognition Systems like Gesture-Based Interactive Whiteboards leverage vision or sensor-based tools to allow presenters to draw and annotate in real-time.

Virtual Reality (VR) and Augmented Reality (AR) Gesture-enabled VR/AR platforms, such as Google Tilt Brush or Microsoft HoloLens, offer immersive ways to draw in 3D spaces, making presentations more engaging.

Smart Projectors Devices like the Sony Xperia Touch and interactive projectors enable presenters to use hand gestures to draw directly on projected screens.

Eye-Tracking and Gaze-Control Systems Devices like Tobii Eye Tracker enable users to control slides or interact with presentation elements using gaze direction.

2.1.3 Successes and Limitations

Gesture-based drawing systems have achieved significant success, enhancing audience engagement by making presentations more interactive and dynamic. Natural gestures, such as pinch-to-draw or swipe-to-erase, align with intuitive human behaviors, simplifying their use and providing a seamless experience. Additionally, these systems promote accessibility by lowering barriers for presenters who may lack proficiency with traditional drawing or digital tools, making them more inclusive and user-friendly.

However, there are notable limitations. Vision-based systems often face accuracy challenges in low-light or cluttered environments, while real-time gesture recognition can introduce latency, disrupting presentation flow. Some systems require users to learn specific gestures, which may not feel intuitive, posing a learning curve. Moreover, sensor-based recognition systems can be costly, limiting their availability to a broader audience.

Gesture-based drawing in presentations represents a promising intersection of technology and communication. While existing approaches and systems have made strides, challenges remain in achieving seamless, accurate, and cost-effective implementations.

2.2 Proposed Solution

To address the challenge of enabling hand gesture-based drawing in presentations, we focused on a combination of sensor-based recognition and real-time hand gesture recognition, an approach that remains relatively underexplored in research but holds significant potential for precision and usability (It is worthy to note that research was done to study each type of recognition individually). We selected the Leap Motion Controller paired with TouchFree software as a tool for sensor-based recognition. TouchFree operates invisibly atop existing user interfaces, offering a cursor controlled through hand gestures and integrated with Windows' input system. This setup allows for quick and intuitive control without requiring modifications to existing software. On the other hand, we deployed a web-based real-time hand gesture recognition system using Tensorflow.js and Fingerpose.

To extend functionality, we designed a ReactJs based web application that displays PDF-formatted slides in the middle of the screen, as PDFs are universally supported by slide applications and websites. The interface features controls and buttons above and below the slides. On the top, there is an input field where the users can insert their PDF-formatted presentations, a color selector to add a sense of personalization, a reset button to go back to the initial presentation (this is mainly for testing purposes), a pair of undo-redo buttons, a button that allows user keyboard input (for precise text annotations.), a button that lets users draw with hand movements, limiting input to a single attempt per activation to reduce errors, next there is a button that lets users download their modified files, and finally, on the bottom, there is a pair of navigation buttons, however slide navigation can also be done using hand gestures by pointing to the direction where the user wants the slides to move.

This user-centered design ensures ease of interaction and ergonomic stability. For the camera setup, we opted for a position below the screen, as it provides optimal stability and ease of use. This position works seamlessly with the Leap Motion Controller, which supports an interaction range of 10 to 60 cm (up to 80 cm maximum). The interaction mechanism is straightforward: users move their index finger (or any other finger, or even their entire hand) to control the cursor. To draw, they bring their finger closer to the screen, triggering a cursor shape change that indicates drawing mode. Once finished, moving the finger away from the screen exits the drawing mode. This intuitive system leverages natural gestures, ensuring an accessible and effective solution for enhancing presentations.

3 Experiments

The evaluation of this idea was conducted using a structured questionnaire designed for students engaging in this novel interaction method. The NASA Task Load Index (NASA-TLX) [4] form was utilized to assess participants' perceived workload and interaction experience. The evaluation process was structured into three key phases: Introduction and Training, Task Execution, and Post-Task Assessment.

3.1 Evaluation plan

Introduction and Training Participants were first introduced to the system, including its objectives and core functionalities. This introductory session provided an essential overview to ensure a clear understanding of the system's purpose. Subsequently, participants underwent a brief hands-on training session to familiarize themselves with the interface and the mechanics of interaction. This preparation phase aimed to minimize initial user errors and enhance the reliability of the subsequent evaluation.

Task Execution Participants were then required to complete a series of predefined tasks that simulated real-world presentation scenarios. These tasks were carefully designed to evaluate the system's usability and participants' ability to perform key functions effectively. The tasks included:

1. Exploring the webpage, moving their hands, and observing the cursor's response to familiarize themselves with the system's controls and interaction flow.
2. Navigating through presentation slides using hand gestures (pointing).
3. Activating the drawing mode and annotating slides using gestures.
4. Drawing shapes such as circles, arrows, or underlines in various colors.
5. Utilizing the different functionalities of the webpage, such as loading a different presentation and downloading the modified presentation.

These tasks provided a comprehensive assessment of the interaction capabilities of the system under different conditions, reflecting its potential practical application.

Post-Task Assessment After completing the tasks, participants were asked to fill out the NASA-TLX survey. This standardized tool allowed participants to rate their perceived workload across six dimensions:

- **Mental Demand:** The cognitive effort required to complete the tasks.
- **Physical Demand:** The extent of physical effort exerted during the interaction.
- **Temporal Demand:** The pressure experienced due to the task's pace or time constraints.
- **Performance:** Participants' self-assessment of how successfully they completed the tasks.
- **Effort:** The overall level of exertion required to accomplish the tasks.
- **Frustration:** The level of irritation or discomfort experienced throughout the process.

This comprehensive evaluation approach ensured a detailed understanding of both the strengths and areas for improvement in the system, aligning with user-centered design principles.

3.2 Results

During the experiments, we noticed some intrinsic problems with this type of interaction, which are: gorilla arms [5], fat finger [6], and target occlusion [7]

Gorilla Arm This phenomenon refers to the physical fatigue and discomfort users experience when interacting with systems that require sustained or repetitive mid-air gestures, such as those enabled by touchless interfaces or vertical touchscreens. This issue arises because holding arms outstretched for extended periods places significant strain on the shoulder and arm muscles, leading to user fatigue. In our implementation, this problem becomes particularly relevant. For example, a presenter drawing for 10 minutes may experience noticeable discomfort and fatigue, impacting both the quality of interaction and overall user satisfaction.

Fat Finger This problem refers to the difficulty users encounter when interacting with touch screens or small interfaces due to the relatively large size of their fingers in relation to tightly packed interactive elements, such as buttons or icons. This can lead to accidental inputs, such as selecting the wrong option or triggering unintended actions, resulting in user frustration and decreased accuracy. In this context, the fat finger issue takes on a unique form. When the touchable screen is calibrated to match the real screen size, users may experience the problem, where the cursor will not be seen because of their finger, making it challenging to determine precisely where the drawing occurs. This lack of precision can lead to unintentional marks or difficulty achieving the desired accuracy.

Target Occlusion It refers to the phenomenon where a user's hand, finger, or input device blocks their view of the target or interactive element during interaction, reducing precision and hindering task performance. This issue is particularly relevant in touchscreens or stylus-operated interfaces, where accurate visual feedback is crucial for effective interaction. In our system, target occlusion presents unique challenges. When users attempt to draw, their hand or even their arm can obstruct the view of the display, making it difficult to see the area being drawn or to verify the accuracy of the input. This obstruction can slow down task execution, increase user frustration, and negatively impact the overall quality of the interaction.

NASA TLX Besides that, the results of the NASA TLX assessment corroborated the insights into various aspects of the interaction, including user satisfaction, device accuracy, quality of drawing, and ease of use. Here's an interpretation based on the provided data:

The relatively low scores for Mental Demand indicate that users found the system cognitively undemanding, which is a positive sign for user satisfaction. However, the high Frustration score suggests significant dissatisfaction, possibly due to challenges in other areas. This mixed feedback highlights the need to address specific pain points to improve overall user satisfaction.

The Performance scores suggest users were not fully satisfied with their ability to complete tasks successfully. Combined with preferences such as selecting Physical Demand over Performance, it appears that accuracy issues may have required users to exert more physical effort to compensate for the system's limitations. This indicates a need to enhance the system's accuracy to reduce reliance on user effort.

The high Physical Demand scores point to a physically strenuous interaction, possibly linked to the quality of drawing. Poor precision or responsiveness in drawing tasks may have required users to exert excessive effort, leading to fatigue. Furthermore, the Effort scores reinforce this observation, suggesting users had to work hard to achieve satisfactory results. Improvements in the drawing interface, such as better gesture recognition and finding an optimized distance from the screen, could enhance the quality of drawing and reduce physical strain.

The variability in Temporal Demand scores indicates that while some users found the pace manageable, others felt pressured, potentially due to interface inefficiencies. The repeated selection of Effort over Temporal Demand and Physical Demand over Temporal Demand suggests that users prioritized managing physical and cognitive strain over timing concerns. These results underscore the importance of optimizing the interface for intuitive use, reducing physical effort, and minimizing frustration.

4 Conclusion and Future Directions

Gesture-based drawing in presentations represents a promising intersection of technology and communication. While existing approaches and systems have made strides, challenges remain in achieving seamless, accurate, and cost-effective implementations. Analyzing the results from the NASA TLX reveal important insights into the usability and user experience of the gesture-based drawing system using a Leap Motion controller and gesture recognition. The assessment highlighted significant challenges in terms of physical demand, where users reported high levels of strain due to prolonged arm movements, indicative of the "gorilla arm" problem. This issue became particularly noticeable during extended drawing sessions, suggesting that ergonomic considerations need to be prioritized to reduce physical fatigue.

Additionally, the fat finger and target occlusion problems were evident, as users struggled with precision when their fingers or hands obscured the drawing area, leading to frustration and errors. These problems underline the importance of interface optimization, such as enhancing visual feedback and improving the calibration of touch targets, to ensure a more intuitive and accurate interaction experience.

The mixed responses to effort and performance further emphasized the need for a balance between system efficiency and user comfort. While users were able to complete tasks, the high levels of effort reported suggest that the system demands more physical and cognitive input than desired. To enhance user satisfaction and performance, future improvements should focus on reducing physical strain, improving gesture recognition accuracy, better drawing functionality (offering users the option to choose between a single-stroke drawing method or a more fluid, continuous drawing experience), and minimizing occlusion. Addressing these issues will not only improve the overall usability of the system but also contribute to a more comfortable and effective gesture-based interaction experience.

Future work could focus on:

- Integrating hand gestures for all of the functions in our system, allowing users to control every feature through intuitive, gesture-based interactions for a more seamless and interactive experience.
- Adding more functionalities to the system, such as an eraser that is launched and controlled using a specific hand gesture.
- Integrating AI for adaptive gesture recognition in varying contexts.
- Reducing hardware dependence through enhanced vision-based solutions (reducing the overall cost of the system).
- Improving the learning curve for our system by providing comprehensive user guides, interactive tutorials, and detailed demonstrations.

This field holds significant potential to revolutionize how we communicate and collaborate, especially in educational, professional, and creative domains.

Our system's code and resources are available on GitHub: https://github.com/abderafii/HCI

References

[1] H. Orovwode, J. A. Abubakar, O. C. Gaius, and A. Abdullkareem, "The use of hand gestures as a tool for presentation.," *International Journal of Advanced Computer Science & Applications*, vol. 14, no. 11, 2023.

[2] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. Yong, J. Lee, *et al.*, "Mediapipe: A framework for perceiving and processing reality," in *Third workshop on computer vision for AR/VR at IEEE computer vision and pattern recognition (CVPR)*, vol. 2019, 2019.

[3] W. Lu, Z. Tong, and J. Chu, "Dynamic hand gesture recognition with leap motion controller," *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1188–1192, 2016.

[4] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research, 1988," *Advances in Human Psychology: Human Mental Workload. Elsevier Science*, 1988.

[5] J. T. Hansberger, C. Peng, S. L. Mathis, V. A. Shanthakumar, S. C. Meacham, L. Cao, and V. R. Blakely, "Dispelling the gorilla arm syndrome: the viability of prolonged gesture interactions," in *Virtual, Augmented and Mixed Reality: 9th International Conference, VAMR 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings 9*, pp. 505–520, Springer, 2017.

[6] P. Huber, "Inaccurate input on touch devices relating to the fingertip," *Media Informatics Proseminar on Interactive Surfaces*, vol. 31, 2015.

[7] M. Li, Y. Zhang, Y. Jia, and Y. Yang, "Advancing multi-object tracking through occlusion-awareness and trajectory optimization," *Knowledge-Based Systems*, vol. 310, p. 112930, 2025.