Mixed Reality for Cultural Heritage: A Pipeline for Spatially Aligned 3D Content in In-Situ Experiences

Federico Martusciello
FrAmeLab, SWEN Research Group
University of L'Aquila
L'Aquila, Italy
federico.martusciello@graduate.univaq.it

Henry Muccini
FrAmeLab, SWEN Research Group
University of L'Aquila
L'Aquila, Italy
henry.muccini@univaq.it

***Abstract*—Mixed Reality can enhance cultural heritage experiences by embedding digital reconstructions and interactive content into the physical environment. Achieving convincing in-situ experiences requires stable spatial alignment between real and virtual elements. This extended abstract presents a pipeline for spatially aligned 3D content placement based on headset-acquired spatial data. The acquired environment is exported as a JSON representation and reconstructed in Unity as a labelled three-dimensional proxy of the real space. This reconstructed environment serves as the reference for anchoring 3D models, spatial interfaces, and interpretive content to physical elements. The proposed approach supports reliable content placement in cultural heritage and exhibition scenarios across single-room and multi-room experiences.**

***Index Terms*—Mixed Reality, cultural heritage, spatial alignment, 3D content placement, in-situ experience.**

I. Introduction

Mixed Reality (MR) is increasingly adopted in cultural heritage (CH) to enrich physical sites with digital reconstructions, annotations, and interactive content. Unlike Virtual Reality, MR preserves the user's perception of the real environment while augmenting it with virtual elements, making it suitable for in-situ experiences where the meaning of a digital reconstruction depends on its relationship with the surrounding space [1], [2]. In heritage contexts, this relationship is critical: ruins, fragmented remains, empty rooms, or museum installations often require visitors to mentally reconstruct missing architectural or historical information. MR can support this process by placing digital objects directly where they are meant to be interpreted. However, the experience is convincing only if virtual content remains spatially coherent with walls, floors, surfaces, and other physical references. Registration error, tracking instability, latency, calibration issues, and manual placement inaccuracies are known limitations in AR/MR systems [3], [4]. This work addresses the problem of spatially aligned 3D content placement by using headset-acquired spatial data to reconstruct the real environment inside Unity. Digital content is then positioned with respect to this reconstructed spatial substrate, rather than authored as part of a fixed virtual scene.

II. Background and Gap

XR technologies have shown significant potential for CH communication, supporting situated interpretation, engagement, and access to reconstructed historical content [1], [2]. Yet accurate 3D reconstruction is not sufficient by itself: virtual content must also be correctly registered with the physical environment. If the alignment is unstable or imprecise, the user perceives the digital layer as detached from the real space, reducing its interpretive value. Registration has been a central issue in AR since early definitions of the field, where virtual 3D objects are expected to be integrated into the real environment in real time [3]. Registration errors may arise from tracking, calibration, latency, modeling inaccuracies, or coordinate-system mismatches [4]. In MR head-mounted displays, additional work is often required to align the virtual scene with the tracking space of the device [5]. Moreover, studies on registration accuracy show that jitter and latency can influence user performance in AR tasks [6]. For CH, this technical problem becomes an interpretive problem. A reconstructed artefact, architectural element, or interface must appear in the correct place relative to the physical site and remain stable during the experience. Although the literature includes several MR applications for heritage, it is often difficult to understand how these systems are technically constructed, how digital objects are anchored to the real environment, and whether their alignment remains stable across sessions. Broader reviews on MR/AR evaluation also highlight the diversity of evaluation methods and scenarios, making comparison and reproducibility difficult [7]. The proposed pipeline contributes to this gap by making the spatial alignment process explicit.

III. Proposed Pipeline

The pipeline follows four stages: spatial acquisition, JSON export, Unity reconstruction, and spatially aligned content placement. First, the MR headset performs a spatial acquisition of the physical environment. Depending on the device and software stack, this process can provide information about room boundaries, walls, floors, surfaces, openings, and other spatial references. Each detected element can be associated with a label or identifier describing its role within the physical space. Second, the acquired spatial data are exported as a structured JSON file. In this workflow, the JSON is not only a configuration file for content loading. It is a machine-readable representation of the real environment, containing both the

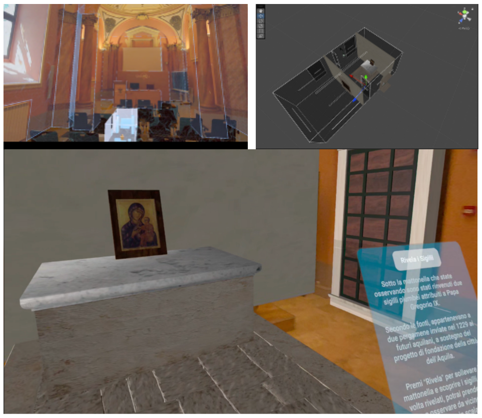

Fig. 1. Spatial alignment workflow: headset-acquired geometry is exported as JSON, reconstructed in Unity, and used as a reference for positioning 3D content and spatial interfaces.

geometric information required to reproduce the room and the labels or IDs associated with its spatial elements. These identifiers allow the MR application to refer to specific parts of the reconstructed environment. Third, the JSON file is imported into Unity. The spatial data are parsed and used to reconstruct a three-dimensional proxy of the real environment. This proxy preserves the spatial relationships between the main elements of the physical room and provides a virtual counterpart of the space acquired by the headset. During reconstruction, the labelled spatial elements are instantiated in the Unity hierarchy and remain accessible through their IDs.

Fourth, digital content is mapped onto the reconstructed environment through an alignment mechanism. A dedicated object named *Aligner* is assigned the ID of a specific spatial element reconstructed from the JSON. The *Aligner* retrieves the corresponding element in the reconstructed room and uses its transform as the reference frame for positioning digital content. As a result, 3D assets, spatial UI elements, annotations, and reconstruction layers can be placed relative to a selected physical element, such as a wall, surface, floor area, or architectural reference.

IV. Spatial Alignment Mechanism

The reconstructed room acts as the main spatial reference for content placement. Digital objects are not positioned in an abstract Unity scene, but in relation to labelled elements derived from the physical environment. Once the JSON has been imported and the room has been reconstructed, each relevant spatial element can be addressed through its ID and used as an anchor for the digital layer. The *Aligner* object receives the ID of the spatial element selected as the alignment reference, searches for the corresponding object in the reconstructed hierarchy, and adopts its transform as the local reference frame for the digital content. Assets placed under, or controlled by, the *Aligner* inherit this spatial relationship. In this way, a reconstructed artefact, an interpretive panel, a spatial UI, or an architectural overlay can be positioned coherently with respect to a specific physical element. This design makes the system adaptable and reusable. Changing the digital content or its placement does not necessarily require rebuilding the MR client or manually redesigning the Unity scene. Instead, the alignment can be modified by changing the selected spatial ID, transform offsets, or associated assets. The same principle also supports multi-room scenarios: each room can be acquired, exported, reconstructed, labelled, and enriched with its own aligned digital content.

V. Preliminary Evaluation

The proposed pipeline was preliminarily evaluated with 40 participants in the same indoor CH environment under consistent environmental and lighting conditions. Participants experienced the same MR application using the proposed spatial alignment pipeline. *Spatial alignment* remained stable throughout all evaluated sessions, yielding a 100% *session-level alignment success rate*, defined as the completion of a session without visually detectable content misalignment or manual realignment. No *tracking loss* was observed, and digital content remained anchored to the intended physical location throughout the experience. Future work will include a quantitative assessment of spatial registration error together with a user study on perceived spatial coherence and overall experience.

VI. Conclusion

This extended abstract presented a pipeline for spatially aligned 3D content placement in in-situ MR cultural heritage experiences. The system uses headset-acquired spatial data exported as JSON to reconstruct a three-dimensional proxy of the real environment in Unity. This reconstructed space becomes the reference substrate for placing 3D models, UI elements, and interpretive layers coherently with the physical surroundings. Preliminary evaluation showed stable spatial alignment throughout all evaluated sessions without requiring manual realignment. By separating spatial acquisition, room reconstruction, and content placement, the proposed approach supports adaptable MR experiences across single-room and multi-room cultural heritage scenarios.

References

[1] M. K. Bekele, R. Pierdicca, E. Frontoni, E. S. Malinverni, and J. E. Gain, "A survey of augmented, virtual, and mixed reality for cultural heritage," *Journal on Computing and Cultural Heritage*, vol. 11, no. 2, pp. 1–36, 2018.

[2] V. Barrile, E. Bernardo, A. Fotia, and G. Bilotta, "A combined study of cultural heritage in archaeological museums: 3D survey and mixed reality," *Heritage*, vol. 5, no. 3, pp. 1330–1349, 2022.

[3] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355–385, 1997.

[4] R. L. Holloway, "Registration error analysis for augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 413–432, 1997.

[5] E. Azimi, L. Qian, N. Navab, and P. Kazanzides, "Alignment of the virtual scene to the tracking space of a mixed reality head-mounted display," arXiv:1703.05834, 2017.

[6] E. D. Ragan, C. Wilkes, D. A. Bowman, and T. Höllerer, "Benefits of registration accuracy in augmented reality," in *Proc. IEEE Virtual Reality Conference*, 2009.

[7] L. Merino, M. Schwarzl, M. Kraus, M. Sedlmair, and D. Weiskopf, "Evaluating mixed and augmented reality: A systematic literature review," in *Proc. IEEE International Symposium on Mixed and Augmented Reality*, 2020.