An Augmented Reality Camera Layout System for Intuitive, Spatially Aware Shot Design

Hyewon Lee
School of Computing, KAIST
Daejeon, Republic of Korea
hyewon0809@kaist.ac.kr

Jee Youn Kim
Industrial Design, KAIST
Daejeon, Republic of Korea
jounkimm@kaist.ac.kr

Christopher Bannon
School of Computing, KAIST
Daejeon, Republic of Korea
cbannon@kaist.ac.kr

Andrea Bianchi
Industrial Design & School of Computing, KAIST
Daejeon, Republic of Korea
andrea@kaist.ac.kr

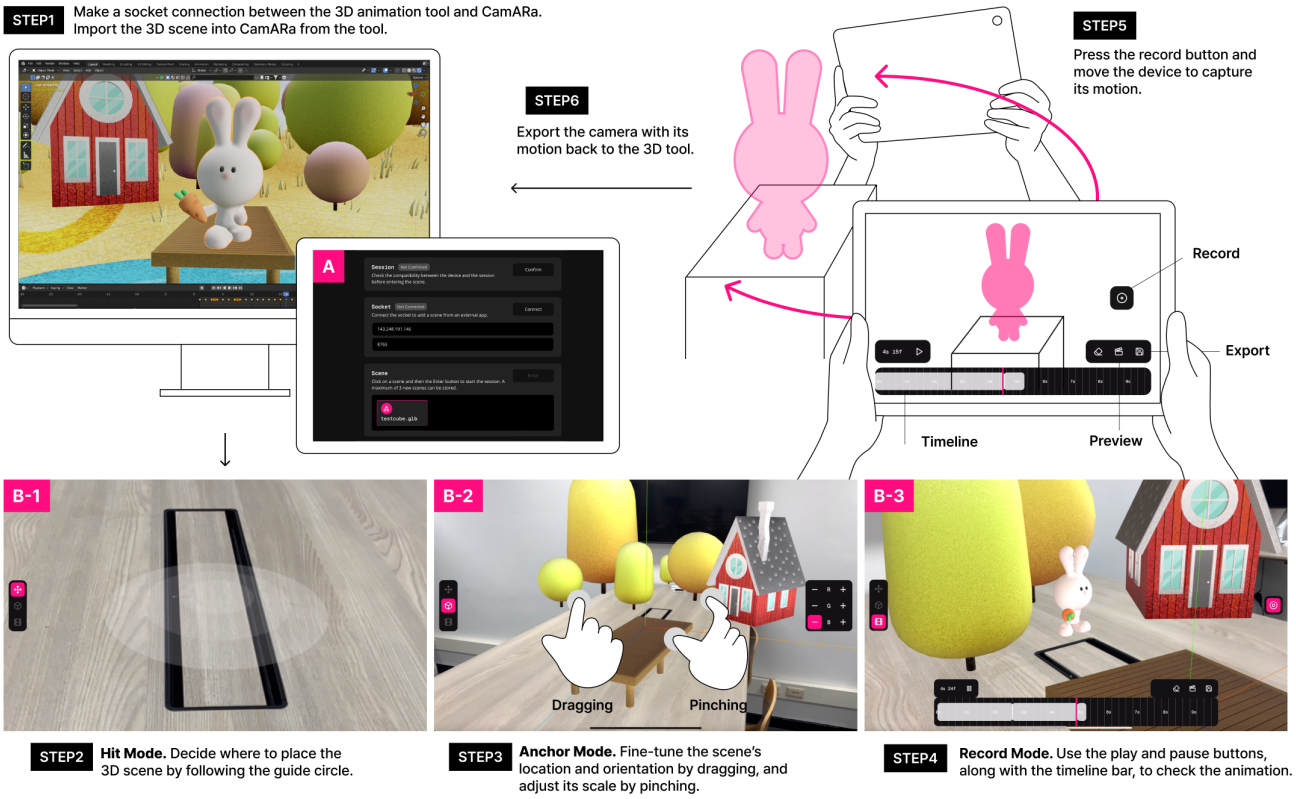

Figure 1: User flow of CamARa (a) Home screen (b) Session screen: (b-1) Hit mode (b-2) Anchor mode (b-3) Record mode

Abstract

Camera layout is a crucial stage in digital production, shaping the narrative and emotional tone of animation, CGI, and previsualization. It involves composing spatial arrangements of scenes and sketching out rough camera movements. Our formative study (N=8) reveals that individual creators often follow non-linear and iterative workflows when they create camera motions, frequently moving between exploration, decision-making, and refinement, which differs from the linear process of studio pipelines. Our findings also highlight the difficulty of achieving desired shots with traditional 2D interfaces, especially for amateurs. We propose CamARa, an augmented camera layout system that leverages augmented reality (AR) for intuitive, spatially aware interaction. By capturing users' embodied movements, CamARa allows rapid exploration and validation of shot ideas through an immersive and engaging experience. Our qualitative study (N=9) demonstrates that CamARa adapts flexibly to users' individual goals and levels of expertise, offering benefits in intuitiveness, time efficiency, and creativity—particularly for non-professionals. It also shows promise in professional contexts such as communication, previsualization, and virtual production by leveraging the affordances of the AR environment.

Preprint, KAIST, Daejeon, Republic of Korea
2025.

CCS Concepts

• **Human-centered computing → Interactive systems and tools**; **Mixed / augmented reality**.

Keywords

camera layout, augmented reality, virtual camera

ACM Reference Format:
Hyewon Lee, Jee Youn Kim, Christopher Bannon, and Andrea Bianchi. 2025. An Augmented Reality Camera Layout System for Intuitive, Spatially Aware Shot Design. In *Unpublished manuscript*. ACM, New York, NY, USA, 14 pages.

1 Introduction

In cinematography, camera layout is a vital phase, as it determines the message and emotion that a scene aims to convey in both live-action films and animation [8]. For any given scene, the style and mood can change significantly based on the numerous combinations of different shots. In digital production for animation, CGI, and previsualization, managing the camera within a 3D scene presents a complex challenge due to its numerous Degrees of Freedom (DOF): pitch, yaw, roll, surge, heave, and sway [5]. This complexity can make it difficult to effectively express ideas, requiring laborious and repetitive tasks, regardless of proficiency. To tackle this issue, prior studies have investigated various control metaphors to overcome the challenges of operating a 6DoF camera with 2D devices such as a mouse [20].

One practical approach, based on the idea of replicating real-world camera functionality, is a virtual camera system [2, 19]. This system allows users to generate camera movements through familiar interactions, such as physically moving a handheld device (HHD) as if operating a real camera [14]. The device's motion is tracked and applied to a virtual camera, with the resulting view rendered on the device's screen. However, these solutions are mainly built for studio pipelines and professional creators, leaving behind amateur individual creators. To incorporate them and understand the workflows of individual creators with varying levels of proficiency, we conducted a formative study with eight creators and observed their personal workflows when designing camera animations for a given 3D scene. We identified that their workflows are highly individualized, influenced by personal preferences, and often involve a complex, iterative, and improvisational process across the exploration, planning, and production stages. This contrasts with the more linear procedures typically followed in studio pipelines. It was found that understanding the spatial position and path of camera movements, as well as comprehensively understanding the entire scene and the relationships between objects, were essential during the design process. Furthermore, the study revealed that creators often struggle with controlling the camera and achieving the desired movements using conventional 2D interfaces, especially non-professional creators. Based on the insights from this study, we envisioned a novel virtual camera system that leverages augmented reality to not only support intuitive camera control but also enhance spatial awareness.

We propose CamARa, an augmented virtual camera layout system for exploring and creating camera movements with mobile AR. It includes importing a 3D scene into an AR environment, creating desired camera movements within AR, and exporting the camera with motion back to the original 3D scene, enabling a smooth transition from AR to existing digital production workflows. Designers can set up 3D scenes using real objects as reference points and tailor the environment to the scene's motion, creating an intuitive and immersive experience. We suggest three types of control strategies; (1) Camera-in-Hand, (2) World-in-Hand, and (3) Hybrid Control; designed to overcome the physical limitations of AR and expand the range of possible camera shots. This system is built as a web-based interface for mobile devices and seamlessly integrates with conventional 2D interface tools through socket connection. We conducted a qualitative user study with nine participants with diverse experience in physical camera and virtual camera. We aimed to answer how do users utilize CamARa in their iterative and improvisational workflow and how does the system support user at which stage and in which role, based on the findings from the formative study. Our findings suggest that CamARa supports flexible, personalized use depending on users' goals and proficiency levels, and highlights its potential not only for non-experts but also for professional contexts such as communication, previsualization, and virtual production.

This paper presents the following contributions:

- Findings from a formative study (N=8), revealing that the camera animation design process is highly individualized and characterized by iterative and improvisational attempts at exploration, planning, and production.
- **CamARa**, an augmented reality-based camera layout system that supports intuitive and spatially aware shot design and its manipulation strategies
- Insights from a user study (N=9) on how CamARa can be integrated into personal workflows across varying levels of proficiency, including the potential of leveraging augmented reality.

2 Related Work

2.1 Automated Camera Layout System

Prior research has explored methods to simplify the complexity of camera layouts through automated systems [10]. These systems automatically generate camera movements for a given 3D scene. Previous studies have examined various computational techniques to create camera motions that consider the 3D environment, including character animation [15, 18]. To achieve a polished camera layout, many systems learn from reference videos using machine learning techniques that follow traditional cinematography principles [11, 12, 23]. However, automated systems can limit individuals' ability to explore camera layouts, restricting their opportunity to express style, creativity, and collaboration. Additionally, these approaches often require expensive computing resources. Our work aims to provide accessible control through personal devices, allowing users the freedom to experiment and express their ideas.

2.2 AR/VR in Previsualization Tool

Previsualization (previs) is a collaborative process that creates preliminary versions of shots, typically in 3D [22]. Through previs, filmmakers visually explore and communicate creative ideas while planning technical solutions, including layout, animation, lighting, and camera work. To provide an immersive and effective previs

experience, researchers leverage AR and VR technologies to control the camera and create camera motion [7, 16, 24]. Immersive VR tools that use head-mounted displays (HMDs) are often less preferred due to their low accessibility and high complexity [13]. Some previous studies have employed handheld devices (HHDs) as proxies for virtual cameras in 3D scenes, capturing the motion of these devices [2, 13, 24]. Unlike prior VR works [2, 14, 19], CamARa leverages the physical context of the real world to create camera motion on top of a virtual camera, enabling users to consider the constraints and possibilities of moving within the surrounding space. Unlike existing AR designs [24], we focus on the interaction of crafting camera motion, and investigate how this approach integrates into creators' personal workflows, including which stages it supports and what roles it plays.

2.3 Embodied Camera Interaction and Spatial Reasoning

Recent research highlights how physical movement, spatial presence, and embodied reasoning significantly enhance user interaction with camera systems in mixed reality environments. For example, Zambaldo et al. [21] demonstrate that embodied control in VR enhances spatial presence and leads to more precise understanding of scenes. Similarly, Nitsche [20] argues that camera control should extend beyond static framing to incorporate spatial navigation and embodied exploration as core elements of meaning-making. These findings show that users construct meaning by moving within the scene and actively interacting with the space, rather than merely observing from fixed viewpoints. Other studies have explored how to enhance scene quality through virtual camera control. For example, Louarn and Christie [1] proposed an automated system that generates multi-shot sequences based on four cinematic rules. Their system connects static keyframes into coherent multi-shot sequences that follow cinematic continuity and enhance narrative flow. In the context of virtual production [6], physical interaction is increasingly re-integrated into digital workflows. This is particularly evident in hybrid systems that combine real and virtual cameras—such as those that integrate LED screens on production stages [4]. These emerging virtual camera systems are not merely replacements for physical cameras, but are increasingly discussed as transformative tools that can reshape the structure of production workflows. Our work builds on these insights by making an AR-based system that allows creators to compose camera shots through physical movement, spatial reference, and real-time visualization. Unlike previous systems that treat physical space as neutral or disconnected, we explore how users can leverage it as a compositional element—turning physical constraints into design opportunities.

3 Formative Study

We conducted a formative study on how independent creators design and produce camera animations for 3D scenes and the challenges they encounter while manipulating cameras.

3.1 Participants

We recruited 8 participants (P1-P8; 6 male and 2 female; age M=28.25 and SD=4.56; see Table 1) through the online community. Among them, three were professional 3D animators, one was a professional 3D motion graphic designer, one served as a cinematographer, and another was a product designer. Additionally, two participants (P7 and P8) were students majoring in industrial design. Three participants (P2, P5, and P6) identified as freelancers. All participants reported having a minimum of three project experiences in creating camera animations using 2D interface tools, including Blender[1], Cinema 4D[2], Keyshot[3], and Maya[4]. To participate in the study, all participants were required to be capable of using Blender.

3.2 Procedure

We initially observed how participants compose camera movements for a given 3D scene (30 minutes), and then we conducted a semi-structured interview (30 minutes) about the task and their prior workflow. We selected Blender as the main tool for the study, which is open source, as it is the most commonly used 3D tool accommodating users of various skill levels for a range of general purposes.

In the first task, participants were instructed to create camera movements within a designated 3D scene titled *Bunny in the Forest*. The scene consists of two animated subjects and static surrounding elements, including trees, a house, and a bridge. The animation script was as follows.

> ***Bunny in the Forest.*** *One sunny day, a bunny was walking in the forest and saw a carrot floating on a lake. As it reached out to grab the carrot, the bunny fell into the water.*

The script was not provided to the participants; instead, we encouraged them to create camera motions based on their individual interpretations, styles, and intentions. There were no restrictions on the number of cameras they could employ. When participants utilized multiple cameras, we requested that they clarify which camera would be rendered for each frame after completing the task. Participants were allowed to freely search on the internet during the task just in case they needed to reference the tool instructions. The first task was performed using the think-aloud method to identify participants' intentions behind their actions throughout the workflow. If their explanations were lacking, we asked additional questions during the interview.

Each participant received approximately $15 as compensation for their one hour of participation, paid in local currency via direct deposit. The studies were conducted via online conference platform and all sessions were recorded.

3.3 Findings

3.3.1 ***F1. Iterative and Improvisational Shot Design Process.*** Through our observations during the free exploration task, we noted that, unlike the linear structure of the studio pipeline, which clearly distinguishes the stages of exploration and planning, individual workflows involve an improvisational iteration among the exploration, planning, and production phases. Every participant modified their shot concepts at least once during the design process, and all except P7 completely replaced their shot ideas or adjusted the number of shots midway through the task.

[1]https://www.blender.org/
[2]https://www.maxon.net/en/cinema-4d
[3]https://www.keyshot.com/
[4]https://www.autodesk.com/products/maya/overview

Table 1: Demographic information of participants in the formative study. Experience and self-rated proficiency were reported in the context of using 2D interface tools for camera layout.

ID	Age	Gender	Current Occupation	Experience	Self-rated Proficiency
P1	30	M	3D Motion Graphic Designer	3 yrs	8
P2	31	M	Cinematographer	<1 yr	3
P3	29	M	3D Animator	3 yrs	4
P4	37	M	3D Animator	5 yrs	8
P5	27	F	3D Animator	3 yrs	2
P6	25	M	Product Designer	<1 yr	7
P7	24	M	Student	<1 yr	3
P8	23	F	Student	<1 yr	5

Except for P3, who primarily works alone, all participants mentioned that they usually create a 2D storyboard in advance. However, all participants also noted that the design changes significantly from the original plan once they begin blocking shots with the camera. P3, P5, and P6 explained that depending on the terrain features and the characteristics of assets, it is tricky to compose the exact same shot from the storyboard. For example, P4 first tried to capture the scene of a bunny walking in the side view, but as the house object obscured the frame, he changed to the other view where the house did not appear.

Several participants (P1, P3, P4, P6, P7, and P8) noted that when they came up with new or more effective shot ideas during the process, they would often create them spontaneously to explore their potential. P2 explained that *"Unlike real shooting environments, working in digital settings allows for repeated revisions, so there's less pressure to plan everything out in advance, like with storyboards."*

3.3.2 ***F2. Highly Individualized and Content-Specific Workflows.*** The shots the participants created during the task were not much different from conventional shots such as pan, tilt, and dolly. However, the number of shots they selected, their combinations, and their working order differed. For example, P1 designed the first and the last frames at first and added the intermediate shots by checking the flow of the animation from start to end. Otherwise, P4 initially divides the animation into three parts and composes representative shots of each in order. He edited in an additional shot that he wanted to highlight and ended up with four shots. P7 designed a continuous one-shot only, using a turntable-like camera motion. The methods were also different based on personal taste. P1 mentioned, *"Having more freedom in camera control allows for more interesting shots. I find creating camera paths as curves too restrictive, so I prefer manually keyframing each movement."* In contrast, P7 explained, *"Manually keyframing every DOF is time-consuming, so I prefer designing camera path curves, which are more intuitive and efficient."*

Even during fine control, after the rough layout is established, P4, P5, and P6 managed the camera view using the mouse. Meanwhile, P1, P3, and P8 adjusted the numeric values for each degree, while P2 and P7 operated in the third-person view. Through the following interview, we confirmed that the workflow has changed depending on what kind of scene they are working on. P6, P7, and P8, who had prior experience in product video production, mentioned that they typically focus on creating and rendering shots that best highlight the product and make the product look most appealing. They also noted that the sequence and combination of shots are considered later during editing using external video editing tools. In contrast, when working on an animation like in the study, P7 and P8 emphasized that narrative becomes more important, and it is essential to consider how to effectively convey the story through the shots when they create the shots.

3.3.3 ***F3. Navigating Between Camera View and Spatial Position.*** All participants frequently switched between camera view and third-person view during the task, or they used a split-window to observe both views simultaneously. P4 and P5 noted that relying solely on the camera view made it challenging to understand the terrain features and difficult to track movement paths. Meanwhile, P1 and P3 expressed that they sometimes struggled with the unexpected camera movements when working exclusively in camera view, even though they were already fully familiar with the 2D interfaces. It is crucial to understand not only the camera's view but also its spatial position and orientation. Additionally, P2 and P3 mentioned that they tended to experience fatigue when continuously switching between the two views, as it was confusing due to the differing degrees of control in each perspective.

3.3.4 ***F4. Composing Shots with Scene-Level Awareness.*** In observing the task process, we confirmed that all participants focused on designing shots that highlight the main subjects. However, it is essential to comprehensively understand and effectively utilize surrounding objects and their relationships. Some amateur participants, specifically P5, P6, and P8, struggled to create the desired shots due to distractions from the surrounding objects. Some objects obstructed the camera's view or blocked its path. Conversely, more experienced participants demonstrated a strong ability to incorporate their surroundings effectively into their shot design. For example, P4 crafted shots that not only emphasized the subject but also enhanced the sense of space by intentionally including nearby trees in the frame. When objects obstructed the camera's path, P4 removed or repositioned them to achieve the desired shot. Additionally, participants P6, P7, and P8 mentioned that they typically explore the scene in various layouts before starting their shot designs.

3.3.5 ***F5. Challenges in Creating Desired Shots.*** All participants expressed that they faced some challenges in controlling the camera and creating its motion in the desired way during the

interview. Even P1 and P4, who are already experts in 3D tools, mentioned that while they generally do not experience critical difficulties with camera control, they sometimes encounter tricky cases—such as unexpected movements caused by gimbal lock or difficulties in managing the timing across multiple degrees of freedom. P3 and P6 struggled with identifying and adjusting the appropriate axes to achieve their intended shots. P5 and P7 noted that creating high-quality shots requires learning and adapting to various tools, such as add-ons. P2 and P8 commonly expressed challenges in precisely controlling the rotation and position of the camera to match their intentions. Specifically, P8 highlighted her difficulty compared to working with a physical camera: "With a real camera, I can make subtle adjustments by physically moving it little by little. But with digital cameras, it's harder — if I want to tilt it slightly at an angle, I have to adjust fixed values along the X or Y axis separately."

3.4 Design Rationale

Based on our formative findings, we envisioned an AR-based camera layout system that enables users to explore and create desired shots intuitively (F5) by physically moving a mobile AR device. To support spatial awareness, the system allows seamless switching between camera and third-person views (F3), and encourages shot composition with an understanding of the entire scene beyond just the subject (F4). We detail the system design in the following section.

4 The CamARa System

4.1 Interface

CamARa consists of two screens: Home and Session. On the **Home Screen (A, Figure 1)**, users can check the compatibility of their device and browser for the AR session, establish a socket connection with a 3D animation tool, and receive a 3D scene from the tool. The **Session Screen (B, Figure 1)** allows users to switch between three modes: *Hit mode*, *Anchor mode*, and *Record mode*.

In **Hit mode (B-1, Figure 1)**, users can detect the floor and place the 3D scene in the real world with the help of a guide circle. **Anchor mode (B-2, Figure 1)** allows users to adjust the orientation, location, and scale of the scene using drag and pinch gestures.

Record mode (B-3, Figure 1) displays a control interface along with a timeline bar, allowing users to play, pause, and interact with the animation. When the user presses the record button, the animation starts plays from the current frame, and the device's movements are captured. These movements are represented on the timeline bar. Users can also preview the recorded camera motions.

4.2 Manipulation Types

CamARa supports three types of camera manipulation strategies: (1) Camera-in-Hand, (2) World-in-Hand, and (3) Hybrid Control. World-in-Hand and Hybrid Control allow users to create dynamic and various shots, overcoming the physical constraints of Camera-in-Hand mode.

4.2.1 Camera-in-Hand. Camera-in-hand (A, Figure 2) is the fundamental method for controlling and creating camera movements. Once users set up the AR scene, they can intuitively create camera motions by physically moving the device within the space, treating the system as a real camera. Since the view provided by the system is identical to the camera view in the digital tool, if users set the scene on a large scale, a large-scale scene indicates that the camera is in proximity to it. In contrast, a small-scale scene implies that the camera is farther away. It not only creates motion by hand but also allows natural effects like handheld and camera shaking.

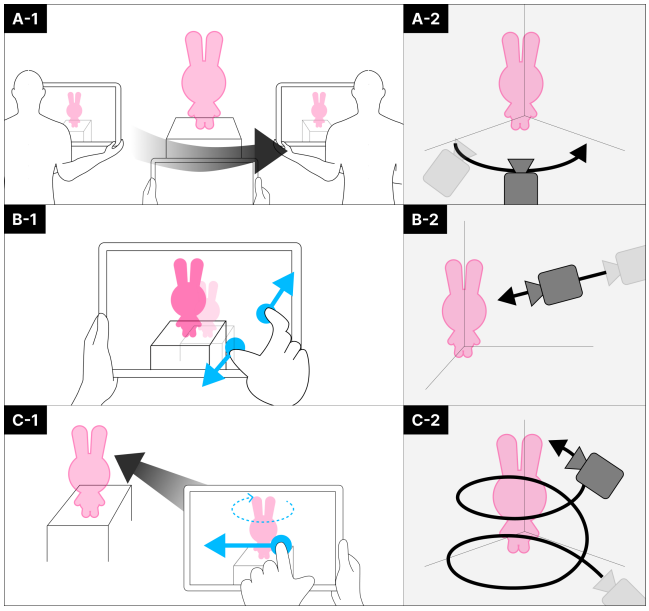

Figure 2: Examples of three types of manipulation in CamARa: (a) Camera-in-Hand: A user moves to one side while holding the camera. (b) World-in-Hand: A user scales up the scene by pinching while recording the camera motion. The camera movement is designed to reach toward the scene. (c) Hybrid Control: A user moves forward into the scene while simultaneously rotating it with a finger. The trajectory is constructed in a spiral shape, allowing the camera to rotate around the scene as it closes.

4.2.2 World-in-Hand. World-in-Hand (B, Figure 2) refers to a method of indirectly creating camera movement by handling the scene itself. Instead of physically moving the device, users can manipulate the scene while maintaining the camera's position and orientation. The camera movements are calculated based on the world frame; therefore, any changes in the world frame will be applied to the camera frame accordingly. If the scene scales up, it creates the motion of the camera moving forward toward the scene. Similarly, if the scene rotates, it simulates the camera rotating around the scene.

4.2.3 Hybrid Control. Hybrid control is the mixing of both camera-in-hand and world-in-hand. As illustrated in C, Figure 2, users can adjust the world frame while physically moving. This approach enables users to create dynamic shots that are difficult to replicate physically.

4.3 Walk-through: Bunny in the Forest

To demonstrate how designers can use physical references for exploring and creating camera motion in AR, we describe a user scenario with two animators, Hana and Yuna, who are using CamARa on their tablets.

Hana and Yuna are working on the animation *Bunny in the Forest*. After they finish the character animation, they open CamARa on their tablets and connect to the 3D animation tool. After each device receives the 3D scene from the tool, they enter AR mode.

4.3.1 ***Explore Shot Combinations.*** Before camera composition, Hana and Yuna discuss what kinds of shots they will capture. They want to visualize each idea to assess its expressiveness and narrative delivery. First, they set up their 3D scene on a table with enough space to move around freely. They adjust the anchor, ensuring that the orientation and scale of the scene fit the space appropriately. They first look around the scene by physically moving in various positions and angles to explore the layouts and movements (Camera-in-Hand), as well as manipulating the scene in different scales and orientations by pinching and dragging (World-in-Hand). Next, they each create camera paths by repeatedly recording and pausing within the animation's timeline, resulting in several shots captured by one camera (A, Figure 3). They iterate multiple times, previewing and refining each other's ideas as needed, and ultimately decide to go with Yuna's concept (B, Figure 3).

4.3.2 ***Take Detailed Shots Immersing in the Scene.*** Hana and Yuna are each responsible for creating two shots. Hana uses a tripod, while Yuna holds a tablet to create camera paths for each shot. At this time, they captured each shot in each camera. When Hana tries to take a shot Y-2, she comes up with a new idea of getting closer to the bunny not far away from it, that Yuna first took, to highlight the bunny's face. To capture it in more expressive way, she adapts hybrid control method. She makes a camera rotation with tripod (Camer-in-Hand) while scale up the scene by pinching (World-in-Hand), so the camera can getting close to the bunny and following the bunny's move in the same time (A, Figure 4). Yuna holding tablet by hand, develop the shots by utilizing surrounding objects like moving relative to the table, and rotate around the chair.

4.3.3 ***Take Dynamic Shots Referencing Physical Space.*** Hana decides to create a more appealing shot of the rabbit running by filming from the rabbit's point of view. However, this was challenging to achieve on the table, and the room was too small for the desired effect. Hana moves to the hallway, repositions the 3D scene, and aligns the animation so the rabbit runs along the floor of the hallway. Hana begins recording his movements facing the carrot, simulating the rabbit's perspective (B, Figure 4). Once Hana and Yuna are satisfied with all the shots, they export the cameras with their motion and proceed to editing and post-processing in the 3D animation tool.

4.4 Implementation

CamARa is built as a progressive web application (PWA) for WebXR[5]-compatible mobile devices. The system's core is implemented using Three.js[6] and TypeScript, with WebXR API integration for AR functionality. Device pose is captured using WebXR's viewer-space reference frame, allowing real-world movement to control the camera in conventional 3D tools. WebXR is supported natively in Chrome for Android, while iOS requires Mozilla's WebXR Viewer. To support interoperability with standard 3D tools, the system supports GLB-based workflows. Users can import 3D scenes in GLB format, create camera motions within the AR environment, and export the animated camera in GLB format back into tools such as Blender or Maya. CamARa features live server connectivity through the WebSocket protocol, enabling real-time interaction with external software.

5 User Study

We conducted a qualitative user study on how individual creators utilize the CamARa system in their workflow. As we confirmed that they have a highly individualized workflow (F2) and iterative and improvisational process (F1), we aimed to identify how this system can be blended into the existing processes at which stage and in which role.

- RQ1. How do users utilize CamARa in their iterative and improvisational workflow?
- RQ2. How does CamARa support users with various levels of proficiency?

To evaluate the questions, we asked users to explore the system freely and conducted the following interview.

5.1 Participants

Through the online community, we recruited 9 participants (U1-U9; 5 female and 4 male; age M=26.67 and SD=3.54; see Table 2). Three participants had more than 3 years prior experience with physical cameras (U2, U3, and U5). The other three participants were virtual camera professionals with over 3 years of experience (U1, U4, and U6). The other three were design major students with approximately 1 year of experience (U7, U8, and U9). We grouped them into three groups, respectively: (1) physical camera experts, (2) virtual camera experts, and (3) non-experts. Three reported them as freelancers (U1, U3, and U4). Two of the participants had previously participated in the formative study (U6 and U8). Like the formative study, all participants reported having at least three project experiences in creating camera animation using 2D interface tools and were all required to be capable of using Blender.

5.2 Procedure

We conducted four stages of study for 90 minutes (A, Figure 6). First, we **outlined the study procedures and delivered system instructions (15 minutes).** Since this task aims to examine how individual creators incorporate the system into their workflow—beyond just creating movements—we also explained how to connect the tool to Blender and share data between the two. Second, we asked users to **use and practice the system freely with Blender (15 minutes).** In this stage we provided a test scene of plain cube that animates for approximately 6 seconds.

[5]https://developer.mozilla.org/en-US/docs/Web/API/WebXR_Device_API

[6]https://threejs.org/

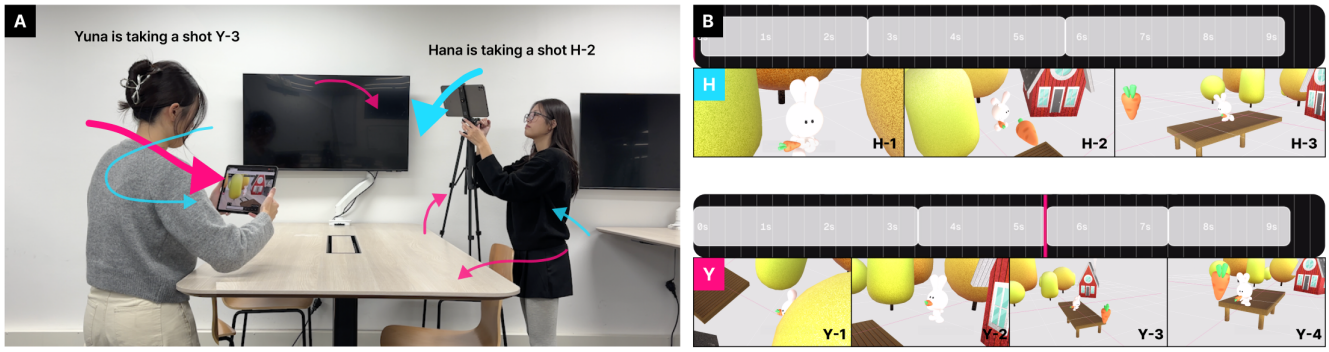

Figure 3: Explore shot combinations with CamARa (a) Two designers capture multiple rough shots in a single camera path to evaluate the coherence, expressiveness, and impact of the shot combination. (b) They preview the scene to assess its overall mood and emotion.

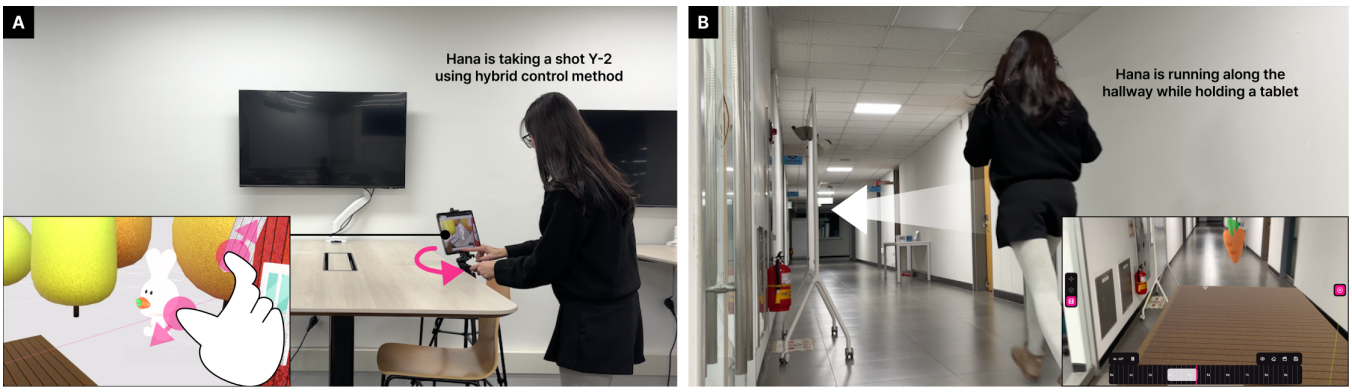

Figure 4: Develop polished shots with CamARa using spatial reference (a) A designer films a shot using the hybrid control method by scaling up the world while simultaneously rotating the camera. (b) A designer utilizes the spatial structure to capture dynamic shots.

Table 2: Demographic information of participants in the user study (U1-U9). Physical camera experience was asked regarding filming experience with professional cameras. Virtual camera experience and self-rated proficiency were reported using 2D interface tools for camera layout. We divide 9 participants into three groups based on their prior experience: (1) Physical Camera Experts (U2, U3, and U5), (2) Virtual Camera Experts (U1, U4, and U6), (3) Non-experts (U7, U8, and U9)

ID	Age	Gender	Current Occupation	Physical Camera Experience	Virtual Camera Experience	Self-rated Proficiency
U1	29	F	Visual Designer	<1 yr	**3 yrs**	6
U2	30	M	3D animator	**3 yrs**	<1 yr	8
U3	30	F	Motion Graphic Designer	**5 yrs**	<1 yr	4
U4	26	F	3D Designer	0 yr	**3 yrs**	7
U5	29	M	PD	**10 yrs**	1.5 yrs	3
U6	29	M	3D Animator	0 yr	**3 yrs**	4
U7	24	M	Student	<1 yr	<1 yr	4
U8	22	F	Student	1.5 yrs	<1 yr	4
U9	21	F	Student	<1 yr	<1 yr	4

As we provided macOS-based devices for the study, some participants spent this stage getting used to the device's settings. We also introduced a few features(e.g., adjusting keyframe density and smoothing the movements) in Blender, which are helpful to clean up and polish the keyframes captured from the system. During the study, we equipped all devices, including a laptop(MacBook Pro m1 chip 16'), tablet(iPad 4th 11'), mouse, and tripod. We provided enough space to move around and two large desks(picnic table size) with chairs in the center((B and C, Figure 6). Two main tasks were conducted in the following 60 minutes.

5.2.1 Free Exploration Task (30 minutes). Participants were asked to create camera animations for the given 3D scene based on their personal interpretation, style, and intentions. The scene was a cartoon-style 10-second-long animation, the same as the formative study. Details were the same as in the formative study. However, we required them to use the system when they felt it was natural and helpful. They were allowed to search on the internet freely during the study. For this task, the authors did not interfere during the process, except in cases where participants asked for help.

5.2.2 Interview (30 minutes). Prior to the interview, we requested participants to assess their experiences with the System Usability Scale (SUS) and Creativity Support Index (CSI). Participants were asked to briefly describe their workflow and their task's outcomes. Additionally, we asked further questions across four categories related to their task experience: (1) workflow integration, (2) spatial reasoning, (3) ideation support, and (4) discussion.

5.3 Data Analysis

During the *Free Exploration Task*, both the screens of the tablet and laptop were recorded, and the area was videotaped to observe participants' physical movements, all with their consent. The *Interview* was also recorded. Two authors conducted semantic coding for each task and complemented each other's codes through peer review. Ultimately, we established the final codes through collaborative discussion.

5.4 Results

5.4.1 Post Survey. This section presents findings from video analysis of participants' user study sessions. We report on the observed behavior patterns and workflows during the task, focusing on how each participant approached capturing the provided 10-second animation. The results are organized chronologically into four stages: task receipt, recording, transmission to Blender, and editing within Blender. To characterize participants' workflows, we first categorized their time spent using the tablet versus Blender, and further segmented their activity within each device to reveal how they structured their shot composition process.

5.4.2 Free Exploration Task. In the free exploration task, participants worked according to their individual preferences, reflecting patterns observed in our formative study. By compiling their choices across each stage of the animation-making process, we were able to analyze distinct tendencies in relation to their backgrounds and working styles. As a result, in the following section, we describe participants' choices at each stage, as summarized in 9.

Previewing the Animation Before Recording. At the beginning of the session, some participants first reviewed the provided animation in Blender before recording (U2, U3, U4), while others previewed it directly on the tablet and immediately began composing the scene layout and framing within the AR view (U1, U5–U9). Several users explored different viewpoints while experimenting with scene scale or layout before committing to a final shot.

Shot Structure and Take Strategy. Participants exhibited different tendencies in how they structured their takes. For instance, U7 repeatedly recorded the same 10-second sequence as a one-take shot to refine a desired framing. Similarly, U6 and U8 each recorded only one or two continuous takes and then proceeded to edit or fine-tune shots entirely within Blender. U1 and U2 used a variation of the one-take method but paused the animation mid-way to record multiple framings in a single continuous session. In contrast, users such as U4, U5, and U9 filmed each sub-shot individually according to their own mental scene breakdown, and later stitched the cuts together during editing.

User Positioning During Recording. Participant movement during recording also varied. U3 and U6 completed their recordings entirely while seated, treating the AR scene like a manipulable object—adjusting its position, size, and angle from a fixed viewpoint. U1 and U7 stood throughout the session, while U2 scaled the scene to room-size and walked into the space to explore immersive close-up shots. Other participants alternated between sitting and standing based on the desired angle or framing. For example, U5, U8, and U9 began seated but stood when shooting low-angle or close-up views.

Camera Angle and Scene Placement. In terms of camera angles, most participants began with a high-angle shot due to initially placing the scene on a tabletop. U7 was the only participant who placed the scene on the floor and immediately started with an aerial perspective. U2 also experimented by positioning the scene on vertical surfaces such as walls. Later, U1, U2, and U3 enlarged the scene to explore standard or low-angle compositions. U5 lifted the scene along the Z-axis to simulate standard eye-level framing. U4 shot from a high angle while standing, but switched to seated shooting for additional takes to achieve standard angles. While U6 and U8 shot primarily from a high angle during AR recording, they performed active reframing and angle composition when editing in Blender.

Camera-in-Hand and World-in-Hand Interaction. We also observed distinct patterns in how participants physically interacted with the system. We categorize this behavior into two modes: world-in-hand, where users manipulated the AR scene by repositioning and scaling it; and camera-in-hand, where users physically moved the tablet to simulate camera movement. U1, U2, U5, U8, and U9 predominantly used camera-in-hand to perform movement-based shots. U3, U4, and U6 combined both modes, and U3 in particular used them in quick succession to simulate dynamic movement effects.

Reviewing Recordings During the Session. Participants differed in how they reviewed recordings. U2, U3, and U5 checked each take by transferring it to Blender immediately after shooting. U1

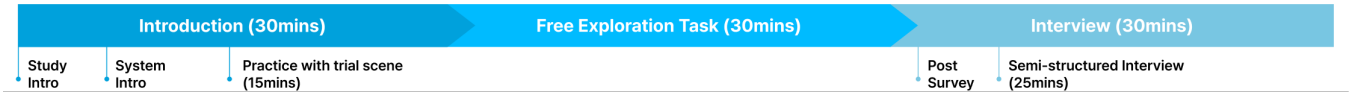

Figure 5: Overview of the user study procedure consisting of three phases: (1) Introduction (30 min), (2) Free Exploration (30 min), and (3) Interview (30 min).

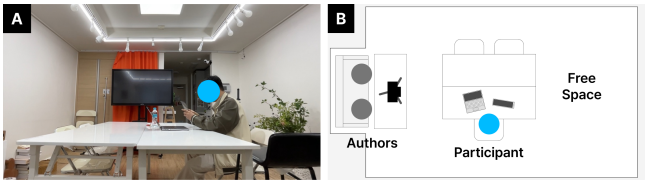

Figure 6: (A) Photograph of the actual study session showing a participant interacting with the system. (B) Floor plan of the setup. Participants were encouraged to freely move within the designated free space during the task.

adjusted camera names and reviewed shots in batch after recording was complete. U4 conducted scene rehearsals before filming and used post-review not to evaluate results, but to plan the next composition. U7 and U9 reviewed takes on the tablet directly after each recording, while U6 and U8 completed fewer than three takes total and reviewed them solely within Blender before proceeding to editing.

Timing of Post-Recording Edits and Keyframe Adjustments. Post-recording workflows also varied in the timing of keyframe edits. U1, U3, U6, U7, and U9 performed all keyframe edits at the end of the session. Among these, U3 and U7 used Blender primarily for playback, not revision. In contrast, U5 and U8 edited each shot immediately after it was recorded. U2, U4, and U6 made keyframe edits in Blender but returned to the tablet to reshoot when unsatisfied with a segment, inserting new frames back into the Blender timeline.

Post-Editing Practices in Blender. All participants performed minor keyframe adjustments to stabilize motion or correct screen framing, but editing depth differed based on their familiarity with Blender. U6 and U9 engaged in more active editing—adding new cuts and repositioning existing ones. U3 focused solely on reducing jitter by minimizing keyframes, while others performed only light refinements to smooth motion and polish composition.

6 Discussion

In this section, we propose a discussion which aligns to the research questions described above.

6.1 Revisiting RQ1: Usage of CamARa in individual's iterative and improvisational workflow

Based on a detailed analysis of participant sessions, we identified four emergent usage patterns that reflect how individuals integrated CamARa into their creative workflows. These patterns are not rigid categories, but rather illustrate diverse patterns of engagement shaped by individual experience, creative goals, and preferred working styles. The findings below are structured around these emergent usage patterns.

6.1.1 ***In-shoot Exploration for New Perspectives.*** Some participants used CamARa as a medium for real-time visual discovery. Rather than planning their shots in advance, they explored spatial arrangements and compositions interactively during filming. U1 and U2 exemplified this behavior. U1 tried several angles for the same 10-second clip, gradually refining the framing as they observed how the scene looked in AR. U2 took a more spatially immersive approach—arranging the animation not just on horizontal surfaces but also on walls, shifting perspectives to explore how the scene might change across different physical placements.

As U1 explained, seeing the AR scene placed on a desk made it *"much more visible than in Blender,"* which led them to rethink their framing mid-session. This exploratory use of the system aligns with creative tool research emphasizing iterative, non-linear workflows[3, 17].

6.1.2 ***Pre-shoot Exploration and Rehearsal.*** For some participants, CamARa served as a rehearsal environment for spatial and temporal composition. Users like U4, U5, and U9 used the AR interface to mentally test compositions and rehearse camera paths before recording their final takes. They reviewed the animation, identified key moments, and rehearsed physical camera movements to align with object motion.

U5 explained that being able to preview movement helped them *"ideate more effectively,"* as they could anticipate how objects would move and position the camera accordingly. This mode of use mirrors practices in animation blocking and previs, where creators iterate quickly through possible compositions before committing to keyframes.

6.1.3 ***Post-shoot Review and Immediate Retake.*** Some participants used CamARa for rapid feedback and iteration. U3 and U7, for example, recorded multiple takes of the same shot, reviewing the results after each and adjusting their approach accordingly. Rather than relying solely on planning or rehearsal, they used playback to identify shortcomings and re-record with refinements in mind.

As U3 described, being able to *"see the animated camera directly with my eyes"* reduced the need for post-editing, allowing them to act more intuitively and treat the AR recording as a form of live-action capture. This strategy reflects an improvisational editing loop, where users iteratively refine camera movements through rapid capture-feedback cycles.

6.1.4 ***In-shoot Navigation for Key Spatial Positions.*** Finally, some participants used CamARa primarily to block out rough camera paths, treating the AR interface as a spatial sketching tool. U6

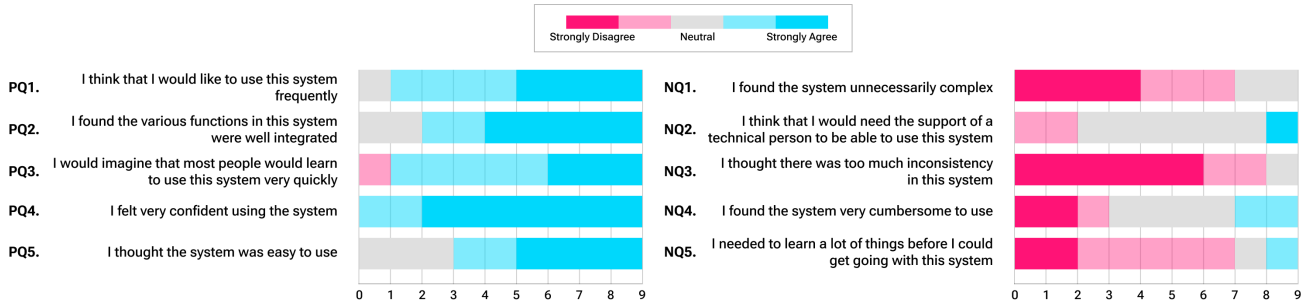

Figure 7: Participant responses to the System Usability Scale (SUS). Positive items (PQ1–PQ5) and negative items (NQ1–NQ5) are presented separately.

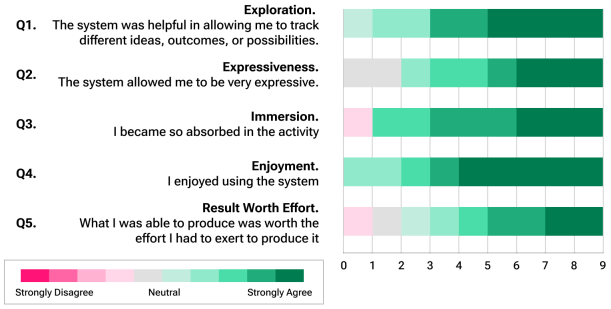

Figure 8: Participant responses to creativity-related questions for CamARa, adapted from the Creativity Support Index (CSI)

and U8 took this approach: they recorded one or two continuous takes covering the full animation, then moved to Blender for selective refinement. In the desktop environment, they retained only key moments and adjusted the framing through detailed keyframe editing.

As U8 explained, AR made it easy to *"rough out"* ideas that would later be finalized in Blender, especially since fine rotational control was more precise on desktop. This approach reflects common workflows in animation and filmmaking, where creators first establish spatial intent before layering in temporal and compositional detail [9].

Overall, our findings show that CamARa supported a range of workflows—from improvisational exploration to rehearsal, live retake, and spatial blocking—demonstrating its flexibility across different creative processes. These behaviors were shaped not only by the system's design but also by participants' prior experiences and the production norms of their respective domains.

6.2 Revisiting RQ2: Adaptive Use of CamARa

In this subsection, we review the types of support that users experience from CamARa, and how these differ depending on various levels of proficiency.

6.2.1 ***Time Efficiency.*** Through the interviews, we confirmed that time efficiency is the most critical factor influencing creators' willingness to adopt the system. Six out of nine participants (U1, U3, U4, U5, U7, and U9) mentioned reduced working time as the main advantage of using CamARa. U4 emphasized that the system significantly reduced trial and error in finding desirable shots, stating, *"Instead of rotating and testing the camera back and forth, I could just shoot and see the result immediately, then try again if needed."* U7 also noted that, unlike expert cinematographers who can visualize and execute a shot perfectly, he often relies on iterative testing, and that CamARa enabled quick exploration and helped determine better options efficiently. U5 found the tool intuitive, noting that it enabled faster results than their current Blender proficiency allows. CamARa particularly benefits non-professional users during the stages of exploration and composition. This aligns with findings from prior previs systems that aim to reduce the technical burden of early camera layout work[16, 22]. None of the professional participants (U2 and U6), who work as 3D animators and are already proficient with 2D interface tools, mentioned advantages of the system in the camera layout stage. However, they described other potential benefits from an expert perspective, which will be discussed in a later section.

6.2.2 ***Intuitiveness.*** Except for the two professional animators (U2 and U6), all seven other participants emphasized the intuitiveness of physically creating camera animations over using traditional 2D interfaces with a mouse and keyboard. U1 noted that in conventional tools, estimating how far the camera had moved based on numerical values felt unintuitive, whereas the physical interaction in CamARa enabled quicker and more embodied spatial awareness. U3 and U4 described it as fascinating to bring virtual camera control into the physical world, explaining that the system made camera movement significantly easier. This echoes earlier camera control metaphors such as 'scene-in-hand' and 'eyeball-in-hand,' which were found to facilitate intuitive navigation[20]. U4 also emphasized that, unlike in existing tools where users must constantly switch between camera view and third-person view while managing multiple control layers, the unified perspective in CamARa allowed her to focus solely on what was being shown. U5 and U9 appreciated the ability to freely and quickly adjust compositions and instantly preview the results. U7 remarked that *"if there had been three metaphorical walls between the user and Blender, this system removed about one and a half of them"*, and noted that users with prior camera handling experience would likely benefit even more. U8 reflected that in traditional interfaces, focusing solely on

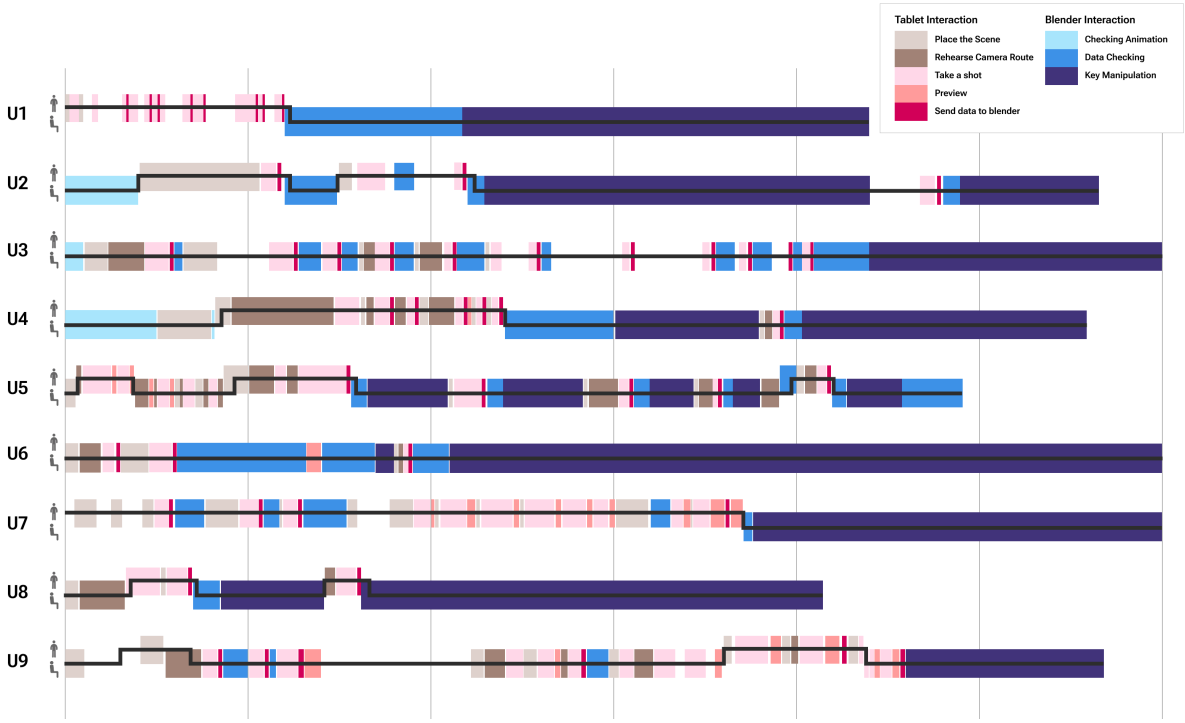

Figure 9: Visualization of participants' workflows during the free exploration task. Each row represents one participant (U1–U9), and the horizontal bars indicate sequential activities on the tablet and in Blender. Color-coded segments distinguish interaction types: tablet-based actions (e.g., scene placement, route rehearsal, shot recording, previewing) and Blender-based actions (e.g., animation review, data checking, and keyframe manipulation). In addition, black lines indicate changes in participants' physical posture—whether they were seated or standing—allowing us to infer their movement style and engagement during the task. This timeline illustrates how participants varied in stage transitions, iteration frequency, physical interaction, and tool usage strategies across the animation-making process.

the camera view often made it difficult to stay aware of the surrounding environment. With CamARa, however, it became easier to step back and consider the surroundings when planning shots.

6.2.3 ***Ideation and Creativity.*** With the exception of the two professional animators (U2 and U6), all other participants reported that CamARa provided strong support in the ideation and exploration stages of their creative workflows. U1 noted that being able to easily view the scene from various angles enabled quicker and more diverse shot development. Similarly, U4 described how the system revealed aspects of the scene that she could not see in Blender, which positively influenced her ideation process. U5 shared a specific moment when, while observing a bunny character running toward the camera using the system, he was inspired to reinterpret the scene with a more aggressive tone—an idea he had not considered when viewing the animation in Blender. He also emphasized that while mastering traditional 2D interfaces can lead to high-quality motion, CamARa offered a different creative perspective that helped generate new ideas. U3 similarly noted that the system did not necessarily make difficult shots easier to produce, but rather stimulated new ideas and encouraged exploration of a wider range of compositions. U7 compared his previous workflow, which relied heavily on mental visualization, with the immediate feedback provided by the system, describing it as a way to *"generate ideas through embodied experience."* U8 explained that using the AR view helped her consider environmental elements—such as rooftops—within the frame, informing spatial decisions earlier in the process. U9 appreciated that the system allowed them to easily create dynamic action shots, such as shaking movements, which would have been more difficult to achieve using conventional tools. These findings echo the importance of creativity support tools that enable rapid, embodied ideation in non-linear workflows[3, 17].

6.2.4 ***Immersion and Engagement.*** Regardless of participants' prior experience or professional background, CamARa received high ratings in terms of immersion and enjoyment (Q3 and Q4, Figure 8). This suggests that user engagement was not strongly influenced by familiarity with animation tools or level of expertise. U1 described the experience as *"really fun because it was easy to use, it sparked a desire to explore more creative applications."* U4, who typically works with 2D interfaces, noted that physically moving with the system was engaging, and that this embodied interaction provided a refreshing departure from her usual workflow. U5 also highlighted immersion as one of CamARa's most compelling qualities. Interestingly, even U6, a professional animator who stated that

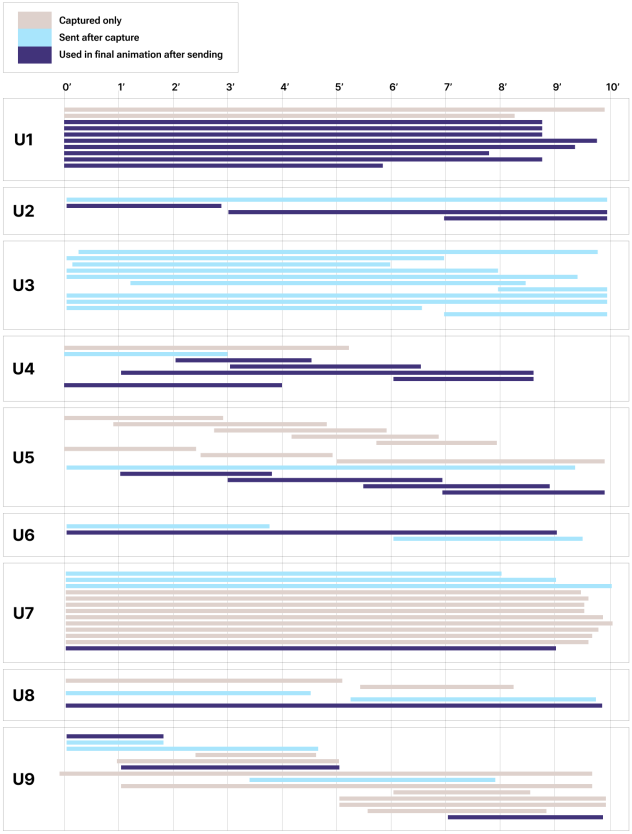

Figure 10: Timeline of recorded camera takes categorized by data transfer and final usage status. Each bar represents a single camera take within the 10-second animation sequence. Colors indicate different outcomes: gray for shots that were recorded but not sent to Blender, light blue for takes that were sent but not included in the final edit, and dark blue for those that were both transferred and used in the final composition. This visualization reveals individual differences in iteration, selectivity, and decision-making during the shot selection process.

he would not adopt the system due to its time inefficiency, acknowledged that physically holding and operating the system provided a sense of realism and presence that keyboard-and-mouse interaction could not offer. The only notably low rating came from U7, who reported discomfort due to the visibility of their hand tremors during interaction, which ultimately diminished their sense of immersion and confidence (Q3, Figure 8).

6.3 Applications

Expert groups (U1-U6; physical camera and virtual camera specialists), with the exception of students, discussed the potential applications of CamARa in various domains. Notably, two 3D animators (U2 and U6), who indicated that they would not incorporate the system into their workflow, suggested alternative use cases that could enhance the augmented reality environments.

6.3.1 ***Communication with Non-Experts.*** U1, U3, and U4 suggested CamARa as a potential communication tool for use with non-experts. U1 described that when users come up with new ideas impulsively, they could use the system to visualize and present those ideas immediately during meetings. Similarly, U4 proposed that CamARa could serve as a sketching tool to share diverse ideas within a team, and additionally noted that it would be beneficial to include a preview sharing feature. U3 pointed out that terminology can vary significantly across domains, and suggested that the system could facilitate communication with non-experts in the 3D field, such as clients.

6.3.2 ***Previsualization and Early Planning.*** U5 explained that he usually plans the scene before moving on to camera work. However, it noted that he will use the system in early planning to check the massing and richness of objects and the overall scene by looking through a tentative camera view. Furthermore, U6 described that creators typically use temporary objects, such as plain cubes, to check the mood of a scene before entering the main modeling stage. He suggested that creators could use physical objects seen through the AR view for blocking purposes. For example, in a scene where the camera looks up at the sky between buildings and follows a fast-moving object, rather than placing rough building models within a 3D scene, one could instead go to an actual street, film a fast-moving object, and immediately get a sense of the composition and feel in real-time.

6.3.3 ***Virtual Production.*** Some participants suggested using CamARa in virtual production workflows, emphasizing its strong use of augmented reality features, which they noted clearly differentiates it from other VR-based virtual camera systems. U1 described the system as particularly beneficial for virtual production grounded in real physical environments. U2 and U6, both professional 3D animators, expressed strong interest in using CamARa for live-action–based digital productions such as hybrid animation, VFX, and Fake Out of Home (FOOH)–style content. In typical workflows for these types of productions, creators usually film the physical environment separately and then track markers to overlay 3D elements on the captured footage. Participants suggested that if CamARa could record the AR background camera view as a separate video layer, it would be possible to utilize both the movement data for post-processing and the camera footage for later compositing.

6.4 System Limitations and Future Directions

6.4.1 Challenges in Creating Stable Movements. Six participants (U1, U4, U6, U7, U8, and U9) mentioned that the handshake was too prominent in their recordings, making them uncomfortable and the result feels unsatisfied. The challenge they mentioned was the primary reason for neutral or agreement on certain negative usability scales (NQ2 and NQ4, Figure 7), as well as the disagreement regarding the value of the effort (Q5, Figure 8). All mentioned that it would be better to have stabilizing features. Particularly, U7 noted, *"I kept recording again to reduce handshake. Having a stabilization feature would improve the result and help me stay motivated."* U3 and U9 also addressed that as the current system records the movements

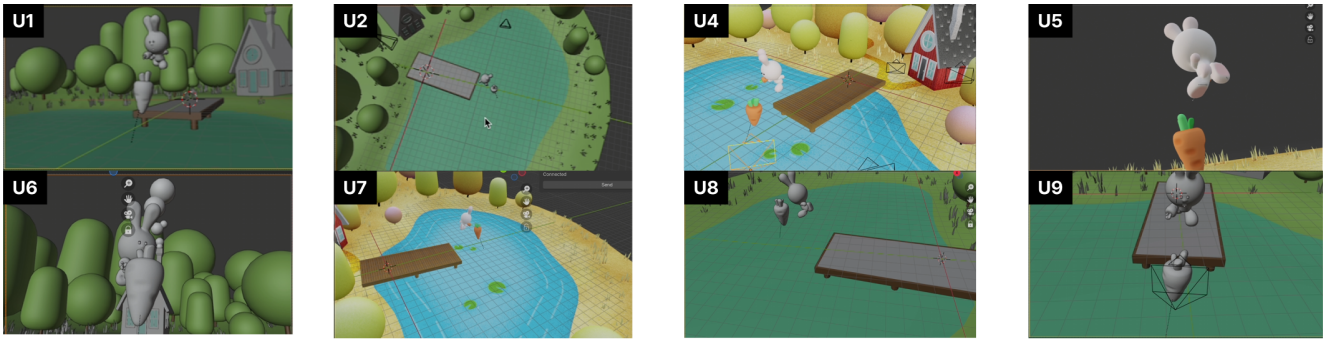

Figure 11: Final screenshots from each participant's output. U3 did not complete the editing process, so a final image is unavailable. Although all participants worked from the same source animation, the resulting scenes reflect distinct interpretations and composition strategies.

immediately, it would be good to have a few seconds of preparation time, like a timer, before starting the capturing.

6.4.2 Limitations of Augmented Reality. One of the inevitable limitations of augmented reality is that it has a fixed FOV value, as it uses built-in cameras on mobile devices. We also observed that many participants utilized and modified FOV settings when designing the shots during the formative study. U3 and U7 mentioned that it would be good to have camera effects such as FOV control or blur.

7 Limitations and Future Work

Need for Personalized Evaluation. The tasks in the studies asked participants to explore a given 3D scene and Blender using supplied macOS-based devices for 30 minutes. Since all the equipment was provided, some participants needed time to familiarize themselves with it. Additionally, several participants expressed that they required more time to fully explore and utilize the system. Specifically, U3 from the user study noted that more diverse attempts could be made if she get enough time, along with more comprehensive feedback and insights. She suggested conducting a study over a longer period, such as a take-home experiment, to allow creators to integrate the system into their existing workflows with their preferred devices and 3D tools (e.g., Blender, Cinema 4D, Keyshot, and Maya), as well as the scenes they typically use.

Rendering Performance of Augmented Reality. Given that our system operates on mobile devices with augmented reality compatibility, it is essential to evaluate its capabilities. The 3D scene utilized in our studies is approximately 14.7MB in size, requiring 127 draw calls, and includes 120,000 vertices and 200,000 triangles. If users want to load particularly heavy scenes with complex textures and animations, they may need to scale these down—such as by reducing object sharpness to low polygons or compressing texture images—to maintain optimal rendering performance. Future improvements may involve scene compression if it is necessary and ongoing performance monitoring.

Lack of Scene Diversity. Our formative study confirmed that various types of 3D scenes—such as animations, products, advertisements, and graphics—affect the workflow and scope of shots that creators compose. However, in the studies, we supplied only a cartoon-style, 10-second animation scene. During the study, one participant (U9) noted that CamARa offers a distinct advantage when she is creating action shots rather than more static, conventional shots. This implies that the utility of the system may vary depending on the style of the shot, influencing workflow based on the specific type of 3D scene. Future research could explore the specific types of shots that yield the most benefit when using the system.

8 Conclusion

This paper presents CamARa, an augmented reality–based camera layout system designed for intuitive and spatially aware shot composition. Based on the findings from a formative study, CamARa provides real-time spatial reference and awareness, enabling immersive interaction for individual creators across varying levels of expertise. Users can create camera animations by physically moving a mobile device while viewing the AR environment overlaid onto the physical world. Our qualitative user study showed that creators used CamARa to address specific challenges in their existing workflows. Rather than following a fixed process, participants adapted the system to their individual styles and goals. The study also confirmed that CamARa benefits non-expert creators in terms of intuitiveness, time efficiency, and creative ideation. Participants across proficiency levels found the system immersive and engaging. Beyond personal use, we identified potential applications of CamARa in professional domains such as communication, blocking, and virtual production, where AR-enhanced interaction can serve as a powerful tool.

Acknowledgments

This work utilizes the 3D asset "House on Lake"[7] by Bogdan Strielecki, which is provided under a CC Attribution-NonCommercial license, in the sample animation scene.

References

[1] Rafid Abdullah, Marc Christie, Guy Schofield, Christophe Lino, and Patrick Olivier. 2011. Advanced Composition in Virtual Camera Control. In *Smart Graphics*, Lutz Dickmann, Gerald Volkmann, Rainer Malaka, Susanne Boll, Antonio Krüger, and Patrick Olivier (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 13–24.

[7] https://skfb.ly/otNyR

[2] Autodesk 2009. *The New Art of Virtual Moviemaking*. Retrieved Jan 21, 2025 from https://images.autodesk.com/apac_grtrchina_main/files/the_new_art_of_virtual_moviemaking_-_autodesk_whitepaper1.pdf

[3] Bill Buxton. 2007. *Sketching User Experiences: Getting the Design Right and the Right Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[4] Philippe Bédard. 2022. Virtual Production and the Transformation of Cameras Mechanical, Virtual, and Actual. *Animation* 17, 2 (2022), 226–243. doi:10.1177/17468477221102498 arXiv:https://doi.org/10.1177/17468477221102498

[5] Marc Christie and Patrick Olivier. 2009. Camera control in computer graphics: models, techniques and applications. In *ACM SIGGRAPH ASIA 2009 Courses* (Yokohama, Japan) *(SIGGRAPH ASIA '09)*. Association for Computing Machinery, New York, NY, USA, Article 3, 197 pages. doi:10.1145/1665817.1665820

[6] Cinzia Cremona and Manolya Kavakli. 2023. *The Evolution of the Virtual Production Studio as a Game Changer in Filmmaking*. Springer International Publishing, Cham, 403–429. doi:10.1007/978-3-031-31360-8_14

[7] Quentin Galvane, I-Sheng Lin, Marc Christie, and Tsai-Yen Li. 2018. Immersive previz: VR authoring for film previsualisation. In *ACM SIGGRAPH 2018 Studio* (Vancouver, British Columbia, Canada) *(SIGGRAPH '18)*. Association for Computing Machinery, New York, NY, USA, Article 4, 2 pages. doi:10.1145/3214822.3214831

[8] Diego M. Garcia. 2023. *Shaping Visual Narratives : Camera Structure in Animation*. Ph. D. Dissertation. https://www.proquest.com/dissertations-theses/shaping-visual-narratives-camera-structure/docview/3098796978/se-2 Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2024-09-25.

[9] Michael Gleicher. 2001. Comparing Constraint-Based Motion Editing Methods. *Graphical Models* 63, 2 (mar 2001), 107–134. doi:10.1006/gmod.2001.0549

[10] Li-wei He, Michael F. Cohen, and David H. Salesin. 1996. The virtual cinematographer: a paradigm for automatic real-time camera control and directing. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96)*. Association for Computing Machinery, New York, NY, USA, 217–224. doi:10.1145/237170.237259

[11] Hongda Jiang, Marc Christie, Xi Wang, Libin Liu, Bin Wang, and Baoquan Chen. 2021. Camera keyframing with style and control. *ACM Trans. Graph.* 40, 6, Article 209 (Dec. 2021), 13 pages. doi:10.1145/3478513.3480533

[12] Hongda Jiang, Bin Wang, Xi Wang, Marc Christie, and Baoquan Chen. 2020. Example-driven virtual cinematography by learning camera behaviors. *ACM Trans. Graph.* 39, 4, Article 45 (Aug. 2020), 14 pages. doi:10.1145/3386569.3392427

[13] Hye-Young Jo, Ryo Suzuki, and Yoonji Kim. 2024. CollageVis: Rapid Previsualization Tool for Indie Filmmaking using Video Collages. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 164, 16 pages. doi:10.1145/3613904.3642575

[14] Thomas Kilkenny. 2023. Virtual Cinematography with Unreal Engine: Lean how Unreal Engine's virtual cinematography framework has given rise to new and advanced techniques in film and animation studios.. In *ACM SIGGRAPH 2023 Talks* (Los Angeles, CA, USA) *(SIGGRAPH '23)*. Association for Computing Machinery, New York, NY, USA, Article 12, 2 pages. doi:10.1145/3587421.3604476

[15] Amaury Louarn, Marc Christie, and Fabrice Lamarche. 2018. Automated staging for virtual cinematography. In *Proceedings of the 11th ACM SIGGRAPH Conference on Motion, Interaction and Games* (Limassol, Cyprus) *(MIG '18)*. Association for Computing Machinery, New York, NY, USA, Article 4, 10 pages. doi:10.1145/3274247.3274500

[16] Thomas Muender, Thomas Fröhlich, and Rainer Malaka. 2018. Empowering Creative People: Virtual Reality for Previsualization. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3170427.3188612

[17] Ben Shneiderman. 2007. Creativity support tools: accelerating discovery and innovation. *Commun. ACM* 50, 12 (Dec. 2007), 20–32. doi:10.1145/1323688.1323689

[18] Jingjing Tang, Liang Song, Jiabao Zeng, and Juncong Lin. 2021. The Virtual Camera Path in 3D Animation. In *Intelligent Computing Theories and Application*, De-Shuang Huang, Kang-Hyun Jo, Jianqiang Li, Valeriya Gribova, and Vitoantonio Bevilacqua (Eds.). Springer International Publishing, Cham, 874–882.

[19] VirtuCamera 2021. *VirtuCamera*. Retrieved Jan 21, 2025 from https://virtucamera.com/

[20] Colin Ware and Steven Osborne. 1990. Exploration and virtual camera control in virtual three dimensional environments. In *Proceedings of the 1990 Symposium on Interactive 3D Graphics* (Snowbird, Utah, USA) *(I3D '90)*. Association for Computing Machinery, New York, NY, USA, 175–183. doi:10.1145/91385.91442

[21] Zheng Wei, Xian Xu, Lik-Hang Lee, Wai Tong, Huamin Qu, and Pan Hui. 2023. Feeling Present! From Physical to Virtual Cinematography Lighting Education with Metashadow. In *Proceedings of the 31st ACM International Conference on Multimedia* (Ottawa ON, Canada) *(MM '23)*. Association for Computing Machinery, New York, NY, USA, 1127–1136. doi:10.1145/3581783.3612580

[22] Hock Hian Wong. 2012. Previsualization: assisting filmmakers in realizing their vision. In *SIGGRAPH Asia 2012 Courses* (Singapore, Singapore) *(SA '12)*. Association for Computing Machinery, New York, NY, USA, Article 9, 20 pages. doi:10.1145/2407783.2407792

[23] Jung Eun Yoo, Kwanggyoon Seo, Sanghun Park, Jaedong Kim, Dawon Lee, and Junyong Noh. 2021. Virtual Camera Layout Generation using a Reference Video. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 580, 11 pages. doi:10.1145/3411764.3445437

[24] Christian Zimmer, Daniel Drochtert, Christian Geiger, Michael Brink, and Rolf Mütze. 2017. Mobile previsualization using augmented reality: a use case from film production. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications* (Bangkok, Thailand) *(SA '17)*. Association for Computing Machinery, New York, NY, USA, Article 24, 5 pages. doi:10.1145/3132787.3132805