PICOVO: Surfaces as Contested Spatial Narratives

Xiaoyue Shan
School of Design & Innovation
China Academy of Art
Hangzhou, China
3231305148@caa.edu.cn

Naizhen Luo
School of Design & Innovation
China Academy of Art
Hangzhou, China
806484912@qq.com

Zhutian Fan
School of Design & Innovation
China Academy of Art
Hangzhou, China
3231305399@caa.edu.cn

Zijin Jiang
School of Design & Innovation
China Academy of Art
Hangzhou, China
2772747958@qq.com

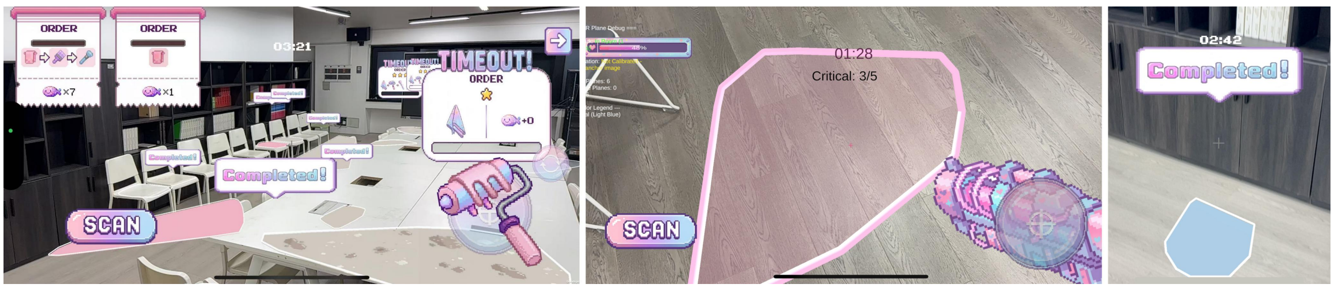

Figure 1: In-game visuals from PICOVO. From left to right, the images show detected physical surfaces transformed into interactive territories, players contesting these surfaces through tool-based actions, and completion feedback embedded directly in the physical environment.

ABSTRACT

PICOVO reframes AR plane detection as narrative substrate: detected surfaces become contested territories that players physically defend. This submission reveals the design process behind this reframing—how iterative prototyping across distinct spatial topologies led us to discover that environmental geometry itself generates emergent story structures. We detail the spatial mechanics we developed, the failures that shaped them, and a demonstration designed to let attendees experience and manipulate the topology–narrative relationship firsthand.

CCS CONCEPTS

• **Human-centered computing** → Mixed / augmented reality; Interaction design theory, concepts and paradigms; • **Applied computing** → Arts and humanities; • **Software and its engineering** → Multiplayer online games.

KEYWORDS

Embodied Interaction, Spatial Computing, AR Game Design, Emergent Spatial Narrative, Mobile AR

SIGGRAPH '26, July 28–August 1, 2026, Denver, CO, USA

ACM ISBN 978-1-4503-XXXX-X/26/07
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 MOTIVATION: SURFACES AS STORY

When AR systems detect a table, they treat it as invisible scaffolding for virtual objects. We asked: what if detected surfaces are the narrative subjects—not the supports?

This question emerged from a frustration. Location-based AR games (Pokémon GO, Ingress) anchor play to GPS coordinates; room-scale platforms use plane detection solely for object placement [4, 8]. Recent work on embodied AR [3, 9, 14] and environmental storytelling [7, 10] establishes that spatial configuration generates narrative meaning. Shin and Woo [12] confirm that spatial trajectories carry narrative intent in AR heritage contexts, and Dong et al. [13] demonstrate that scene semantics can guide automatic placement of authored story content in AR environments. Yet these systems author narratives *into* space rather than letting space author narrative.

Trajectories are designed, and content is pre-written. We wanted to explore whether physical topology alone could produce distinct story structures without any scripted content.

2 THE DESIGN JOURNEY: WHAT WORKED, WHAT DIDN'T

2.1 First Attempt: Free Tool Access

Our initial prototype let players freely switch tools anywhere via a UI menu. Set in a speculative 2036 where surfaces are programmable, 2–4 players compete as coating workers across randomly detected planes.

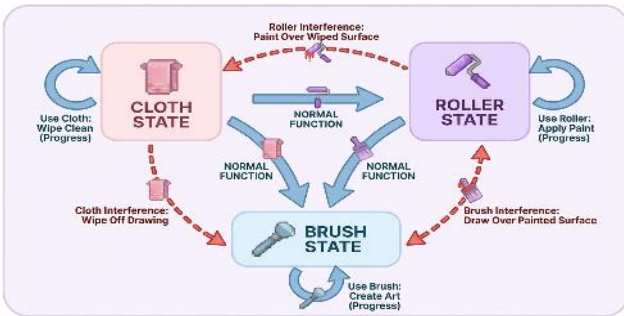

Figure 2: Tool State Machine. Sequential tools enable asymmetric interference under limited usage.

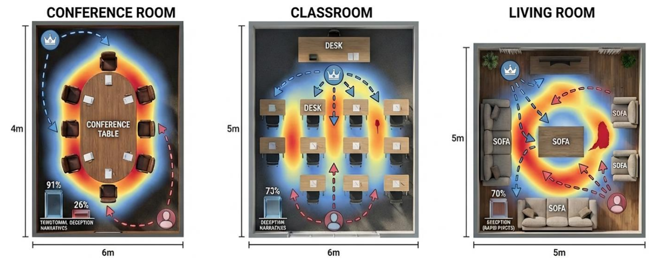

Figure 3: Movement Heatmaps Across Room Topologies Conference rooms (5.2m) yield territorial play (91%), classrooms (3.4m) enable deception (73%), living rooms (1.8m) produce rapid pivots (70%).

ARKit/ARCore finds horizontal and vertical surfaces as "dirty" work orders (small/medium/large tier); environmental topology determines the play space—furniture-dense rooms yield fragmented territories while open spaces create contested zones. Orders require sequential tool application (Cloth→Roller→Brush), with asymmetric interference (Figure 2): Cloth strips progress, Roller overwrites with player color, Brush only completes. Early playtests were mechanically functional but narratively dead—players optimized silently, rarely spoke, and never referenced space. We had built a spatial game that produced no spatial stories.

2.2 The Workbench Breakthrough

We introduced a shared virtual workbench for tool switching, a single spatial constraint that transformed player behavior. Players now had to physically leave claimed surfaces to change tools, and each departure created a moment of vulnerability that opponents could exploit. The distance between the workbench and a surface became a source of dramatic tension. A player coating a distant table had to choose between completing the task and risking the loss of progress during a long return trip.

In the first session using the workbench, a player shouted, “Don’t touch my table,” marking the first instance of spontaneous possessive language observed after weeks of testing. Physical displacement had become a narrative engine. Players began narrating their own actions by framing departures as sacrifices, returns as rescues, and opponents’ approaches as invasions.

2.3 Parameter Tuning Through Failure

Early workbench placement was arbitrary. When placed too close (<1.5 m) to surfaces, tool switches felt costless and vulnerability disappeared. When placed too far (>6 m), players abandoned distant surfaces entirely—strategic complexity collapsed.

We converged on 2–5 m as the productive range where departure created real risk but return remained viable. Similarly, the 1.5 m interaction radius emerged from testing: smaller radii forced awkward precision; larger radii eliminated the legible body-language of threat.

2.4 Three Emergent Spatial Patterns

Through this iterative process, we formalized three design patterns, drawing on Dourish [3], Benford [1], and Mueller et al.'s framework for human–computer integration [14], where the body itself becomes the primary interface between player intent and spatial computation:

(1) Territorial Inscription. Progress visualization anchored to real geometry makes labor and vulnerability legible. Persistent, high-contrast surface states turn detected planes into owned territory.

(2) Vulnerability Rhythm. Mandatory workbench trips create polyrhythmic movement. Vulnerability duration = traversal distance / walking speed (~1.2 m/s). This coupling of distance and time is the core mechanism converting room layout into dramatic structure.

(3) Legible Threat. The 1.5 m interaction radius makes opponent proximity an unambiguous spatial signal. Players read intention through body movement rather than screen indicators.

These patterns require physical topology: room layout determines workbench–surface distances (vulnerability duration), surface distribution (territorial fragmentation), and movement visibility (threat legibility).

3 IMPLEMENTATION INSIGHTS

Built in Unity 2022.3 with AR Foundation 5.1, targeting iPhone 14 Pro and Galaxy S22 (>56 fps). Co-localization uses a printed AR image marker (<3 cm positional, <2° rotational error).

The most challenging technical issue was cross-device plane consistency. Each device detects planes independently, producing different boundaries and detection timing. LiDAR-enabled devices are approximately 1.5 seconds faster. After several failed attempts at plane matching algorithms, we developed a Shadow Plane Protocol. The host device broadcasts authoritative plane data, including centroid, extent, alignment, and boundary vertices, via Photon PUN2. Remote devices then instantiate simplified colliders based on this broadcast, eliminating plane inconsistency. A two-second stability filter prevents flickering caused by unstable detections.

4 WHAT THE SPACES TOLD US

We ran 36 sessions (N=72, ages 18–58) across three topology types to understand what role space played.

Does physical space matter at all? Compared to a stationary baseline with identical mechanics but no physical movement, PICOVO produced dramatically different engagement. Physical movement increased from 31 m to 147 m per session ($F(1,69)=168.4$, $p<.001$). Narrative engagement (adapted from Busselle & Bilandzic [2], $\alpha=.79$) rose from 2.7 to 5.4. Spontaneous narrative language—utterances about ownership, threat, and reversal—increased sixfold (5.1/min vs. 0.8). Unprompted replay requests: 31/72 PICOVO vs. 1/72 baseline.

Do different topologies produce different stories? Yes, systematically (Figure 3). Post-session interviews (~340 pages, Cohen's $\kappa=.86$) revealed topology-dependent narrative patterns:

- ***Conference rooms*** (6 large planes, mean distance 5.2 m): 91% territorial narratives, only 26% deception. Long distances amplified ownership but prevented effective feinting.
- ***Classrooms*** (12 mixed planes, mean distance 3.4 m): 73% deception narratives. Moderate distances enabled deliberate misdirection—"I faked going to the wall."
- ***Living rooms*** (mixed plane sizes, mean distance 1.8 m): 70% deception via rapid pivots; multi-surface cascading reversals.

Mean distance correlated with territorial framing ($r=.59$) and inversely with deception ($r=-.41$). Environmental confounds (room formality, object semantics) exist; preliminary furniture-rearrangement sessions ($n=8$) suggest spatial metrics dominate, but controlled manipulation studies are needed.

5 DEMONSTRATION: EXPERIENCING THE PROCESS

The demonstration is the research process itself, as attendees discover topology and narrative relationships through embodied play.

Stage 1: Uninformed Play (6 min). Four attendees play PICOVO in a pre-configured space with mixed distances (~2 m / ~5 m), given only task-completion and sabotage instructions. A shared display shows real-time movement trails, tool-switch frequency, and live transcription of spatial language—making the process legible as it unfolds.

Stage 2: Process Reveal (8 min). We replay key moments from Stage 1, highlighting correlations between distance and narrative behavior: long-distance stations producing territorial language and defensive sprints, close-range stations producing rapid deception and feints. Attendees then manipulate topology via a tablet interface—dragging the workbench position, toggling surfaces on/off, and adjusting interaction radii. A lightweight predictive overlay (heuristic model trained on our playtest data) updates expected narrative tendencies in real time, showing how each spatial change shifts the balance between territorial and deception narratives. One modified configuration is briefly tested, letting attendees verify the prediction against their own experience.

Stage 3: Attendee-Led Design (6 min). Attendees physically rearrange furniture and place a new AR anchor. The system scans the resulting topology, visualizes spatial metrics, predicts dominant narrative tendencies, and exposes internal processes—plane synchronization events, FSM state transitions, vulnerability timers—during live play. The three design patterns become directly observable as spatial parameters change.

Setup: 4 phones, 1 tablet, 2 displays, ~5×5 m area with movable furniture.

Core insight: By playing, modifying, and replaying space, attendees experience firsthand how spatial configuration alone generates distinct narrative behaviors—without scripted story.

6 DISCUSSION: TOPOLOGY SHAPES NARRATIVE

PICOVO's contribution is a discoverable principle that physical topology determines emergent narrative structure, a principle we reached not through theoretical reasoning but through iterative making and breaking. The three parameterized patterns, territorial inscription, vulnerability rhythm, and legible threat, emerged from observing players across dozens of failed and successful spatial configurations. These patterns are transferable beyond games, including museum AR in which exhibits become contested territories, collaborative design review where CAD surfaces carry ownership stakes, and physical therapy AR where body part targets generate movement based narratives.

Our process taught us that spatial narrative design is fundamentally about constraint placement: the workbench's position matters more than any content we could script. The most scripted version of PICOVO was narratively inert; the most spatially constrained version produced stories we never anticipated. Future work includes controlled topology manipulation within constant environments, longitudinal studies with pre-existing social groups, and accessibility-centered adaptations (gesture-based claiming, adjustable interaction radii) for users with mobility limitations.

REFERENCES

[1] Benford et al. 2009. The Frame of the Game. CHI '09. ACM.
[2] Busselle & Bilandzic. 2009. Measuring Narrative Engagement. Media Psychology 12(4), 321–347.
[3] Dourish. 2001. Where the Action Is. MIT Press.
[4] Dow et al. 2007. Presence and Engagement in an Interactive Drama. CHI '07. ACM.
[5] Engberg & Bolter. 2020. The Aesthetic Experience of Virtual and Augmented Reality. DHQ 14(1).
[6] Isbister. 2016. How Games Move Us. MIT Press.
[7] Jenkins. 2004. Game Design as Narrative Architecture. First Person, MIT Press.
[8] Lindley et al. 2008. Stirring Up Experience Through Movement in Game Play. CHI '08. ACM.
[9] Mueller et al. 2020. Towards Designing Bodily Integrated Play. TEI '20. ACM.
[10] Salen & Zimmerman. 2004. Rules of Play. MIT Press.
[11] Braun & Clarke. 2006. Using Thematic Analysis in Psychology. Qualitative Research in Psychology 3(2), 77–101.
[12] Shin & Woo. 2023. How Space is Told: Linking Trajectory, Narrative, and Intent in AR Storytelling. CHI '23. ACM.Dong et al. 2022. Interactive AR Storytelling Guided by Scene Semantics. ACM Trans. Graph. 41(4).
[13] Dong et al. 2022. Interactive AR Storytelling Guided by Scene Semantics. ACM Trans. Graph. 41(4).
[14] Mueller et al. 2022. Human – Computer Integration: Towards Integrating the Human Body with the Computational Machine. Found. Trends HCI 16(1), 1–64.