PicassoAR: A Mixed Reality Approach on guided Sketching on Magic Leap 2

Alen Bišanović
ETH Zurich
Switzerland
abisanovic@ethz.ch

Haleema Ramzan
ETH Zurich
Switzerland
hramzan@ethz.ch

Han Xi
ETH Zurich
Switzerland
hxi@ethz.ch

Abstract

We present ***PicassoAR****, a novel Mixed Reality (MR) platform designed to simplify the process of learning to draw by providing real-time guidance through a Magic Leap 2 (ML2) headset. Our system transforms user-selected or captured images into a series of stepwise outlines, which are then projected onto a physical surface for tracing. To achieve a stable overlay, PicassoAR leverages marker-based tracking and continuously updates the projected geometry as users move around the scene. The most compute-intensive vision tasks are offloaded to a lightweight server, ensuring that the ML2 retains sufficient resources for smooth rendering and user interaction. Experimental results demonstrate the effectiveness of our method in providing intuitive, interactive, and immersion-rich drawing experiences.*

1. Introduction

Mixed Reality (MR) and Augmented Reality (AR) technologies have recently gained traction as powerful learning aids, particularly in creative fields such as painting, sculpture, and drawing. By providing an interactive and visually augmented experience, MR has the potential to seamlessly merge virtual guidance with physical media to improve educational outcomes. Systems that employ MR in creative contexts often focus on enabling inexperienced users to develop their artistic skills more quickly by offering overlays or real-time feedback.

However, many MR drawing systems rely on computationally expensive tasks running directly on head-mounted devices, which limits their responsiveness and scalability. As these devices can have restricted CPU and GPU capabilities, the challenge lies in balancing real-time user interaction and advanced vision-based features. This can be especially critical when running complex image-processing pipelines for tasks like contour extraction or segmentation.

We introduce **PicassoAR**, a lightweight yet robust solution for guided sketching, specifically designed for Magic Leap 2 (ML2). Inspired by prior works on AR-based drawing aids [3], our approach employs a server-assisted pipeline to convert any user-provided or captured image into easy-to-follow outlines. PicassoAR anchors these outlines onto a physical surface using robust marker-tracking capabilities, ensuring stable alignment throughout the sketching process. By offloading tasks such as contour extraction and edge detection to a Python Flask server, PicassoAR avoids overburdening the headset, ensuring a smooth experience for the end user.

2. Related Work

AR-based drawing assistants have been explored in various contexts, enabling users to overlay digital instructions onto a physical medium. Early methods often used straightforward camera feeds and planar tracking, but these approaches struggled with drift and lacked precise alignment. With recent advances in marker-based tracking, particularly with devices such as Magic Leap and HoloLens, researchers have achieved more robust registration between real and virtual elements, leading to more accurate overlays in complex environments [2].

Moreover, the importance of offloading computational tasks for vision applications has been widely acknowledged in the mobile and wearable computing literature. Techniques such as edge computing or server-based processing can significantly reduce the load on resource-constrained devices and enable more sophisticated algorithms to be executed without sacrificing user experience. Our approach builds upon these findings by adopting a Python Flask server for image-to-outline transformations. This architecture leverages the server's more powerful computing resources to handle computationally expensive tasks, all while maintaining high frame rates on the ML2.

3. Methodology

PicassoAR's architecture is divided into three distinct components:

1. The ML2 headset,

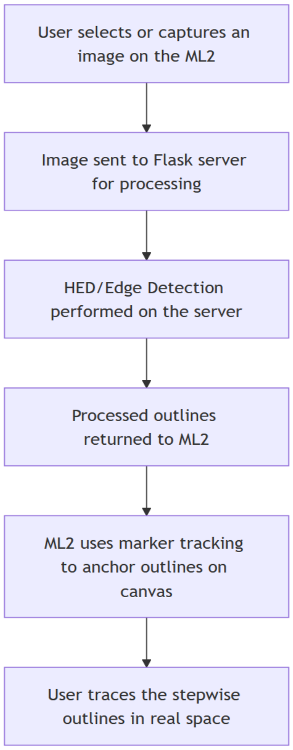

Figure 1. **System Pipeline.** The user selects or captures an image on the ML2 (1). The image is sent to a Flask server for preprocessing (2). The server returns stepwise outlines (3), which are projected onto a canvas anchored to a physical marker (4). Users then trace these outlines in real space, receiving immediate visual feedback (5).

2. The marker-based anchoring subsystem,

3. A lightweight server for image processing.

An overview of this pipeline is shown in Figure 2.

3.1. Magic Leap 2 Headset

The ML2 device is responsible for user interaction, camera capture, and real-time rendering of the AR content. It prompts the user to either capture or upload an image, then forwards this image to the Flask server. Once the server returns an annotated or processed image, the ML2 composites the result onto a virtual plane in the physical environment. The device's onboard sensors are continually used to track the user's head position and orientation. By maintaining a rapid rendering cycle and leveraging the ML2's built-in spatial awareness, our system ensures that the user sees a stable overlay on the physical canvas.

3.2. Marker-Based Anchoring

To ensure stable alignment between digital outlines and the real-world drawing surface, PicassoAR relies on either

Figure 2. **Marker Detection for Anchoring.** The camera feed from the ML2 is used to detect and localize ArUco (or QR) markers. This ensures that the virtual outlines remain stable despite user movements, changes in perspective, or minor occlusions.

QR or ArUco markers placed on or near the paper. By scanning and identifying these markers, the ML2 localizes the correct position and orientation of the canvas in 3D space. As the user moves around, the system updates the pose of the virtual plane to maintain alignment with the marker, allowing the overlaid outlines to remain fixed on the actual surface. This approach has proven to be significantly more robust than methods relying purely on planar feature tracking, especially under challenging lighting conditions [2] or when the canvas is partially occluded by the user's hand.

3.3. Server-Assisted Image Processing

Initially, the image processing pipeline utilized Sobel filters but yielded suboptimal results in providing clear and instructive outlines. Consequently, we now leverage Holistically-Nested Edge Detection (HED) [1], which delivers finer and more continuous contours. Although Contrast Limited Adaptive Histogram Equalization (CLAHE)-based preprocessing is available as an option for more complex images, it is disabled by default; our current focus remains on providing a drawing guide primarily for relatively simple or medium-complexity images.

By offloading these computations to a Python Flask server, PicassoAR enables complex vision algorithms to run efficiently without overloading the headset's GPU and CPU. The pipeline also includes an option for stepwise outline generation, allowing the user to cycle through increasing levels of detail. This stepwise approach helps learners follow a drawing in progressive stages, enabling them to visualize key structures before adding finer details.

4. Experiments

4.1. Experimental Setup

We conducted several experiments to evaluate the stability of alignment and overall system performance. These tests involved users walking around the canvas and occasionally blocking markers to gauge how the system would recover from brief occlusions. Sub-centimeter drift was observed during typical user movements, and the alignment was quickly restored once the marker became visible again.

For performance metrics, we measured the average frame rate and network latency with and without server offloading. When offloading was disabled and all computations were done on the ML2, the average frame rate dropped significantly to around 30 FPS. With the server offloading enabled, the ML2 consistently maintained 65–70 FPS, indicating a clear advantage in responsiveness.

4.2. User Study

4.2.1 Study Goals

We conducted a short user study to investigate whether PicassoAR serves as a useful tool for aiding the learning process in sketching. Our primary hypothesis was:

> **Hypothesis: PicassoAR's real-time overlays and stepwise outlines can significantly improve a novice user's ability to follow drawing guidelines with minimal alignment issues.**

4.2.2 Setup & Participants

We conducted a pilot user study with a sample size of two, given the time and resource constraints of this project. Each participant was asked to use the app from start to finish. This meant that the user would interact with the interface, pick an image of their choice, and sketch it with the help of PicassoAR on the physical canvas placed in front of them.

- **Participant 1:** A 25-year-old art student with basic experience in digital art tools. They had moderate proficiency in sketching but had never used an MR device before.

- **Participant 2:** A 30-year-old hobbyist who had no formal art training and was new to MR tools. They primarily relied on tutorials and step-by-step guides in digital media for casual learning.

4.2.3 Quantitative Results

We evaluated the following metrics for our quantitative analysis:

- **UI interaction:** Number of errors or mis-clicks the user made while navigating the ML2 interface. This helped us understand how intuitive the user interface was.

- **Task completion time:** The total duration from when the user started the app's main menu to when they completed their final traced sketch. This included the time taken to choose images, align the canvas, and finalize their drawing.

- **Alignment accuracy:** Estimated alignment error in centimeters while sketching. We tracked how well the virtual outlines remained on the physical surface, as subjectively noted by the users and supplemented by our own measurements of drift.

- **Overall satisfaction:** A rating out of 10, in which participants rated how content they were with their overall experience in terms of ease of use, clarity of instructions, and quality of the final sketch.

- **Likelihood to recommend:** A rating out of 10 indicating if the participant would recommend this app to someone looking to learn or practice drawing.

Table 1 summarizes the key quantitative findings from our study:

Table 1. Quantitative Metrics for PicassoAR (P1 = Participant 1, P2 = Participant 2).

Metric	P1	P2
UI interaction (errors)	1	2
Task completion time (min)	12	16
Alignment accuracy (cm drift)	0.5	0.8
Overall satisfaction (out of 10)	8	9
Likelihood to recommend (out of 10)	8	9

4.2.4 Qualitative Evaluations

Both participants found the concept of real-time overlay helpful in guiding their sketches.

Participant 1: They appreciated the intuitive interface and step-by-step outlines, which allowed them to focus on specific features of the drawing. Although they found alignment stable for the most part, they noted occasional difficulty in aligning the virtual outlines perfectly with the physical canvas when the environment lighting dimmed significantly. The participant suggested implementing a quick calibration option to help adjust the marker detection threshold under low-light conditions.

Participant 2: As a newcomer to MR tools, they were initially skeptical about how comfortable it would be to draw while wearing the headset. After a brief onboarding, they adapted quickly and expressed enthusiasm about the guided aspect of the system. The participant found the progressively detailed outlines particularly helpful for tackling complex shapes in smaller segments. They suggested having a brief, built-in tutorial or user manual that can guide first-time users through the basics of marker placement and movement around the canvas.

5. Discussions

5.1. Quantitative Analysis

From Table 1, both users made few interaction errors, indicating a relatively user-friendly interface. The task completion time ranged from 12 to 16 minutes, which included everything from loading the app to finishing the outline tracing. Given the preliminary nature of this study and the small sample size, these times seem reasonable. Alignment errors were reported in the range of 0.5 to 0.8 cm, which our participants described as only mildly noticeable. This suggests that the marker-based tracking approach is sufficiently robust for general drawing tasks, although further refinements could be made to improve the experience in challenging lighting conditions.

5.2. Qualitative Analysis

The participant feedback confirmed that PicassoAR offers a robust and immersive sketching experience. Despite occasional challenges such as lighting variations, both participants reported that they were able to produce satisfactory sketches by following the virtual outlines. The minor drift in alignment did not heavily impact the user experience, especially once they learned to quickly adjust their viewpoint or reacquire the marker.

Additionally, participants requested more intuitive tutorials and adaptive features, such as brightness and contrast controls for the marker-detection camera feed. This feedback aligns with our roadmap to incorporate real-time lighting adaptation and deeper visual analytics for guiding sketches at different expertise levels.

6. Conclusion & Future Work

We have presented PicassoAR, a Mixed Reality platform for guided sketching on Magic Leap 2, powered by off-device vision processing. By merging stable anchoring with advanced contour extraction methods, PicassoAR delivers a responsive and intuitive drawing experience. Our user study, although small in scale, demonstrates the promise of this system in helping individuals learn and practice drawing skills through an immersive interface.

6.1. Areas of Further Work

We have identified several key areas in which the app may be improved or extended, ensuring a more seamless learning experience for users:

- **Lighting Adaptation:** Implementing automatic calibration for varying lighting conditions. In low-light scenarios, the marker-tracking algorithm could adapt its thresholding dynamically, improving robustness and reducing the likelihood of alignment loss.
- **Guided Tutorial:** Introducing an interactive onboarding workflow for first-time users. This could include a sample sketch and short instructional prompts, helping them understand how to align markers, move around the canvas, and toggle between outline steps.
- **Expert Feedback:** Integrating real-time drawing analysis and corrective suggestions. For instance, the system could compare user-drawn lines with ideal contours in real time and highlight significant discrepancies. Generative AI could also be employed to offer style-based tips or aesthetic suggestions.
- **Sequential Tutorials:** Developing multi-stage guides where edges are revealed in progressive segments. This would require semantic understanding of the image, allowing the system to logically group contours (e.g., major shapes first, followed by finer details).
- **Collaborative Sketching:** Enabling multiple ML2 users to share the same virtual environment, facilitating group or instructor-led drawing sessions. This could open possibilities for real-time feedback or competitive sketching activities in educational settings.

These extensions aim to make PicassoAR an even more comprehensive and versatile tool for creative learning, further bridging the gap between traditional drawing techniques and emerging Mixed Reality technologies.

References

[1] S. Xie and Z. Tu. Holistically-Nested Edge Detection. In *ICCV*, pages 1395–1403, 2015.

[2] *Marker Tracking API Documentation*, Magic Leap Developer Resources.

[3] *Best Practices for AR UI*, Unity Documentation.

[4] *MLSDK Documentation*, Magic Leap Developer Resources.

[5] *ML Hub Tutorials*, Magic Leap Developer Resources.