Allegheny College
Computer And Information Science Department

Senior Thesis

A Movement-Based Digital Painting Interface for Embodied Artistic Expression

by

Chi Pham

ALLEGHENY COLLEGE

DEPARTMENT OF COMPUTER AND INFORMATION SCIENCE

Project Supervisor: **Emily Graber**
Co-Supervisor: **Byron Rich**

Abstract

This thesis investigates how full-body movement can be materialized into digital brushstrokes through computer vision, proposing Gesturism—a browser-based digital painting interface that leverages real-time body tracking as a creative tool. Rooted in theories of embodied interaction, gestural communication, and computational media, this research critiques the constraints of conventional digital painting tools that limit creative expression to the hand and wrist. Instead, it offers a novel system that integrates full-body kinesthetic data with customizable brush settings, thereby enabling artists, performers, and non-artists alike to draw with the entire body. The project synthesizes methodologies from human-computer interaction, creative coding, and movement studies. A user study (N=14) evaluated the system's usability and expressiveness across diverse demographics, including dancers, athletes, and digital painters. The data reveals differentiated modes of interaction: some users approached the tool as a space for choreographic improvisation, while others prioritized visual precision. These findings affirm the value of movement not only as input but as material for design—dynamic, relational, and expressive. Beyond prototyping, this thesis contributes a conceptual and technical framework for designing with gesture as a medium, examining how movement data can be both structurally legible and aesthetically open-ended. An accompanying exhibition and film documented how viewers engaged with Gesturism in situ, revealing social, cultural, and perceptual layers of participation. The project foregrounds co-participation, affective feedback, and embodied authorship—offering an alternative path forward for digital art systems grounded in the whole body.

Table of contents

1 Introduction **7**
- 1.1 Background and Motivation . . . 7
- 1.2 Computer Vision in Interactive Art . . . 8
- 1.3 Purpose of the Study . . . 8
  - 1.3.1 Novel Properties . . . 9
- 1.4 Ethical Considerations . . . 10

2 Related work **11**
- 2.1 Computer Vision . . . 11
  - 2.1.1 Computer Vision Architecture . . . 11
- 2.2 Conceptual Frameworks: Artists working with Gesture and Vision . . . 15
  - 2.2.1 Golan Levin . . . 15
  - 2.2.2 Zachary Lieberman . . . 17
- 2.3 Digital Painting . . . 19
- 2.4 Embodied Interaction and Creativity . . . 22
  - 2.4.1 A framework of key design features of a motion-sensing drawing tool. 23
- 2.5 Computational Motion-sensing Drawing Application . . . 25
- 2.6 Materializing Movement for Digital Interaction . . . 26
  - 2.6.1 Understanding movement data . . . 28

3 Method of approach **31**
- 3.1 Computer Vision Module . . . 31
  - 3.1.1 Comparision of TensorFlow PoseNet and MediaPipe BlazePose . . . 31
  - 3.1.2 Implementation and Justification . . . 31
  - 3.1.3 Compatibility Considerations . . . 32
- 3.2 Environment Setup . . . 32
- 3.3 Body Part Integration . . . 34
  - 3.3.1 Getting started with Mediapipe and Error Handling . . . 34
  - 3.3.2 The Mediapipe Model . . . 36
  - 3.3.3 Full Body Integration . . . 37
  - 3.3.4 Debugging and Landmark Visualization . . . 37
  - 3.3.5 The Customization Menu . . . 38
  - 3.3.6 File functionality . . . 42
- 3.4 Gesture-Based Control (Under Development) . . . 42

4 Experiments **46**
- 4.1 Experiment Method . . . 46
- 4.2 Study Procedure . . . 46
- 4.3 Gesturism User Experience Questionnaire . . . 56
  - 4.3.1 Consent and Demographics . . . 56
  - 4.3.2 Prompt Completion Confirmation . . . 56
  - 4.3.3 Usability and Interface Interactions . . . 57
  - 4.3.4 Customization and Enjoyment . . . 57
  - 4.3.5 System Responsiveness and Suggestions . . . 58
- 4.4 Data Collection . . . 58
- 4.5 Recruitement . . . 59
- 4.6 Inclusion Requirments . . . 59

4.7 Evaluation 60
4.7.1 Participants 60
4.7.2 Results 62

5 Threats to Validity **67**

6 Conclusion **68**
6.1 Conclusion 68
6.2 Visual Design 68
6.3 Exhibition 70
6.4 Future Work 74
6.4.1 Advancing UI Design 74
6.4.2 Enhancing Technical Accuracy 75
6.4.3 Expanding Exhibition Potential 75
6.4.4 Back-End Implementation 75
6.5 Ethical Considerations 76

List of Figures

1 Gesturism Proposed Functionalities 10
2 Fundamental steps of a computer vision for digital image processing (Source: Vergara, Osslan & Cruz-Sánchez, Vianey & Ochoa, Humberto & Nandayapa, Manuel & Flores Abad, Angel. (2014). Automatic Product Quality Inspection Using Computer Vision Systems. 10.1007/978-3-319-04951-9_7.) 12
3 Steps involved in visual feature extraction (Source: Varshney, Priyanka & Farooq, Omar & Upadhyaya, Prashant. (2014). Hindi viseme recognition using subspace DCT features. International Journal of Applied Pattern Recognition. 1. 257. 10.1504/IJAPR.2014.065768.) 13
4 Feature extraction using computer vision algorithms (Source: Sophia, Gino & Sharmila, Ceronmani. (2020). Computer vision algorithms for dominant contact lens feature extraction using fuzzy-logic-based classifications. Soft Computing. 24. 10.1007/s00500-020-04791-1.) 14
5 Yellowtail (1998) 16
6 2003 | Tmema (Golan Levin and Zachary Lieberman) with Jaap Blonk and Joan La Barbara 17
7 Lieberman's Reflection Studies| Exploratorium Museum Exhibit 19
8 I/O Brush 2004 (MIT Tangible Media Group - Kimiko Ryokai, Stefan Marti and Professor Hiroshi Ishii) 21
9 Flowers and People, Cannot be Controlled but Live Together 22
10 Carolee Schneemann, Meat Joy, Whitechapel Gallery during SHOP I 23
11 One of Frank Gilbreth's light painting (1914) 25
12 Metaplay communication 26
13 Bill T. Jones '21' performance 27
14 `Manifesto` performance 28
15 Each x-y-z point can be visualized by Sync as one or several or all marks . . 30
16 MediaPipe Pose Module: 33 Body Landmarks 32
17 Gesturism Tech Stack 33
18 Mediapipe integration in Gesturism 36
19 The complete flow of Gesturism 37
20 The pose tracking status 38
21 The color palette implemented in Gesturism 40
22 An abstract flower sketch created by researcher (March 2025) 41
23 Comparison between brushstrokes 4px and 20px 41
24 Tracking Status Active 43
25 Tracking Status Paused 44
26 Sample of saved artwork filename 59
27 User Experience Distribution in Digital Painting 61
28 Familiarity with Digital Painting Interfaces 62
29 Prompt 1. Draw a Green Circle with your RIGHT WRIST on the RIGHT SIDE of the canvas 63
30 Prompt 2. Draw a BLUE CIRCLE with your LEFT INDEX on theTOP of the canvas 64
31 Prompt 3. Draw a RED CIRCLE with your RIGHT KNEE with 10PX STROKE WIDTH 64
32 FREEFROM RESPONSES. PAINT A SCENE WITH A RAINBOW 65

33 Gesturism logo breakdown 68
34 Color Palette 69
35 Exhibition setup 70
36 The documentation 71
37 New proposed user interface 75

List of Tables

1 Frameworks for key deisgn festures of a motion-sensing drawing tool 24
2 A schema for identifying semantic properties of movement dynamics 28
3 A proposed framework for mapping gesture vocabulary to UI functionality . . 44
4 Table of participator goals, strategies to achieve them and implementation in the User Research Study 46

1 Introduction

In the intersection between art and technology, digital media tools have reshaped how we understand and perform the act of painting. From traditional brushwork to advanced algorithmic generation, the practice of visual art now spans a continuum of material and immaterial, static and dynamic, analog and computational. This research positions itself within this continuum, asking how we might reclaim the body as a central actor in the creation of digital artworl. It centers on a key proposition: movement is not merely a means to manipulate a cursor or stylus, but a primary expressive medium in itself.

To explore this proposition, this thesis introduces *Gesturism*, a full-body gesture-based digital painting system that captures, translates, and visualize human movement using computer vision. The project draws from interdisciplinary domains of computer vision, interaction design, digital painting software, performance art, and embodied cognition.

1.1 Background and Motivation

The integration of new media and digital technologies has simplified the painting process, while emphasizing the unique qualities that define painting as an art form.[20] With digital tools, artists can explore novel ideas, experiment with diverse visual techniques, and expand the boundaries of their creative expression. In addition, incorporating digital media technologies into painting has led to considerable breakthroughs in the creative methods and expressions used in paintings. Today's artists have access to various digital resources, including 2D and 3D programs, film special effects, and animation production. These all offer the technical support necessary to create digital artworks that mimic the painting process. These advancements have broadened the aesthetic possibilities and opened up new ways for realizing that promise.

However, common contemporary digital drawing input devices, such as mouse and tablet, only allow users to draw with limited body movement, mainly just finger or wrist movement in a sitting posture.This emphasis aligns with a long-standing tradition of seated, stationary art creation, neglecting the role that full-body motion can play in artistic expression. Most digital drawing designs seem to neglect the fact that body movement can help reveal or articulate one's inner feelings, which is essential for creativity to flourish. [27] This research begins from a counter-position: that movement, in its broadest sense, is a form of thinking, feeling, and making.

Research on embodied creativity underscores the critical relationship between physical movement and cognitive processes, particularly in enhancing creativity and fostering emotional engagement. An expanding body of experimental studies reveals the potential benefits of unstructured, interactive, or spontaneous motions—such as gestures, dance, shifting body postures, physical object manipulation, and drawing—on creative performance.[13] Research suggests that fluid movement, such as freely navigating a space, can stimulate the generation of original ideas during brainstorming sessions.[14] In daily life, we use our bodies to navigate, negotiate, and communicate nonverbally. We adapt our posture, the way we move our limbs, and how we distribute our weight based on the environment, the people we interact with, and our intentions. Goffman (1959) characterizes these adjustments in posture, gestures, and gaze as part of a social performance. Today, designers can tap into movement data through widely accessible sensors, such as the iPhone or Kinect. Moreover, the open-source community has democratized software development with tools like openFrameworks and Processing, enabling further exploration. However, despite these

technological advancements, few resources exist within interaction design that meaningfully engage with full-body movement data. This gap presents an opportunity to explore how technology and interaction design can evolve to incorporate the whole body, moving beyond the conventional use of fingers swiping on screens.

1.2 Computer Vision in Interactive Art

This study addresses that gap by focusing on computer vision (CV), a subset of artificial intelligence that enables machines to understand visual input. CV technologies are now ubiquitous in fields ranging from autonomous vehicles to facial recognition, but they also present profound implications for creative practice. In particular, CV allows us to transform physical movement into digital information, which can then be visualized, analyzed, or manipulated in real time.

Figure 1 illustrates how CV systems typically operate: through a pipeline of image acquisition, feature extraction, and pattern recognition, often implemented using convolution neural networks. These systems detect deges, track landmarks, segment objects, or infer motion to interpret the visual world.

One of the pioneering works in this domain is Myron Kruger's Videoplace, widely regarded as the first interactive artwork to incorporate computer vision. Developed between 1969 and 1975, Videoplace was motivated by Kruger's belief that the entire human body should play a central role in our interaction with computers. The installation invited participants to stand before a backlit wall facing a projection screen, where their silhouette was digitized and analyzed in real time. In response, the system generated interactive graphics, such as "critters" that climbed up the participant's projected form or colored loops that connected their fingers.[17] Through Videoplace, Kruger demonstrated the potential of computer vision not just as a technological tool but as a medium for artistic expression and human-computer symbiosis. His work laid the foundation for subsequent generations of interactive art, where the body itself became a vital component of the creative process.

Historically, the implementation of computer vision was restricted to experts in signal processing and artificial intelligence, with application largely confined to military and law enforcement contexts. Today, interactive-media artists and student programmers are increasingly incorporating computer vision into their creative practices, leading to a surge of innovation in art, games, and other interactive systems.[19] Open-source libraries like OpenCV, frameworks like TensorFlow, and tools such as OpenPose have played a crucial role in the exploration of new modalities of interaction. These tools allow artists to track gestures, facial expressions, and full-body motion, facilitating the development of artworks that respond to users' physical presence in increasingly nuanced ways.

1.3 Purpose of the Study

This project explores how movement can be approached as a fundamental element in interaction design, specifically investigating how movement data can be leveraged for digital interaction. To design with movement as a medium, it is essential to understand the properties and particularities of such data as a dynamic and expressive design material.

The purpose of this study is to develop and evaluate a digital painting interface that utilizes computer vision technology to incorporate full-body movement into the creative process. This approach addressed the limitations of conventional digital painting tools, which often constrain users to localized hand and wrist movements. By integrating concepts

from the field of expressive movement and motion-sensing technology, this study introduces a whole-body, kinesthetic painting prototype, named Gesturism, offering performers and artists an alternative medium for creative expression through motion.

To support the creative needs of performers and artists, the research also explores advanced digital brush stroke stylization and customizable interface settings, aiming to align with industry-standard tools such as Adobe Sketch and Procreate. The study places equal emphasis on technical feasibility and artistic innovation, ensuring the prototype meets both functional and expressive criteria.

There are three research questions this research aims to answer: 1. What considerations must be made for gesture-based web application? 2. How do we find the balance between tracking accuracy and real-time performance in translating movement into digital brushstrokes?

1.3.1 Novel Properties

The researcher emphasized the following novel properties of this kinaesthetic drawing system:

1. **Dual Mode Functionality**: The system incorporates two distinct modes tailored to different user needs: The Performer Mode, which emphasizes the quality of movement, and the Painter Mode, which focuses on the resulting artwork. This dual-mode design enables users to engage with the system either as performers, exploring the expressive potential of their movements, or as painters, concentrating on the final visual output.
2. **Customizability and Versatility**: The system offers extensive customization options, allowing users to adjust the tracking system and map control to various body parts. This flexibility enables a broad range of user scenarios, supporting diverse creative expressions both in terms of the process (movement) and the product (the final artwork).
3. **Enhanced Control and Creative Expression**: The system prioritizes a high degree of control over brush stylization, providing users with a range of options to tailor the digital brush to their preferences. This increased control enhances users' ability to creatively express themselves, ultimately leading to the production of high-quality artwork.

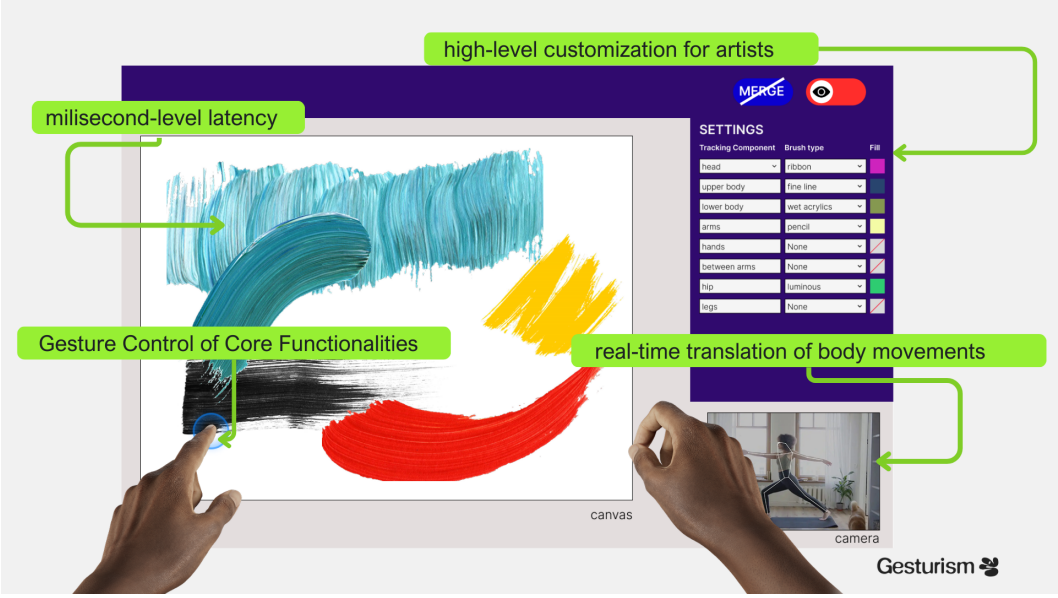


Figure 1: Gesturism Proposed Functionalities

1.4 Ethical Considerations

This project prioritizes user privacy and data protection by ensuring that no personal data or body movement information is stored or retained beyond the duration of the interaction. The system does not collect or store any user data on external servers, and all movement data used for interaction is processed in real-time, with no long-term storage. Users will only be able to export their artwork and associated data to their local machines, ensuring full control over their personal content. The system's design avoids any tracking of user identities, and no data is shared with third parties.

2 Related work

2.1 Computer Vision

Computer vision (CV), as a subfield of artificial intelligence, refers to the computation process of intepreting understanding visual information from the world. This capabilitiy has become increasingly important to interactive systems development. Computer vision aims to replicate and extend human visual perception using mathematical models, machine learning, and image processing techniques. The practical application of CV technologies in real-time systems has expanded considerably with the advacement of hardware capabilities, enabling responsive installations that detect, analyze, and react to human movement with increasing accuracy and granularity [25]. At its core, computer vision operates through a pipeline that typically involves image acquisition, preprocessiong, feature extraction, and decision making. Key low-level operations include edge detection, color segmentation, and optical flow, while higher-level tasks include object detection, pose estimation, and scene understanding.

Modern CV systems leverage machine-learning- particularly deep learning- to learn complex representations directly from raw data. CNN (convolution neural networks) have proven hightly effective in learnign spatial hierachies of features, making them central to tasks like facial recognition and human pose estimation. Libraries such as Open CV, first introduced by Bradski (2000) have made these tools widely accessible to developers. Open CV provides optimzed implemtation of essential algoritms including template matching, background subtraction and Haar cascade classifiers [11]. More recently, frameworks like TensorFlow and PyTorch have facilitated the training and deployment of deep learning modesl on consumer hardware, allowing for real-time performance in interactive installations.

2.1.1 Computer Vision Architecture

Computer vision systems follow a sequential processing model that transforms raw pixel data into semantically meaningful outputs (demonstrated in Figure 1). This begins with image acquisition, the process by which sensors capture visual input from the physical world. Typically, this is performed using CCD or CMOS cameras that digitize light into image data, yielding frames that represent scenes in 2D pixel arrays. The quality of this acquisition is contigent upon several variables including resolution, frame rate, lighting conditions, and color representation models such as RGB or HSV.

Once an image has been captured, the stage of preprocessing begins to standardize and enhance it for analysis. Preprocessing operations such as noise filtering, histogram equalization, and edge enhancement mitigate environmental inconsistencies and prepare the image for feature extraction. These transformations ensure that critical structures within the image can be detected with great reliability. For real-time system, this technique helps mitigate visual clutter and variable lighting with sigificant noise. [26]

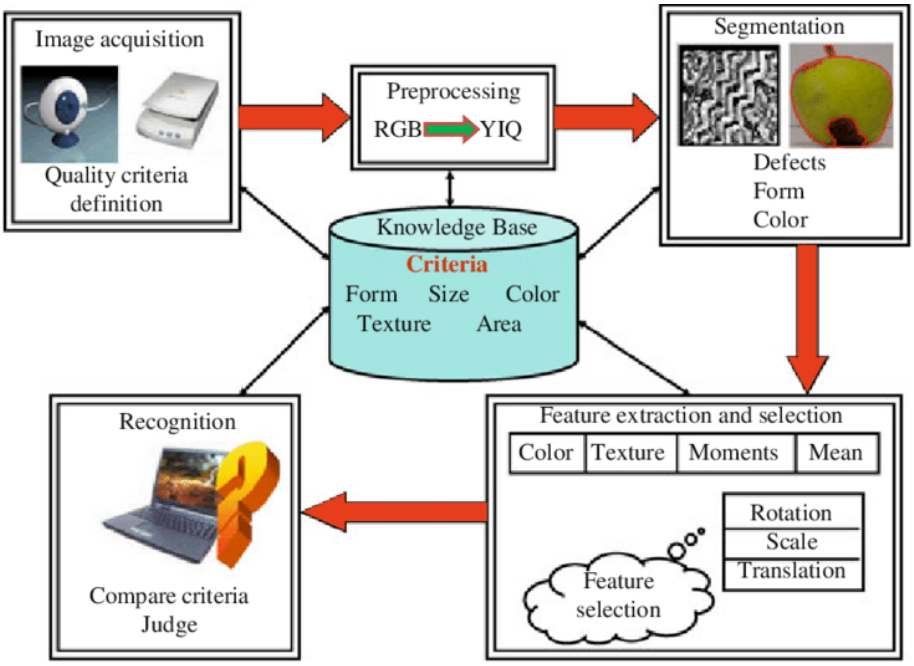


Figure 2: Fundamental steps of a computer vision for digital image processing (Source: Vergara, Osslan & Cruz-Sánchez, Vianey & Ochoa, Humberto & Nandayapa, Manuel & Flores Abad, Angel. (2014). Automatic Product Quality Inspection Using Computer Vision Systems. 10.1007/978-3-319-04951-9_7.)

Following preprocessing, the system enters the phase of feature extraction (Figure 2). Here, it identifies and isolates salient patterns such as edges, corners, or blobs. Feature detectors like Sobel operators for edges, or Harris and Shi-Tomasi methods for corners, provide localized representations of image structures. More robust descriptors such as SIFT (Scale-Invariant Feature Transform) and SURF (Speeded-Up Robust Features) enable recognition across variations in scale and orientation. These extracted features became the foundation for higher-level interpretation tasks such as object detection or gesture classification. [24]

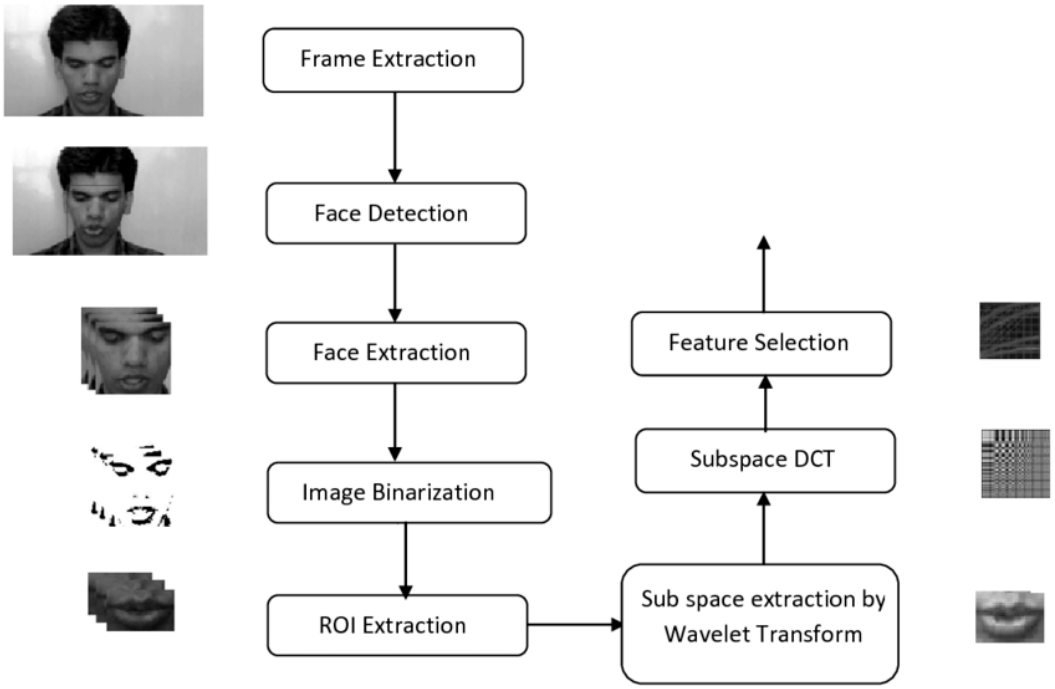


Figure 3: Steps involved in visual feature extraction (Source: Varshney, Priyanka & Farooq, Omar & Upadhyaya, Prashant. (2014). Hindi viseme recognition using subspace DCT features. International Journal of Applied Pattern Recognition. 1. 257. 10.1504/IJAPR.2014.065768.)

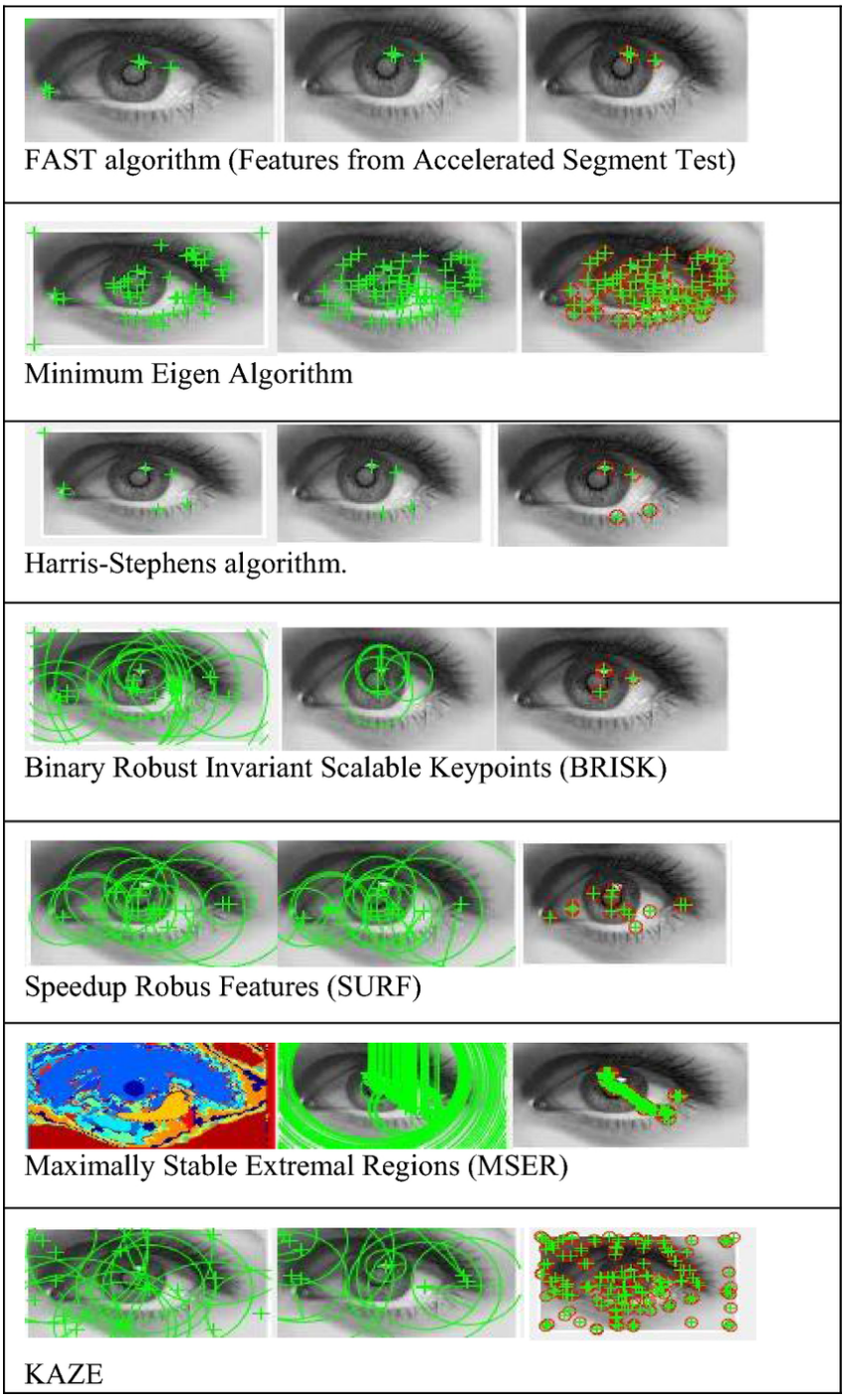


Figure 4: Feature extraction using computer vision algorithms (Source: Sophia, Gino & Sharmila, Ceronmani. (2020). Computer vision algorithms for dominant contact lens feature extraction using fuzzy-logic-based classifications. Soft Computing. 24. 10.1007/s00500-020-04791-1.)

Object detection and recognition includes mapping low-level featuers to specific semantic categories, either through rule-based matching or learned models. In traditional systems, template matching and Haar cascades were employed for face detection and similar taks, However, deep learning has since redefined this stage with convolutional neural networks (CNNs) learning hierachical representations of visual data that allow for more accurate and scalable detection. These models can recognize static objects and also human body parts, facial expressions and abstract postures.

Motion tracking is particularly relevant in performative contexts. This step measures temporal continuity by tracking the movement of pixels or features across frames. Optical flow algorithms, which compute the apparent motion of brightness patterns, provide dense motion fields. Meanwhile, predictive models like "Kalman filter" estimate future positions of tracked objects by integrating noise and uncertainty. These tracking capabilities provides the techniques for *Gesturism* to follow body parts continuously and with precision, forming the basis for expressive drawing.

A critical advancement in this domain is pose estimation, basically the ability to predict human joint locations from images or video. Models such as PoseNet and BlazePose leverage deep neural networks to estimate coordinates of skeletal keypoints, enabling full-body tracking with high accuracy and low latency. In practice, these systems represent the body as a set of interconnected landmarks (e.g., wrists, elbows, knees), offering a skeletal model that is both expressive and analytical tractable. [3] The next chapter will discuss the model I implemented *Gesturism*.

2.2 Conceptual Frameworks: Artists working with Gesture and Vision

The artistic deployment of CV spans a variety of interaction paradigms: from motion-triggered audio feedback in sound installations, to projection mapping based on body location, to dynamic generative visuals informed by hand gestures. This has transformed CV from a tool of analysis into a material for design. Yet, artistic contexts also expose the limitations of these systems. Real-time applications require not only accurate tracking but also robustness to oclusion, lighting changes, and unpredictable movements, which can affect system performance. Environmental conditions significantly impact the fidelity of CV-based systems. Lighting variance can compromise edge detection and depth sensing, while occlusion (e.g. crossing limbs or multiple participants) challenges model inference. Latency, often introduced by processing or network delays, can diminish the responsiveness of the interface, breaking the illusion of real-time interactivity. Furthermore, pose estimation models are typically trained on benchmark datasets like COCO or MPII that may not capture the stylistic diversity or cultural specificity of movement vocabularies present in performative art [21]. The deployment of CV in creative contexts demand both technical adaptation and conceptual sensitivity.

2.2.1 Golan Levin

Golan Levin frequently considers the aesthetics of communication between systems, users, gestures and media. Levin is a pivotal figure in media art, foregrounding human input as an expressive, generative source, and often leverage computer vision to translate the body into an interface.

In 1998, Levin created a reponsive environment, titled *Yellowtail*, capturing the user's

cursor movements and extends them in time, generating flowing, calligraphic trails (Figure 5). Althought not reliant on full-body computer vision, *Yellotail* models the concept of procedural drawing, turning motion itself into visual materials. The work emphasizes rhythm, curvature, and speed as aesthetic parameters. *Yellowtail* offers a framework for considering motion as a performance, treating the body as a dynamic artifact. [7]

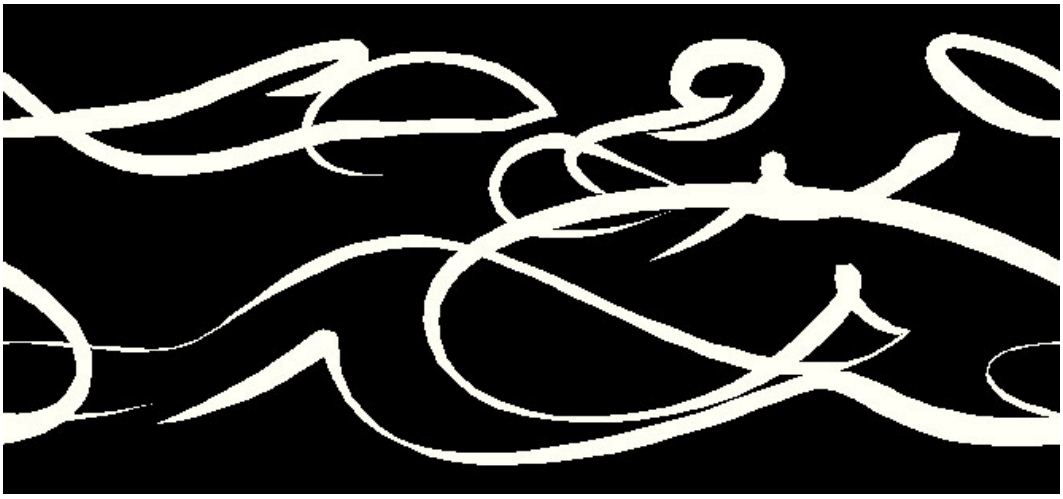


Figure 5: Yellowtail (1998)

In collaboration, with Zachary Lieberman, *Messa di Voce* (2003) translates spoken word and bodily gesture into animated graphics in real time, allowing performers to manipulate visual elements through their physical and vocal presence (FIgure 6). The project illustrates a multi-modal interaction framework, tracking and intepreting the intent behind gestures. [4]

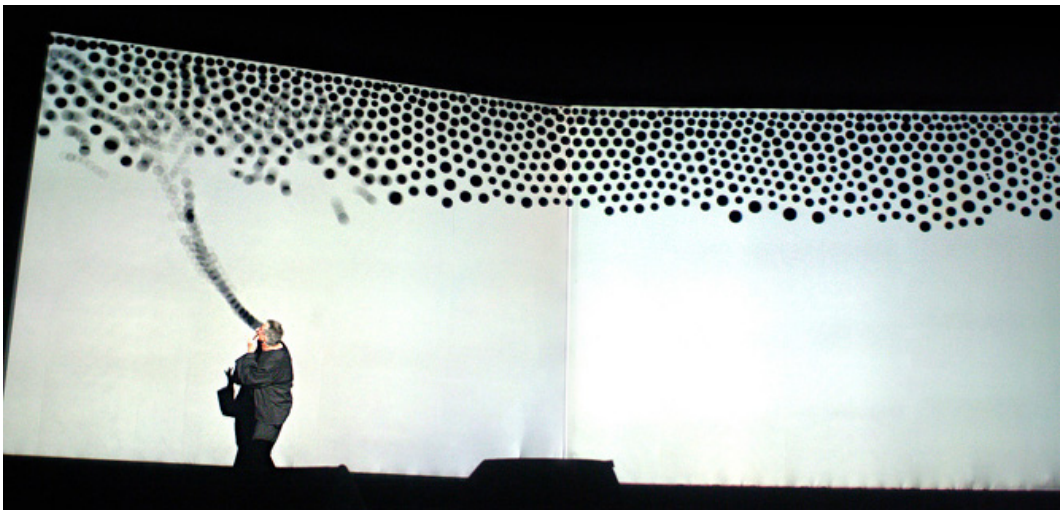
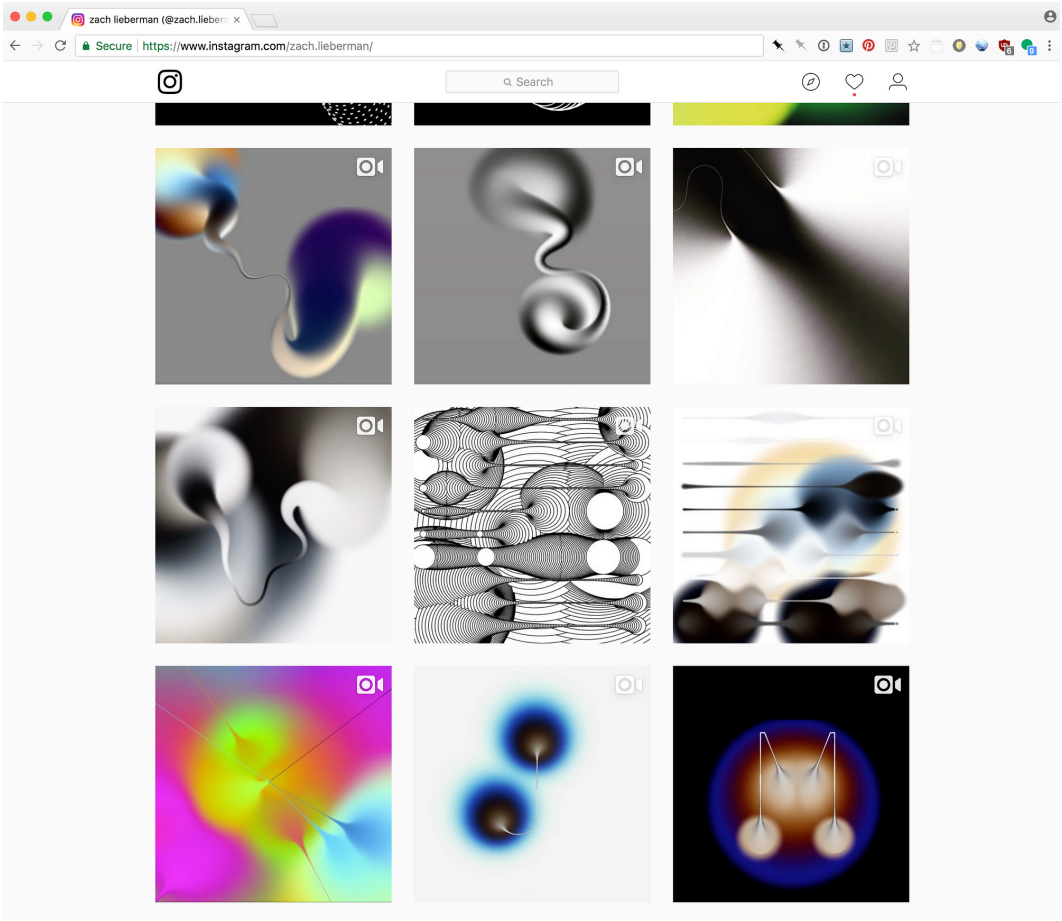


Figure 6: 2003 | Tmema (Golan Levin and Zachary Lieberman) with Jaap Blonk and Joan La Barbara

2.2.2 Zachary Lieberman

Zachary Lieberman is a creative coder, educator and artist who thinks about the boundaries of perception and interaction. With the poetic use of code to visualize the tangible, transforming invisible forces (breath, sound, light) into visible, manipulable phenomena. Lieberman's *Daily Sketches* is a series of algorithmic visual experiments posted online the explore iteratively the use of gesture, interaction and abstraction. Many of these sketches incorporate facial tracking, contour analysis, and skeletal detection to create visuals that respond to subtle shifts in the human body (Figure 6).



ß Lieberman's *Reflection Studies* (Figure 7) and other collaborative installations investigate mirrored motion, feedback loops and participatory image-making. These systems often utilize CV to create environments where user presence is simultaneously mirrored and abstracted. In theses contexts, co-participation becomes central to the meaning of the work. Viewers are no longer passive spectators but are drawn into a generative system where their gestues co-author evolving visual outcomes. The reciprocal nature of feedback in these environments repositions the user as a co-creator, accomodating the nuances of individual expression into their visual feedback. It encourages a reframing of the artwork as a space of negotiation between artist, audience, and algorithm.



Figure 7: Lieberman's Reflection Studies| Exploratorium Museum Exhibit

2.3 Digital Painting

From the 1990s to the early twenty-first century, digital media experienced rapid technological advancements, transitioning from the "digital revolution" to the social media era. While the foundations of digital technologies were laid decades earlier, these innovations became increasingly pervasive in the late twentieth century. This period saw hardware and software become more refined and affordable, coinciding with the advent of the World Wide Web in the mid-1990s, which introduced a new dimension of global connectivity [1].

In the twenty-first century, the term "digital art" has evolved into an umbrella category encompassing a diverse array of artistic practices and works, reflecting the medium's versatility. Within this scope, digital painting emerges as a distinct form that retains foundational elements of traditional art—such as line, shape, value, texture, and color—while adhering to guiding principles like balance, emphasis, unity, and variety. However, the digital medium also introduces novel affordances that distinguish it from traditional media. These include precision and control, enabling the exact replication of shapes and lines for enhanced accuracy; iterative flexibility, allowing for non-destructive modifications through undo and redo functions; and efficiency, facilitating faster workflows. [20]

A recurring theme in the evolution of digital painting is the pursuit of realism. Techniques such as physically-based simulations and data-driven approaches strive to replicate the tactile qualities of traditional media, such as oil paints and watercolors. This drive to digitally capture the subtle nuances of physical materials underscores the value placed on achieving a level of visual fidelity comparable to traditional art forms. Despite significant progress in the field, replicating the textural qualities of natural media remains a challenge.

Existing digital painting methodologies can be broadly categorized into three approaches. Procedural approaches rely on heuristics designed to mimic specific natural media effects based on developers' intuition. For instance, commercial tools like Adobe Photoshop offer a

range of effects for digital artists but often struggle to achieve the visual fidelity of natural media. Moreover, achieving a convincing level of realism can demand considerable user effort.

The development of digital painting techniques has attracted considerable scholarly attention. Technological advancements, such as the development of Sketchpad in 1964 and the graphical interfaces created at Xerox PARC in the 1970s, pioneered in making digital tools more accessible. This progress eventually led to the release of MacPaint in 1985, which significantly expanded the possibilities for digital painting.[10] These early innovations paved the way for a variety of painting software, including Corel Photo Paint, MetaCreations, Painter 3D, Adobe Photoshop, and JASC's PaintShop Pro, which enabled users to replicate traditional painting techniques on a computer screen.[8]

For higher fidelity, simulation-based approaches model the physical interactions between the virtual brush, canvas, and pigment medium. Techniques such as deformable 3D brushes and paint fluid dynamics simulations have been successfully applied to replicated particular natural media, including water-color, oil paint and pastels. These approaches have yielded impressive results, with commercial applications like Microsoft's Fresh Paint incorporating advanced simulation techniques. [28] However, these simulations are typically tailored to specific media, limiting their adaptability to other artistic tools.

When diving deeper into motion-sensing drawing applications,two primary types are defined: Tangible User Interfaces (TUI) and Kinaesthetic User Interface (KUI). TUI-based systems requires users to manipulate a physical devices to interact with digital canvases. In 2004, MIT Tangible Media Group developed a new drawing tool, I/O brush, which synthesized colors, textures and movement found in everyday materials simply by touching the objects. I/O Brush looks like a regular paintbrush, but the devices consists of a camera and sensors implemented inside (FIgure 6). Outside of the digital canvas, the brush can pick up color, texture, movement from the brushed surface. On the canvas, artists can create artwork with the special "ink" from the materials they collects.



Figure 8: I/O Brush 2004 (MIT Tangible Media Group - Kimiko Ryokai, Stefan Marti and Professor Hiroshi Ishii)

KUI-based system leverage the user's body as the primary controller, enabling freeform drawing gestures in the air. These interfaces, such as finger-tracking systems of depth-sensing cameras, create a more immersive and intuitive experience by eliminating the need for physical tools. The absence of handheld tools reinforces the notion of unmediated embodiment - an appaorch that *Gesturism* adopts by enabling the body to act as the primary interface of creation.

The interaction of my system is inspired by TeamLab's installation: *Flowers and People, Cannot be Controlled by Live Together (2017)* (FIgure 9). This immersive environment is continously shaped by the real-time presence and gestures of visitors As individuals walk through the room, virtual flowers bloom and scatter, forming an evolving ecosystem that responds fluidly to the body's proximity and motion. Using depth-sensing cameras and finger-tracking systems, this installation places users as the center of life and depth as flowers would bloom around their positions and butterflies would wither when touched. The installation underscore movement not as a byoroduct but as a creative act in generating, shaping and performing visual experiences.



Figure 9: Flowers and People, Cannot be Controlled but Live Together

2.4 Embodied Interaction and Creativity

Embodied interaction refers to the ways in which physical movement and sensory input shape human cognition and creativity. Grounded in the theory of embodied cognition, which posits that the mind is deeply interconnected with the body, studies have demonstrated the significance of physical engagement in enhancing creative processes. Movements such as gestures, body posture shifts, and spontaneous interactions with physical objects have been shown to stimulate divergent thinking and foster emotional engagement, which are crucial for creativity. A systematic review across 20 studies on experimental research on embodied creativity shows that embodied movement robustly enhances creativity across nearly all studies (90%), with no studies showing detrimental effects. [13] Research consistently demonstrates a significant connection between embodied movement and creativity, with experimental studies showing that enacting specific bodily movements can amplify creative output[12].

The significance of embodied interaction in artistic practice is perhaps most powerfully exemplified in performance art, where the body serves as the central site of meaning-making. Carolee Schneemann's seminal work *Meat Joy* (1964) is a radical exploration of the corporeal as material. In this performance, as Schneemann described it as an "erotic rite", semi-nude performers engaged in spontaneous, chaotic physical contact involving raw fish, poultry, meat, paint and paper, celebrating the body untethered from normative social constratins. Rejecting the sanitized, cerbral models of art and reclaiming a corporeal, pre-linguistic form of communication, the performance elevate the body from a passive subject to an active instrument of authorship.



Figure 10: Carolee Schneemann, Meat Joy, Whitechapel Gallery during SHOP I

Trisha Brown's *Drawing Performances* invite a fundamental reconsideration of what constitutes both drawing and dance. Her work, which involves attaching drawing instruments to the body and executing a choreographed movement vocabulary, raising questions about the disciplinary boundaries between visual and performing arts. In this hybrid practice, the student becomes not merely a site for production, but a stage for inquiry, as Bruce Nauman's assertion that art is simply "whatever an artist does in the studio." Rather than trying to fix these works within a rigid disciplinary category, Brown's approach challenges the viewer to dwell in the ambiguity between movement and mark, process and artifact. Brown's gestures inscribe time and motion into a visual surface, capturing fleeting bodily movement and transforms it into visual composition.

2.4.1 A framework of key design features of a motion-sensing drawing tool.

The application of interactive systems and the resulting creativity outburst has also been studied in art therapy. Interactive digital art installations can elicit strong emotional responses and impact the psychological well-being of the user. A notable example of this is the development of a kinesthetic drawing application designed specifically for art therapists. This system leverages markerless motion-sensing technology, the Kinect, to enable whole-body interaction to help enhance a client's ability to express creativity. The combination allows clients, particularly those struggling with verbal communication, to express emotions and experiences that might be difficult to articulate in words. The full body engagement provides a more comprehensive picture of the client's state of mind and energy level compared to traditional drawing methods. Although the primary purpose of this application is therapeutic, facilitating the healing process rather than producing aesthetically pleasing or

technically proficient artwork, the features of the kinaesthetic drawing system are very close to what I'm aiming to achieve. The system recognizes movements of different body parts, including hands, head, and chest, translating them into strokes on the virtual canvas, allowing for a wider range of movement and expressive possibilities. Utilizing a virtual canvas projected onto a screen, the system allows users to switch between drawing and non-drawing modes by stepping in and out of a designated "drawing zone'. In terms of customization, a range of stroke types is provided, including normal, ribbon, fluid, smoke, and spray, each with unique visual characteristics inspired by both traditional and digital media. Some stroke types respond dynamically to the user's movement speed, offering visual feedback of the embodied interaction. The notable difference is this system automatically records the drawing process, capturing both the visual evolution of the artwork and the user's body movements. This allows therapists and clients to review the session and gain further insights into the creative process and emotional expression. [27]

Wang research also provides a solid foundation for establishing a design framework for future researchers developing motion-sensing drawing tools for art therapy. The author proposes a matrix framework categorizing the key design features of a motion-sensing drawing tool.

Table 1: Frameworks for key deisgn festures of a motion-sensing drawing tool

Feature	Qualities
Move	**Range**: Options for small, medium, and large movement ranges allow customization for different client needs and spaces.
Move	**Parts**: Assigning strokes to various body parts like hands, head, chest, and legs encourages full-body engagement.
Visual	**Stroke Types**: Different strokes (normal, ribbon, fluid, smoke, spray) enhance expression and engagement.
Visual	**Life Period**: Fading strokes promote intuitive drawing and focus on the present moment.
Visual	**Effects**: Thickness, behavior (e.g., wildness, calmness), and particle interactions enrich the experience.
Interface	**Commands**: Accessible controls like voice commands, hand gestures, or touchless interaction make the interface easy to use.
Interface	**Colors**: A full-color palette offers creative freedom and helps therapists interpret symbolic meanings.
Interact	**Therapist Modes**: Clients can draw alone, with therapists, or collaboratively to fit therapy goals.
Interact	**Save & Record**: Recording and saving features allow adding labels and notes for therapist analysis and discussion.
Interact	**Embodiment**: The system reflects strokes directly linked to body movement.

2.5 Computational Motion-sensing Drawing Application



Figure 11: One of Frank Gilbreth's light painting (1914)

The concept of drawing in the air has long captured human imagination. One of the earliest recorded examples of this idea can be tracked back to Frank Gilbreth's light painting experiments in the 1910s and 1920s, which used motion trails to analyze movement efficiency. this desire to visualize movement dynamically laid the groundwork for later technological advancements in motion-sensing drawing.

During the early stages of digital art, Myron Krueger pioneered interactive environments where users could engage with virtual images. His work, *Metaplay* introduced a virtual space that responded to human emotion, demonstrating the potential of computational systems to bridge physical movement with digital interaction [18]. Krueger created a virtual space (Figure 12) for users to interact with a reponse to a visual image. *Metaplay* consisted of two rooms connected by a closed-circuit video feed that allows Krueger to interact with participants with his drawing on a tablet from a different space. *Metaplay* draws attention to the virtual interaction between the participants, the media, and the artists.

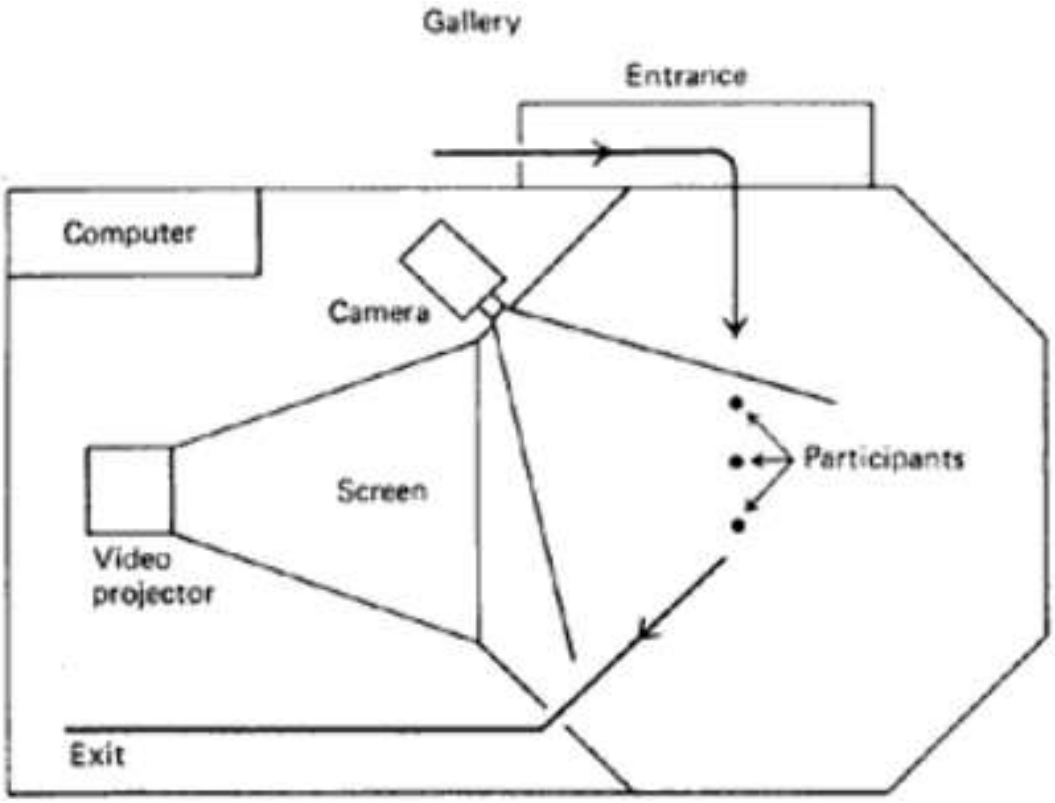


Figure 12: Metaplay communication

2.6 Materializing Movement for Digital Interaction

In 2019, Google Creative Lab collaborated with renowned choreographer Bill T. Jones to create *Body, Movement, Language*, a suite of interactive experiments that explore the intersection of dance, technology, and identity. Utlizing Google's PoseNet - a machine learning model capable of real-time human pose estimation through a web browser - the project invited users to engage with Jone's iconic solo *21* by mirroring his movements and reflecting on the embodied experience of dance in a digital context.

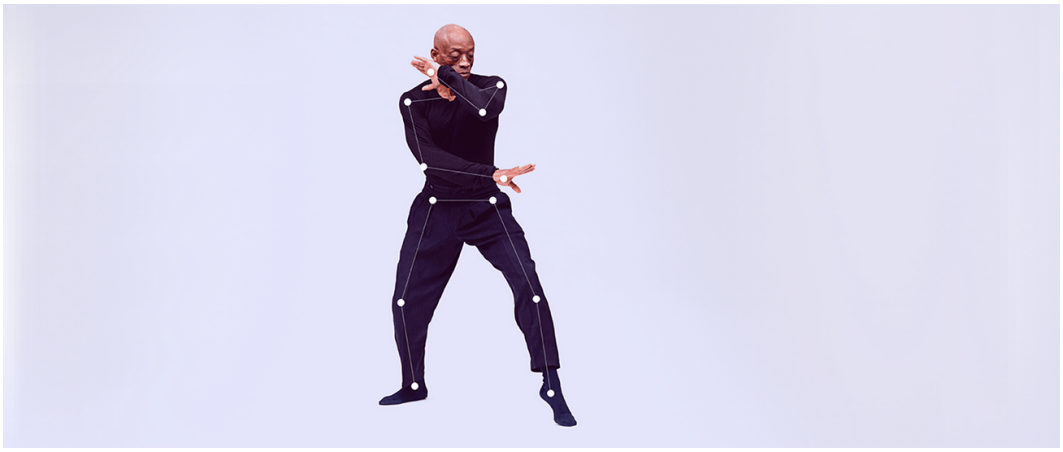


Figure 13: Bill T. Jones '21' performance

Jones approached this collaboration with a critical eye, questioning how technology could serve the expressive depth of human movement. He was particularly interested in whether AI could evoke genuine human emotion, stating "Maybe it's ego, but I want to be the one to know how to use PoseNet to make sombody cry". This perspective aligns with my vision for *Gesturism*, to design experiences where technology fades into the background, amplifying what makes us most human: our capacity to feel, create, and connect. The project comprises four experiments: *Manifesto* (Figure 14), *The Game, Hold That Thought, and Naming Things: Approaching 21.* Each experiment integrates movement and language, allowing users to interact with Jone's choreography and narratives. For instance, *Manifesto* enables users to create a trail of words that respond to their movements in real-time, effectively composing a personal statement through embodied interaction.



Figure 14: `Manifesto` performance

2.6.1 Understanding movement data

The rich communicative potential of full body movement for digital interactions Movement, as Sheets-Johnstone(1999b) critiques, is often reduced to an object in motion, whereas in reality, it encompasses a dynamic and communication process. [23] Building on Ervin Goffman's concept of the performative nature of human behavior, movement is often intentional and interpretable by both humans and technology. The intentionality of movement signifies that actions are not just physical events but imbued with purpose, shaping the way in which they are perceived and interpreted. Full-body movement communication is inherently complex and multimodal, incorporating elements like posture, glances, and utteraces. This complexity arises from the semiotic and somatic aspects; movements carries physical information about the performer (i.e. age, gender, race) and is contextualized by spatial and temporal factors.

Drawing from social semiotics, Hansen and Morrison (2014) designed the *Movement Schema*, provided a structure approach for analyzing full-body movement in interaction design. [16]

Table 2: A schema for identifying semantic properties of movement dynamics

Semiotic Resource	Core Modalities	Salient Characteristics	Visual Description
Communicative Movement	Velocity	Speed the rate of change in x-y-z number	Size of Mark related to the change in velocity
	Position	Location of velocity	Placing of Mark in the visualized body

Semiotic Resource	Core Modalities	Salient Characteristics	Visual Description
	Repetition	Continuation of Movement	Rhythm in the visual trail
	Frequency	Use of Timing	Structure in the visual trail

The *Movement Schema* provided a deeper analytical framework for my thesis, where I consider movement as a dynamic and communication process.

Velocity is an essential yet often overlooked element of human experience. When analyzing movement data, velocity offers a perspective on how motion is executed rather than merely tracking position. The communicative power for movement lies in the manner in which a body part transitions rather than its static locations. By emphasizing the rate of change, we can visualize the expressive and dynamic qualities of gestures, providing insights into their affective and intentional dimension.

Position. The significance of position in movement analysis extends beyond the traditional understanding of posture of overall pose. Human communication is inherently multimodal, engaging several semiotic systems simultaneously. Thus, movement data must account for the relative positioning of body parts during motions, rather than treating the body as a singular, static entity. This perspective allows for an enriched interpretation of meaning, shifting the focus from absolute location to the relational dynamics of movement.

Repetition and Frequency. Movement gains meaning through its recurrent and contextual patterns. Inspired by Goffman's notion of framing and turn-taking, repetition and frequency provide a framework for understanding the structuring of movement over time. [15] Certain gestures acquire significance through their reiteration or contrast within an interaction. Communication relies on the ability to recognize and adapt to recurring movement patterns, ensuring mutual understanding. By incorporating repetition and frequency into movement analysis, we can capture the emergent patterns that shape gestural communication.

2.6.1.1 Sync(2014) Based on the *Movement Schema*, Sync was created to make movement data and, by extension, movement more accessible for designers and other non-programmers to explore its potential in digital interactions. Sync enables a visual reading, identification, and interpretation of data drawn from full-body movement. Sync allows designers to "see" the data and make informed decisions in linking movement data to communication, function, and aesthetic. This kind of setting is what needed for interaction design to appreciate movement as a material.

With Sync, x-y-z points may be visualized in different ways: vertical and horizontal lines, circles, and ribbons. Each visualization choice will foreground different qualities of the movement data. The rate of change in an x,y,z, number is indicated by a change in the size of the visual mark representing that point. An increase of speed will extend the marks, extend the size of the circle, or extend the length of the lines or the width of the ribbons. The marks are continuously drawn, which allows for a visual trail where repetition may build up visually and relationally.

Sync provides designer access to algorithms, allowing them to explore movement data through play, by twisting, shifting, and shaping parameters. This elasticity opens up a material understanding of the possiblities of movement data that emerges through use.

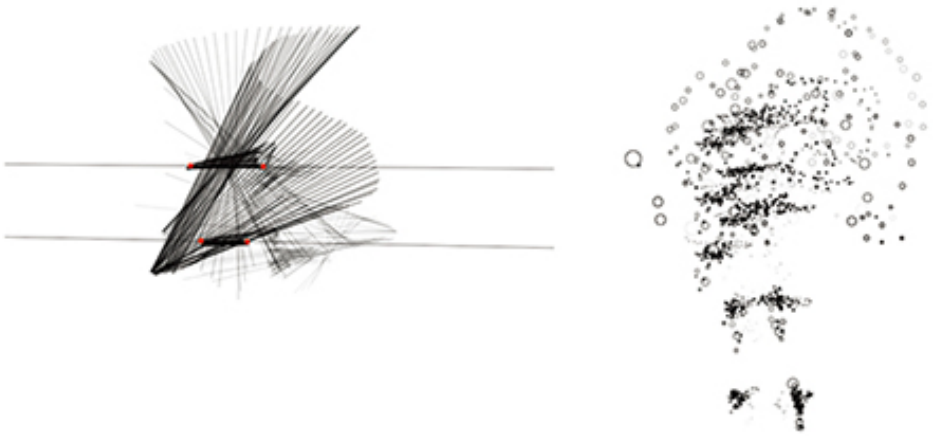


Figure 15: Each x-y-z point can be visualized by Sync as one or several or all marks

3 Method of approach

The development of Gesturism involved multiple stages including the research and implementation of a computer vision module, integration of body landmarks, front-end development and user interface design, and key functionalities for a digital painting application. Until this phase, the aesthetics of Gesturism has not been considered. I kept all visuals utilitarian and functionality-focused: the result is intentionally not beautiful, pointing the users' focus in their experience and drawing experience for the Experiment (Chapter 4). The visual design is discussed in Chapter 5.

3.1 Computer Vision Module

To effectively capture body movements and translate them into artistic gestures, selecting an optimal computer vision module was crucial. Two prominent frameworks considered for this purpose were PoseNet and MediaPipe. While both frameworks offer pose estimation capabilities, differences in performances, possibilities, and precision influenced the final decision.

3.1.1 Comparision of TensorFlow PoseNet and MediaPipe BlazePose

PoseNet is a lightweight neural network model designed for real-time human pose estimation in web browsers and mobile applications [22]. It effectively identifies 17 key body points and has been widely adopted for browser-based interactive applications. While PoseNet offers solid performance in controlled conditions, several limitations emerged when assessing its performance in Gesturism's requirements. On the other hand, MediaPipe BlazePose, developed by Google Research, has demonstrated superior performance in both speed and accuracy. According to [9] , BlazePose achieved a faster inference speed than PoseNet when testing in web-based environments. Furthermore, BlazePose tracks 33 body landmarks (Figure 16), offering a more comprehensive representation of the human body. This expanded set of landmarks provided finer control, aligning well with Gesturism's goal of intuitive gesture-based painting controls. Given these findings, MediaPipe BlazePose was chosen for Gesturism due to its higher inference speed and improved accuracy in real-world conditions.

3.1.2 Implementation and Justification

In the implementation phase, the Pose module from MediaPipe was selected as the primary tracking tool. The Pose module specifically tracks 33 body landmarks, which balances between precision and simplicity. This choice allows for comprehensive user control without overwhelming users with excessive data points. While MediaPipe Holistic offers even greater tracking capabilities with 543 landmarks (33 pose landmarks, 468 facial landmarks, and 21 hand landmarks per hand), this extensive data set was deemed excessive for Gesturism's intended functionality. The Pose module's focused landmark set streamlined the development process while maintaining sufficient granularity for accurate gesture-based interactions.

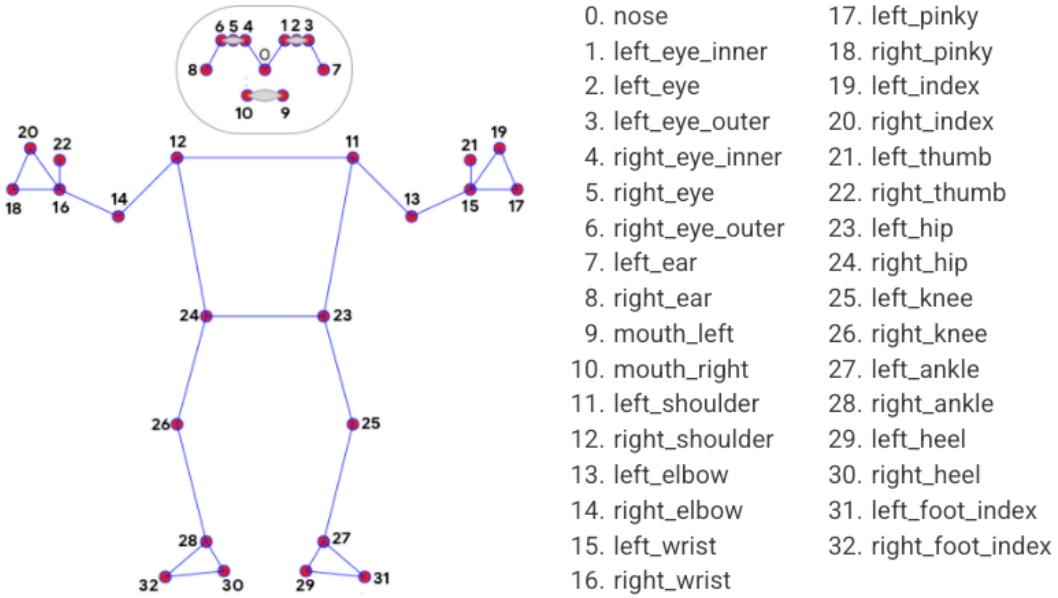


Figure 16: MediaPipe Pose Module: 33 Body Landmarks

3.1.3 Compatibility Considerations

While MediaPipe offered strong pose tracking capabilities, it is important to note that it has known compatibility issues with certain modern frontend frameworks, such as Vite (i.e. TypeError: Pose is not a constructor). To address these challenges, a workaround was implemented in the subsequent development phase. The next section outlines the specific adjustments made to ensure integration of MediaPipe within the Gesturism framework.

3.2 Environment Setup

The development environment of Gesturism was carefully designed to optimize performance, maintainability, and scalability. This section outlines the technology stack, installation procedures, and configurations essential for setting up the system.

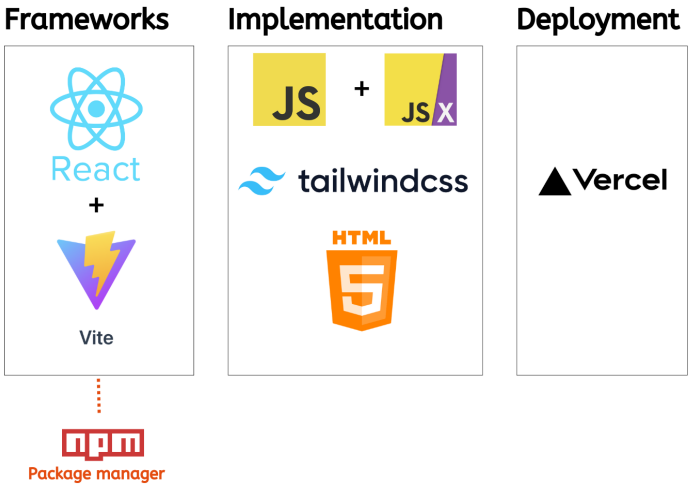


Figure 17: Gesturism Tech Stack

The development of Gesturism required a robust and efficient front-end framework capable of handling real-time user interactions. To start, I selected ReactJS [2] for its component-based architecture, facilitating modular and maintainable development. Vite [5] was preferred over Create React App (CRA) for its enhanced performance, and optimized build process. For package management, Node Package Manager (NPM) is used.

```
npm create vite@latest my-gesturism-app --template react
cd my-vite-project
npm install
npm run dev
```

Command line interfaces include `vite`, `vite build`, `vite preview` to launch the development server with hot module replacement, generate an optimized production build, and serve the production build locally for testing, respectively.

For styling, TailwindCSS [6] , an utility-first CSS framework that enables rapid UI development while maintaining consistency in styling. TailwindCSS offers a more efficient styling approach compared to traditional CSS by allowing developers to style components directly in HTML using predefined utility classes. This eliminates the need for writing custom styles in separate stylesheets, reducing redundancy and improving maintainability. Example of Traditional CSS Styling include something like this:

```
/*styles.css*/
.button {
background-color: blue;
color: white;
padding: 10px 20px;
border-radius: 5px;
font-weight:bold;
}
```

```
{/* Component using CSS */}
<button className = "button"> Click Me </button>
```

Equivalent styling with TailwindCSS removes the need for separate stylesheets, making styling more scalable and reducing CSS bloat. Additionally, Tailwind's utility classes are optimized for performance, improving page load times and development speed.

```
{/* PoseDrawing.jsx */}
<button className="bg-blue-500 text-white px-4 py-2 rounded-md font-bold">
Click Me
</button>
```

Gesturism utilizes both JavaScript and JSX within the React framework. While JavaScript manages logic and applications state, JSX enables declarative UI creation by allowing developers to write HTML-like syntax within JavaScript. React embraces the fact that rendering logic is inherently coupled with other UI logic: how events are handled, how the state changes over time, and how the data is prepared for display. Hence, JSX produces React "elements" as a syntax extension to JavaScript.

In the example below, a button is declared then used inside JSX by wrapping it in curly braces. JavaScript is used to handle logic.

```
const handleExportDrawing = () => {
alert("Button clicked!");
};

<button
className = "absolute botton-4 right-40 text-white"
onClick = {handleExportDrawing}
>
Export PNG
</button>
```

3.3 Body Part Integration

3.3.1 Getting started with Mediapipe and Error Handling

Gesturism relies on MediaPipe for real-time body tracking and gesture recognition. The MediaPipe library provides a fast and lightweight solution for real-time pose estimation using ML models. However, developers using Vite in their tech stack should note a common issue might arise where `npm run dev` functions correctly in a local development environment, but `npm run build` faults, displaying errors like `Pose is not a constructor`. This is likely due to how Vite handles ES module exports when bundling for production. To mitigate this, a custom plugin `mediapipe_workaround` is implemented within Vite's configuration. This plugin explicitly re-exports necessary modules from `@mediapipe/pose` and `@mediapipe/camera_utils` to avoid obfuscation issues. The function below ensures that Vite preserves the required exports:

```
// Minimal workaround for MediaPipe components
function mediapipe_workaround() {
  return {
    name: 'mediapipe_workaround',
    load(id) {
      const MEDIAPIPE_EXPORT_NAMES = {
        'pose.js': [
          'Pose',
          'POSE_CONNECTIONS',
          'VERSION',
        ],
        'camera_utils.js': [
          'Camera',
        ],
        'drawing_utils.js': [
          'drawConnectors',
          'drawLandmarks',
        ],
      };

      let fileName = path.basename(id);
      if (!(fileName in MEDIAPIPE_EXPORT_NAMES)) return null;

      let code = fs.readFileSync(id, 'utf-8');
      for (const name of MEDIAPIPE_EXPORT_NAMES[fileName]) {
        code += `exports.${name} = ${name};`;
      }
      return { code };
    },
  };
}
```

To use this workaround, add the function to `vite.config.js` and include it in the plugins array:

```
export default defineConfig({
  plugins: [
    mediapipe_workaround(),
    react(),
  ]});
```

The specific modules used in this project include: **mediapipe/pose:** This module offers 33 key points of body tracking, balancing accuracy and efficiency. **mediapipe/camera_utils:** Facilitates access to webcam feeds and streams them as input to the model **mediapipe/drawing_utils:** Provides helper functions for rendering detected pose landmarks onto the canvas

3.3.2 The Mediapipe Model

The `Pose` model is initialized with specific parameters to optimize its performance. `POSE_CONNECTIONS` defines the connectivity between different body landmarks for visualization. `modelComplexity` determines the trade-off between inference speed and accuracy. A higher complexity value leads to improved pose estimation accuracy at the cost of increased computational load. `minDetectionConfidence` set a stale threshold to ensure accurate detection of body parts. `minTrackingConfidence` maintains consistency and avoids flickering in the detected key points.

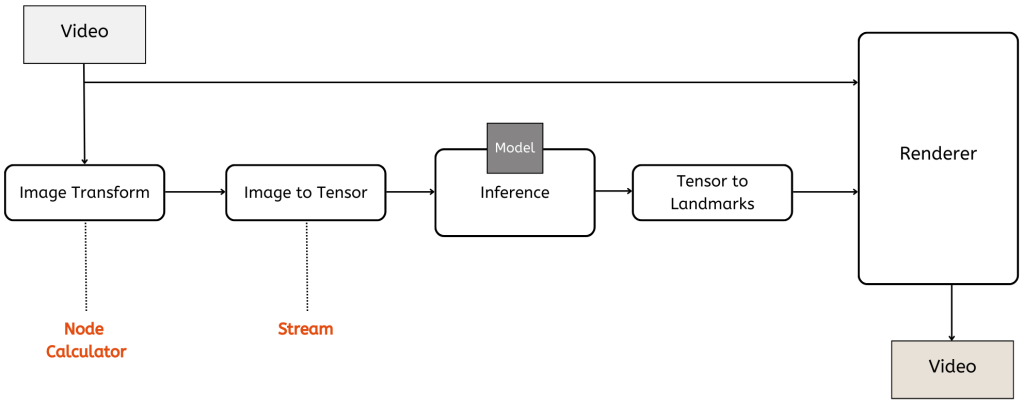


Figure 18: Mediapipe integration in Gesturism

The above diagram illustrates the Mediapipe Pose processing pipeline, which converts a live video feed into pose landmarks through a series of actions. Initially, the video input undergoes an image transformation stage, where frames are preprocessed for better model compatibility. These transformed images are then converted into tensors, which serve as numerical representations that the model can process. The core of the pipeline is the inference step, where a trained deep-learning model analyzes the input tensors and predicts key body landmarks. These raw tensor outputs are then translated into landmark coordinates that correspond to human body joints. Finally, the landmark data is rendered back onto the original video stream for real-time visualization of detected poses.

Once the pose landmarks are identified, the system applies a smoothing algorithm to reduce noise and ensure fluid motion. The detected positions are then mapped onto the virtual canvas, where the custom creation setting, to determine the final visual output.

3.3.3 Full Body Integration

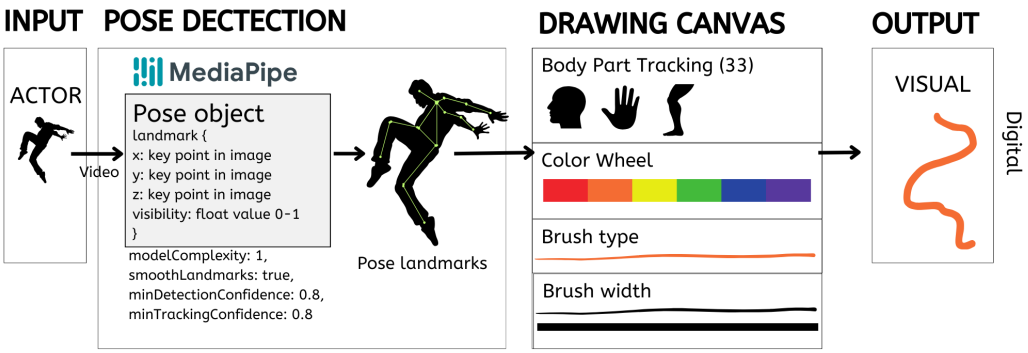


Figure 19: The complete flow of Gesturism

The Mediapipe framework employs a deep learning models to extract skeletal joint positions, allowing for full-body tracking. Gesturism integrates all 33 landmark points in the `Pose` module, categorize into five main body regions:

Head (0-10): Includes key facial and head points, such as the nose, eyes, and mouth. **Left Arm (11,13,15,17,19,22):** Covers shoulder, elbow, wrist, and figure joints **Right Arm (12,14,16,18,20,22) Left Leg (23,25,27,29,31):** Extends from the hip to the ankle, including the foot **Right Leg (24,26,28,30,32)**

Each of these body segments can be individually controlled, granting users the ability to focus on particular regions.

3.3.4 Debugging and Landmark Visualization

To further enhance usability, the application incorporates Render Debug Info features, which visually confirms active tracking points and provides real-time feedback on detected body parts. This feature serves as a real-time validation where users can confirm which landmarks are currently detected and being tracked, as well as assistance in optimizing positioning for better tracking accuracy.

```
const renderDebugInfo = () => {
   if (!poseData) return null;

   return (
     <div className="absolute bottom-4 left-4 bg-black/70 text-white p-4
     rounded-lg text-sm max-w-xs">

    <h3 className="font-bold mb-2">Tracking Status</h3>
```

```
        {Object.entries(trackingSettings).map(([bodyPart, settings]) => {
          if (!settings.isEnabled) return null;

          const landmark = poseData.landmarks[settings.selectedLandmark];
          const isVisible = landmark && landmark.visibility > 0.5;
          const landmarkName = LANDMARK_NAMES[settings.selectedLandmark];

          return (
            <div key={bodyPart} className="flex items-center justify-between
             mb-1">
              <span className="mr-2">
                {bodyPart} ({landmarkName}):
              </span>
              <span className={`font-bold ${isVisible ? 'text-green-400':
              'text-red-400'}`}>
                {isVisible ? ' ' : ' '}
              </span>
            </div>
          );
        })}
      </div>
    );
  };
```



Figure 20: The pose tracking status

3.3.5 The Customization Menu

To further enhance usability, Gesturism includes a Customization Menu that offers comprehensive control of different functionality on the front-end. This menu was designed to

offer flexibility in interaction by enabling users to selectively activate or deactivate specific body parts for tracking, modify visual elements such as stroke color and width, and choose different brush styles.

3.3.5.1 Body Part Selection Users were provided with the ability to toggle tracking for different body parts, including the head, arms and legs and all 33 tracking points. Each body part could be independently enabled or disabled, customized with different aesthetics to offer flexibility. This feature is particularly useful for tailoring the system to different use cases, such as emphasizing hand movements for precision-based tasks or full body interactions for expressive motion-based application. Using React's `useState`, this system ensures that the setting is tracked dynamically.

```
const handleToggleBodyPart = (bodyPart) => {
    onSettingsChange({
      ...trackingSettings,
      [bodyPart]: {
        ...trackingSettings[bodyPart],
        isEnabled: !trackingSettings[bodyPart].isEnabled
      }
    });
  };
```

In addition, when exporting an artwork, the file name includes information about the enabled body parts and landmarks. This ensures better organization and traceability of generated visuals. The naming convention follows a structure format where the active body parts and their corresponding landmarks are listed in the filename, making it easier to reference specific tracking configuration used in the creation of each drawing.

```
const handleExportDrawing = () => {
    if(!drawingCanvasRef.current) return;

    // Get all enabled landmarks from trackingSettings
    const enabledLandmarks = Object.values(trackingSettings)
      .filter(settings => settings.isEnabled)
      .map(settings => settings.selectedLandmark)
      .sort((a, b) => a - b);

    // Create filename with enabled landmark indices
    const landmarkString = enabledLandmarks.join('_');

    // Create and trigger download
    const link = document.createElement('a');
    const dataURL = drawingCanvasRef.current.toDataURL('image/png');

    link.download = `${landmarkString}.png`;
    link.href = dataURL;

    document.body.appendChild(link);
```

```
    link.click();
    document.body.removeChild(link);
  };
```

3.3.5.2 Color Seclection A color selection tool was integrated, enabling users to define the color of the drawing strokes associated with each body part. This feature was facilitated through the SketchPicker component from react-color, allowing real-time selection and preview of hues.



Figure 21: The color palette implemented in Gesturism

3.3.5.3 Brush Type Selection The menu also supports customizing different body parts with different brush types. The current implementation supports only fine line strokes, which characteristic is precise continuous strokes. The feature was intended to support different artistic styles and enhance expressivity within the drawing interface.



Figure 22: An abstract flower sketch created by researcher (March 2025)

3.3.5.4 Brush Width Adjustment To further refine the drawing experiences, Gesturism also implemented a slider-based control to adjust the width of the brush stroke, ranging from 1 to 20 pixels.

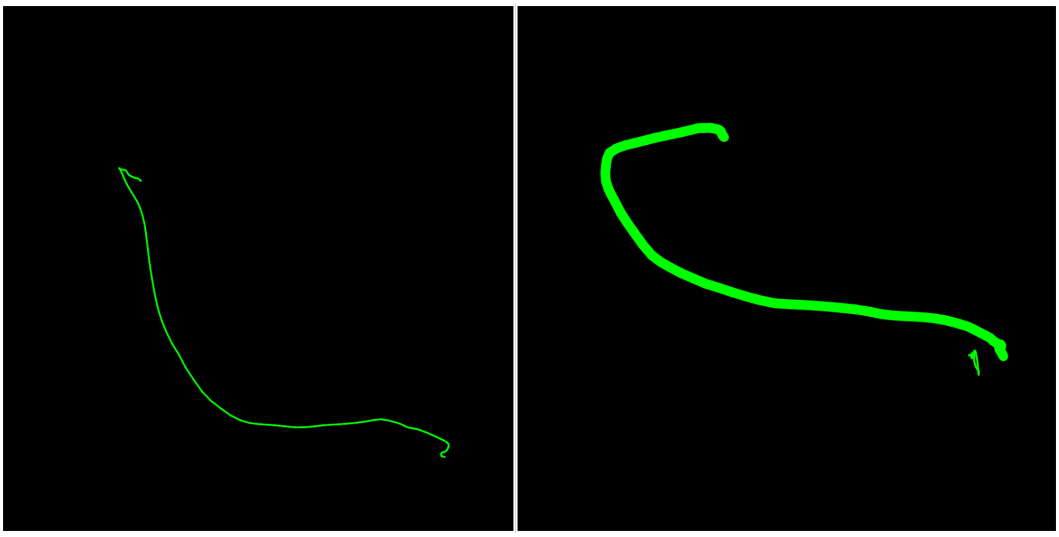


Figure 23: Comparison between brushstrokes 4px and 20px

Once users complete the selection of different brush strok settings and correponding body part, the canvas is updated in real time.

3.3.5.5 Canvas Clear Gesturim provides a functionality to clear the canvas. The functionality is added when researchers realize the overwhelming effect when looking at an artwork with too many unintentional strokes. This functionality allows users to start clean.

3.3.6 File functionality

In the process of creating the artwork, there are several supplement functions to offer users the flexibility of artwork manipulation. One limitation of this version is that there is no user authentication and database implementation, which is a response to concerns about privacy issues. Hence, in the current implementation, the researcher explore how users can store their files and still have neccessary information about their artwork without providing their personal information. When exporting a drawing, the file name now includes information about the enabled body parts and landmarks. This ensure tracesability of generated visuals.

```
// Get all enabled body parts and landmarks from trackingSettings
  const enabledParts = Object.entries(trackingSettings)
    .filter(([_, settings]) => settings.isEnabled)
    .map(([part, settings]) => `${part}-${settings.selectedLandmark}`)
    .join('_');

  // Create and trigger download
  const link = document.createElement('a');
  const dataURL = drawingCanvasRef.current.toDataURL('image/png');

  link.download = `drawing_${enabledParts}.png`;
  link.href = dataURL;
```

3.4 Gesture-Based Control (Under Development)

To further expand the hands-free and immersive qualitites of Gesturism, a gesture recofnition feature is currently under development. This functionality explores the use of full-body post estimiation, particularly hand and arm movements, to provide users with alternative modes of interacting with the application. Instead of requiring users to manually navigate the interface to change settings, specific body gestures can be used to trigger commands such as adjusting visualization parameters and start/stop tracking module.

In the current prototype, the system leverages MediaPipe Pose to detect a specific gesture- a right-hand fist raise to shoulder level with a bent elbow- to toggle the tracking state between active and paused. This gesture is selected due to its intuitive association with intention but upon launching, proved to be high risk of accifental activation. To ensure robustness, the gesture must be detected consistently across multiple frames (n 5) with a cooldown period of one second to prevent rapid toggling. This mechanism is implemented through a real-time analysis of pose landmarks (e.e. shoulder, elbow, wrist) and their geometric relationships.

```
const isWristRaised = rightWrist.y <= rightShoulder.y + 0.1;
const isBent = shoulderToWrist < (shoulderToElbow + elbowToWrist) * 0.8;
const isToSide = rightWrist.x < rightShoulder.x;
```

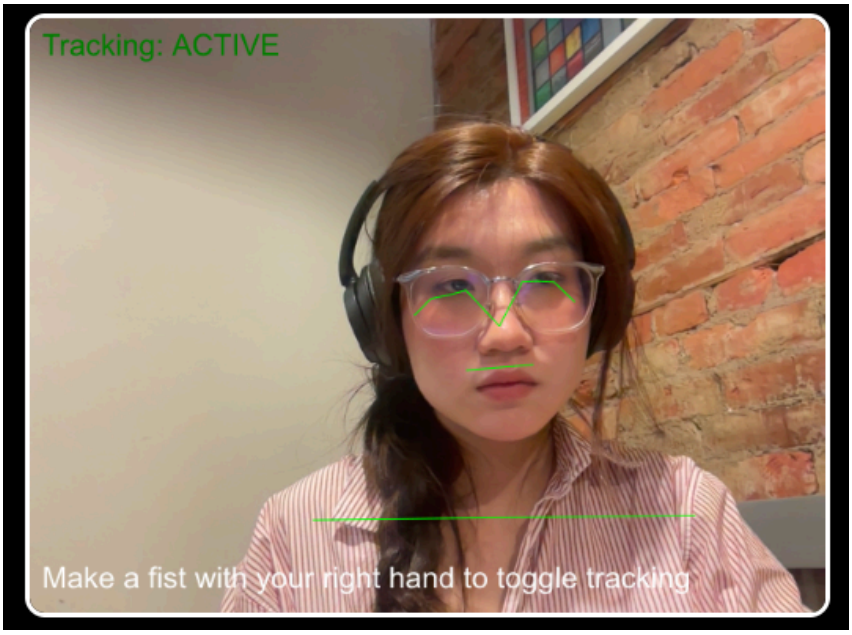


Figure 24: Tracking Status Active

Upon detection, the system provides immediate visual feedback to inform the user of the change in tracking status. Additionally, instructions and progress indicators are overlaid on the preview canvas to guide users through the interaction process. This feedback mechanism ensures accessibility and transparency in user interaction.



Figure 25: Tracking Status Paused

The long-term objective of this gesture-based control is to support a broader set of gestural commands for functionalities such as clearing the canvas, switching brush styles, and exporting artwork, thereby eliminating the need to switch between creative actions and UI navigation. A discussion of candidate gestures for these operations is ongoing, with future iterations aiming to include user-defined gesture mapping to enhance personalization and accessibility.

Table 3: A proposed framework for mapping gesture vocabulary to UI functionality

Gesture Type	Intended Functionality	Interaction Modality	Strengths	Potential Challenges
Right-Hand Fist (Raised)	Toggle Tracking	Static pose with relative arm positioning	High detectability, high false positive	May fatigue arm; may misfire during expressive movement
Double Hand Raise	Clear Canvas	Symmetric full-body gesture	Symbolic of reset, low accidental trigger	Needs wide camera view; occlusion risk
Left Arm Cross to Right Shoulder	Switch Brush Style	Diagonal arm movement	Visually distinct, culturally associated with change	Requires fine tracking of crossing movement

Gesture Type	Intended Functionality	Interaction Modality	Strengths	Potential Challenges
Two-Hand Open Palm Forward	Export Artwork	Static open pose with both arms	Intuitive and metaphorically "offering"	May be confused with start gestures in performance
Rapid Wrist Flick (Right)	Change Color	High velocity micro-movement	Fast interaction, minimal body movement	Prone to noise; high misfire chance without robust filtering

4 Experiments

4.1 Experiment Method

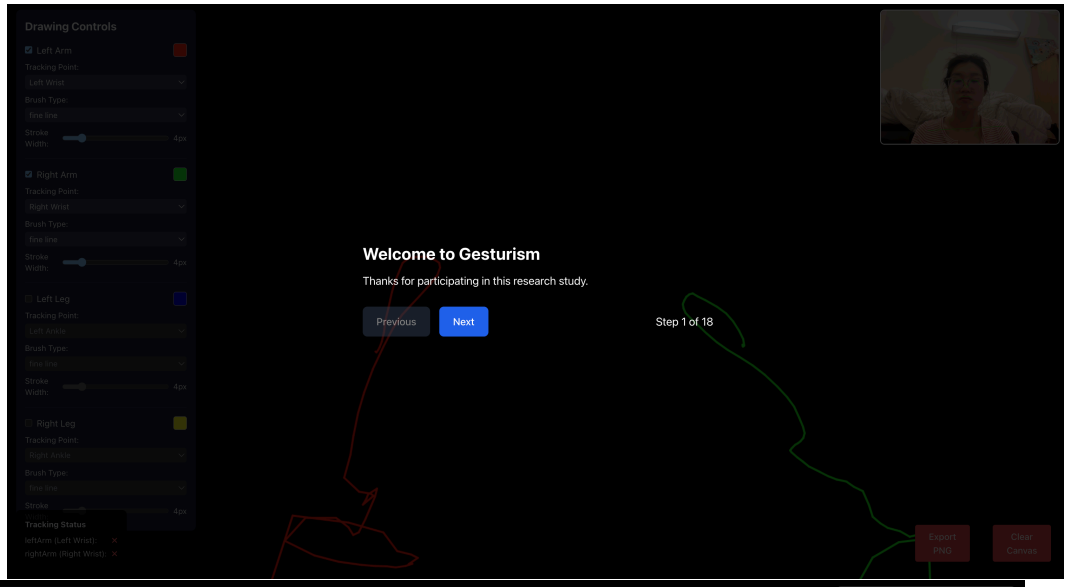
This study evaluates the user experience of Gesturism, a movement-based digital painting interface using computer vision for interaction. In particular, the usability, interaction fluidity, and interface intuitiveness of Gesturism will be evaluated. This study will answer RQ1: What considerations must be made for gesture-based web application?

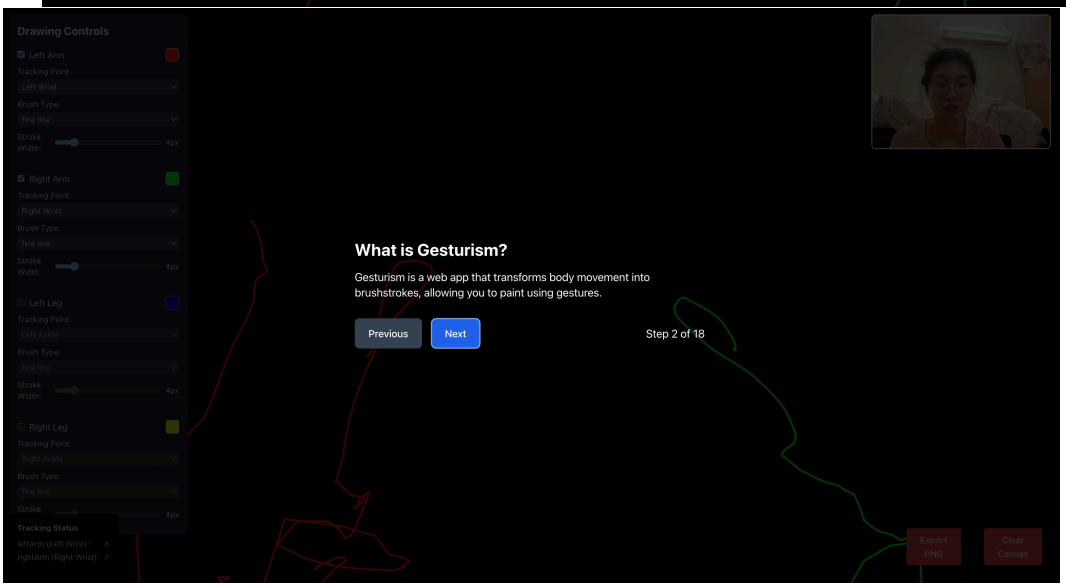
4.2 Study Procedure

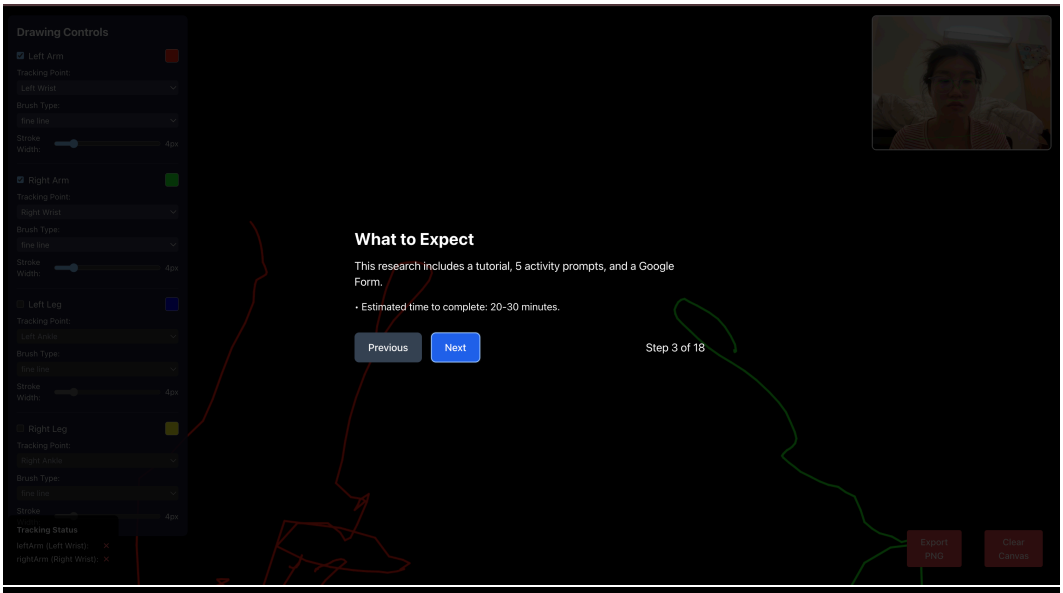
Table 4: Table of participator goals, strategies to achieve them and implementation in the User Research Study

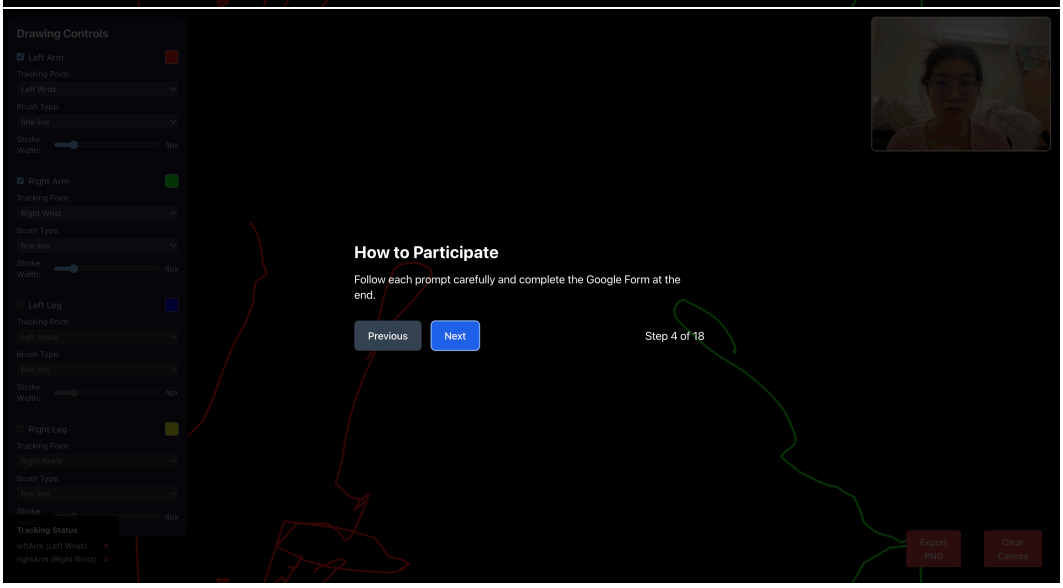
Participatory Goal	Strategy	Implementation
Make experience informed and approachable	Walk participants through a digital tour with breakdown steps of navigation	Develop 18 steps of guided tour, focusing on the configuration menu and setting requirements
Facilitate comfort over using Gesturism	Provide initial practice prompts	Develop four instructional prompts to draw simple shapes with detailed configuration requirements
Facilitate open exploration of Gesturism and movement	Give them an open-ended prompts with little to none instructions, yet the prompt need to be easy enough for universal understanding	Create a freeform prompt that encourages users to use a combination of settings.
Facilitate reflection on Gesturism interaction	Encourage comprehensive feedback on their experience	Provide a series of multiple choice and short-answer questions asking about opinions on functionality and interface ratings, and reflection on this new form of art creation

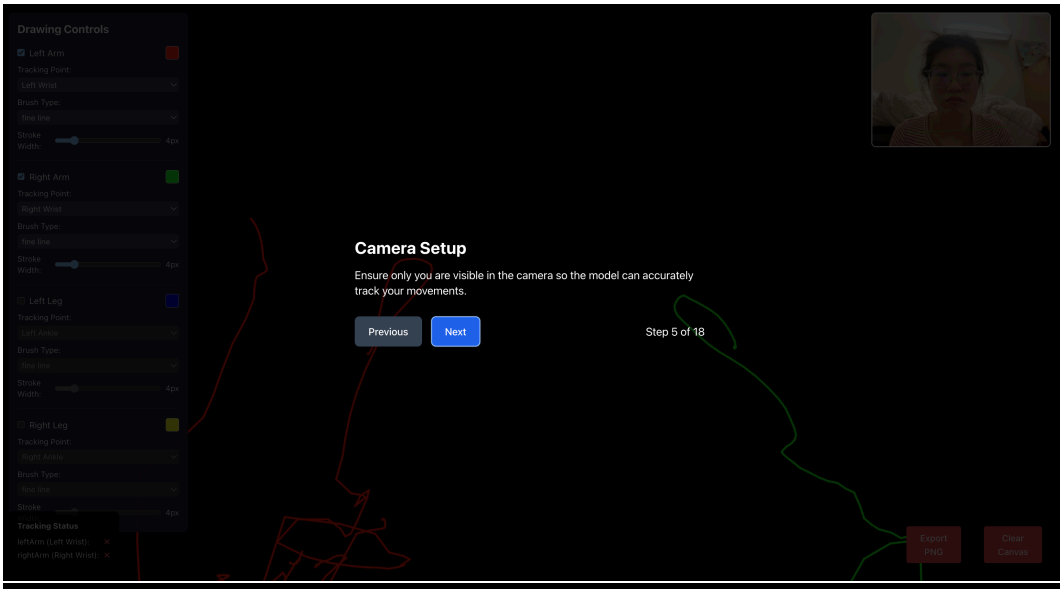
After self-selecting for participation and following the link on the recruitment flyer, participants will be taken to the Gesturism web application where they will follow an 18-step guided tour covering optimal camera positioning (e.g., standing five feet away, camera at hip height) and computer requirements. This tour will also introduce key features, including body-part tracking, brush customization, and stroke control.

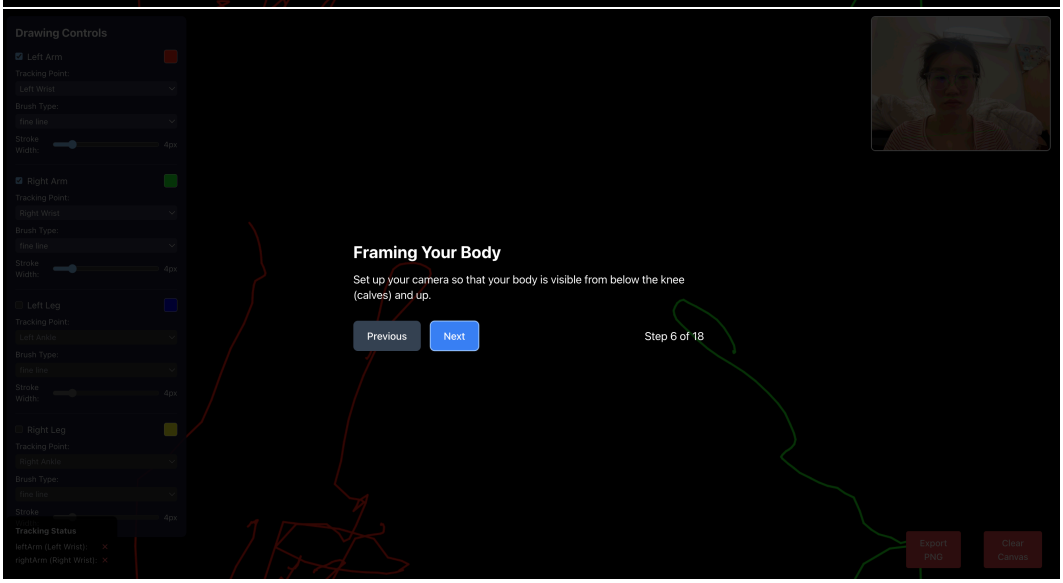

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist):
rightArm (Right Wrist):
Welcome to Gesturism
Thanks for participating in this research study.
Previous
Next
Step 1 of 18
Export PNG
Clear Canvas

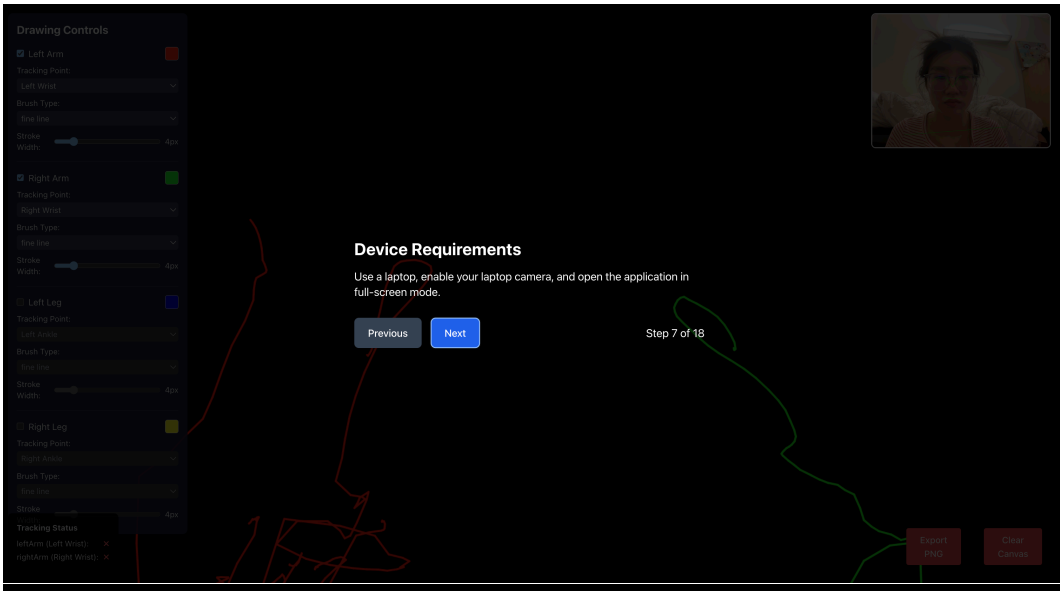

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist):
rightArm (Right Wrist):
What is Gesturism?
Gesturism is a web app that transforms body movement into brushstrokes, allowing you to paint using gestures.
Previous
Next
Step 2 of 18
Export PNG
Clear Canvas

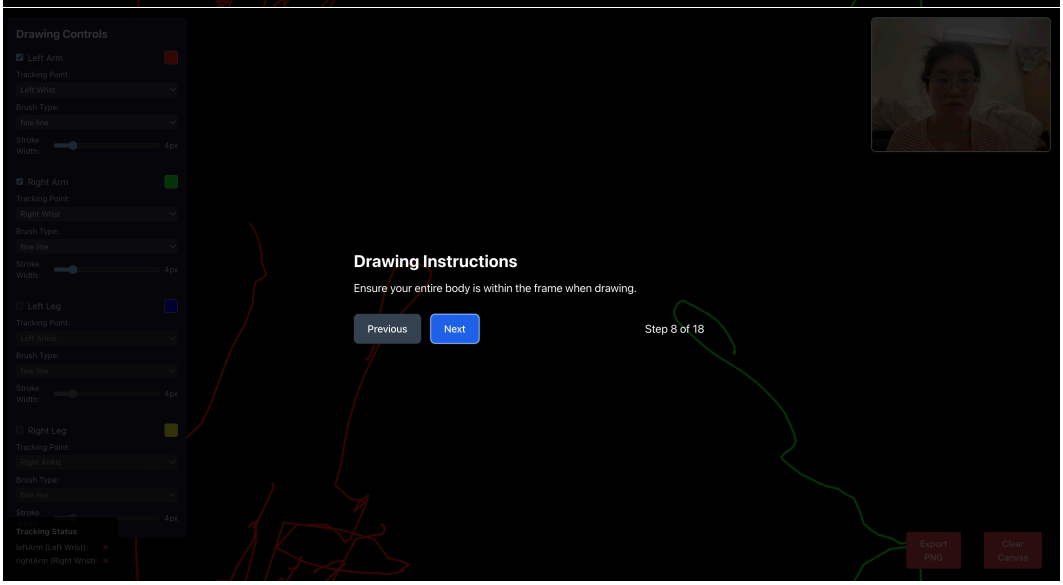

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist): ×
rightArm (Right Wrist): ×
What to Expect
This research includes a tutorial, 5 activity prompts, and a Google Form.
• Estimated time to complete: 20-30 minutes.
Previous
Next
Step 3 of 18
Export PNG
Clear Canvas

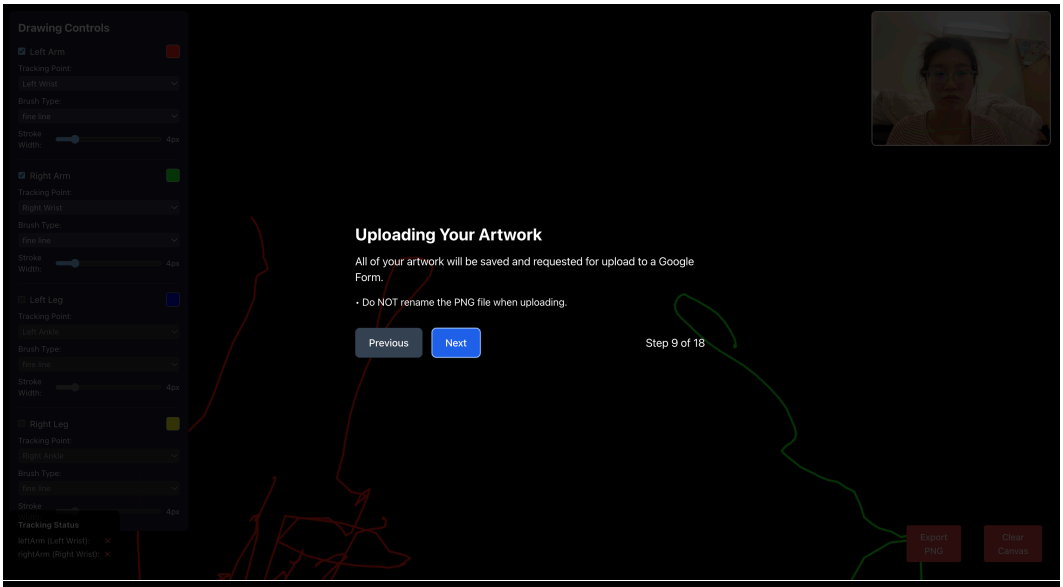

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist): ×
rightArm (Right Wrist): ×
How to Participate
Follow each prompt carefully and complete the Google Form at the end.
Previous
Next
Step 4 of 18
Export PNG
Clear Canvas

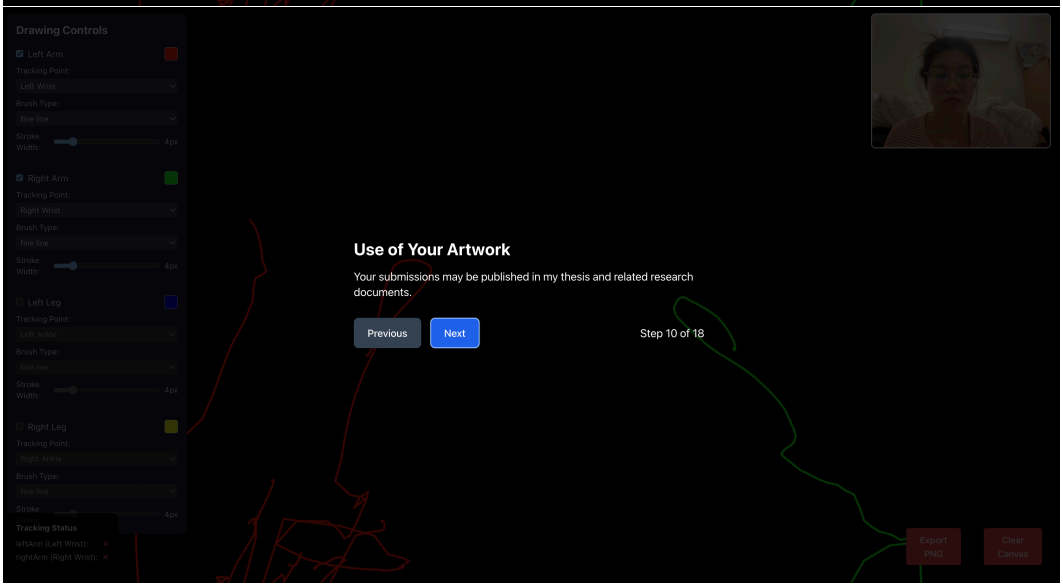

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist): ×
rightArm (Right Wrist): ×
Camera Setup
Ensure only you are visible in the camera so the model can accurately track your movements.
Previous
Next
Step 5 of 18
Export PNG
Clear Canvas

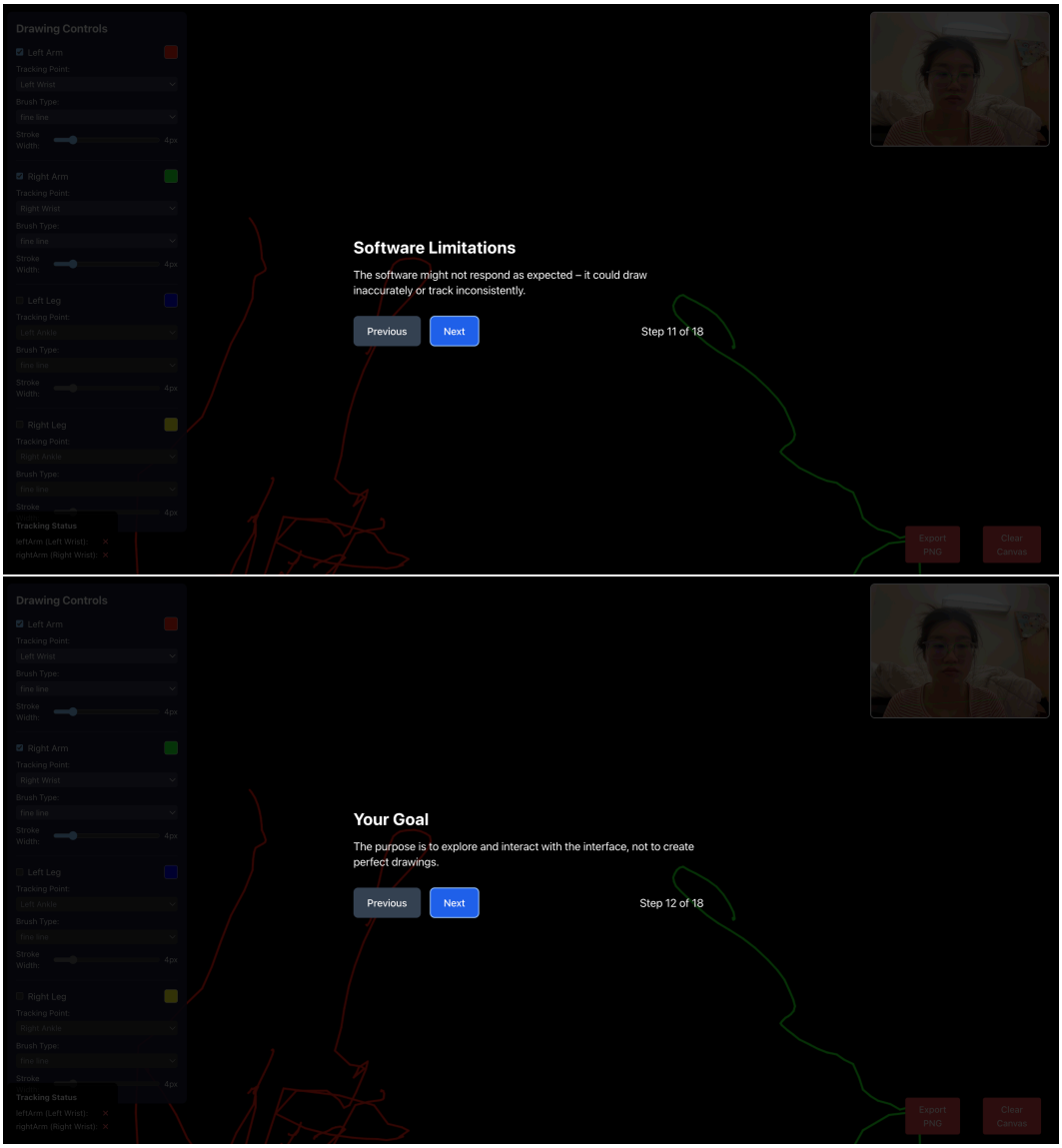

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist): ×
rightArm (Right Wrist): ×
Framing Your Body
Set up your camera so that your body is visible from below the knee (calves) and up.
Previous
Next
Step 6 of 18
Export PNG
Clear Canvas

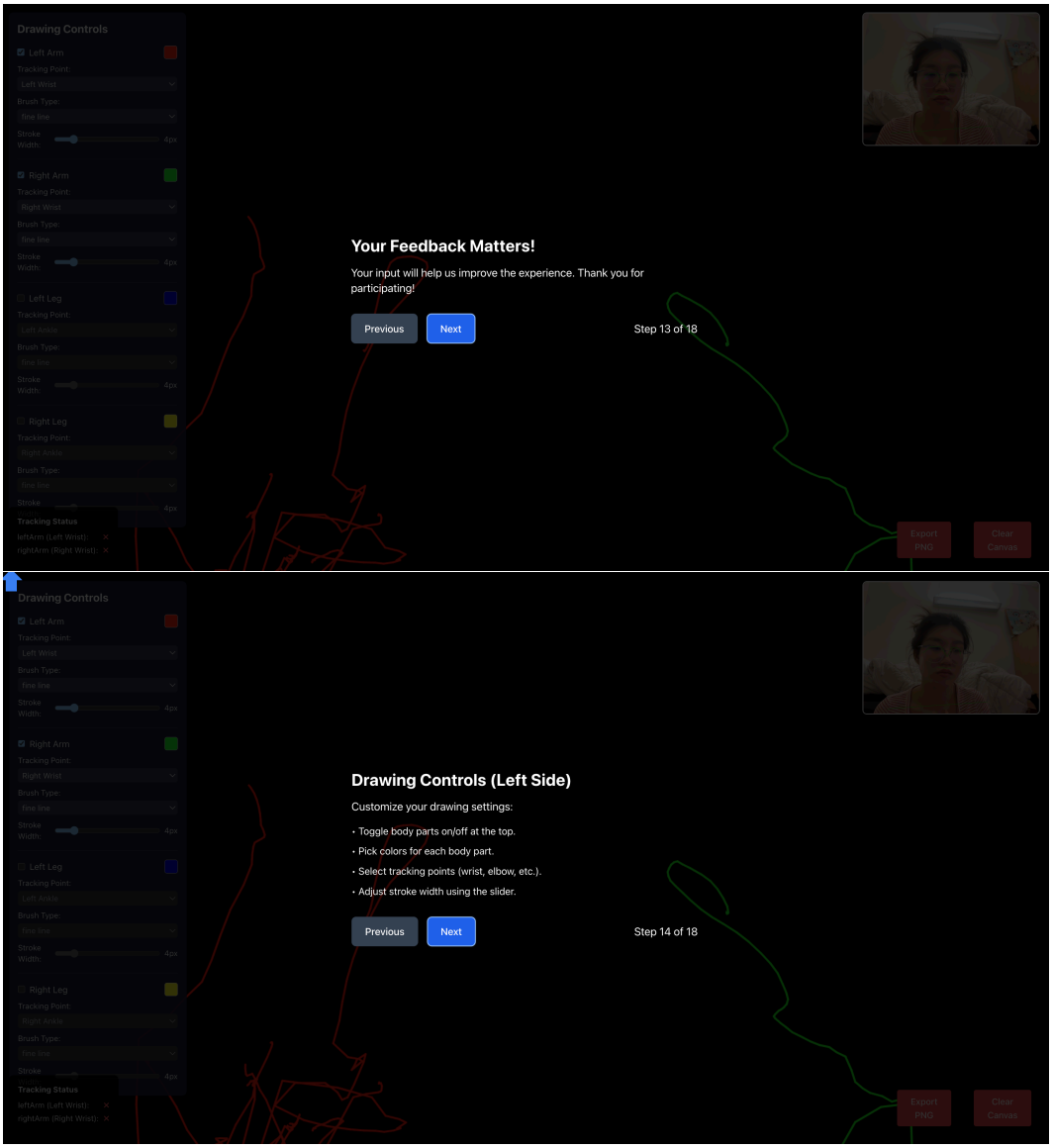

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist): ×
rightArm (Right Wrist): ×
Device Requirements
Use a laptop, enable your laptop camera, and open the application in full-screen mode.
Previous
Next
Step 7 of 18
Export PNG
Clear Canvas


Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist): ×
rightArm (Right Wrist): ×
Drawing Instructions
Ensure your entire body is within the frame when drawing.
Previous
Next
Step 8 of 18
Export PNG
Clear Canvas

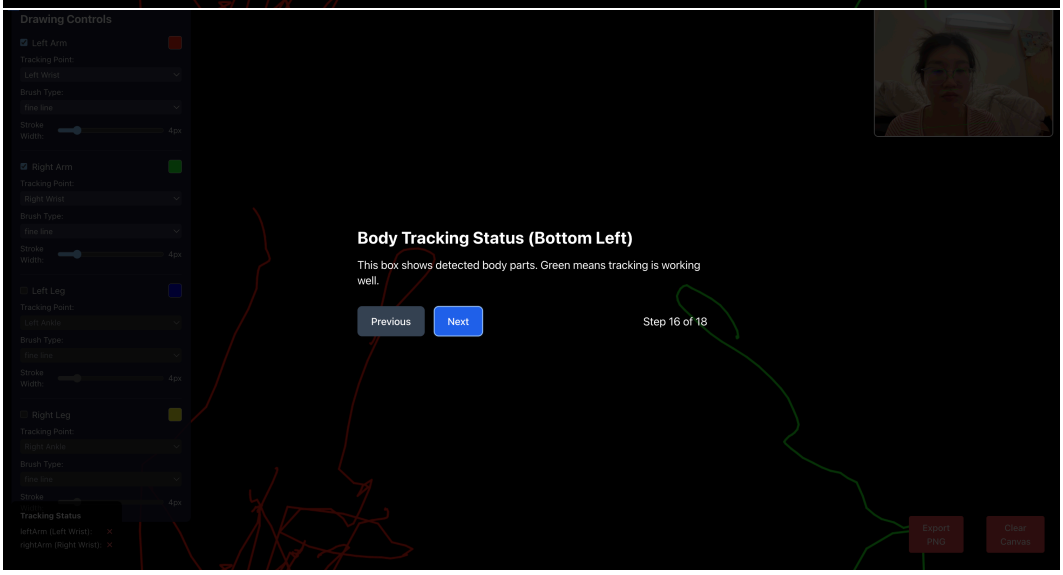

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist):
rightArm (Right Wrist):
Uploading Your Artwork
All of your artwork will be saved and requested for upload to a Google Form.
• Do NOT rename the PNG file when uploading.
Previous
Next
Step 9 of 18
Export PNG
Clear Canvas

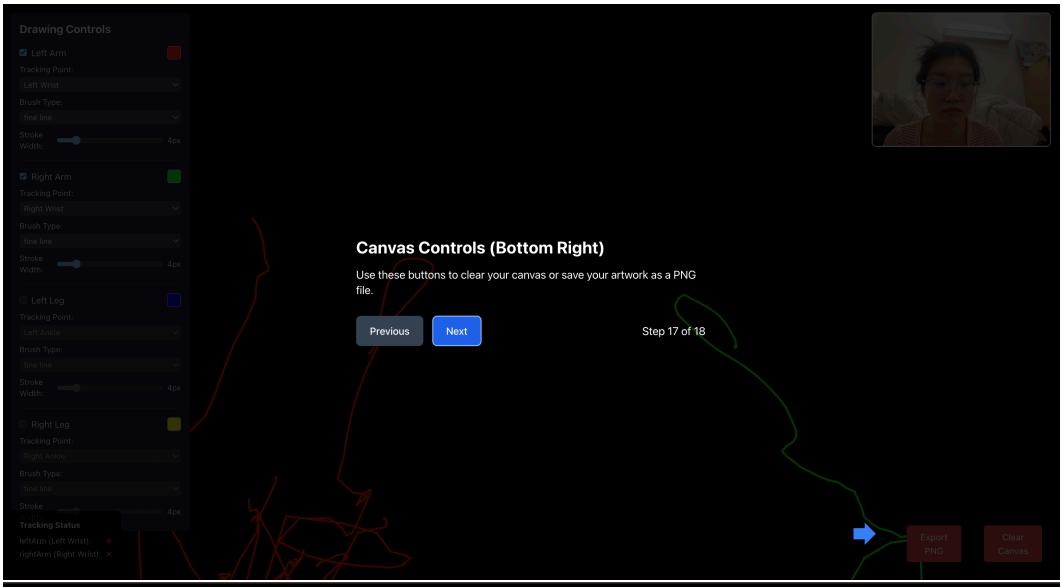

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist):
rightArm (Right Wrist):
Use of Your Artwork
Your submissions may be published in my thesis and related research documents.
Previous
Next
Step 10 of 18
Export PNG
Clear Canvas

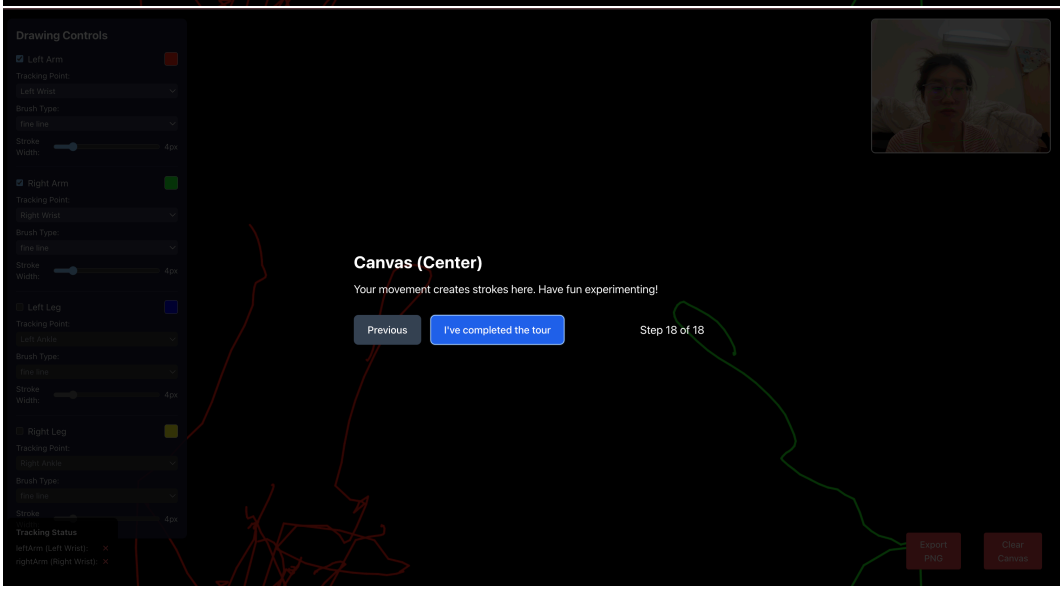

Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist):
rightArm (Right Wrist):
Software Limitations
The software might not respond as expected – it could draw inaccurately or track inconsistently.
Previous
Next
Step 11 of 18
Export PNG
Clear Canvas
Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist):
rightArm (Right Wrist):
Your Goal
The purpose is to explore and interact with the interface, not to create perfect drawings.
Previous
Next
Step 12 of 18
Export PNG
Clear Canvas


Drawing Controls
Left Arm
Right Arm
Left Leg
Right Leg
Tracking Status
Your Feedback Matters!
Your input will help us improve the experience. Thank you for participating!
Previous
Next
Step 13 of 18
Export PNG
Clear Canvas
Drawing Controls
Left Arm
Right Arm
Left Leg
Right Leg
Tracking Status
Drawing Controls (Left Side)
Customize your drawing settings:
• Toggle body parts on/off at the top.
• Pick colors for each body part.
• Select tracking points (wrist, elbow, etc.).
• Adjust stroke width using the slider.
Previous
Next
Step 14 of 18
Export PNG
Clear Canvas


Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist):
rightArm (Right Wrist):
Camera Preview (Top Right)
This window shows your movements in real-time. Ensure your body is fully visible in the frame!
Previous
Next
Step 15 of 18
Export PNG
Clear Canvas


Drawing Controls
Left Arm
Tracking Point:
Left Wrist
Brush Type:
fine line
Stroke Width:
4px
Right Arm
Tracking Point:
Right Wrist
Brush Type:
fine line
Stroke Width:
4px
Left Leg
Tracking Point:
Left Ankle
Brush Type:
fine line
Stroke Width:
4px
Right Leg
Tracking Point:
Right Ankle
Brush Type:
fine line
Stroke
4px
Tracking Status
leftArm (Left Wrist):
rightArm (Right Wrist):
Body Tracking Status (Bottom Left)
This box shows detected body parts. Green means tracking is working well.
Previous
Next
Step 16 of 18
Export PNG
Clear Canvas





Participants will follow the 18-step guided tour described above before engaging in two structured activity prompts:

Part 1: Draw basic geometric shapes with specific brush settings, color, body parts, and canvas position (e.g., circle with right wrist, in green, right side, 20px) to familiarize users with the interface.

PROMPT 1: Draw a GREEN CIRCLE with your RIGHT WRIST on the RIGHT SIDE OF THE CANVAS PROMPT 2: Draw a BLUE CIRCLE with your LEFT INDEX on the TOP OF THE CANVAS PROMPT 3: Draw a RED CIRFLE with your RIGHT KNEE with 10PX STROKE WIDTH PROMPT 4: Draw a YELLOW CIRCLE with your NOSE with 20PX STROKE WIDTH

Part 2: respond to an open-ended creative prompt. FREEFROM PROMPT: PAINT

A SCENE WITH A RAINBOW

Part 3: reflect on the experience by answering questions

4.3 Gesturism User Experience Questionnaire

4.3.1 Consent and Demographics

1. Consent Confirmation

Mark only one oval.

- Agree

By selecting "Agree", I confirm that I am at least 18 years old and have read and understood the information provided.

2. What is your dominant hand?

Mark only one oval.

- Right
- Left
- Both

3. Number of years you have used a digital painting application (e.g., Procreate, Adobe Fresco, etc.)

Mark only one oval.

- 0 years
- 1–2 years
- 3–5 years
- 5+ years

4. How familiar are you with digital painting interfaces?

Mark only one oval (1 = Not familiar, 5 = Advanced)

- 1
- 2
- 3
- 4
- 5

4.3.2 Prompt Completion Confirmation

5–9. Uploads and completion confirmation for each prompt:

- Prompt 1: Draw a GREEN CIRCLE with your RIGHT WRIST on the RIGHT SIDE of the canvas
- Prompt 2: Draw a BLUE CIRCLE with your LEFT INDEX on the TOP of the canvas
- Prompt 3: Draw a RED CIRCLE with your RIGHT KNEE using a 10px stroke
- Prompt 4: Draw a YELLOW CIRCLE with your NOSE using a 20px stroke
- Freeform Prompt: Paint a scene with a rainbow

For each prompt, users are asked:

- Have you completed this prompt?

Mark only one oval.

- Complete

4.3.3 Usability and Interface Interactions

10. Were the given instructions/guided prompts clear?
Mark only one oval (1 = Not clear at all, 5 = Very clear)
- 1
- 2
- 3
- 4
- 5

11. Which body part was hardest to use for drawing?
Mark only one oval.
- Left Arm components
- Right Arm components
- Left Leg components
- Right Leg components
- Head components

12. Which part of the canvas was hardest to reach?
Mark only one oval.
- Top
- Bottom
- Middle
- Left
- Right

13. Was the export process (saving your artwork) easy to understand and use?
Mark only one oval.
- Yes
- No

4.3.4 Customization and Enjoyment

14. If No, please explain:
Open-ended response

15. Were customization options (brush size, color, stroke width) easy to use?
Mark only one oval.
- Yes
- No

16. If No, please explain:
Open-ended response

17. How enjoyable was the "scene with a rainbow" drawing experience?
Mark only one oval (1 = Very unintuitive, 5 = Very intuitive)
- 1
- 2
- 3
- 4
- 5

4.3.5 System Responsiveness and Suggestions

18. How would you rate the responsiveness of the system?
Mark only one oval (1 = Very unresponsive, 5 = Very responsive)
- 1
- 2
- 3
- 4
- 5

19. Was there any part of painting the rainbow scene that was frustrating?
Open-ended response

20. What feature did you like the most about Gesturism?
Open-ended response

21. What part of the user interface or experience would you improve?
Open-ended response

22. Do you have any additional feedback or suggestions for improving Gesturism?
Open-ended response

4.4 Data Collection

To track participant compliance, they will be asked to export their painting after each prompt. Each exported artwork file will be automatically assigned an encoded label based on the tracked body part and prompt (e.g., drawing with the right wrist generates 16.png). This ensures participant anonymity while allowing validation of task completion. The exported artwork will be submitted to DropBox file request links attached in a Google Form, along with quantitative (Likert-scale) and qualitative (open-ended) feedback on usability, instruction clarity, and user experience. Participant ratings of familiarity with other digital painting software will also be collected. No personal identifying information will be collected in the Google Form. The artwork files' content and metadata do not contain location, IP address, or any personal identifying information.

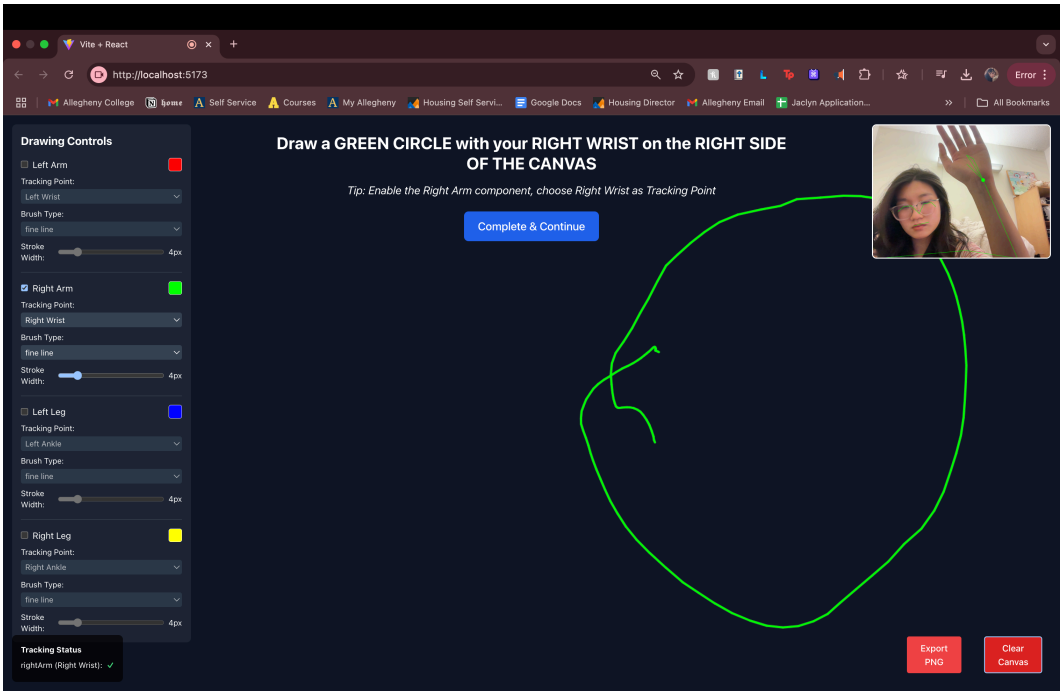


Figure 26: Sample of saved artwork filename

The data from the anonymous Google form will be collected and stored using the principal researchers' Allegheny College Google Workspace account. Principal researchers will be the only two with access to the original Google form, but the participant responses, including artwork, may be shared in thesis writing and research dissemination materials. The Google accounts that do have access to the original form data are password protected. The computers belonging to the PIs are also password protected. Participants will not have access to their individual results. The results of the study will be included in senior thesis document writing. The full data set will not be shared.

4.5 Recruitement

Recruitment occured via flyers and outreach to the ACM Student Chapter at Allegheny College. Flyers will be posted in public locations on Allegheny's campus. The recruitment material will contain a link to the Gesturism web application. The recruitment material will list participant inclusion criteria.

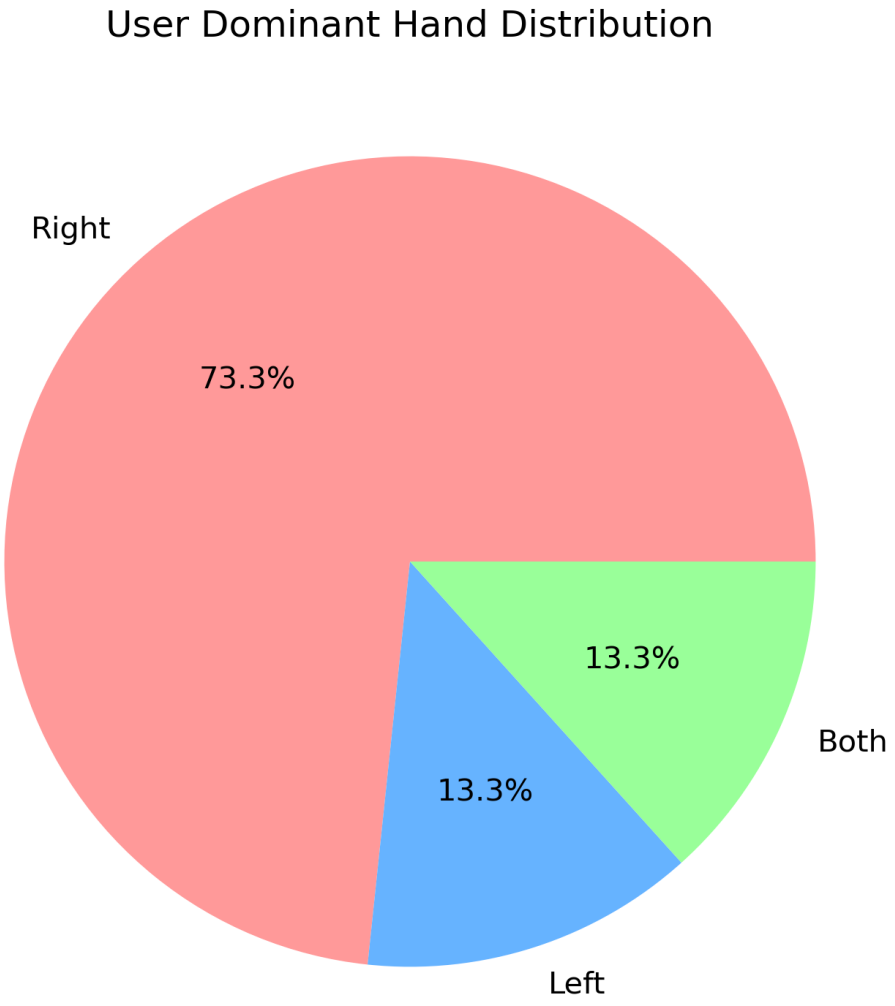
4.6 Inclusion Requirments

Participants must be 18+ years old and have a laptop with internet connection. The participants must be able to move comfortably and safely in daily life, and must be able to follow instructions in English.

4.7 Evaluation

4.7.1 Participants

We received fourteen participants, addressing each individual's experience of Gesturism. The user demographic consists primarily of right-handed participants (73%), with a small portion identifying as left-handed (13%) or ambidextrous (13%).



The majority of users (60%) reported having no prior experience with digital painting applications such as Procreate or Adobe Fresco, while 33% had 1-2 years of experience, and only 7% had used such applications for 3-5 years.

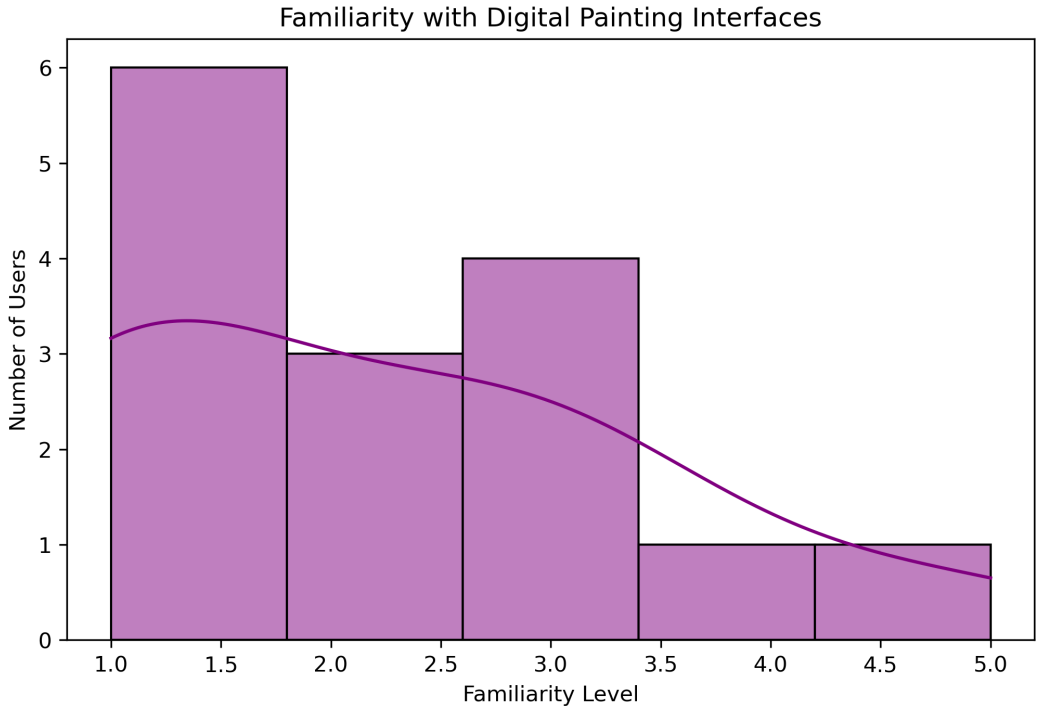


Figure 27: User Experience Distribution in Digital Painting

In terms of familiarity with common digital painting interfaces, responses varied widely: 40% of users rated themselves at the lowest familiarity level (1), while 26% selected level 3, and the remaining participants were more evenly distributed across levels 2 (20%), 4 (7%), and 5 (7%). This suggests that the user base skews toward beginners, with only a small percentage having significant prior experience in digital painting.

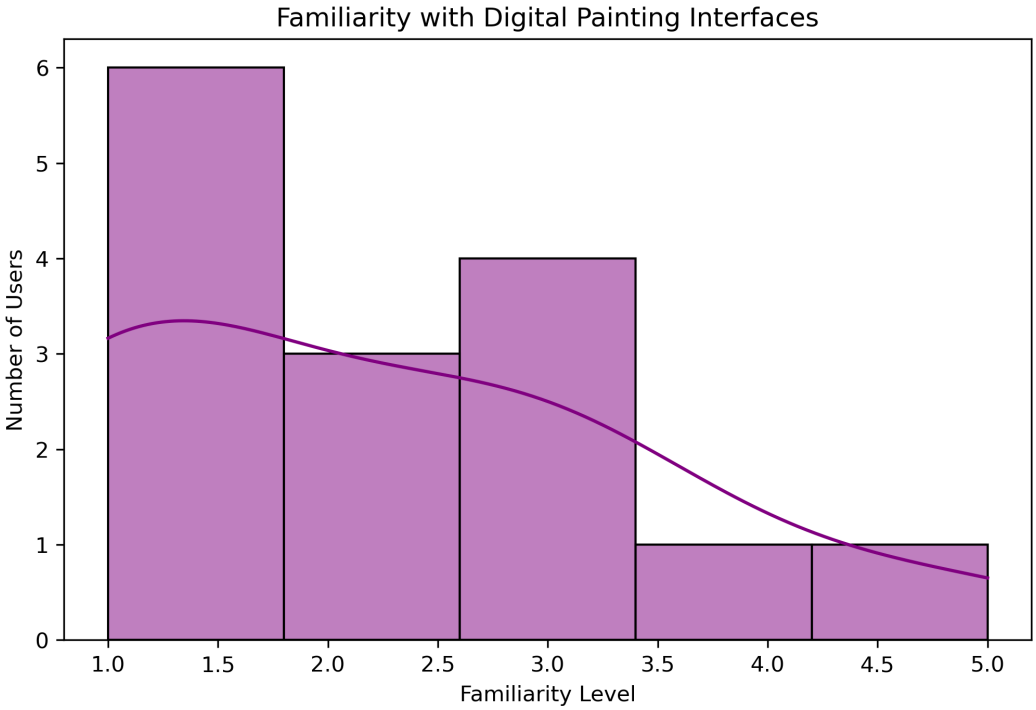


Figure 28: Familiarity with Digital Painting Interfaces

I wanted to test my hypothesis that there is a correlation between user's familiarity with traditional digital painting interfaces (i,e. Procreate, Adobe Sketch) with liking this new medium of creation. However, after performing the Pearson Correlation, the analysis of the relationship between users' familiarity with common digital painting interfaces and their enjoyment of the Gesturism experience revealed a moderate positive correlation ($r = 0.39$). Since this correlation was not statistically significant ($p = 0.1563$), which may be attributed to the limited number of participants, we cannot confidently conclude a strong relationship between these two factors based on our current sample.

The demographic profile suggests a participant base that is relatively inexperienved with conventional digital art platforms, underscoring the necessity for clear onboarding procedures and suggesting an opportunity for *Gesturism* to serve as an accessible entry point into digital creativity.

4.7.2 Results

With the implementation of the system's 18-step guided tour to reduce usage confusion, every participant successfully completed the tour and all five prompts, overally demonstrating the accessibility and functionality of the interface despite general unfamiliarity with related tools. Hence, the base line usability in the design is achieved. Differences in execution quality between prompts reveal some subtler insights into how bodily movement interacts with interface design.

In evaluating the artwork results, prompts involving upper-body joints such as right wrist and left index consistently yeileded cleaner, centralized visual compositions. In con-

trast, prompts requireing the use of the knee or nose reflect both ergonomic constraints and the technical sensitivity of the computer vision system in tracking peripheral landmarks. While participants attempted to remain visible to the camera, several noted frustration with gestures failing to register when using lower limbs or when facing slightly away from the camera. These limitations also show in written feedback, with 6 out of 14 participants citing issues with calibration and gesture fidelity.

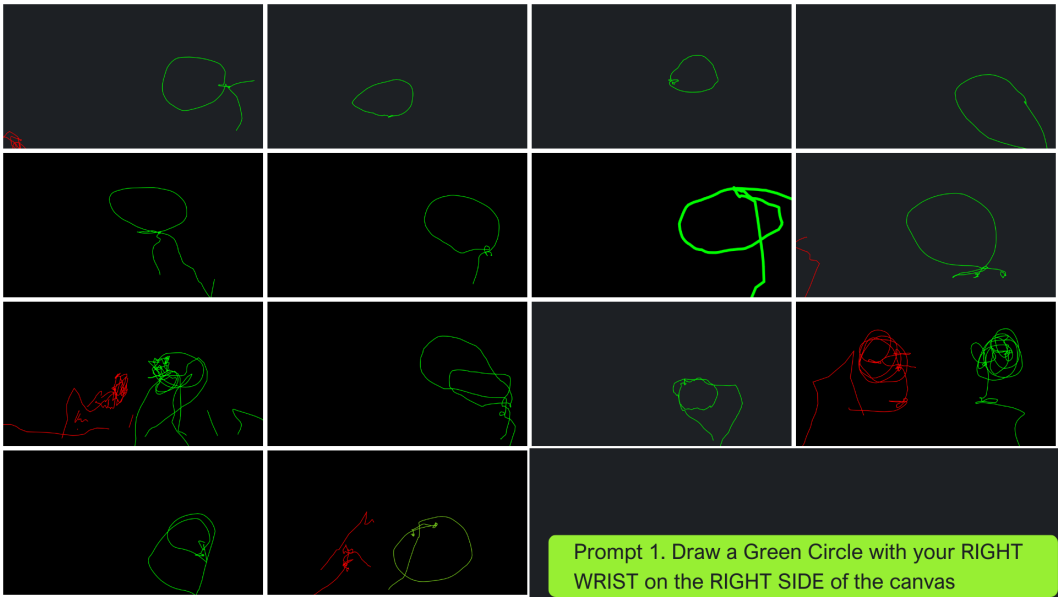


Figure 29: Prompt 1. Draw a Green Circle with your RIGHT WRIST on the RIGHT SIDE of the canvas

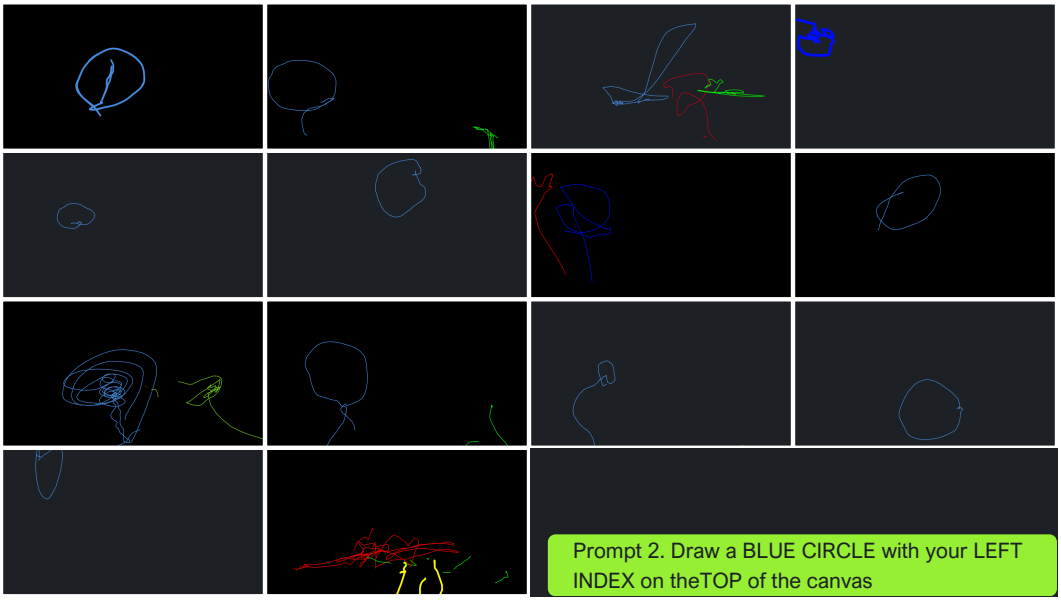


Figure 30: Prompt 2. Draw a BLUE CIRCLE with your LEFT INDEX on theTOP of the canvas

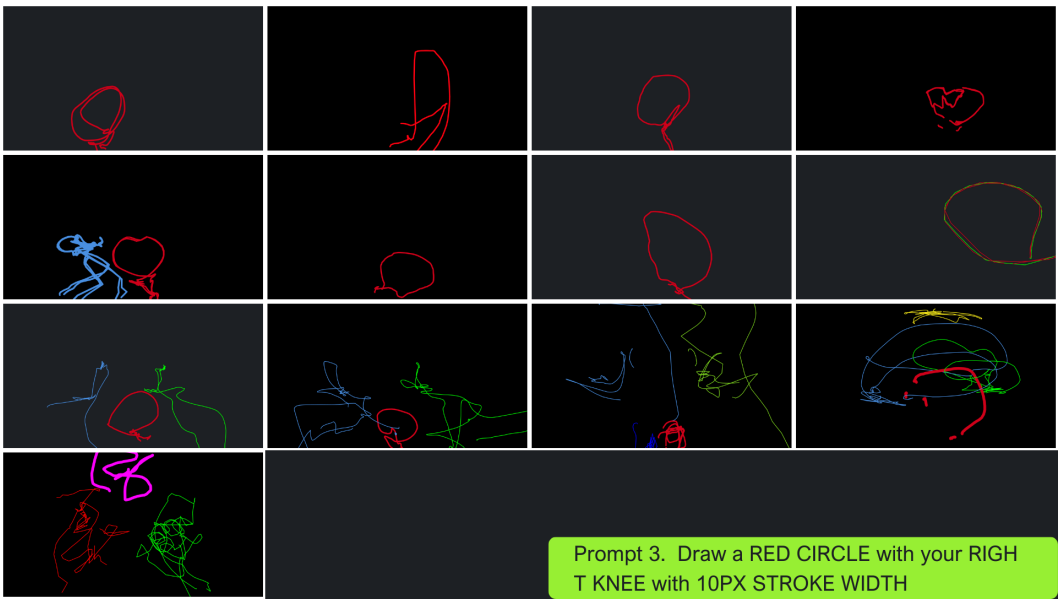


Figure 31: Prompt 3. Draw a RED CIRCLE with your RIGHT KNEE with 10PX STROKE WIDTH

Regarding user experience, the majority of responses convey a high level of satisfacetion with the overall experience. When asked to rate the clarity of guided prompts on a 5-point

Likert scale, 79% of participants rated the instructions as either a 4 or 5, and no rate below 3.

Furthermore, 93% of users agreed that customization features such as color wheel, brush width slider and body part selection were intuitive to use. Qualitative feedback described the interface as “clean”, “engaging” and “easy to navigate”.

However, the clarity of certain interface features received mixed responses. While most participants adapted to the customization settings, three noted that the location of controls (e.g, stroke width slider, landmark toggle) was not immediately apparent and required trial and error.

The freeform rainbow scene prompt served as a litmus test for open-ended exploration. All participants submitted work for this task, and feedback was positive. Eleven of the fourteen described the task as “enjoyable” or “fun”, and 10 out of 14 rated as a 4 or 5 in terms of intuitiveness. Interestingly, even participants who had expressed minor frustration in the guided prompts described the freeform task as “freeing” or “expressive”. This distinction suggests that *Gesturism* thrives as a system of open experimentation, opening the possibilities creative improvisation with technology advancement.

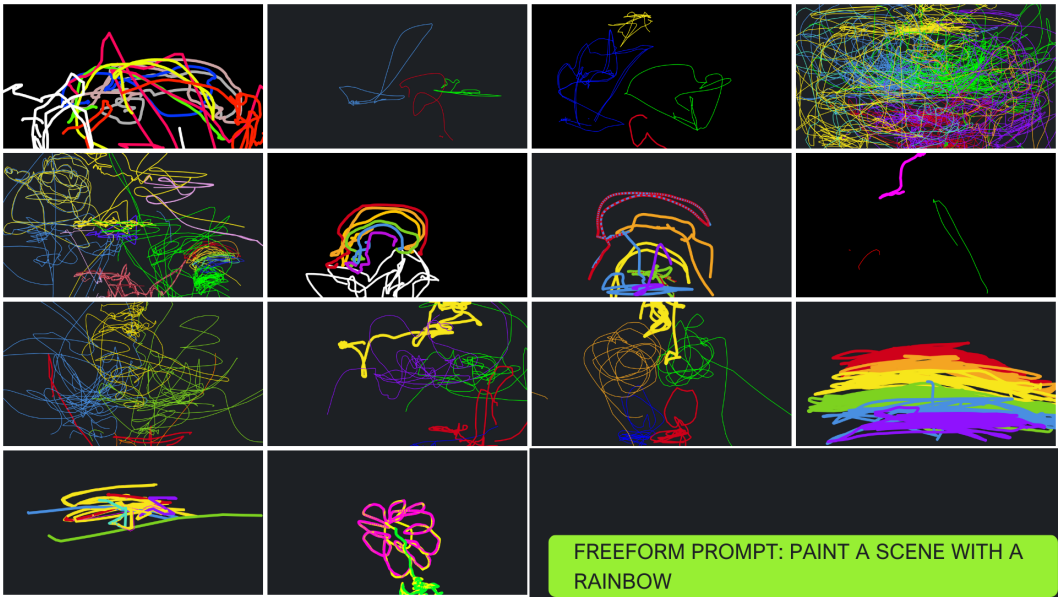


Figure 32: FREEFROM RESPONSES. PAINT A SCENE WITH A RAINBOW

When asked about the system’s responsiveness, 85% of users rated it as a 4 or 5, though multiple users observed a performance dip when switching setting smed-drawing or using fast, abrupt movements. These issues point to computational bottlenecks that could be addressed thorugh algorithm optimization or asynchronous UI rendering. Additonal performance-related concerns stemmed from physical space constraints: several users noted the difficulty of drawing near the bottom of the canvas due to their webcam height or limited room space, conditions that adversely affected knee-based and full-body prompts.

From a sentiment analysis standpoint, the tone of participant feedback leaned toward positive. The majority of open-ended responses hightlighted novelty, embodiment, and

personal expression. Some feedback phrases such as “I liked feeling the movement”, “it made me want to dance” reflect teh resonance between user experience and the system’s conceptual ambition. At the same time, the most frequently cited negative sentiments such as “frustrated”, “confused”, and “annoyed” were not directed at the creative aspect, but at specific functional breakdowns, involving gesture misrecognition or unclear calibration feedback.

These findings offer a layered portrait of *Gesturism*’s interaction design. The system achieves its goal of expanding digital painting into an embodied modality, yet it reveals limitations of gesture recognition threshold, interface scaffolding and physical environment adaptation. The study confirms that users are willing to tolerate a degree of unpredictability when the system delivers creative agency and expressive output.

5 Threats to Validity

The evaluation is subjected to several potenttial threats to validity, which mush be acknowledged to contextualize the findings. The participant pool for this study was primarily college students recruited through flyters and department outrech within a single academic institution. This demographic skew limits the generalizability of the results, as participants may share similar technological literacy, educational background, or cultural context. Additionally, although inclusion criteria focused on basic motor mobility and English comprehension, the sample lacked diversity in age, artistic experience, and cultural familiarity with gestural expression. These factors may have influenced participants' comfort with movement or interpretation of the prompts. In addition, the study was conducted in uncontrolled environments, typically participants' personal spaces, introducing variability in camera quality, room lighting, and available physical space. Since the system's pose estimation relies on real-time computer vision, environmental inconsistencies could result in unevent system performance.

6 Conclusion

6.1 Conclusion

Gesturism explores how larger bodily gestures like arms, legs, head can be captured in real time using computer vision, and translated into visual mark-making. Based on both strucutred feedback and open-ended responses from participants, this system successfully opens new pathways for embodied creativity. Many users described the expereice as playful and surprisingly intuitive. For users without a background in digital art, the system lowered the barrier to entry by offering a direct and physical way to engage with digital media. For users with prior experience, *Gesturism* was seen as a gresh and freeing alternative to traditional methods, especially when combined with the ability to track different limbs or adjust stroke dynamics.

The study demonstrates the value of embodied interfaces not just as a technical novelty but as a meaningful creative tool. Participant responses reveal taht *Gesturism* invites users to consider their body as a medium and not a controller, expanding the sense of authorship beyond hand-eye coordination to full-body expression. As a system, it has both practical and conceptual implications that allow users to reimagine how gestures can be visualized and preserved.

6.2 Visual Design

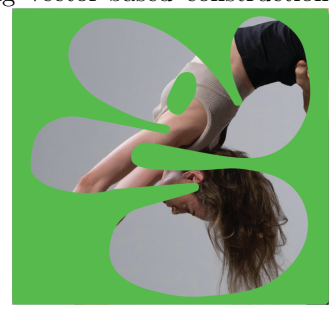
The visual identity of *Gesturism* is currently under development as an extension of the technical implementation. The concept of *Gesturism* explore the fluid motion ethos. Through iterative design exploration using Adobe Illustrator, I developed a cohesive identity system grounded in both design principles and theoretical underpinings of movement-based creativity.

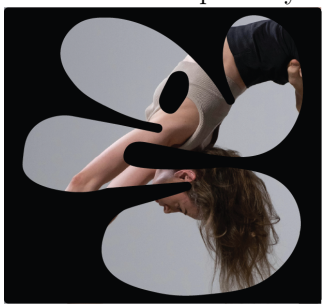
The core logo design features a custom wordmark integrated with an abstract glyph that references a gestural stroke. The glyph forms the lowercase letter “g” and the organic aesthetic of movement.



Figure 33: Gesturism logo breakdown

The logo was prototyped and finalized in Illustrator using vector-based construction to maintain adaptability across screen sizes and interfaces.





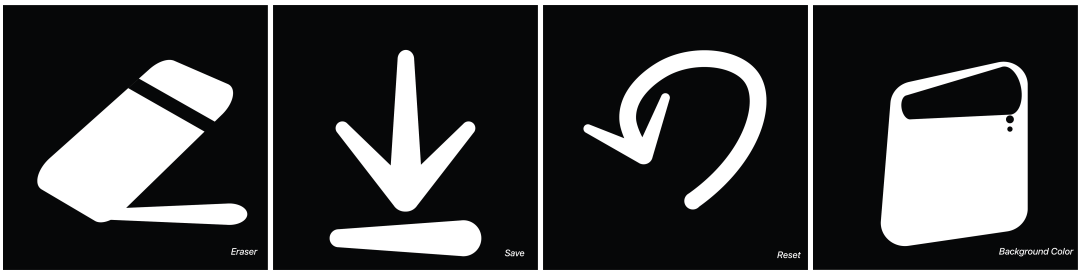


Color played a key role in translating the experience of embodied creativity into visual terms. A balance palette was established that includes high contrast monochromes (black and white) with a series of accent hues of seafoam green (#71CAC9), azure blue (#68C5DD), and lime green (#59BA4B). The inclusion of grayscale midtones ground the palette for a neutral backdrop that allows gesture-driven strokes to remain in the visual foreground.



Figure 34: Color Palette

The chosen typeface for headers is geometric yet soft, with rounded terminals that remain clean and modern. I also designed a custom iconography system that depict stylized limbs and directional actions (e.g, save, reset, menu). These icons are rendered in bold, monochrome silhouettes to reduce cognitive loads for users navigating the interface via body-based control.



6.3 Exhibition

The *Gesturism* exhibition was designed to be an immersive, participatory experience in which audience members could step into the interface and engage with the painting system directly. The installation used a webcam to track the users' body, projecting the generated visual output onto the opposite wall. The looping feedback between human movement, computer vision, and projected outcome was engineerd to create a live, responsive visual field that blurred the line between viewers and creator.

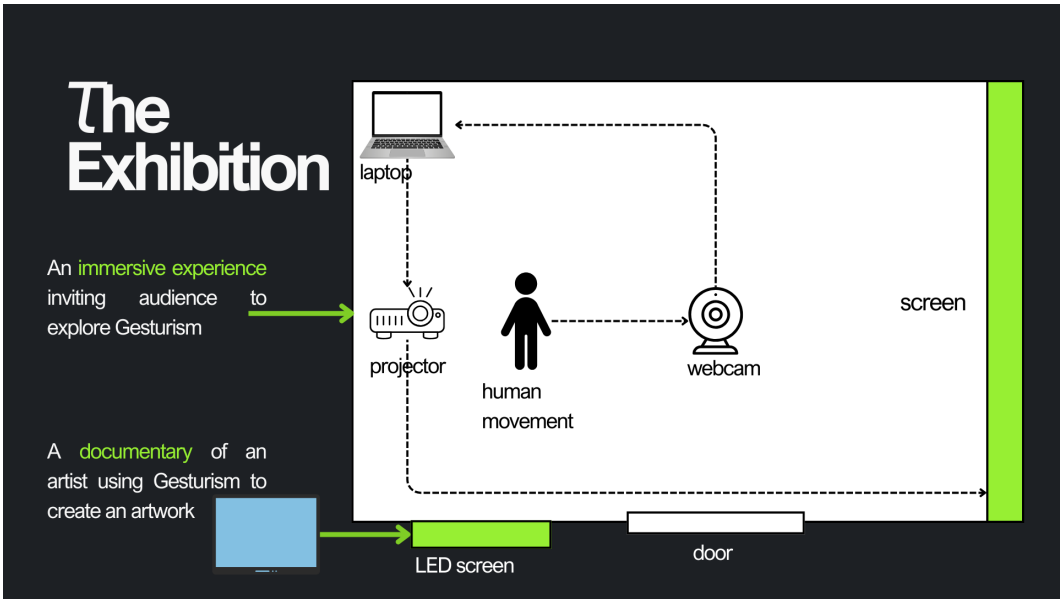


Figure 35: Exhibition setup

A short documentary was also featured in the space to invite and prompt users to enter the immersive exhibtiion. The documentation featured a painter interacting with *Gesturism*, reflecting on the discussion of movement contraints in modern society.

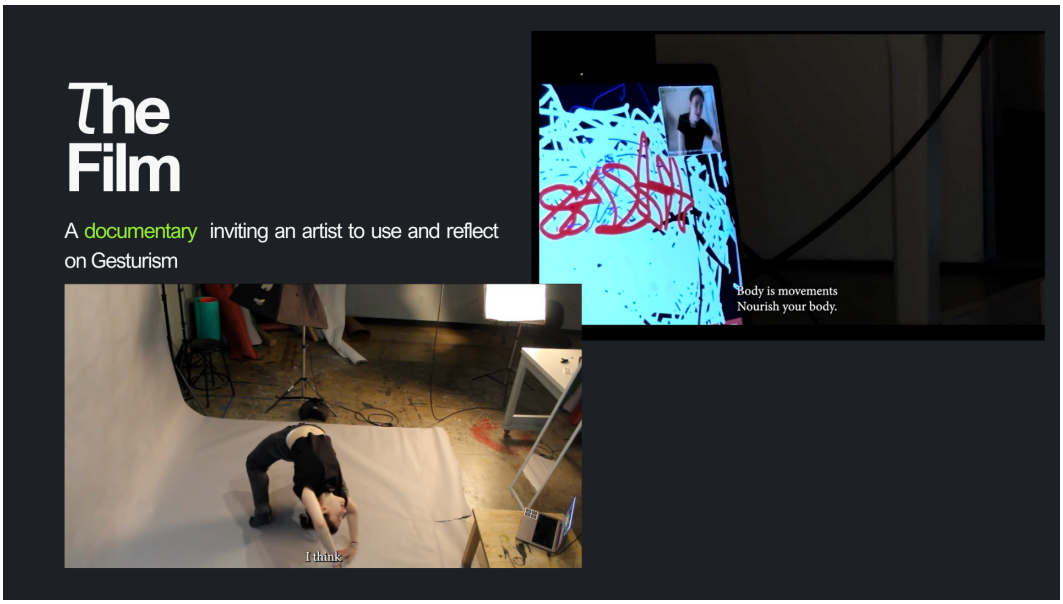
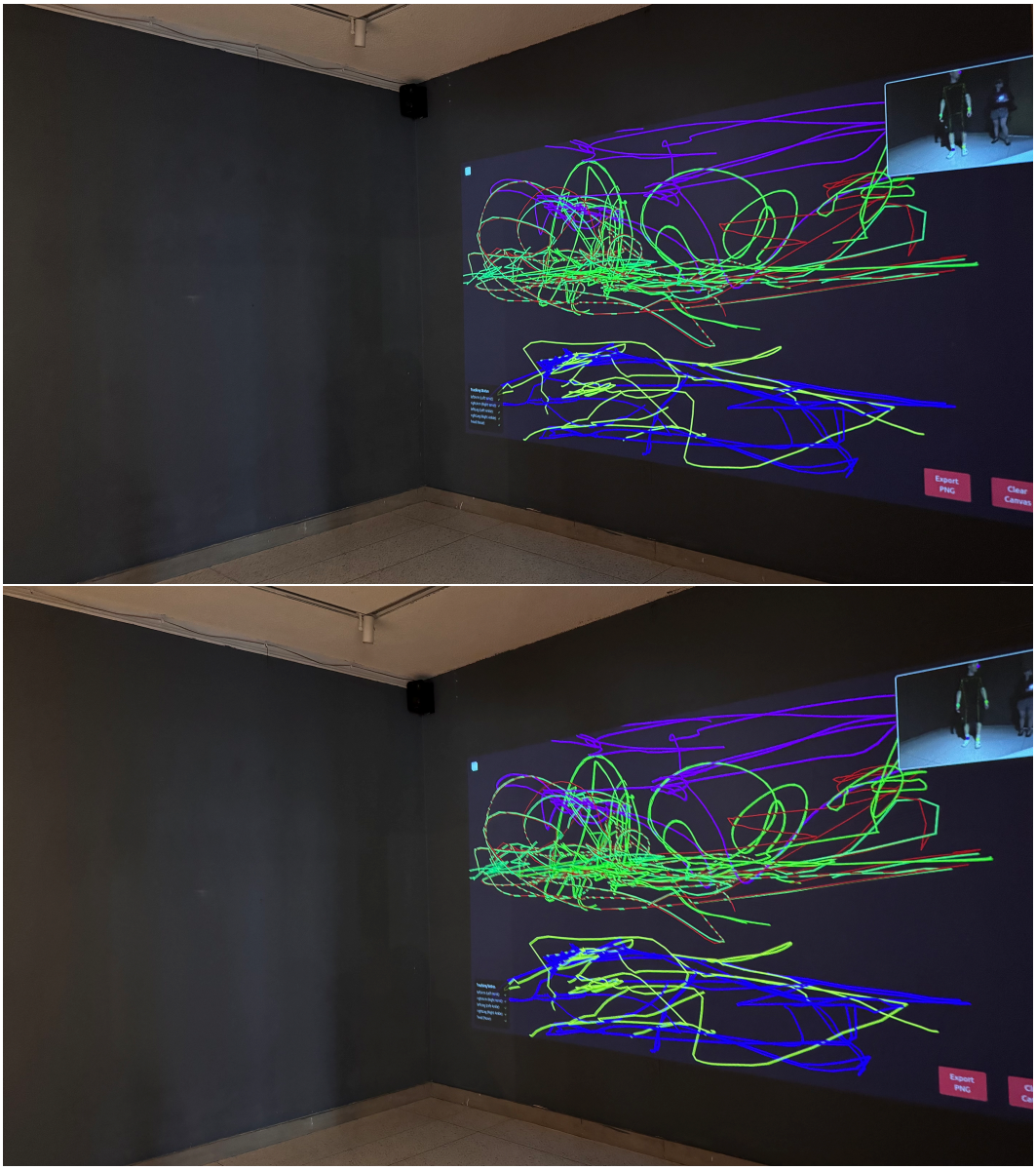
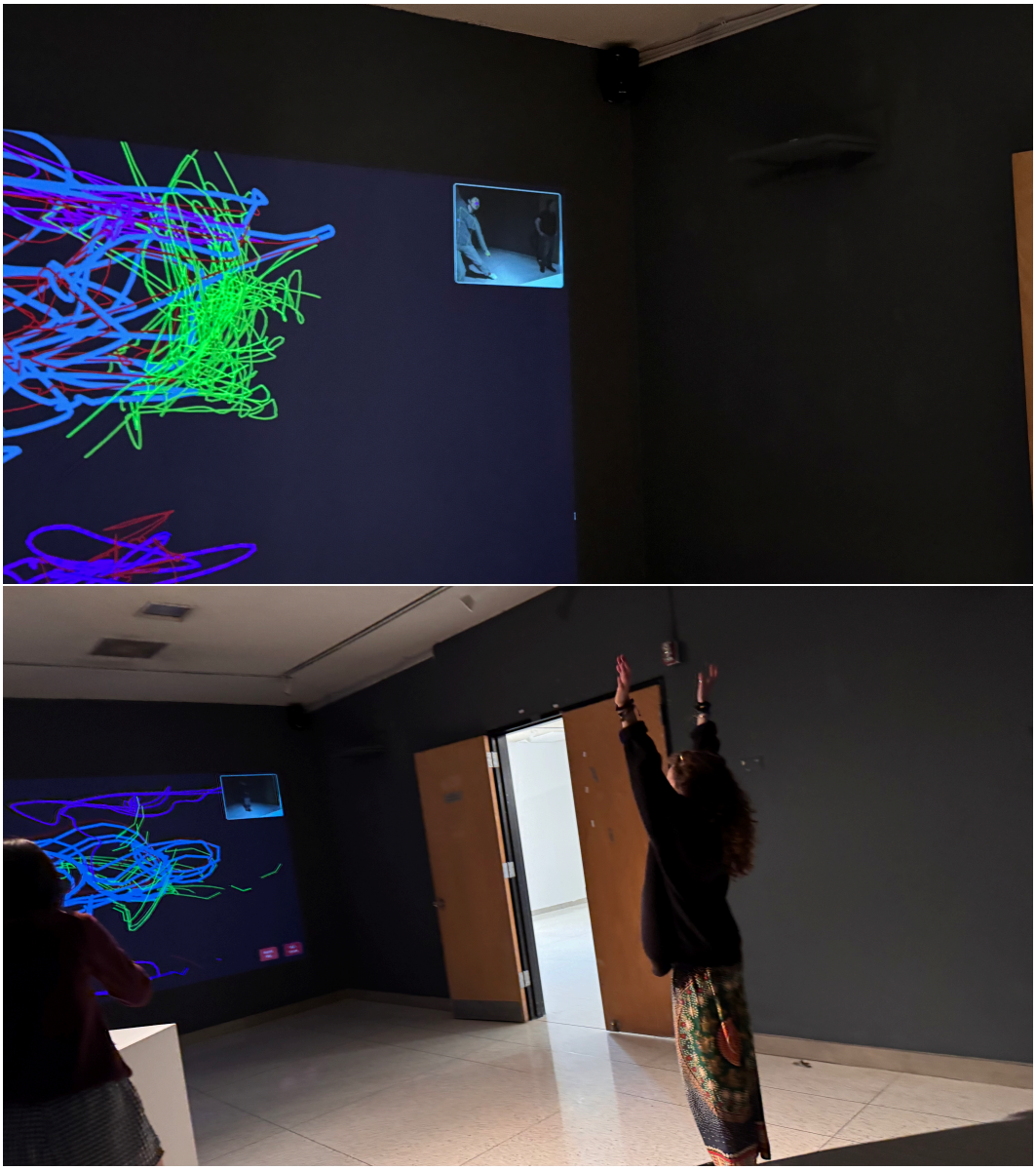


Figure 36: The documentation

Participant responses during the exhibition highlighted distinct patterns in engagement based on their background. Dancers explored fluidity and repetition, treating the interface as an extension of choreographic language. They often exhibition the limitation of the landmark tracking component with variants of complex postures and movement. Artists focused more on composition and control, often revisitng the interface to test layering and color blending. They are particularly interested about the accuracy of movement tracing in order to maintain the accurate aesthetic vision. Athletes and non-artists tended to test the limits of the system by exaggerating motion or testing the speed. These divergent modalities of interaction reinforced *Gesturism* potential of different development paths.


Export PNG
Clear Canvas
Export PNG





6.4 Future Work

6.4.1 Advancing UI Design

Currently, Gesturism's interface is designed with a strong emphasis on functionality. However, to align with user expectations, future work will retain its minimalistic ethos but introduces key improvements aimed at enhancing interaction clarity and functional accessibility. New features considered **expandable menu system** to minimize interface clutter during drawing sessiong, **real-time camera preview window** for body alignment and calibration. Alternatively, **eraser and background color toggles** would be integrate to expand functionalities.

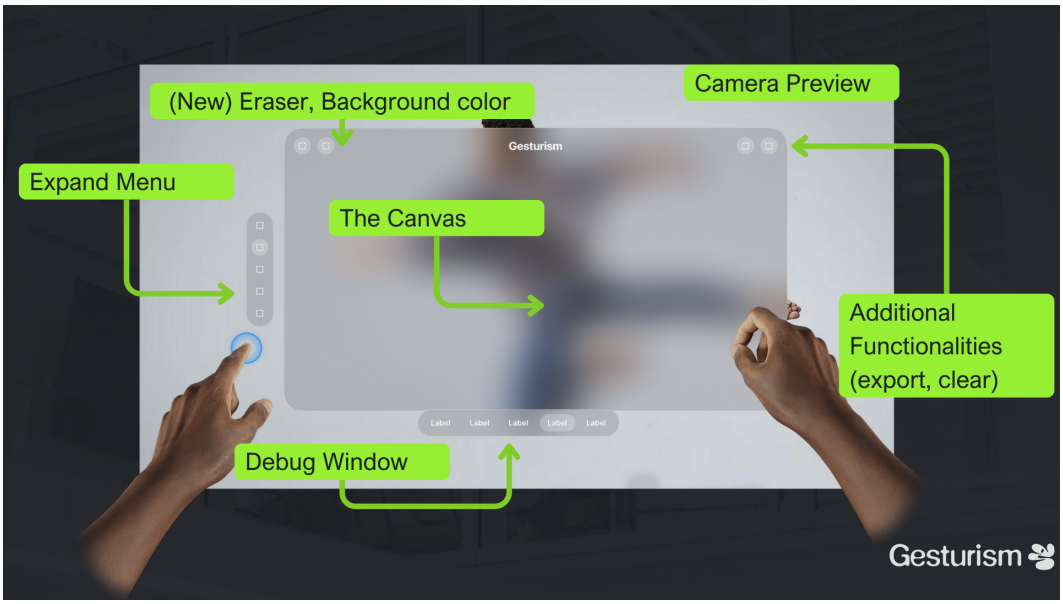


Figure 37: New proposed user interface

6.4.2 Enhancing Technical Accuracy

To improve the reliability of gesture-based painting, future iterations will focus on refining tracking accuracy based on user feedback. One major development will be the implementation of an eraser functionality, allowing users to modify their artwork dynamically. Furthermore, gesture recognition will be expanded to support interaction with key menu elements, such as opening the color palette or saving the image. These enhancements aim to make the system more intuitive and responsive to users' creative intentions.

6.4.3 Expanding Exhibition Potential

Insights gained from faculty discussions during the midterm review highlight the potential for integrating depth sensors into Gesturism. This development would allow for spatial painting, enabling users to create artwork in three-dimensional space rather than on a flat canvas. Another key enhancement is the introduction of realistic brush stroke textures, which will provide a more authentic and expressive painting experience, further bridging the gap between digital and traditional art.

6.4.4 Back-End Implementation

A robust back-end infrastructure will be essential for Gesturism's long-term usability. Future work will include user authentication to facilitate personalized experiences and cloud-based storage for artwork preservation. Additionally, a screen recording feature will be introduced to capture the painting process, offering valuable insights for artists and educators. The system will also explore different modes tailored for distinct user groups, such as dancers and painters, expanding Gesturism's versatility beyond static visual art.

6.5 Ethical Considerations

As Gesturism develops into a more advanced tool, ethical implications must be addressed. Future research will explore issues such as data privacy, especially concerning gesture tracking and user authentication. Additionally, considerations around accessibility and inclusivity will be prioritized to ensure that the system remains usable for diverse populations, including individuals with disabilities. Another key aspect involves assessing the impact of AI-driven gesture recognition on creative agency, ensuring that technological mediation does not compromise artistic expression. By addressing these future directions, Gesturism will continue to refine its capabilities, enhance its artistic potential, and establish itself as a valuable tool at the intersection of technology and creative expression.

References

[1] 2003. https://search.worldcat.org/title/1378381428

[2] 2015. https://react.dev/

[3] 2018. *Tensorflow.org* (2018). https://blog.tensorflow.org/2018/05/real-time-human-pose-estimation-in.html

[4] 2020. https://www.flong.com/archive/projects/messa/index.html

[5] 2024. https://vite.dev/guide/

[6] 2025. https://tailwindcss.com/

[7] 0xStc. 2021. Yellowtail (1998-; SVG/NFT version, 3.2021) – Golan Levin. https://www.creativeapplications.net/nft/yellowtail-1998-svg-nft-version-3-2021-golan-levin/

[8] G Annum. 2014. Digital painting evolution: A multimedia technological platform for expressivity in fine art painting. *Journal of Fine and Studio Art* 4 (2014), 1–8. https://doi.org/10.5897/JFSA2014.0025

[9] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, et al. 2020. BlazePose: On-device Real-time Body Pose Tracking. In *arXiv preprint arXiv:2006.10204*.

[10] Anthony M. Blatner, James A. Ferwerda, Benjamin A. Darling, and Reynold J. Bailey. 2011. TangiPaint: A Tangible Digital Painting System. *Color and Imaging Conference* 19 (01 2011), 102–107. https://doi.org/10.2352/cic.2011.19.1.art00022

[11] Gary Bradski. 2000. The Opencv Library. *Doctor Dobbs Journal* 25, 11 (Nov 2000). https://www.researchgate.net/publication/233950935_The_Opencv_Library

[12] M. Csikszentmihalyi. 2009. *Creativity: Flow and the Psychology of Discovery and*. HarperCollins. https://books.google.com/books?id=aci_Ea4c6woC

[13] Emily Frith, Stephanie Miller, and Paul D. Loprinzi. 2020. A Review of Experimental Research on Embodied Creativity: Revisiting the Mind–Body Connection. *The Journal of Creative Behavior* 54, 4 (2020), 767–798. https://doi.org/10.1002/jocb.406 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jocb.406

[14] Emily Frith and Stephanie E Miller. 2024. Creativity in motion: examining the impact of meaningful movement on creative cognition. *Frontiers in Cognition* 3 (08 2024). https://doi.org/10.3389/fcogn.2024.1386375

[15] Erving Goffman. 1974. Frame Analysis: An Essay on the Organization of Experience. *Philosophy and Phenomenological Research* 39, 4 (1974), 601–602. https://doi.org/10.2307/2106908

[16] Lise Amy Hansen and A Morrison. 2014. Materializing Movement-Designing for Movement-Based Digital Interaction. https://www.semanticscholar.org/paper/Materializing-Movement-Designing-for-Movement-Based-Hansen-Morrison/34f1f37d25218dc10ebe44e6450634cbbda7f64e

[17] Myron W. Krueger. 1977. Responsive environments. In *Proceedings of the June 13-16, 1977, National Computer Conference* (Dallas, Texas) *(AFIPS '77)*. Association for Computing Machinery, New York, NY, USA, 423–433. https://doi.org/10.1145/1499402.1499476

[18] Myron W. Krueger, Thomas Gionfriddo, and Katrin Hinrichsen. 1985. VIDEOPLACE—an artificial reality. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Francisco, California, USA) *(CHI '85)*. Association for Computing Machinery, New York, NY, USA, 35–40. https://doi.org/10.1145/317456.317463

[19] Golan Levin. 2006. Computer Vision for Artists and Designers: Pedagogic Tools and Techniques for Novice Programmers. *AI and Society* 20, 4 (2006), 462–482. https://doi.org/10.1007/s00146-006-0049-2

[20] Heng Liu, Yao Xu, and Feng Chen. 2023. Sketch2Photo: Synthesizing photo-realistic images from sketches via global contexts. *Engineering Applications of Artificial Intelligence* 117 (2023), 105608. https://doi.org/10.1016/j.engappai.2022.105608

[21] Alexander Mathis, Pranav Mamidanna, Taiga Abe, Kevin Cury, Venkatesh Murthy, Mackenzie Mathis, and Matthias Bethge. 2018. Markerless tracking of user-defined features with deep learning. (04 2018). https://doi.org/10.48550/arXiv.1804.03142

[22] George Papandreou, Tyler Zhu, Nima Kanazawa, et al. 2018. PersonLab: Person Pose Estimation and Instance Segmentation with a Bottom-Up, Part-Based, Geometric Embedding Model. In *Proceedings of the European Conference on Computer Vision (ECCV)*.

[23] Ilene Serlin. 2013. Maxine Sheets-Johnstone: The Primacy of Movement—Expanded Second Edition (Advances in Consciousness Research). *American Journal of Dance Therapy* 35 (12 2013). https://doi.org/10.1007/s10465-013-9164-z

[24] Richard Szeliski. 2010. *Computer Vision: Algorithms and Applications* (1st ed.). Springer-Verlag, Berlin, Heidelberg.

[25] Richard Szeliski. 2016. *Computer Vision*. Springer London. https://doi.org/10.1007-978-1-84882-935-0

[26] Osslan Vergara, Vianey Cruz-Sánchez, Humberto Ochoa, Manuel Nandayapa, and Angel Flores Abad. 2014. *Automatic Product Quality Inspection Using Computer Vision Systems*. 135–156. https://doi.org/10.1007/978-3-319-04951-9_7

[27] J. Wang. 2012. *A Whole Body, Kinaesthetic Digital Drawing Tool for Art Therapy*. Simon Fraser University. https://books.google.com/books?id=2Xx20AEACAAJ

[28] Nadine Zahreddine. 2023. The Impact of Digital Art on Traditional Arts. *ISSN: 2435-9475 – The Barcelona Conference on Arts, Media & Culture 2023: Official Conference Proceedings* (Nov 2023), 19–31. https://papers.iafor.org/submission70949/