

# IGUANA: Immersive Guidance, Navigation, and Control for Consumer UAV

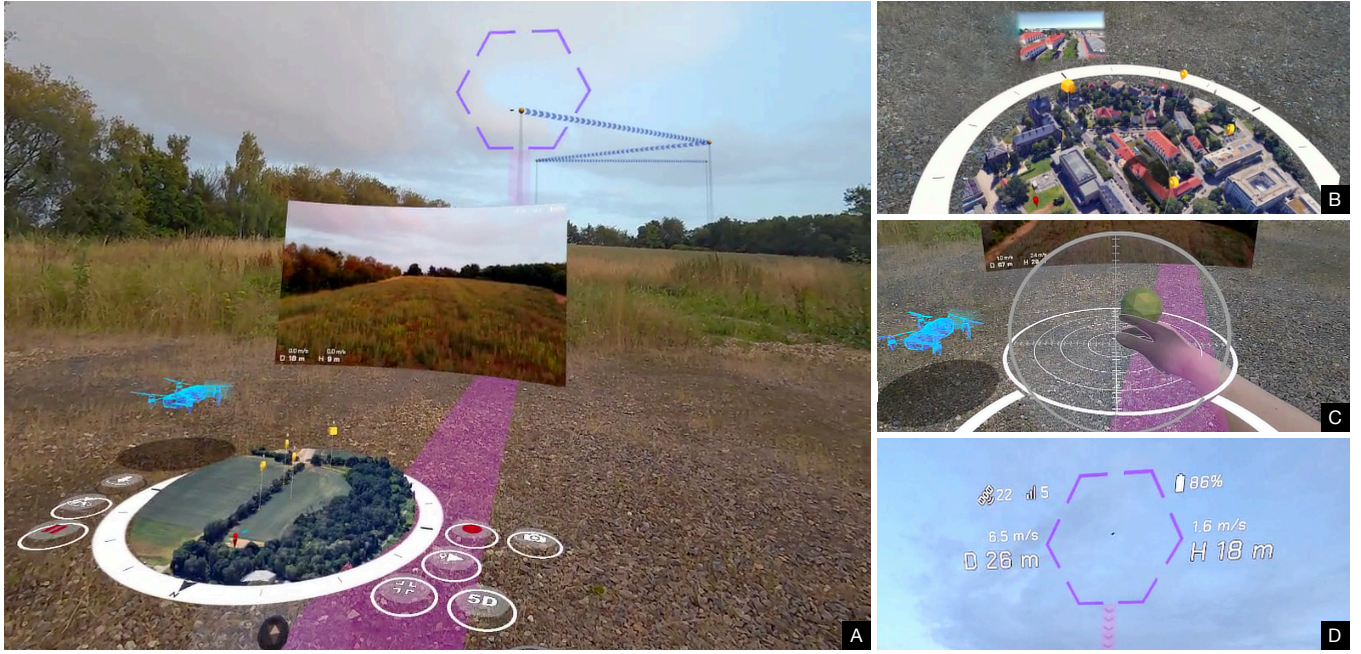
Victor Victor\*  
Chair of Software Technology  
Technische Universität Dresden  
Dresden, Germany  
victor.victor@tu-dresden.de

Tania Krisanty\*  
Chair of Computer Graphics and  
Visualization  
Technische Universität Dresden  
Dresden, Germany  
tania.krisanty@tu-dresden.de

Matthew McGinity  
Immersive Experience Lab  
Technische Universität Dresden  
Dresden, Germany  
matthew.mcginity@tu-dresden.de

Stefan Gumhold  
Chair of Computer Graphics and  
Visualization  
Technische Universität Dresden  
Dresden, Germany  
stefan.gumhold@tu-dresden.de

Uwe Aßmann  
Chair of Software Technology  
Technische Universität Dresden  
Dresden, Germany  
uwe.assmann@tu-dresden.de



**Figure 1: Screenshots of IGUANA captured on Meta Quest 3 headset. (A) Dashboard in 3D map interface mode. (B) High-level control using the 3D map interface. (C) Low-level control using the virtual ball interface. (D) Spatial overlay around the UAV.**

## Abstract

As the markets for unmanned aerial vehicles (UAVs) and mixed reality (MR) headsets continue to grow, recent research has increasingly explored their integration, which enables more intuitive, immersive,

\*Both authors contributed equally to this research.

VRST '25, Montreal, QC, Canada

© 2025 Copyright held by the owner/author(s).

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *31st ACM Symposium on Virtual Reality Software and Technology (VRST '25)*, November 12–14, 2025, Montreal, QC, Canada, <https://doi.org/10.1145/3756884.3766033>.

and situationally aware control systems. We present IGUANA, an MR-based immersive guidance, navigation, and control system for consumer UAVs. IGUANA introduces three key elements beyond conventional control interfaces: (1) a 3D terrain map interface with draggable waypoint markers and live camera preview for high-level control, (2) a novel spatial control metaphor that uses a virtual ball as a physical analogy for low-level control, and (3) a spatial overlay that helps track the UAV when it is not visible with the naked eye or visual line of sight is interrupted. We conducted a user study to evaluate our design, both quantitatively and qualitatively, and found that (1) the 3D map interface is intuitive and easy to use,

relieving users from manual control and suggesting improved accuracy and consistency with lower perceived workload relative to conventional dual-stick controller, (2) the virtual ball interface is intuitive but limited by the lack of physical feedback, and (3) the spatial overlay is very useful in enhancing the users' situational awareness.

## CCS Concepts

- **Human-centered computing** → *Mixed / augmented reality*;
- **Computing methodologies** → *Mixed / augmented reality*.

## Keywords

Unmanned Aerial Vehicle, Drone, Spatial User Interface, 3D User Interface, Mixed Reality, Augmented Reality

### ACM Reference Format:

Victor Victor, Tania Krisanty, Matthew McGinity, Stefan Gumhold, and Uwe Aßmann. 2025. IGUANA: Immersive Guidance, Navigation, and Control for Consumer UAV. In *31st ACM Symposium on Virtual Reality Software and Technology (VRST '25)*, November 12–14, 2025, Montreal, QC, Canada. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3756884.3766033>

## 1 Introduction

Consumer Unmanned Aerial Vehicles (UAVs) commonly carry cameras for aerial photography, surveying, and monitoring. Their ability to fly over obstacles, water, and rough terrain makes them well-suited for observing hard-to-reach or large-scale areas, enabling tasks that would otherwise be dangerous, costly, or infeasible.

While UAV motion is intuitive (objects move up/down, turn, and tilt in 3D), producing that motion with conventional controllers is not. Users must map intended actions onto limited rotational or translational axes, constrained by the hardware. Dual-stick controllers assign two axes per stick, often requiring simultaneous coordination of both. Users must also operate a gimbal-mounted camera, creating additional complexity. Even basic tasks demand managing both vehicle and camera orientation, shifting attention from *what to do* to *how to do it*.

Higher-level autonomy, such as waypoint missions, reduces manual navigation by following user-defined Global Positioning System (GPS) waypoints. However, authoring missions remains tedious and error-prone since users still have to manually enter coordinates and altitudes for each waypoint. As fine-tuning at each waypoint is still common, configuring camera orientation (e.g., pitch) and working without immediate contextual feedback or previews often forces reliance on external tools like Google Earth [16].

Maintaining a visual line of sight further conflicts with monitoring a controller-mounted live camera feed, increasing cognitive load and reducing situational awareness [4]. Novices are especially affected by this problem. While First-Person View (FPV) goggles offer immersive video, they also limit situational awareness [36].

These challenges present opportunities for Mixed Reality (MR) solutions. Waypoint planning can benefit from World in Miniature (WIM) interface [40], a scaled 3D representation shown to be safer and more efficient than joysticks or 2D touchscreens [33]. Prior work often uses Virtual Reality (VR) controllers for placement, which can feel cumbersome compared to hand tracking.

Hand-tracking improves waypoint placement [34], contextual previewing, and route planning while reducing reliance on numeric coordinates and external tools. Headset-based MR also yields better virtual object placement experiences than handheld Augmented Reality (AR) [43]. A 3D user interface can replace dual-stick limitations by unifying two sticks and a dial into a natural, intuitive virtual ball control. MR can also enhance situational awareness with a persistent, headset-visible indicator of the UAV's real-world location, even when out of sight.

In this work, we propose a design for an MR-based guidance, navigation, and control system for consumer UAVs. We present IGUANA as a proof-of-concept implementation of this design and evaluate its usability and acceptance in a user study. This study addresses how IGUANA supports task performance in UAV control, how users experience IGUANA in terms of usability and satisfaction, and whether IGUANA enhances situational awareness during UAV operation. The IGUANA source code is publicly available to facilitate future research and development [42].

## 2 Related Work

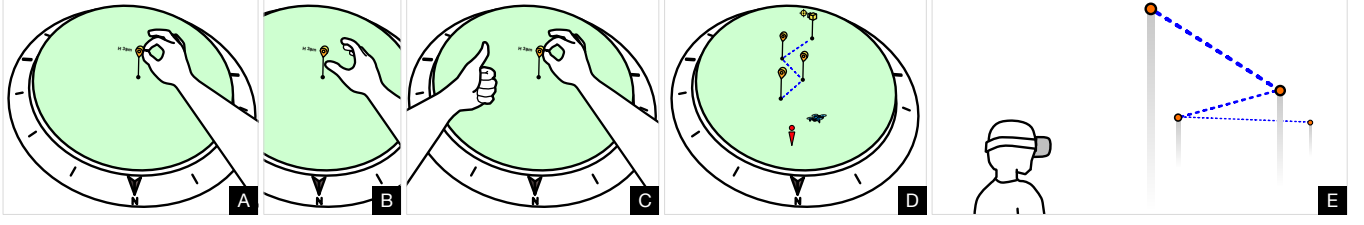
**High-Level Control.** We define high-level control as supervisory or task-level inputs that specify goals, focusing on the desired outcome rather than the specific control actions required to achieve it. Path planning is a common high-level UAV control method, which can be implemented in a grid-based fashion [43] or using waypoints [11, 26, 32–34]. A 3D map [11, 25, 26, 34, 37] or 3D environment [5, 46], acting as WIM, is employed to provide spatial context. Some works reconstructed the 3D map from the UAV's camera [28, 31, 34], while others used pre-existing 3D maps, such as 3D satellite views of the environment [11] or Mapbox [30, 37]. Voice commands [18, 29], gaze with hand-gestures [15, 28], or a combination of both [19] are other methods of high-level control, providing intent- or goal-based control.

Using voice commands requires some understanding of the environment, such as predefined landmarks [19] or object detection [18], which may not be practical. On the other hand, some works demonstrate path planning efficiency gains over traditional 2D touch-screen interface [14, 32, 33, 43].

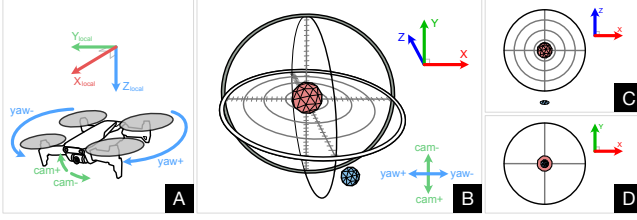
In this work, we propose a 3D map interface with waypoint path planning that utilizes photorealistic 3D tiles from Google Earth [5, 37, 39], providing a rich real-world spatial context for outdoor environments.

**Low-Level Control.** Low-level control enables users to directly manipulate the UAV's movement and orientation, providing more granular control over its flight. The common methods include the use of spatial controllers [3, 5, 8, 11, 27, 41], dual-stick controllers [22], or hand-gestures [2, 15, 23, 28, 31]. Some works have proposed using a car-steering metaphor [9], a VR Treadmill [38], body posture and voice [29], or hand poses [1]. Although the car steering or VR Treadmill provides an intuitive and arguably precise control, they are impractical, especially for outdoor usage. Additionally, the hand poses may add a mental load by requiring users to remember the different gestures needed for each command.

The hand-gesture methods that interact directly with the UAV may not be suitable for outdoor use due to the limited interaction distance and the need for a line-of-sight [15, 28, 31]. In contrast,



**Figure 2: 3D map interface.** (A) Pinch to select a marker. (B) Release to place the marker. (C) Alternatively, confirm placement with a thumbs-up gesture. (D) Overview with user marker (red), UAV (blue), waypoint and virtual camera markers (yellow), and the planned trajectory (dashed line). (E) Spatial overlay of waypoints and the planned trajectory.



**Figure 3: (A) UAV local coordinate system. (B) Virtual ball interface aligned to the world coordinate system; red sphere controls UAV flight direction and speed; blue sphere controls UAV yaw and camera pitch. (C) Top view and (D) front view of the virtual ball interface.**

hand-gesture methods that enable the UAV to mimic the hand’s pose [2, 23] are better suited for outdoor use. The disadvantage of hand-gesture control is the lack of physical feedback, which may affect the control precision and user confidence [20, 38, 47].

In this work, we propose a virtual ball interface that enables more natural and intuitive control, acting like a joystick with a rest position when not in use [2, 23].

*Spatial Overlay.* Atanasyan et al. [4] made a significant contribution to the field of spatial overlays for UAV operation, with UAV attitude and position indicators. When waypoint control is involved, previous studies have demonstrated the benefits of displaying the planned waypoint path in a spatial overlay [14, 18, 44, 48].

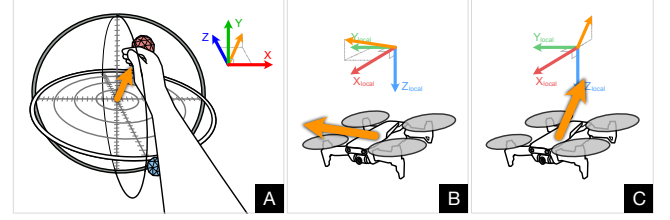
In this work, we propose a spatial overlay to enhance situational awareness, building on previous work that shows the position of the UAV in the air [4, 48], attitude and telemetry information [4, 28], as well as the waypoint path [14, 18, 44, 48].

### 3 IGUANA

#### 3.1 Design

IGUANA features three key components: high-level control via 3D map interface, low-level control via virtual ball interface, and enhanced situational awareness via spatial overlay.

*3.1.1 High-Level Control via 3D Map Interface.* High-level control refers to planning and executing waypoint missions for the UAV. This is achieved through a 3D map interface that displays a spatially accurate, scaled-down 3D reconstruction of the real world. The 3D map is displayed in a floating circular cutout in front of the user,

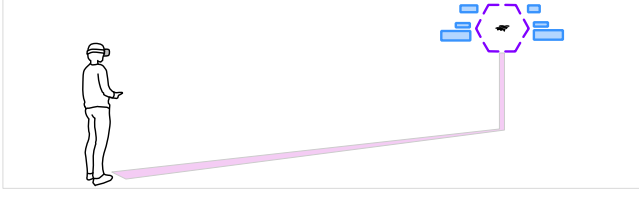


**Figure 4: (A) Push on the virtual ball diagonally toward the user’s forward upper right. Resulting motion in (B) drone-centric and (C) user-centric control modes.**

which can be panned, rotated, and zoomed (see Fig. 1-A). On top of the map, the user can place, move, and orient waypoint markers using pinch gestures (see Fig. 2-A) and release gestures (see Fig. 2-B). An additional thumbs-up gesture (with the other hand) is added to help release the marker more precisely (see Fig. 2-C). There is also a virtual camera marker (see Fig. 2-D, top-most marker), in addition to the regular waypoint markers, which displays a live preview of the scene to assist with pose adjustment (see Fig. 1-B, floating window at the top). The virtual camera marker has an additional reticle helper that can be pulled to adjust the UAV’s yaw and camera pitch. When a hand is near a waypoint or virtual camera marker, height information is displayed to assist with altitude planning, particularly to avoid collisions with the terrain (see Fig. 2-A, B, and C). A UAV marker and a user marker are also displayed on the map (see Fig. 2-D, two bottom-most markers).

*3.1.2 Low-Level Control via Virtual Ball Interface.* Low-level control refers to direct manipulation of a UAV’s attitude (roll, pitch, yaw, and thrust) to achieve a desired flight direction. To abstract this complexity, we introduce a control metaphor based on a virtual ball, serving as a physical analogy for the UAV. The interface consists of two spheres: a big sphere for translational movement (forward/backward, left/right, up/down) and a small sphere for yaw and camera pitch (see Fig. 3-A and B, big sphere is colored red, small sphere is colored blue). The big sphere has 3degrees of freedom (DoF) movement, bounded within a larger transparent sphere, while the small sphere has 2DoF, constrained to move along a vertical circular path (see Fig. 3-C and D). To initiate translational control, the user reaches out and pushes the big sphere in the desired direction. The small sphere, however, is manipulated using a





**Figure 5: Spatial overlay with horizontal and vertical leading lines (magenta), UAV frame (purple), and telemetry (blue).**

pinch gesture. It is possible to control both spheres simultaneously using both hands, combining translational and rotational control.

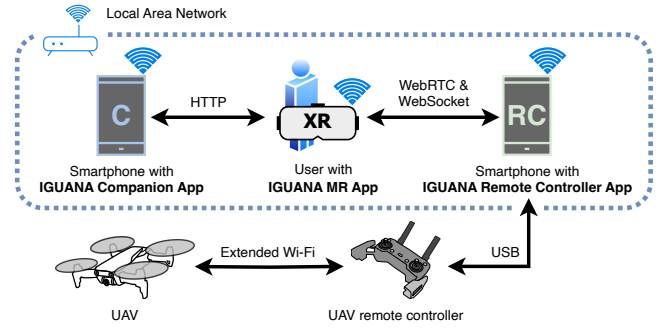
We implemented two interchangeable control schemes for the virtual ball interface: the *drone-centric* and *user-centric* control. In the *drone-centric* control, the virtual ball interface is aligned with the UAV's local coordinate system. Pushing the big sphere in a certain direction will cause the UAV to move in the same direction (see Fig. 4-A and B), as it places the user in a virtual position behind the UAV. The *user-centric* control, on the other hand, aligns the virtual ball interface with the world coordinate system. Pushing the big sphere in a certain direction will cause the UAV to move in that direction relative to the user, regardless of the UAV's current orientation (see Fig. 4-A and C). This control scheme is implemented to help reduce the cognitive load associated with heading-relative controls, especially when controlling the UAV when it is positioned in front of the user, rather than through the camera stream.

**3.1.3 Enhanced Situational Awareness via Spatial Overlay.** Spatial overlay refers to augmenting the real world with virtual information anchored to the UAV's actual position. This helps users maintain awareness of the UAV's position, especially when it is too far away or beyond the line of sight (see Fig. 5, purple frame). To help track the UAV's position, we added horizontal leading lines that extend from the user's feet to the UAV's ground position and vertical leading lines that extend from the ground to the UAV's altitude (see Fig. 5, magenta lines). The overlay displays its telemetry information, such as ground distance, ground speed, altitude, vertical speed, GPS signal strength, and battery level (see Fig. 1-D and 5, blue panels). As the user plans the waypoint missions with the 3D map interface (see Fig. 2-D), we also visualize the planned waypoint path in the real world, providing context for the UAV's intended trajectory (see Fig. 2-E). Walker et al. [44] found that this visualization effectively communicates the intended flight path.

## 3.2 Implementation

Based on the system architecture depicted in Figure 6, we implemented IGUANA using: Da-Jiang Innovations (DJI) Mavic Air, Meta Quest 3, Samsung Galaxy Note20 Ultra for remote controller application, and iPhone 16 Pro Max for companion application. IGUANA, however, is designed to be compatible with any consumer UAV that supports waypoint missions and telemetry via a mobile SDK.

**3.2.1 Mixed Reality Application.** The MR application is built using Unity 6 with Meta Extended Reality (XR) SDK v77. Figure 1-A shows the dashboard with the 3D map interface and the UAV marked with a spatial overlay in the background.



**Figure 6: IGUANA system architecture.**

**3D Map Interface.** We implemented the 3D map interface using Cesium-based Google's Photorealistic 3D Tiles [7] (see Fig. 1-B). The UAV marker updates in real-time, based on its actual GPS position and altitude value. The user marker position is determined by the iPhone's location. Users can interact with the waypoint and virtual camera markers using pinch gestures for both positioning and orientation adjustments. The virtual camera marker renders a live view of the 3D tiles environment into the viewport and represents the camera mounted on the UAV, with its rotation and field of view constrained to match it.

After placing the waypoints, the user can start the mission by pressing the start mission button. Waypoint data is then transmitted to the remote controller application, which translates each marker's position and orientation into flight parameters, and then forwards it to the UAV. Each position is converted into GPS coordinates and an altitude. The rotation around the X-axis sets the camera's pitch, while rotation around the Y-axis determines the UAV's heading. One important note is that the altitude value sent to and received from the UAV is defined relative to the takeoff point, not in World Geodetic System 1984 (WGS84). Therefore, it is necessary to first obtain the altitude of the takeoff location by querying the height of the 3D tiles where the UAV is located.

**Virtual Ball Interface.** The virtual ball interface (see Fig. 1-C) is implemented using Meta XR Interaction SDK v77. To support intuitive interaction, we provide visual and auditory feedback. When the user pushes the ball, it changes color to indicate that it is being pushed, and reverts to its original color upon release. Additionally, sound cues vary based on how far the ball is pushed, compensating for the lack of haptic feedback. The distance of the virtual ball from its resting position is mapped to the UAV's pitch, roll, and thrust.

**Spatial Overlay.** To overlay the UAV's position accurately, we use Cesium georeference in real-world scale mode, without loading terrain or imagery tiles. This enables overlays to be positioned based on the UAV's latitude, longitude, and altitude in the WGS84 system. To ensure correct spatial alignment, we first set the user's position as the origin of Cesium's georeference. At initialization, the user holds an iPhone running the Companion Application until the headset detects the AprilTag. AprilTag is a robust marker-based tracking system for accurate 6DoF pose estimation [24]. The headset computes the AprilTag's pose in its coordinate system and simultaneously queries the iPhone's geographic location.

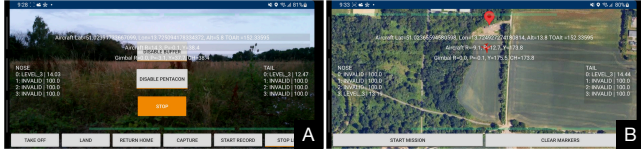


Figure 7: Screenshots from IGUANA remote controller application showing (A) the camera and (B) the map interfaces.



Figure 8: (A) Screenshot of the companion application running on an iPhone. (B) IGUANA MR application scanning the AprilTag displayed on the companion application.

**3.2.2 Remote Controller Application.** The remote controller application is built in Java using the DJI Mobile SDK v4.18. It connects to the UAV remote controller via Universal Serial Bus (USB) and to the headset via Wi-Fi. It has two main interfaces, switchable by tapping the screen, namely the camera interface (see Fig. 7-A) and the map interface (see Fig. 7-B). The camera interface displays the live feed from the UAV's camera and includes all the important buttons for emergency stop, takeoff, landing, return to home, and a safety override that blocks control input from the virtual ball interface in the MR application. On the other hand, the map interface is useful for showing the UAV's position on the map and monitoring the waypoint mission, as it displays every waypoint marker as it is sent from the headset.

**3.2.3 Companion Application.** The companion application is built in Swift (see Fig. 8-A and B). It connects to the headset via Wi-Fi and provides the phone's GPS and compass data as IGUANA's position and orientation. Additionally, it displays an AprilTag, which is used as a pose reference for the headset.

**3.2.4 Communication Protocol.** Communication between the headset and the Android phone is handled via a WebSocket and Web Real-Time Communication (WebRTC), a protocol suggested by previous work [45]. High-frequency data, such as telemetry updates at 10 Hz, is sent from the phone to the headset via WebRTC data channel. Commands are sent from the headset to the phone via WebSocket, ensuring reliable delivery and correct execution order. To minimize latency and payload size, data is transmitted as raw bytes, where the first byte indicates the message type, and the number and structure of subsequent bytes vary depending on the type. This compact format ensures fast and consistent communication for frequent updates. For real-time preview and situational awareness, the UAV's live camera feed was streamed to the headset via WebRTC. This protocol was preferred over the UAV's built-in

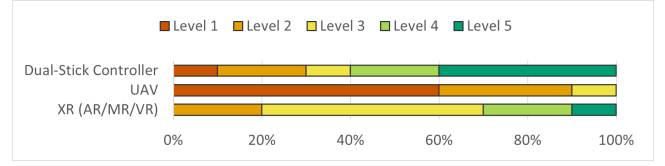


Figure 9: Participants' domain-specific experience level.

Real-Time Messaging Protocol (RTMP) streaming due to its lower latency.

## 4 User Study

We evaluated IGUANA in a within-subjects study where participants completed the same tasks twice: once with IGUANA and once with a baseline. For the baseline, referred to as RC, we used the DJI Mavic Air remote controller and DJI GO 4 [13] on a Samsung Galaxy S10. Although IGUANA's 3D map interface is designed for complex environments, such as dense urban terrain, we conducted the study in a safe, open field to minimize risk.

### 4.1 Objectives

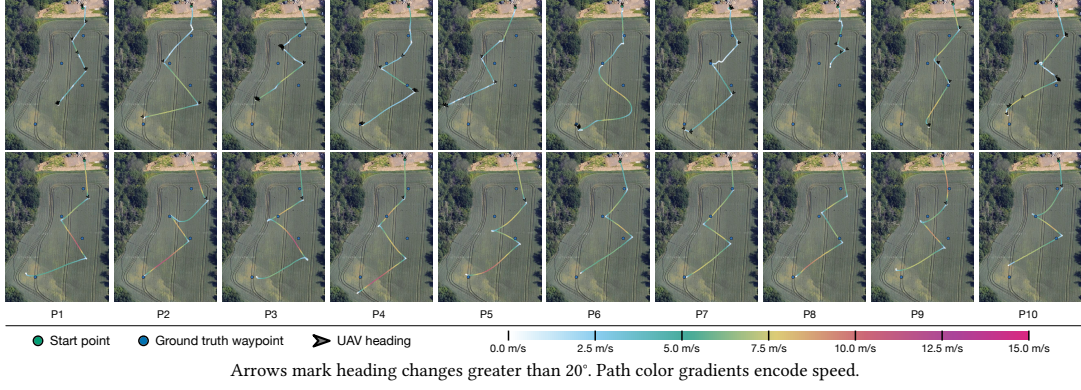
The main objectives were to evaluate the usability and effectiveness of IGUANA in comparison to conventional control interfaces for consumer UAVs. We aimed to assess the following aspects: (1) **Task Performance:** Measure the time taken and accuracy in completing a specific UAV control task using both IGUANA and the RC. (2) **User Experience:** Gather qualitative feedback on the overall user experience, including ease of use, intuitiveness, and satisfaction with the control methods. (3) **Situational Awareness:** Evaluate how well IGUANA enhances situational awareness during UAV operation.

### 4.2 Study Design

**4.2.1 Quantitative Measure.** Participants completed Task A for this quantitative measure. To evaluate objective performance, we utilized flight logs generated automatically by the DJI system for each flight.

**Task A: Exploration.** For the exploration task, participants were given a printed map marked with four waypoints. They were instructed to pilot the UAV through each of the four waypoints on the printed map, while avoiding crashing into the terrain below, and to capture a photo of a designated tree at the final waypoint. Using IGUANA, participants completed this task by entering the 3D map interface, placing waypoint markers, adjusting the virtual camera angle as needed, and then initiating the waypoint mission. Using the RC, participants manually flew the UAV through each waypoint using the dual-stick controller and adjusted the camera gimbal at the final waypoint to take a photo.

**Self Assessment.** We also collected subjective data using post-task and post-test questionnaires. We used the Single Ease Question (SEQ) [35] to measure task difficulty, the NASA Task Load Index (NASA-TLX) [17] to assess workload, the System Usability Scale (SUS) [6] for usability, and the Technology Acceptance Model (TAM) [10] for user acceptance.



**Figure 10: Paths performed for Task A. Top row: RC. Bottom row: IGUANA 3D map interface.**

**4.2.2 Qualitative Measure.** Participants completed Task B as part of the qualitative evaluation. Unlike Task A, which focused on objective performance metrics, Task B was designed to assess subjective impressions, preferences, and perceived ease of use when interacting with the user-centric control mode.

*Task B: Going Home.* This task began once the UAV had reached the participant’s final waypoint from the exploration task, located on their far right. Participants were instructed to move the UAV from the far right to the center (far in front of them) and then straight toward their position, forming an upside-down capital ‘L’ trajectory. Using IGUANA, participants used the virtual ball interface in user-centric mode. In this mode, the task can be completed by pushing the virtual ball to the left until the UAV reaches the center, then pulling it towards themselves until the UAV returns to the participant.

*Interview.* Additionally, a semi-structured interview was conducted after the study to gather qualitative feedback.

### 4.3 Participants

Ten participants (2 female, 8 male; age 20–42,  $M = 29.8$ ,  $SD = 6.3$ ) took part. Experience in XR (AR/MR/VR), UAV, and dual-stick controller was self-rated on a 1–5 scale (1 = beginner, 5 = expert); distributions are shown in Figure 9. Each participant received a compensation of €20 for ~60 minutes of participation.

### 4.4 Apparatus

Apart from the hardware listed in Section 3.2, the study setup included: GL.iNet AXT1800 router, iPhone 15 Pro as a mobile hotspot for the router, MacBook Pro for headset casting, Windows 11 laptop running DJI Assistant 2 (Consumer Drones Series) flight simulator [12], and Samsung Galaxy S10 running DJI GO 4 [13].

### 4.5 Procedure

Each session began with an indoor tutorial where participants practiced using each system, with the UAV connected to the Windows laptop via a USB cable, providing simulated telemetry. Following the tutorial, the session moved to an open field (approx. 100m x 250m) with ground elevation ranging from 215m to 310m, surrounded by trees. Participants then used the same setup, with the UAV placed

on the ground. An umbrella was used as needed to prevent sunlight from interfering with hand tracking.

We applied counterbalancing to avoid learning effects in our within-subjects design. Half of the participants started with RC, while the other half started with IGUANA. After completing each task, participants filled out the SEQ and NASA-TLX questionnaires. After completing all tasks for one system, participants filled out SUS and TAM questionnaires.

Once all questionnaires are completed, the semi-structured interview begins. We asked open-ended questions about their experience with both our and regular control systems, such as what they liked and disliked about each system, what aspects they found intuitive or confusing, and how they felt during interaction. Finally, we debriefed the participants and gave them the compensation.

### 4.6 Data Analysis

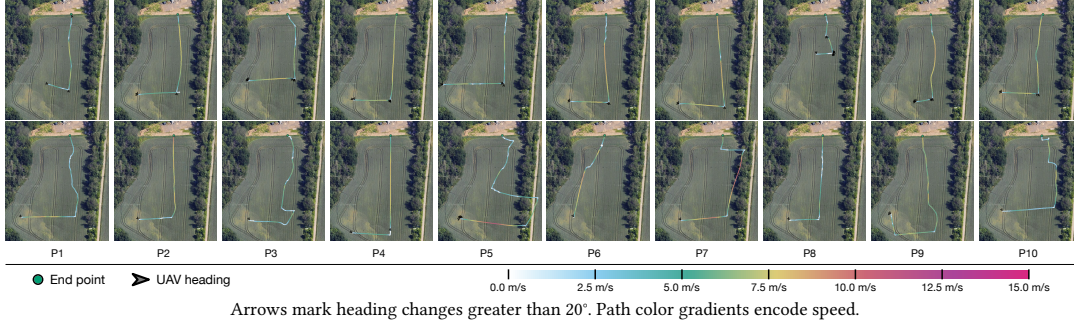
For each participant, we collected two flight logs: one from their IGUANA flight and one from their RC flight. These logs contain detailed telemetry data, including position, velocity, and control inputs. From this data, we derived a performance score by comparing the actual flight trajectory to the predefined waypoints and calculated the task completion time. Photos taken were also collected to verify task completion.

*Task A: Exploration.* The performance score was calculated using Equation (1).

$$S = \frac{\overline{D}_{wp} + \begin{cases} \left(1 - \frac{T - T_{\min}}{T_{\max} - T_{\min}}\right) + P & \text{if } D_{wp_4} < \delta, \\ 0 & \text{otherwise} \end{cases}}{3}, \quad (1)$$

where  $\overline{D}_{wp}$  is the average closest distance score,  $D_{wp_4}$  is the closest distance score to the final waypoint,  $T$  is the participant’s completion time,  $T_{\min}$  and  $T_{\max}$  are the minimum and maximum completion times across all participants, and  $P$  is a binary photo indicator (1 if a photo was taken, 0 otherwise). A waypoint was considered passed if the UAV’s trajectory came within a predefined distance threshold ( $\delta = 10$  m). Completion time and photo were only factored into the score if the final (farthest) waypoint was successfully reached,





**Figure 11: Paths performed for Task B. Top row: RC. Bottom row: IGUANA virtual ball interface (user-centric).**

ensuring that the object was captured from a reasonable distance to be considered valid.

## 5 Results

To support our analysis, we developed a custom JavaScript tool that parses the flight logs and visualizes the UAV's trajectory on a map, along with relevant telemetry data (e.g., speed, heading). The resulting visualizations are shown in Figure 10 and 11, illustrating that for Task A, all participants produced similar trajectories with the 3D map interface (see Fig. 10, bottom row), in contrast to the more varied paths observed with RC (see Fig. 10, top row).

### 5.1 Quantitative Results

To prepare the data for statistical analysis, we first assessed the normality of the paired differences for all comparisons using Shapiro-Wilk tests and Q-Q plots. The assumption of normality was met for all variables ( $p = .061-.978$ ) except for Satisfaction in the TAM ( $p = .003$ ), Physical Demand in the NASA-TLX for Task A ( $p = .001$ ), and Temporal Demand in the NASA-TLX for Task B ( $p = .034$ ). Therefore, Wilcoxon signed-rank tests were used for these three comparisons, and paired t-tests were used for the rest. To control the family-wise error rate, we applied Holm-Bonferroni corrections ( $\alpha = .05$ ) within three predefined families of tests: SUS and TAM ( $m = 6$ ), Task A measures including SEQ, NASA-TLX, completion time, average closest distance to waypoints, and performance score ( $m = 10$ ), and Task B measures including SEQ and NASA-TLX ( $m = 7$ ). All analyses were conducted in JASP [21].

SUS scores were generally higher for IGUANA ( $M = 80.5$ ,  $SD = 2.98$ ) than for RC ( $M = 68.25$ ,  $SD = 5.89$ ) (see Fig. 13-A). A paired-t test results in a non-significant difference ( $t(9) = 2.293$ ,  $p_{adj} = .192$ ), indicating usability was comparable to RC.

Paired t-tests on the TAM components revealed that IGUANA received significantly higher ratings than RC for Perceived Usefulness ( $t(9) = 5.609$ ,  $p_{adj} = .002$ ) and Behavioral Intention to Use ( $t(9) = 5.238$ ,  $p_{adj} = .003$ ) (see Fig. 13-B). No significant differences were found for Perceived Ease of Use ( $t(9) = 1.832$ ,  $p_{adj} = .3$ ) or Trust ( $t(9) = 1.222$ ,  $p_{adj} = .506$ ), nor for Satisfaction (Wilcoxon signed-rank test;  $W = 4$ ,  $p_{adj} = .85$ ). This suggests that IGUANA is perceived as more useful with higher intent to adopt, while ease of use, trust, and satisfaction are comparable to RC.

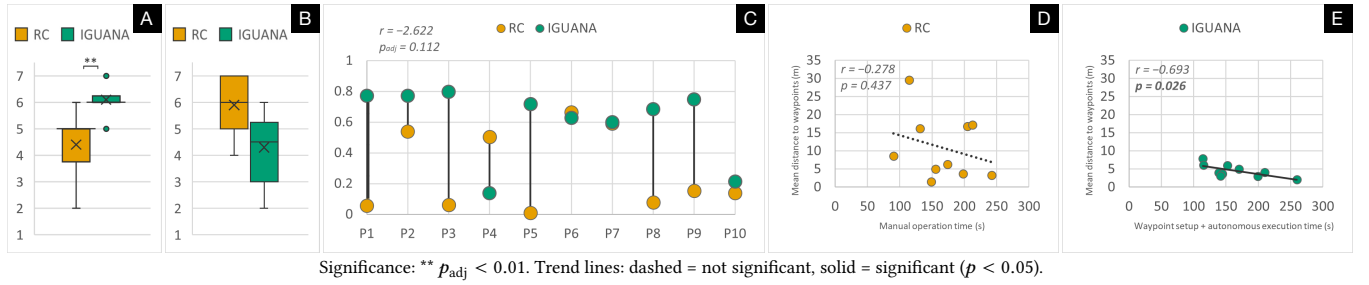
For Task A, we first analyzed completion time and the average closest distance to waypoints as separate metrics. Paired t-tests

showed no significant differences between the 3D map interface and RC for either completion time ( $t(9) = 0.195$ ,  $p_{adj} = 1$ ) or mean distance ( $t(9) = 2.097$ ,  $p_{adj} = .195$ ). For the photo-taking subtask, IGUANA achieved an 80% success rate, while only three of ten participants succeeded with RC. We then examined the relationship between these two metrics using Pearson correlation. For IGUANA, a strong, significant negative correlation ( $r = -0.693$ ,  $p = .026$ ) showed that a longer, more deliberate strategy was linked to greater accuracy (shorter mean distance). This downward trend is visible in the scatter plot (see Fig. 12-E). In contrast, the RC showed no such relationship ( $r = -0.278$ ,  $p = .437$ ) (see Fig. 12-D). We then analyzed performance scores which we calculated using Equation (1). A paired t-test did not detect a significant difference in performance scores between IGUANA's 3D map interface and RC ( $t(9) = -2.622$ ,  $p_{adj} = .112$ ). The results are shown in a dumbbell plot (see Fig. 12-C), which connects each participant's scores across the two interfaces.

Paired t-tests on the SEQ scores revealed opposite results for the two tasks (see Fig. 12-A and B). Participants found the 3D map interface significantly easier for Task A compared to RC ( $t(9) = 4.636$ ,  $p_{adj} = .008$ ). In contrast, there was no significant difference in difficulty for the virtual ball interface in Task B ( $t(9) = -2.667$ ,  $p_{adj} = .156$ ).

For Task A, paired t-tests on the NASA-TLX scores showed that participants experienced significantly lower workload with the 3D map interface compared to RC (see Fig. 14-A). This was observed on Mental Demand ( $t(9) = -5.027$ ,  $p_{adj} = .007$ ), Temporal Demand ( $t(9) = -4.038$ ,  $p_{adj} = .015$ ), Effort ( $t(9) = -4.793$ ,  $p_{adj} = .008$ ), and Frustration ( $t(9) = -4.355$ ,  $p_{adj} = .012$ ). They also reported significantly higher subjective Performance with the 3D map interface ( $t(9) = 4.836$ ,  $p_{adj} = .008$ ). No significant difference was found for Physical Demand (Wilcoxon signed-rank test;  $W = 14$ ,  $p_{adj} = 1$ ).

For Task B, paired t-tests on the NASA-TLX scores revealed that the virtual ball interface resulted in significantly lower subjective Performance compared to RC ( $t(9) = -3.482$ ,  $p_{adj} = .049$ ) (see Fig. 14-B). While other workload ratings were numerically higher for the virtual ball interface, no statistically significant differences were found for Mental Demand ( $t(9) = 0.506$ ,  $p_{adj} = 1$ ), Physical Demand ( $t(9) = 1.8$ ,  $p_{adj} = .525$ ), Effort ( $t(9) = 1.542$ ,  $p_{adj} = .612$ ), Frustration ( $t(9) = 1.562$ ,  $p_{adj} = .612$ ), or Temporal Demand (Wilcoxon signed-rank test;  $W = 10$ ,  $p_{adj} = 1$ ).



**Figure 12: Boxplots of SEQ scores for (A) Task A and (B) Task B. (C) Dumbbell plot comparing performance scores, each line connects a participant's paired scores across RC and IGUANA. Scatter plots showing the relationship between completion time and accuracy (mean distance to waypoints) for (D) RC and (E) IGUANA.**

## 5.2 Qualitative Results

A descriptive quality analysis approach was used to organize and summarize the interview results. By systematically reviewing the feedback, we were able to identify key themes such as the usefulness of the 3D map interface, the usability challenges of the virtual ball interface, and the mixed responses to the spatial overlay.

Participants consistently praised the 3D map interface, describing it as "the best option," "intuitive," and "easy to use." *P4* remarked, "I like the declarative approach in the 3D map interface better than the RC's imperative style." Similarly, *P3* liked that "there is no micro-managing with the 3D map interface." Features such as the height indicator and preview for the camera waypoint were described as helpful. Overall, the most positively received aspect was the fact that the participants did not need to manually control the UAV.

As for waypoint placement, several suggestions were recorded. *P9* observed that having "a ground shadow for the waypoint marker would make it better" and, along with *P10*, suggested adding movement scaling for finer control. A separate input for height was also suggested by several participants (*P6*, *P7*, *P9*).

Although all participants understood the intent behind combining multiple axes controls into the virtual ball interface and acknowledged its intuitiveness, some found it harder to use than a standard dual-stick controller (*P8*, *P10*). The highest complaints were the lack of physical feedback from the virtual ball, resulting in them frequently looking down to make sure they were performing the correct gesture. The visual and auditory feedback alone were not enough to make participants feel confident (*P8*). Additionally, a short tutorial session was perceived as insufficient to familiarize them with the hand pose and movement, which is very new to them. *P10* noted, "Maybe it's easier to grab the ball," and with *P7* added, "It needs an axis locking option." However, seven participants believed that with enough exposure, the virtual ball interface could be as easy to use as the RC (*P2*, *P3*, *P5*, *P6*, *P7*, *P9*, *P10*).

All participants agreed that the spatial overlay helped them to quickly locate the UAV in the sky. As *P7* pointed out, "I find myself constantly looking at the map when using the RC." However, participants expressed mixed opinions regarding the amount and placement of additional information on the overlay. Two participants did not find it necessary to display anything beyond distance and height (*P1*, *P5*). Three participants suggested moving the additional information closer to the ground (*P4*, *P5*, *P9*). Additionally,

four participants complained about the constant size of the overlay, suggesting that it should scale with the distance (*P4*, *P5*, *P7*, *P10*).

## 6 Discussion

### 6.1 IGUANA Proposed Interfaces

Our user study revealed that the 3D map interface was particularly well-received. Participants consistently praised the interface for relieving them from any manual control when flying, suggesting that automatic navigation is still preferable to manual control for guiding the UAV to follow a predefined path. They rated it as more usable and useful with higher adoption intent, and also experienced significantly less workload overall. Performance scores were numerically higher, though this difference was not statistically significant after correction, suggesting that a larger sample may be needed to further investigate this effect. Using Google's Photorealistic 3D Tiles helps users navigate and understand spatial relationships in the 3D WIM reconstructed from real ones [5, 37, 39]. This resulted in a consistency in trajectory, as all participants produced similar paths using the 3D map interface. Participants found the 3D map interface intuitive and easy to use in its current form but suggested enhancements for finer control, including: (1) **marker movement scaling**, which allows for adjustable, fine-grained control of marker placement, (2) **marker ground shadow**, which shows the marker's projected position on the ground level, and (3) **separate input for height**, which decouples the position and the height of a marker, allowing precise adjustment for each.

The spatial overlay was considered a primary feature for this type of MR application, confirming previous works [4, 14, 28, 48], particularly when users needed to track moving objects. Showing the planned waypoint path in the spatial overlay helps users see clearly where the UAV is going to go [14, 44, 48]. Observing the participants' behavior when completing tasks using RC, eight participants looked down at the screen most of the time, reducing their situational awareness [4]. In contrast, with the IGUANA interface, the spatial overlay was positioned between the user and the UAV, allowing more direct visual engagement. One technical limitation encountered was the sensor precision of the UAV's GPS location and height, both of which can drift over time. To compensate, we implemented a relatively large overlay area. While this mitigated the drift issue, it also resulted in a less precise location indicator. Consistent with prior findings [4], participants' concerns about the



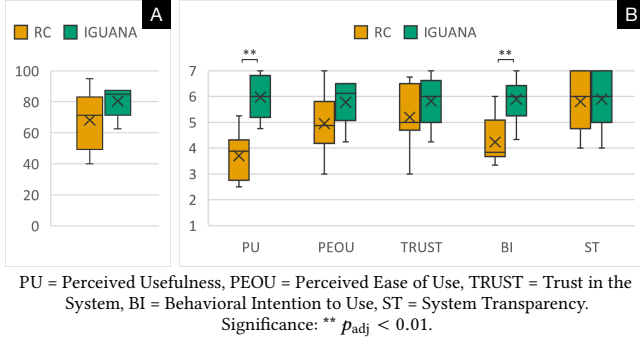


Figure 13: Boxplots of (A) SUS scores and (B) TAM scores.

constant overlay size suggest that adaptive scaling could provide a more intuitive visual cue, producing a sense of distance at a glance without needing to read the distance or height information first.

The difficulties participants experienced with the virtual ball interface indicate that physical feedback remains crucial in aiding control [20, 38, 47], particularly for first-time use. To improve the virtual ball interface, participants proposed: (1) **axis locking** control for more precise directional control [2], and (2) **grab-based interaction** allowing users to grab and hold the ball in hand rather than push it, making an impression that the UAV is in the user's hand [5, 23]. While some participants strongly agreed that, with sufficient practice, the virtual ball interface could become as effective as a traditional dual-stick controller, further study is needed to validate this.

## 6.2 Limitations

The small sample size of 10 participants may not capture the full range of user experiences and preferences. To address this, we employed a within-subjects design, which, however, may have introduced learning effects. As participants became more familiar with the task over time, this could have potentially skewed the results in favor of the second system used.

The use of a single UAV model (DJI Mavic Air) may also limit the generalizability of our findings. Different UAVs may have different control schemes, capabilities, and user interfaces, which could influence user interaction and performance.

Furthermore, this study primarily focused on qualitative assessments of user experience, which, while valuable for understanding subjective impressions, may not fully capture objective metrics such as speed or reliability.

## 6.3 Future Work

Optional haptic feedback could help first-time users feel more confident with the virtual ball interface and could be removed once trust is established.

Future experiments could explore different UAVs equipped with improved sensors and communication protocols, offering better GPS accuracy, more precise altitude estimation, and greater communication range.

Additionally, future user studies should consider adopting a between-subjects design. This may introduce variability in the

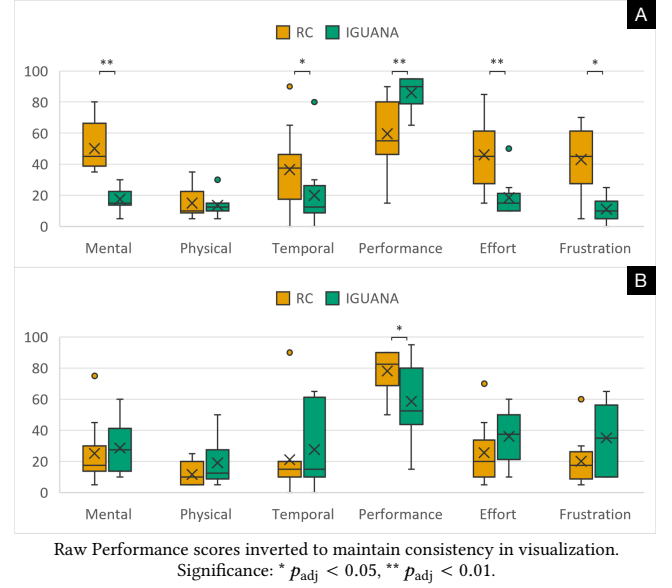


Figure 14: Boxplots of NASA-TLX scores for (A) Task A and (B) Task B.

results, as different participants may have varying levels of experience and familiarity with UAV operation. However, the effect can be mitigated by having a larger and more diverse participant pool.

## 7 Conclusion

In this paper, we present IGUANA, an MR-based interface for UAV guidance, navigation, and control. IGUANA offers (1) high-level control via a 3D map interface, (2) low-level control via a virtual ball interface, and (3) enhanced situational awareness through a spatial overlay.

We conducted a user study to evaluate the usability of IGUANA and found that (1) the 3D map interface was intuitive and easy to use, suggesting better accuracy and greater consistency with lower perceived workload compared to conventional dual-stick controller, (2) the virtual ball interface was intuitive but required practice due to the lack of physical feedback, and (3) the spatial overlay was useful for instantly spotting the UAV and enhancing situational awareness. With further refinement, MR-based interfaces could become a viable alternative to conventional UAV controllers.

## Acknowledgments

This work was funded by

- the European Social Fund (ESF Plus) and the German Federal State of Saxony within the project ProSECO (100687967),
- the European Union, co-financed by tax revenues based on the budget adopted by the Saxon State Parliament, as part of an ESF Plus scholarship (100670474) awarded to T. Krisanty,
- the German Federal Ministry of Education and Research (BMBF, SCADS22B) and the Saxon State Ministry for Science, Culture and Tourism (SMWK) by funding the competence center for Big Data and AI “ScaDS.AI Dresden/Leipzig”, and
- DFG grant 389792660 as part of TRR 248 - CPEC.

## References

- [1] Isaac Osei Agyemang, Isaac Adjei Mensah, Sophyani Banaamwini Yussif, Fiasam Linda Delali, Bernard Cobinnah Mawuli, Bless Lord Y. Agbley, Collins Sey, and Joshua Berkohd. 2023. Gesture Control of Micro-drone: A Lightweight-Net with Domain Randomization and Trajectory Generators. (2023). doi:10.48550/ARXIV.2301.12470
- [2] Mike Allenspach, Till Kötter, Rik Bähnmann, Marco Tognon, and Roland Siegwart. 2023. Design and Evaluation of a Mixed Reality-based Human-Robot Interface for Teleoperation of Omnidirectional Aerial Vehicles. In *2023 International Conference on Unmanned Aircraft Systems (ICUAS)*. 1168–1174. doi:10.1109/ICUAS57906.2023.10156426
- [3] Chanun Asavasirikulkij and Muhammad Hanif. 2023. Human Workload Evaluation of Drone Swarm Formation Control Using Virtual Reality Interface. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23)*. Association for Computing Machinery, New York, NY, USA, 132–136. doi:10.1145/3568294.3580057
- [4] Alexander Atanasyan, Rodrigo Gabriel Cortés Estévez, and Jürgen Rossmann. 2023. Augmented Reality-Based Drone Pilot Assistance Using Consumer-Grade Hardware. In *2023 IEEE 13th International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*. 123–128. doi:10.1109/ICCE-Berlin58801.2023.10375669
- [5] Julio Betancourt, Baptiste Wojtkowski, Pedro Castillo, and Indira Thouvenin. 2023. Exocentric Control Scheme for Robot Applications: An Immersive Virtual Reality Approach. *IEEE Transactions on Visualization and Computer Graphics* 29, 7 (July 2023), 3392–3404. doi:10.1109/TVCG.2022.3160389
- [6] John Brooke. 1996. SUS: A 'Quick and Dirty' Usability Scale. In *Usability Evaluation In Industry*. London: Taylor & Francis, 189–194. doi:10.1201/9781498710411-35
- [7] Cesium. 2023. Google Maps Platform Photorealistic 3D Tiles. Retrieved 2025-01-23 from <https://cesium.com/learn/photorealistic-3d-tiles-learn/>
- [8] Haodong Chen, Fen Liu, Yuanlin Yang, and Wei Meng. 2022. MultiVR: Digital Twin and Virtual Reality Based System for Multi-people Remote Control Unmanned Aerial Vehicles. In *2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. 647–652. doi:10.1109/ICARCV57592.2022.10004244
- [9] Jinsai Cheng, Saifuddin Mahmud, Misbahuddin Mohammed, Aditi Singh, and Jong-Hoon Kim. 2024. Design of a Novice-Friendly Drone Control System. In *2024 IEEE 14th Annual Computing and Communication Workshop and Conference (CCWC)*. 0184–0190. doi:10.1109/CCWC60891.2024.10427804
- [10] Fred D. Davis. 2007. On the Relationship Between HCI and Technology Acceptance Research. In *Human-Computer Interaction and Management Information Systems: Foundations* (1st edition ed.). Routledge, 409–415.
- [11] P. Dayani, N. Orr, A. Thomopoulos, V. Saran, S. Krishnaswamy, E. Zhang, N. Hu, D. McPherson, J. Menke, A. Yang, and K. Vetter. 2021. Immersive Operation of a Semi-Autonomous Aerial Platform for Detecting and Mapping Radiation. *IEEE Transactions on Nuclear Science* 68, 12 (Dec. 2021), 2702–2710. arXiv:2103.10057 [cs] doi:10.1109/TNS.2021.3122452
- [12] DJI. 2018. DJI Assistant 2 - Download Center. Retrieved 2025-09-02 from <https://www.dji.com/downloads/softwares/assistant-dji-2>
- [13] DJI. 2018. DJI Go 4. Retrieved 2025-01-23 from <https://www.dji.com/downloads/djiapp/dji-go-4>
- [14] Yousef Emami, Kai Li, Luis Almeida, Sai Zou, and Wei Ni. 2025. On the Use of Immersive Digital Technologies for Designing and Operating UAVs. arXiv:2407.16288 [cs] doi:10.48550/arXiv.2407.16288
- [15] Okan Erat, Werner Alexander Isop, Denis Kalkofen, and Dieter Schmalstieg. 2018. Drone-Augmented Human Vision: Exocentric Control for Drones Exploring Hidden Areas. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1437–1446. doi:10.1109/TVCG.2018.2794058
- [16] Google. 2018. Google Earth. Retrieved 2025-01-23 from <https://earth.google.com/web>
- [17] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, Peter A. Hancock and Najmedin Meshkati (Eds.). Human Mental Workload, Vol. 52. North-Holland, 139–183. doi:10.1016/S0166-4115(08)62386-9
- [18] Robin Helmert, Tobias Hardes, and Enes Yigitbas. 2023. Design and Evaluation of an AR Voice-based Indoor UAV Assistant for Smart Home Scenarios. In *Proceedings of the 2023 ACM Symposium on Spatial User Interaction (SUI '23)*. Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3607822.3614522
- [19] Baichuan Huang, Deniz Bayazit, Daniel Ullman, Nakul Gopalan, and Stefanie Tellex. 2019. Flight, Camera, Action! Using Natural Language and Mixed Reality to Control a Drone. In *2019 International Conference on Robotics and Automation (ICRA)*. 6949–6956. doi:10.1109/ICRA.2019.8794200
- [20] Roman Ibrahimov, Evgeny Tsykunov, Vladimir Shirokun, Andrey Somov, and Dmitry Tsetserukou. 2019. DronePick: Object Picking and Delivery Teleoperation with the Drone Controlled by a Wearable Tactile Display. *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (Oct. 2019), 1–6. doi:10.1109/RO-MAN46459.2019.8956344
- [21] JASP Team. 2025. JASP (Version 0.95.1)[Computer Software]. <https://jasp-stats.org/>
- [22] Basaran Bahadır Kocer, Harvey Stedman, Patryk Kulik, Izaak Caves, Nejra Van Zalk, Vijay M. Pawar, and Mirko Kovac. 2022. Immersive View and Interface Design for Teleoperated Aerial Manipulation. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 4919–4926. doi:10.1109/IROS47612.2022.9981656
- [23] Konstantinos Konstantoudakis, Kyriaki Christaki, Dimitrios Tsiakmakis, Dimitrios Sainidis, Georgios Albanis, and Petros Daras. 2022. Drone Control in AR: An Intuitive System for Single-Handed Gesture Control, Drone Tracking, and Contextualized Camera Feed Visualization in Augmented Reality. *Drones* 6, 2 (Feb. 2022), 43. doi:10.3390/drones6020043
- [24] Maximilian Krogus, Acshi Haggenmiller, and Edwin Olson. 2019. Flexible Layouts for Fiducial Tags. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, Macau, China, 1898–1903. doi:10.1109/IROS40897.2019.8967787
- [25] Bryson Lawton and Frank Maurer. 2022. A Case for Enhancing UAV Ground Control Stations with Cross Reality. <https://api.semanticscholar.org/CorpusID:251395739>
- [26] Bryson Lawton and Frank Maurer. 2022. Immersive Maps for Drone Control: A Case for Improving Multi-UAV Ground Control Station Maps with Extended Reality. <https://api.semanticscholar.org/CorpusID:251395764>
- [27] Nguyen Thanh Trung Le, Howe Yuan Zhu, and Hsiang-Ting Chen. 2021. Remote Visual Line-of-Sight: A Remote Platform for the Visualisation and Control of an Indoor Drone Using Virtual Reality. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*. ACM, Osaka Japan, 1–3. doi:10.1145/3489849.3489910
- [28] Chuhao Liu and Shaojie Shen. 2020. An Augmented Reality Interaction Interface for Autonomous Drone. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 11419–11424. doi:10.1109/IROS45743.2020.9341037
- [29] Chun-Yu Lo and June-Hao Hou. 2022. A Natural Human-Drone Interface For Better Spatial Presence Experiences. In *CAADRIA 2022: Post-Carbon*. Sydney, Australia, 99–108. doi:10.52842/conf.caadria.2022.1.099
- [30] Mapbox. 2010. Mapbox | Maps, Navigation, Search, and Data. Retrieved 2025-09-08 from <https://www.mapbox.com>
- [31] Luca Morando and Giuseppe Loianno. 2024. Spatial Assisted Human-Drone Collaborative Navigation and Interaction through Immersive Mixed Reality. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*. 8707–8713. doi:10.1109/ICRA57147.2024.10611351
- [32] Nelson Andres Sanchez O., Santiago Felipe Munera D., and Fernando De La Rosa R. 2023. Virtual Reality and Human-Drone Interaction Applied to the Construction and Execution of Flight Paths. In *2023 International Conference on Unmanned Aircraft Systems (ICUAS)*. 144–151. doi:10.1109/ICUAS57906.2023.10155889
- [33] Jesse Rawlins Paterson, Jiwoong Han, Tom Cheng, Paxtan Huish Laker, David Livingston McPherson, Joseph Menke, and Allen Y Yang. 2019. Improving Usability, Efficiency, and Safety of UAV Path Planning through a Virtual Reality Interface. In *Symposium on Spatial User Interaction (SUI '19)*. Association for Computing Machinery, New York, NY, USA, 1–2. doi:10.1145/3357251.3362742
- [34] Sanket A. Salunkhe, Pranav Nedunghat, Luca Morando, Nishanth Bobbili, Guanrui Li, and Giuseppe Loianno. 2025. Intuitive Human-Drone Collaborative Navigation in Unknown Environments through Mixed Reality. <https://api.semanticscholar.org/CorpusID:277502310>
- [35] Jeff Sauro and Joseph S. Dumas. 2009. Comparison of Three One-Question, Post-Task Usability Questionnaires. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Boston MA USA, 1599–1608. doi:10.1145/1518701.1518946
- [36] Oleg Sautenkov, Selamawit Asfaw, Yasheerah Yaqoot, Muhammad Ahsan Mustafa, Aleksey Fedoseev, Daria Trinitatova, and Dmity Tsetserukou. 2024. FlightAR: AR Flight Assistance Interface with Multiple Video Streams and Object Detection Aimed at Immersive Drone Control. arXiv:2410.16943 [cs] doi:10.48550/arXiv.2410.16943
- [37] Kamil Sedlmajer, Daniel Bambušek, and Vítězslav Beran. 2019. Effective Remote Drone Control Using Augmented Virtuality. In *3rd International Conference on Computer-Human Interaction Research and Applications*. 177–182. doi:10.5220/0008349401770182
- [38] Nassim Sehad, Bilel Cherif, Ibrahim Khadraoui, Wassim Hamidouche, Faouzi Bader, Riku Jäntti, and Mérouane Debbah. 2023. Locomotion-Based UAV Control Toward the Internet of Senses. *IEEE Transactions on Circuits and Systems II: Express Briefs* 70, 5 (May 2023), 1804–1808. doi:10.1109/TCSII.2023.3257363
- [39] Luljeta Sinani, Anna Bahnmüller, Frank Steinicke, Andreas Gerndt, and Georgia Albuquerque. 2025. Drone Teleoperation Interfaces: Challenges and Opportunities with XR Integration. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3706599.3719777
- [40] Richard Stoakley, Matthew J. Conway, and Randy Pausch. 1995. Virtual Reality on a WIM: Interactive Worlds in Miniature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '95)*. ACM Press/Addison-Wesley Publishing Co., USA, 265–272. doi:10.1145/223904.223938

- [41] Tarik Taleb, Nassim Sehad, Zinelaabidine Nadir, and Jaeseung Song. 2023. VR-Based Immersive Service Management in B5G Mobile Systems: A UAV Command and Control Use Case. *IEEE Internet of Things Journal* 10, 6 (March 2023), 5349–5363. doi:10.1109/JIOT.2022.3222282
- [42] Victor Victor and Tania Krisanty. 2025. IGUANA. <https://github.com/PiBu-Labs/IGUANA>
- [43] Tim Wächter and Matthias König. 2024. Indoor Drone Path Planning with Real/Virtual Obstacles by Mixed Reality. In *2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 600–603. doi:10.1109/ISMAR-Adjunct64951.2024.00172
- [44] Michael Walker, Hooman Hedayati, Jennifer Lee, and Daniel Szafir. 2018. Communicating Robot Motion Intent with Augmented Reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, Chicago IL USA, 316–324. doi:10.1145/3171221.3171253
- [45] Daniar Estu Widiyanti, Krisma Asmoro, and Soo Young Shin. 2024. HoloGCS: Mixed Reality-Based Ground Control Station for Unmanned Aerial Vehicle. *Virtual Reality* 28, 1 (Feb. 2024), 40. doi:10.1007/s10055-023-00914-9
- [46] Baptiste Wojtkowski, P. Castillo, and I. Thouvenin. 2020. A New Exocentric Metaphor for Complex Path Following to Control a UAV Using Mixed Reality. *ArXiv* (Feb. 2020). <https://api.semanticscholar.org/CorpusID:211126655>
- [47] Grigoriy A. Yashin, Daria Trinitatova, Ruslan T. Agishev, Roman Ibrahimov, and Dzmitry Tsetserukou. 2019. AeroVr: Virtual Reality-based Teleoperation with Tactile Feedback for Aerial Manipulation. In *2019 19th International Conference on Advanced Robotics (ICAR)*. 767–772. doi:10.1109/ICAR46387.2019.8981574
- [48] Stefanie Zollmann, Christof Hoppe, Tobias Langlotz, and Gerhard Reitmayr. 2014. FlyAR: Augmented Reality Supported Micro Aerial Vehicle Navigation. *IEEE Transactions on Visualization and Computer Graphics* 20, 4 (April 2014), 560–568. doi:10.1109/TVCG.2014.24