drones

MDPI

Article

A UAV–3DGS–VR Workflow for Scenario-Comparable Immersive Review in Heritage Landscapes

Xintong Li [1], Wenqi Sheng [2], Yixuan Tang [1], Yingwen Yu [3] and Yuyang Peng [3,*]

[1] Huaqing College, Xi'an University of Architecture and Technology, Xi'an 710043, China; lixintonglw@outlook.com (X.L.); tangyixuanlw@outlook.com (Y.T.)
[2] Department of landscape architecture, Rhode Island School of Design, Providence, RI 02903, USA; wsheng@risd.edu
[3] Department of Architecture and the Built Environment, Delft University of Technology, 2628 BL Delft, The Netherlands; christinayu@tudelft.nl
* Correspondence: y.peng-1@tudelft.nl

Highlights

What are the main findings?

- The UAV–3DGS–VR workflow created reliable and comparable immersive scenes.
- The approach combined visual evaluation with behavioral evidence for scenario assessment.

What are the implications of the main findings?

- UAV-based VR can be used as a practical tool for comparative design review.
- Behavioral data can complement subjective judgment in the comparative interpretation of scenario compatibility in heritage landscapes.

Abstract

Unmanned aerial vehicles (UAVs) are widely used for documentation, surveying, and 3D modeling in the built environment, yet their outputs often remain difficult to reuse for immersive comparison of alternative construction scenarios. This study presents a low-cost UAV-to-3DGS-to-VR workflow for constructing scenario-comparable immersive environments for built-environment review. The workflow combines multi-angle UAV imagery, point-cloud-based geometric anchoring, 3D Gaussian Splatting (3DGS), and Unity-based virtual reality (VR) to transform drone-captured reality into a reusable scene for controlled scenario comparison. The workflow is demonstrated in Middenbeemster, the central town of the Beemster polder World Heritage property. One present-condition scene (M0) and three alternative construction scenarios (M1 to M3) were created within a shared spatial reference. Reconstruction quality was assessed using PSNR and SSIM, and the VR scenes were further evaluated through eye-tracking, head-motion recording, and subjective ranking. The results indicate that the workflow can generate visually reliable and directly comparable immersive scenes from UAV data in this case study. Behavioral and subjective findings showed a consistent pattern, with M1 appearing more compatible than M2 and M3 in this pilot evaluation. The study contributes a pilot UAV-based workflow that links reality capture, immersive scenario comparison, and supplementary behavioral evidence within one process.

Keywords: drone photogrammetry; 3D Gaussian Splatting (3DGS); virtual reality; eye-tracking; head-tracking; world heritage site; scenario-based landscape management

Academic Editors: Arnadi Murtiyoso, Lorenzo Teppati Losè and Elisabetta Colucci

Received: 18 March 2026
Revised: 18 May 2026
Accepted: 21 May 2026
Published: 23 May 2026

1. Introduction

Unmanned aerial vehicles (UAVs) have become a routine platform for rapid, low-cost acquisition of high-resolution site imagery, supporting documentation, surveying, and 3D modeling in the built environment [1–3]. In construction and planning, however, decisions often depend less on producing a single accurate model and more on repeatedly comparing an existing condition against multiple plausible alternatives [4]. Many UAV-enabled pipelines still end in static deliverables such as orthomosaics, point clouds, or heavy mesh models, which are effective for measurement and archival purposes but inconvenient for fast scenario iteration and immersive stakeholder review [5].

Rather than addressing UAV-based decision support in general, this study targets a narrower methodological gap: current UAV-to-VR pipelines rarely reuse a single captured baseline to compare multiple proposals under matched immersive viewing conditions, and they rarely integrate lightweight behavioral measures as supplementary evidence for interpreting those comparisons [6,7]. This gap is particularly consequential in heritage-sensitive contexts [8], where small differences in placement, height, and typology can trigger disagreement, and where review processes benefit from both comparability and traceable evidence. This paper develops and demonstrates an integrated UAV-to-VR workflow for heritage-sensitive comparative review. The workflow combines three components: (1) construction of a reusable immersive baseline from UAV-derived data, (2) controlled comparison of alternative massing scenarios within a shared spatial reference, and (3) optional behavioral evaluation through eye-tracking and head-movement recording as supplementary evidence during VR-based review. The work is guided by three research questions:

RQ1: How can UAV-derived data be transformed into immersive, scenario-comparable environments for reviewing construction alternatives in a heritage-sensitive context?

RQ2: How do alternative construction scenarios affect visual attention, head-movement behavior, and subjective judgment during VR-based review?

RQ3: To what extent can behavioral measures complement subjective ranking in the comparative interpretation of scenario compatibility?

To illustrate the workflow in a realistic setting, we apply it to an expansion project in Middenbeemster, the central town of the Beemster polder World Heritage property (UNESCO), and use this case to demonstrate how UAV capture can be transformed into a scenario-comparable immersive review environment. This paper makes three contributions. It provides a reproducible UAV-to-VR pipeline that turns drone imagery into a reusable immersive baseline for scenario review. It introduces a scenario authoring strategy that enables direct comparison across alternatives by layering proposals onto a shared baseline under controlled VR viewing conditions. Finally, it demonstrates the workflow in a World Heritage planning case and shows how lightweight behavioral signals can complement subjective ranking for more evidence-informed review.

2. Related Works

Research related to this study mainly comes from three areas: UAV-based reality capture, lightweight scene representation for interactive visualization, and VR-supported comparative review in the built environment [9]. The following subsections summarize the main developments in each area and clarify the research gap addressed in this paper.

2.1. Drones in Documentation, Surveying, and 3D Modeling

Unmanned aerial vehicles have become a common tool for in situ documentation of the built environment [6,10,11]. Compared with conventional ground-based approaches,

UAVs offer a practical combination of mobility, flexible viewpoints, and relatively low operational cost, making them especially effective for capturing sites that are spatially extensive, visually complex, or difficult to access [12,13]. A single UAV campaign can often provide both geometric and visual information [14], which is why drones have increasingly been adopted in workflows related to as-built documentation, heritage recording, and digital-twin construction [15].

Another important advantage of UAV capture is its ability to generate dense multi-view imagery with repeatable coverage [16]. This gives UAV-based workflows a dual value. On the one hand, the collected imagery supports photogrammetric reconstruction and point-cloud generation for measurement-oriented tasks [17,18]. On the other hand, the same imagery preserves contextual and appearance-related information that is highly relevant for communication and visual evaluation. In this sense, UAVs already provide more than survey data alone; they also produce rich visual material that can potentially support immersive and interactive representations of place [19,20].

However, the dominant outputs of UAV workflows remain largely static. Orthomosaics, point clouds, and textured meshes are extremely useful for measurement, inspection, and archival purposes, but they are not always convenient when multiple alternative designs need to be compared quickly and repeatedly [21,22]. In planning and construction review, the key challenge is often not simply to reconstruct the existing condition, but to make that condition reusable as a baseline onto which proposals can be inserted, modified, and compared [23,24]. This is where a gap begins to emerge between the strengths of UAV reality capture and the requirements of immersive scenario-based evaluation.

2.2. 3D Gaussian Splatting as a Scenario-Comparable Visualization Substrate

Recent developments in 3DGS have opened a new possibility for bridging this gap [25,26]. Unlike conventional mesh-based pipelines, which often require substantial modeling, texturing, and optimization effort to remain visually convincing in interactive environments, 3DGS represents scenes in a way that is better suited to real-time rendering and free-viewpoint navigation [27,28]. This makes it especially relevant for workflows in which the existing condition needs to be explored interactively and reused across multiple alternatives [29]. Rather than treating reconstruction as the final deliverable, 3DGS makes it possible to treat the reconstructed scene as an active visualization substrate [30].

This role is particularly important in scenario-based design review. When several alternatives must be evaluated under comparable conditions, the baseline scene needs to remain visually stable while only the proposed intervention changes [31]. In this regard, 3DGS is attractive because it can preserve appearance cues related to material, surface texture, and lighting while still supporting real-time viewing [32]. This is a major advantage over many conventional static representations, where changing the proposal often requires rebuilding or heavily re-exporting the surrounding environment, making iteration slow and inconsistent [33].

The relevance of 3DGS becomes even stronger when considered together with UAV capture. Because drone imagery naturally provides dense multi-view observations, UAVs offer a suitable input modality for 3DGS reconstruction [34]. Recent work has shown that UAV imagery, when combined with point-cloud information, can support photorealistic visualization and interactive rendering in heritage-related contexts [35]. This suggests that 3DGS is not only a technical alternative to mesh reconstruction, but also a practical means of turning UAV-derived reality capture into a reusable baseline for immersive review [36]. Compared with mesh-based pipelines, 3DGS reduces remeshing and texture-optimization overhead for repeated scenario review. Compared with NeRF-style representations, it is generally easier to deploy for interactive VR comparison, although it remains less suitable for precise geometric editing. Existing studies typically emphasize either scene

representation or review performance. Kerbl et al. focused on real-time radiance-field rendering, Kim and Lee evaluated 3DGS against other reconstruction methods in VR, and Ma et al. examined immersive VR for end-user design review. Fewer studies operate these strands as one UAV-based workflow for reusable baseline construction and controlled scenario comparison.

2.3. Virtual Reality in Built-Environment Contexts

Virtual reality is increasingly used as a medium for communicating and evaluating spatial change in the built environment. Its main value lies in the ability to present design alternatives at human scale, allowing users to experience enclosure, visual dominance, continuity, and occlusion more directly than with 2D drawings or static rendered views [37,38]. For this reason, VR has been adopted in urban design, participatory planning, construction review, and heritage-related communication [39,40]. In these contexts, VR does not simply improve visualization quality; it changes the mode of evaluation by allowing stakeholders to inspect space through embodied exploration [41].

An additional strength of VR is its capacity to standardize the presentation of alternatives. When multiple scenarios are experienced through the same viewing positions, navigation logic, and rendering settings, differences in user judgment can be attributed more confidently to the scenarios themselves rather than to inconsistencies in presentation [42]. This makes VR particularly useful for comparison-based tasks, such as identifying which development option is less intrusive, more acceptable, or more compatible with an existing setting [43]. In heritage-sensitive contexts, where judgments may hinge on subtle differences in massing or placement, this comparability is especially important [44].

Recent developments also show that VR can provide more than immersive viewing alone. With the increasing availability of eye-tracked head-mounted displays, VR environments can now record gaze and head-motion data during evaluation [45]. These behavioral signals offer a lightweight physiological layer that can complement subjective responses by showing how participants allocate attention and how they explore different scenes. Although VR, eye-tracking, and head-motion analysis are all increasingly mature in their own right, they are still not frequently integrated with UAV-derived reality capture and scenario authoring in one coherent workflow [46]. As a result, the potential of VR as an environment for comparative review remains only partially realized in many planning and construction applications [47].

2.4. Summary

In summary, prior research has established the value of UAVs for site capture, 3DGS for interactive scene rendering, and VR for spatial review and evaluation. What remains insufficiently addressed is their integration into a practical workflow for immersive, scenario-comparable review. This is the gap addressed in the present study, particularly in the context of heritage-sensitive planning, where realistic visualization, controlled comparison, and lightweight behavioral evidence are all important. The gap addressed here is therefore not immersive visualization in general, but the methodological problem of reusable baseline construction, controlled scenario comparison, and behavior-informed interpretation within one workflow.

3. Methods

This study follows a drone-3DGS-VR workflow to create scenario-comparable immersive environments for construction preview and review. We combine AHN LiDAR point clouds as a stable geometric context with multi-angle drone imagery to reconstruct a present-condition 3DGS scene and to author alternative construction scenarios. The workflow is demonstrated on a case in the Beemster polder, and the resulting

environments are implemented in Unity for VR presentation and eye-tracking and head-motion measurement (Figure 1).

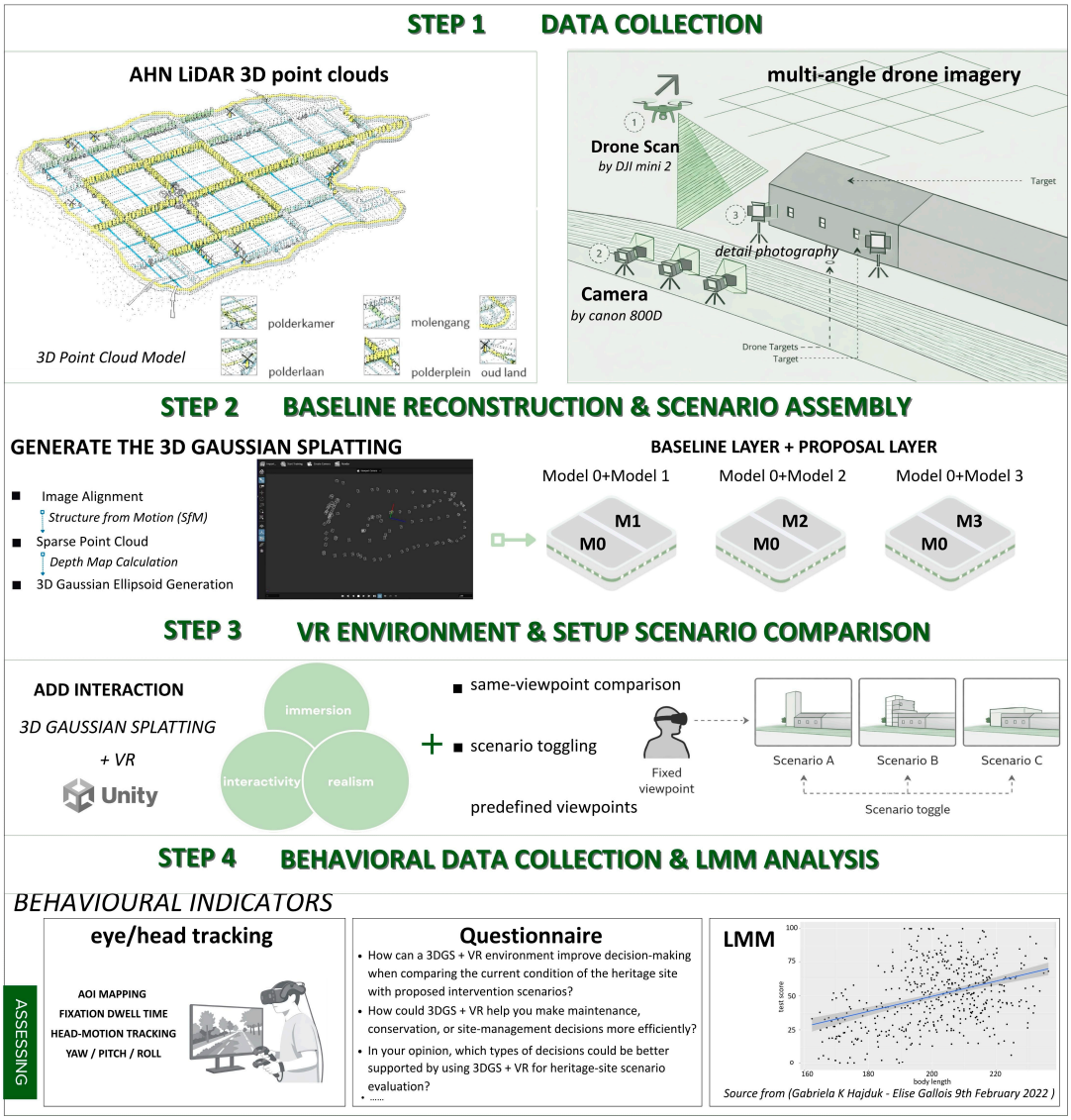

Figure 1. Overview of the methods, including the DJI Mini 2 UAV (DJI, Shenzhen, China) and Canon EOS 800D camera (Canon Inc., Tokyo, Japan); The LMM example shown in the lower/right panel is adapted from Hajduk and Gallois [Hajduk, G.K.; Gallois, E. Introduction to Linear Mixed Models. Our Coding Club. 2022].

3.1. Case Study

The case is situated in the Beemster polder, with Middenbeemster representing the central town within a World Heritage cultural landscape [48] (Figure 2). The site is visually characterized by an open polder setting, long sightlines, low building heights, and a highly legible spatial structure, all of which make changes in massing, placement, and skyline readily perceptible. This also makes the area visually fragile: new development is not easily absorbed by surrounding topography or dense urban fabric, and even moderate interventions may alter the perceived relationship between the historic town and its wider landscape. Against ongoing expansion pressure, the site therefore offers a relevant and timely case for testing immersive, scenario-based review in a heritage-sensitive planning context.

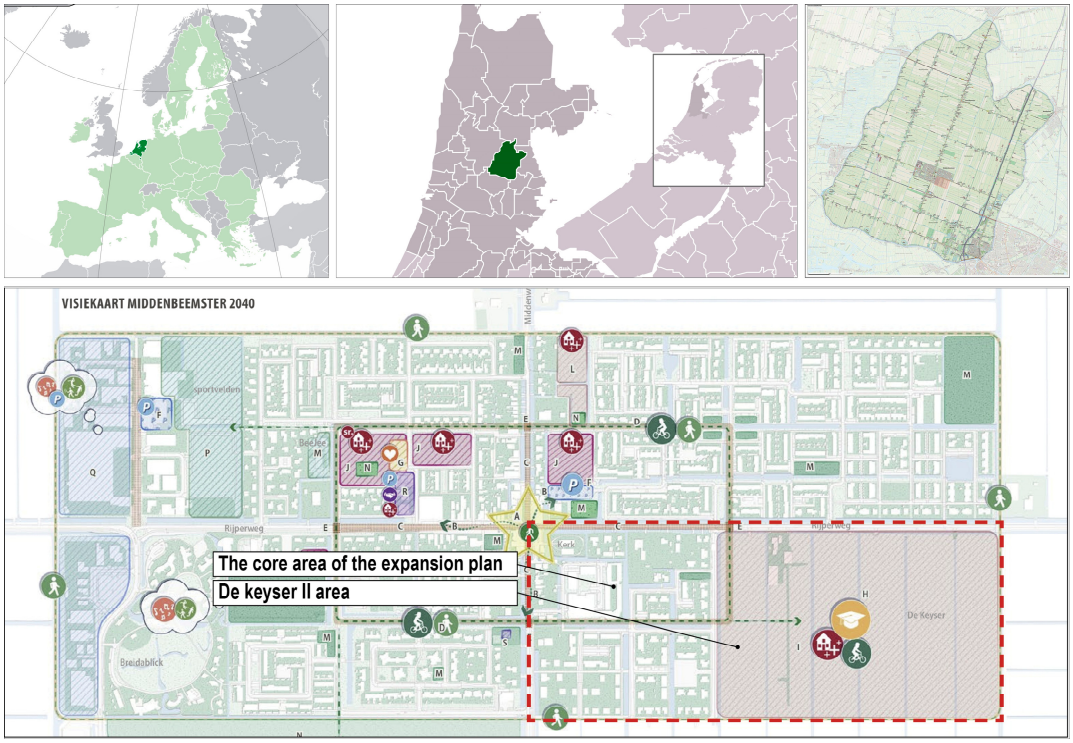

Figure 2. Middenbeemster expansion plan and Beemster Polder: the location of the targeted study area.

3.2. Data Collection and Preprocessing

The data used in this study consisted of a UAV image dataset supported by an AHN LiDAR point cloud. The UAV imagery served as the primary source for scene reconstruction and immersive visualization, while the AHN point cloud was used as auxiliary geometric reference data for scale control, spatial alignment, and contextual anchoring. In this way, the workflow remained UAV-driven, with the point cloud mainly supporting geometric consistency across reconstruction and scenario insertion.

The UAV survey was conducted using a DJI Mini 2 equipped with an onboard RGB camera. The camera uses a 1/2.3-inch CMOS sensor, 12 MP effective resolution, a 24 mm equivalent lens, fixed f/2.8 aperture, and 4000 × 3000 px still-image output. The survey was conducted under clear-sky daylight conditions during the late morning to early afternoon period to maintain stable illumination and reduce long-shadow effects. The images were acquired at an average altitude of approximately 40 m above ground level, corresponding to an estimated GSD of approximately 1.6 cm/pixel. Image acquisition followed a multi-pass configuration. A nadir grid flight was first conducted to provide uniform site coverage, followed by oblique passes along streets, plot boundaries, and façade-facing corridors to improve coverage of vertical surfaces and reduce occlusions near the scenario insertion area. The survey targeted approximately 70–80% forward overlap and 60–70% side overlap, with overlap kept above 60% across the retained image set. Camera settings were kept broadly consistent within each flight segment to reduce radiometric variation under stable daylight conditions. A total of 205 raw images were captured, of which 16 images were excluded during quality screening. After quality screening, 189 images were retained as the final image set. The full set was used for camera-pose estimation to ensure stable image alignment, after which 170 images were used for 3DGS optimization, and 19 images were held out from the optimization stage for image-based validation (Table 1).

Table 1. UAV acquisition and image-dataset parameters.

Category	Parameter	Value/Description
UAV platform	UAV model	DJI Mini 2
Camera	Sensor	1/2.3-inch CMOS RGB sensor
	Effective resolution	12 MP
	Image size	4000 × 3000 px
	Lens	24 mm equivalent
	Aperture	Fixed f/2.8
Acquisition condition	Weather/illumination	Clear-sky daylight
	Acquisition period	Late morning to early afternoon
Flight configuration	Average flight altitude	Approx. 40 m above ground level
	Estimated GSD	Approx. 1.6 cm/pixel
	Flight strategy	Multi-pass configuration
	Nadir coverage	Nadir grid flight for uniform site coverage
	Oblique coverage	Oblique passes along streets, plot boundaries, and façade-facing corridors
Image overlap	Forward overlap	Approx. 70–80%
	Side overlap	Approx. 60–70%
Image dataset	Raw images	205
	Excluded images	16 blurred, unstable, or poorly exposed images
	Final retained images	189
Image alignment	Camera-pose estimation	Full retained set: 189 images
3DGS optimization	Training images	170
Image-based validation	Held-out validation images	19
Validation logic	Validation split	Held-out images excluded from 3DGS optimization

To simulate potential construction interventions, four conditions were defined for scenario comparison. M0 represents the present-site condition in Middenbeemster. Based on this reference scene, three alternative development patterns were constructed using locally recognizable building prototypes. The three proposal scenarios were designed to remain broadly comparable in overall development scale, while differing in spatial organization, vertical expression, and settlement pattern. M1 represents a dispersed low-rise residential pattern, M2 represents a more centralized linear configuration with shed-like characteristics, and M3 represents a centralized community-center-type configuration with a stronger vertical expression. These four conditions provide a compact but diverse scenario set for examining how alternative construction forms may be perceived in a heritage-sensitive setting. Figure 3 shows the study area together with the prototype references used to guide scenario design. To ensure comparability, the three proposal scenarios were also quantified using a shared set of geometric indicators, including footprint, height, volume, site coverage, number of building blocks, and setbacks (Table 2), so that differences in visual response could be interpreted against explicit design parameters rather than qualitative labels alone.

Table 2. Quantitative descriptors of proposal scenarios.

Scenario	Footprint (m^2)	Max. Height (m)	Eave Height (m)	Ridge Height (m)	Volume (m^3)	Site Coverage (%)	No. of Building Blocks	Mean Setback (m)
M1	860	7.2	3.5	7.2	4950	18.4	6	14.5
M2	910	8.4	4.2	8.4	6180	19.5	1	8.6
M3	895	10.8	5.0	10.8	6540	19.2	2	5.6

Figure 3. Case study sites and prototype references in/around Beemster Polder.

The UAV survey was designed specifically to support view-consistent 3D reconstruction and immersive scene generation. Image acquisition followed a multi-pass flight strategy to improve coverage of both horizontal and vertical surfaces [49]. A nadir flight was first conducted in a grid pattern to provide uniform site coverage and strong image overlap for stable camera alignment. This was followed by one or more oblique flights along streets, plot boundaries, and facade-facing corridors, with the camera tilted to increase parallax on vertical surfaces that are less reliably captured from nadir views alone. Additional short-range oblique sweeps were then used to reduce occlusions caused by vegetation, narrow corridors, and edge conditions, particularly in those areas where scenario massing would later be inserted. This acquisition logic was important because small gaps or weakly reconstructed foreground elements become much more noticeable under immersive VR viewing than in conventional static visualization.

Throughout the UAV campaign, camera settings were kept consistent within each flight segment, and basic acquisition metadata were recorded to support repeatability in later processing and scenario updates. The resulting image set therefore functioned not only as documentation of the existing condition, but also as a reconstruction-ready dataset optimized for downstream immersive use.

During preprocessing, the UAV imagery was screened to remove frames that could negatively affect reconstruction quality, such as severely blurred, unstable, or poorly exposed images. The remaining images were then organized into a consistent multi-view dataset for camera pose estimation and 3DGS optimization. In parallel, the AHN point cloud was clipped to the study area and prepared as the common geometric reference

used to align the reconstructed scene and the inserted proposal layers within the same coordinate framework.

3.3. 3DGS Reconstruction Pipeline and Scenario Assembly

Image alignment and camera pose estimation were first performed in RealityCapture 1.4, and the aligned image set together with the estimated camera parameters was subsequently imported into Postshot v0.4.43 for 3DGS training. This workflow was adopted to improve camera registration stability and to separate the structure-from-motion stage from the neural reconstruction stage. Postshot supports training from externally imported camera poses and sparse points, allowing the radiance-field optimization to start directly from the aligned dataset.

The 3DGS model was trained in Postshot v0.4.43 using the imported aligned images and camera poses. Training was run for 30,000 iterations, with densification enabled at regular intervals during the early and middle optimization stages. The input images were processed at half-resolution to balance reconstruction fidelity and computational efficiency (Table 3).

Table 3. Key reconstruction and VR implementation parameters.

Component	Key Parameters
SfM alignment	RealityCapture 1.4; 189 retained UAV images; aligned cameras and camera poses exported for 3DGS training
3DGS training	Postshot v0.4.43; 170 training images; 19 held-out validation images; 30,000 iterations
Training setting	Half-resolution images; densification enabled during early and middle optimization
Validation	PSNR and SSIM calculated on held-out validation views
Unity implementation	Unity 2022; constant 3DGS baseline with scenario-specific proposal layers
Scenario comparison	Fixed viewpoints; controlled scenario switching; identical lighting, rendering, and interaction settings
Behavioral recording	HTC VIVE Pro Eye (HTC Corporation, New Taipei City, Taiwan); Tobii Pro Lab (Tobii Pro AB, Stockholm, Sweden, version 24.21); 120 Hz; semantic AOI mask render pass in Unity 2022

The clipped AHN point cloud was used as an external geometric reference for scale control and contextual anchoring. Because the AHN LiDAR point cloud and the UAV-derived reconstruction were acquired at different times, vegetation, temporary objects, and other seasonally variable surfaces were not used for alignment validation. Instead, the alignment and residual assessment were based on temporally stable hard-surface and built-environment features that could be identified in both datasets, including roof corners, road-edge features, and site-boundary points. The final transformation was then propagated to the 3DGS baseline and the derived scenario layers so that all reconstructed conditions shared a consistent coordinate framework. Registration quality was assessed using the root mean square error (RMSE) of the final alignment.

The reconstructed 3DGS scene was then treated as the persistent reference scene (M0) for all subsequent design comparisons. Scenario assembly was implemented compositionally: the shared reference scene remained unchanged, while different proposal layers were added to create the three intervention conditions (M1, M2, and M3). This structure ensured that all scenarios were directly comparable within the same spatial and visual context.

To evaluate whether the reconstructed 3DGS scene was visually reliable enough for immersive scenario comparison, we performed image-based validation on held-out views using peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). A subset of images was reserved for testing and excluded from optimization. The rendered outputs from the trained 3DGS model were then compared against the corresponding reference images to assess appearance fidelity and structural consistency. These metrics were used not as the primary goal of the study, but as a technical validation step to confirm that the reconstructed scene preserved sufficient visual quality to serve as the common reference environment for subsequent VR-based comparison.

3.4. VR Environment Construction in Unity

We built the immersive review environments in Unity 2022 by integrating the reconstructed baseline 3DGS layer with scenario-specific planned massing layers. The Unity scene was organized so that the baseline environment remained constant across all alternatives, while only the proposal layer differed between scenarios. To ensure controlled comparison, lighting configuration, rendering settings, and interaction rules were held constant throughout all conditions.

To support direct comparison between alternatives, the VR environment was structured around a same-viewpoint scenario comparison mechanism. Participants reviewed different design options from the same spatial context, allowing proposal differences to be judged against an identical baseline environment. Scenario switching was implemented as a controlled toggle between alternatives, so that comparisons could be made rapidly within the same immersive context without reloading scenes or altering system parameters.

Navigation was implemented using standard VR locomotion methods suitable for review sessions. Participants could move through teleportation among predefined locations and, where required, through movement constrained within a designated viewing zone.

3.5. Human-Subject Evaluation

To assess the comparative-review potential of the proposed workflow, we conducted a human-subject evaluation in immersive VR. The evaluation was designed to capture both behavioral responses during scenario viewing and subjective judgments after exposure to the full set of alternatives. Together, these outputs were intended to support a final comparative interpretation of scenario compatibility in the heritage-sensitive planning context.

3.5.1. Participant Recruitment

30 participants for the VR perception study were recruited using convenience sampling. Eligibility criteria required normal or corrected-to-normal vision and the ability to use a VR headset comfortably. Individuals reporting strong susceptibility to motion sickness or other medical conditions that could make VR use uncomfortable were excluded or advised not to participate. All participants provided informed consent prior to the experiment, and the study procedures complied with institutional ethics and data protection requirements. Participant characteristics are summarized in Table 4.

Table 4. The participants information of VR perception experiment.

Group	Features	Count	Native
Age	18–25	11	2
	26–30	14	3

	31–35	5	0
Gender	Male	16	3
	Female	14	2
Education	Bachelor's	13	2
	Master's	14	3
	PhD	3	0

Note: Native refers to participants who had lived in the Netherlands for five years or more and were familiar with the polder landscape context.

3.5.2. VR Procedure and Behavioral Measures

The VR sessions were conducted using an HTC VIVE Pro Eye headset with integrated eye-tracking, recorded through Tobii Pro Lab at 120 Hz, with a nominal accuracy of approximately 0.5–1.1°. Before the formal trials, each participant completed headset fitting and eye-tracking calibration, followed by a short practice session to become familiar with locomotion and scenario switching.

During the formal experiment, participants viewed each scenario from the same predefined virtual locations, so that observation position remained constant across conditions while head and eye movement were left unrestricted. Each scenario was presented for a fixed duration, and the order of presentation was randomized or counterbalanced to reduce sequence effects. Standard quality-control procedures were applied to exclude trials with unsuccessful calibration or insufficient tracking quality.

Participants evaluated the four conditions from three predefined viewpoints. Each viewing interval lasted 20 s, followed by 5 s for scenario switching. Scenario order was controlled using a Latin-square counterbalancing scheme. The full session lasted approximately 20 min including calibration and practice. The key procedural settings are summarized in Table 5.

Table 5. Key parameters of the VR experimental procedure.

Category	Parameter	Value
Viewing design	Number of scenarios	4 (M0 to M3)
	Number of predefined viewpoints	12 (3 viewpoints × 4 scenarios)
	Viewpoint logic	Fixed review viewpoints representing the principal visual approaches to the site
Timing	Viewing duration per scenario at each viewpoint	20 s
	Total session length	Approximately 20 min
	Session components	Headset fitting, eye-tracking calibration, practice, formal viewing, and final ranking task
Order control	Scenario order strategy	Latin-square counterbalancing
Interaction control	Participant position during comparison	Fixed at the same virtual viewpoint within each comparison sequence
	Allowed movement	Head and eye movements unrestricted; no free walking during formal trials
Eye-tracking	Device	HTC VIVE Pro Eye
	Recording software	Tobii Pro Lab
	Sampling frequency	120 Hz
Quality control	Gaze-validity threshold	≥80%
	Tracking-loss threshold	≤20% of the viewing interval

Exclusion criteria	Failed calibration, gaze-validity ratio below 80%, or tracking loss exceeding 20%

Eye-tracking and head-tracking signals were recorded continuously throughout each viewing interval. To characterize visual attention, we extracted area-of-interest (AOI) fixation dwell-time composition across semantic regions, including proposed massing and contextual background. AOIs were defined through a semantic mask render pass in Unity, allowing pixel-consistent assignment of gaze samples to AOI labels. AOI labels were interpreted relative to the scenario condition. In the proposal scenarios, the Building AOI captured visible built-form elements in the reviewed scene, including inserted proposal massing where present; in M0, the same AOI referred to existing built-form surfaces only. Similarly, the Intervention Site Area AOI referred to the geographically corresponding development area across all conditions. In M0, this AOI represented the pre-intervention site surface, whereas in M1 to M3 it represented the same area after scenario insertion. Accordingly, M0 was used as the present-condition behavioral reference, while comparisons among M1, M2, and M3 were used to interpret proposal-induced differences in visual attention. To characterize exploratory behavior, we also extracted head-tracking dispersion, defined as the standard deviation of yaw, pitch, and roll during each viewing interval. Together, these measures captured both attention allocation and embodied scene exploration during immersive scenario review.

3.5.3. Subjective Ranking

After viewing the complete set of scenarios, participants completed a brief subjective preference task. They were asked to rank the scenarios from most acceptable to least acceptable in terms of their perceived negative impact on the heritage setting. This forced-ranking task was intended to provide a lightweight, comparative outcome that resembles practical review situations while avoiding the burden and redundancy of a longer questionnaire-based survey.

3.5.4. Statistical Analysis

To analyze scenario effects on behavioral measures, we used linear mixed-effects models (LMMs) to account for repeated observations within participants. Scenario was specified as a fixed effect, and participant-level random intercepts were included to account for baseline differences between individuals. Where supported by the data, random slopes for scenario were added to capture inter-individual variation in sensitivity across alternatives. Model assumptions were checked through residual inspection, and post hoc contrasts were performed for pairwise comparisons between scenarios.

3.5.5. Integrated Comparative Review Output

The final output of the workflow was designed as a structured comparative review package rather than as a visualization alone. This package combined three layers of evidence: (1) the scenario-comparable immersive scenes (M0 to M3) presented in VR, (2) behavioral indicators derived from eye-tracking and head-tracking during scenario viewing, and (3) a subjective ranking outcome reflecting participants' overall judgment of perceived impact on the heritage setting. Together, these components enabled the scenarios to be interpreted not only in terms of how they looked, but also in terms of how they were visually engaged with and how they were ultimately judged.

In this sense, the workflow was intended to support a synthetic assessment of construction alternatives. The immersive scene provided the common spatial reference, the behavioral measures provided supportive evidence about attention allocation and

exploration, and the ranking task provided an explicit comparative outcome. The final delivery was therefore a structured basis for comparing scenario compatibility and identifying which intervention pattern was more or less appropriate for a heritage-sensitive landscape context.

4. Results

This section reports the results in direct correspondence with the three research questions introduced earlier. Sections 4.1 and 4.2 address RQ1 by examining whether the proposed UAV-to-3DGS-to-VR workflow can generate visually reliable and scenario-comparable immersive environments for reviewing construction alternatives. Section 4.3 addresses RQ2 by analyzing how different construction scenarios shape participants' visual attention and head-movement behavior during VR-based review. Section 4.4 addresses RQ3 by examining whether these behavioral responses align with subjective scenario ranking and can therefore support the comparative assessment of scenario compatibility in a heritage-sensitive context.

4.1. Drones-to-3DGS Scenario Visualization Results

A present-condition scene (M0) and three proposal scenarios (M1 to M3) were successfully generated for immersive comparison. Once M0 was established, the three intervention scenarios could be represented within the same shared environment, making the scene set suitable for repeatable comparison across alternatives (Figure 4).

Figure 4. 3DGS scenario (M0–M1/M0–M2/M0–M3).

Qualitatively, the reconstructed scenes preserved stable environmental appearance across conditions. The main visible differences were concentrated in the proposal areas, while the surrounding polder townscape remained perceptually consistent. This was particularly important in the present heritage-sensitive setting, where small differences in massing, continuity, and skyline expression can strongly affect perception. As a result, the scene set functioned as a scenario-comparable immersive environment rather than as a series of separately reconstructed models.

The visual quality of the scenes was also sufficient for immersive review. Oblique UAV coverage improved the continuity of facades, street edges, and foreground elements,

reducing the visual incompleteness that would have been more noticeable in VR under a nadir-only acquisition. Although some limitations remained in occlusion-prone areas, especially near vegetation edges, the overall scene continuity was adequate for direct side-by-side comparison of the scenarios.

Spatial-alignment validation first supported the comparability of the scenario set. The final stable-feature UAV/SfM–AHN registration achieved a 3D RMSE of 0.144 m, with a mean absolute residual of 0.126 m, a 95th-percentile residual of 0.218 m, and a maximum residual of 0.287 m. Because the AHN and UAV datasets were acquired at different times, these residuals were calculated on stable built-environment features rather than vegetation or other seasonally variable surfaces. These results indicate that the baseline reconstruction and scenario layers were sufficiently aligned for comparative VR review (Table 6).

Table 6. Spatial alignment validation.

Parameter	Value/Description
Reference and CRS	AHN LiDAR point cloud; Amersfoort/RD New, EPSG: 28992
Alignment features	Stable built-environment features: roof corners, road edges, and site-boundary points
Check correspondences	6 correspondences distributed around the study area and scenario insertion zone
Refinement	Rigid ICP; 6-DoF transformation; scale adjustment disabled
Accuracy	3D RMSE = 0.144 m; mean absolute residual = 0.126 m; 95th percentile = 0.218 m; maximum = 0.287 m
Scenario consistency	Final transformation propagated to M0 baseline and M1–M3 scenario layers

In addition to spatial alignment, image-based validation was used to assess the visual usability of the reconstructed 3DGS scene. On held-out views, the model achieved a PSNR of 25.21 dB and an SSIM of 0.71, indicating satisfactory visual fidelity and structural consistency for immersive rendering (Table 7). Together, the spatial-alignment and image-based validation results suggest that the reconstructed scene preserved both the spatial consistency and the main appearance cues necessary for use as the common visual reference in subsequent scenario comparison.

Table 7. Image-based validation of 3DGS reconstruction quality.

Method	Dataset	Split	Images	PSNR ↑ (dB)	SSIM ↑	Quality
Depth Regularization	Beemster	test	19	25.21 (25.207728)	0.71 (0.705251)	Good

Note: ↑ indicates that a higher value represents better performance.

4.2. VR Environment Quality and Controlled Scenario Comparison

All four conditions (M0 to M3) were successfully presented within a single Unity-based VR environment. Across conditions, the surrounding environmental context remained constant, and only the proposal content varied between scenarios. This provided a stable basis for controlled immersive comparison and ensured that differences observed during review could be attributed to scenario content rather than to changes in the broader scene.

Scenario comparison in VR was also spatially consistent. Participants viewed the alternatives from the same virtual locations, which reduced contextual drift and made the comparison more direct and interpretable. Fast switching between scenarios further

supported this consistency, allowing participants to compare alternatives within the same immersive frame rather than treating them as separate experiences.

The VR environment achieved stable interactive performance suitable for review tasks. Loading time remained short, rendering operated within a real-time frame range, and latency remained low enough to support comfortable immersive viewing. Memory usage and GPU overhead were also kept within a manageable range for the scene size, indicating that the UAV-to-3DGS workflow can support practical immersive scenario review without relying on excessively heavy visual assets (Table 8).

Table 8. Runtime performance of the 3DGS-based VR environment.

Aspect	Loading Time	Real-Time FPS	Memory Usage	GPU Processing Overhead	Visual Clarity Depth	Latency
3DGS	1 s	60–90	13.6 MB	12–25%	continuous	~10 ms

4.3. Participant Eye-Tracking and Head-Tracking Results

Eye-tracking and head-tracking data were successfully collected across all four conditions. Because M0 represents the present-site condition without inserted proposal massing, its AOI values should be interpreted as attention to corresponding existing scene components, including the pre-intervention site area, rather than as attention to new construction. The analysis therefore distinguishes two comparative levels: M0 provides a behavioral reference for the existing site, while M1 to M3 are compared to identify how alternative intervention forms redirected attention and shaped exploratory viewing. Among the three intervention scenarios, M1 produced the weakest proposal-focused attentional response, whereas M2 and especially M3 produced stronger attention to built-form elements and more concentrated viewing patterns.

4.3.1. Gaze Allocation Across Scenarios

The gaze results indicate that scenario form affected the distribution of visual attention during immersive review. Because M0 contains no inserted proposal massing, its Building and Intervention Site Area values should be interpreted as attention to existing built elements and to the corresponding pre-intervention site area. The main proposal-related interpretation therefore focuses on differences among M1, M2, and M3, with M0 serving as the present-condition reference. Figure 5 illustrates the aggregated gaze heatmaps and the corresponding semantic AOI masks used for gaze-sample assignment across the proposal scenarios.

Among the intervention scenarios, M1 produced the weakest proposal-focused attentional response. Total dwell time was approximately 3.8 s, building-related dwell was approximately 167 ms, and dwell on the intervention-site AOI was approximately 88 ms. This pattern suggests that the dispersed low-rise configuration introduced limited additional visual competition with the surrounding landscape context. In this condition, attention remained less concentrated on built-form intervention elements than in the other proposal scenarios.

M2 generated a stronger proposal-related gaze response. Total dwell time increased to approximately 5.5 s, building-related dwell rose to approximately 568 ms, and dwell on the intervention-site AOI reached approximately 239 ms. These values indicate that the centralized linear configuration directed more attention toward the reviewed intervention area and its built-form expression than M1. This supports the interpretation that M2 was visually more assertive within the inherited landscape structure.

M3 showed the strongest built-form attentional capture among the proposal scenarios. Building-related dwell reached approximately 1.8 s, substantially higher than in M1 and M2. However, dwell time on the Intervention Site Area AOI did not increase monotonically; it was lower than in M2 and showed wider uncertainty. This indicates that M3's effect was not simply a stronger fixation on the geographically defined site area. Rather, participants appeared to respond to a broader reorganization of scene hierarchy caused by the centralized and vertically stronger built form. Figure 6 further illustrates these differences in visual exploration, showing how fixation distributions and scanpaths became more concentrated around proposal-related built forms in the more visually assertive scenarios.

The LMM results support this descriptive pattern (Table 9). Scenario condition affected gaze allocation, and post hoc contrasts indicated that M2 and M3 elicited higher building-focused dwell than M1, although the contrast between M2 and M3 depended on the AOI category considered. Taken together, the gaze evidence should be interpreted as an ordering of proposal-induced perceptual intensity among the intervention scenarios: M1 generated the weakest built-form attention, M2 produced an intermediate increase, and M3 produced the strongest built-form attentional shift. This section therefore reports behavioral differences in attention allocation, while the evaluative interpretation of scenario compatibility is addressed together with subjective ranking in Section 4.4.

Table 9. Summary of LMM-estimated eye-tracking metrics across the present condition and proposal scenarios.

Metric	Scenario	Mean	95% CI
Building	M1	167.21	[115, 242]
	M2	567.81	[351, 917]
	M3	1788.73	[957, 3342]
	M0	1360.28	[1255.31, 1465.25]
Intervention Site Area	M1	87.74	[56, 138]
	M2	239.35	[142, 403]
	M3	112.54	[31, 402]
	M0	2721.5	[2482.44, 2960.56]
Polder land	M1	8.28	[6, 11]
	M2	17.72	[11, 27]
	M3	125.87	[46, 341]
	M0	350.89	[307.24, 394.54]
Sky	M1	250.73	[169, 371]
	M2	125.8	[88, 180]
	M3	563.25	[338, 938]
	M0	224.66	[188.51, 260.81]
Total	M1	3808.48	[3064, 4734]
	M2	5524.26	[4201, 7264]
	M3	6134.06	[3297, 11,412]
	M0	10,049.74	[9676.23, 10,423.25]

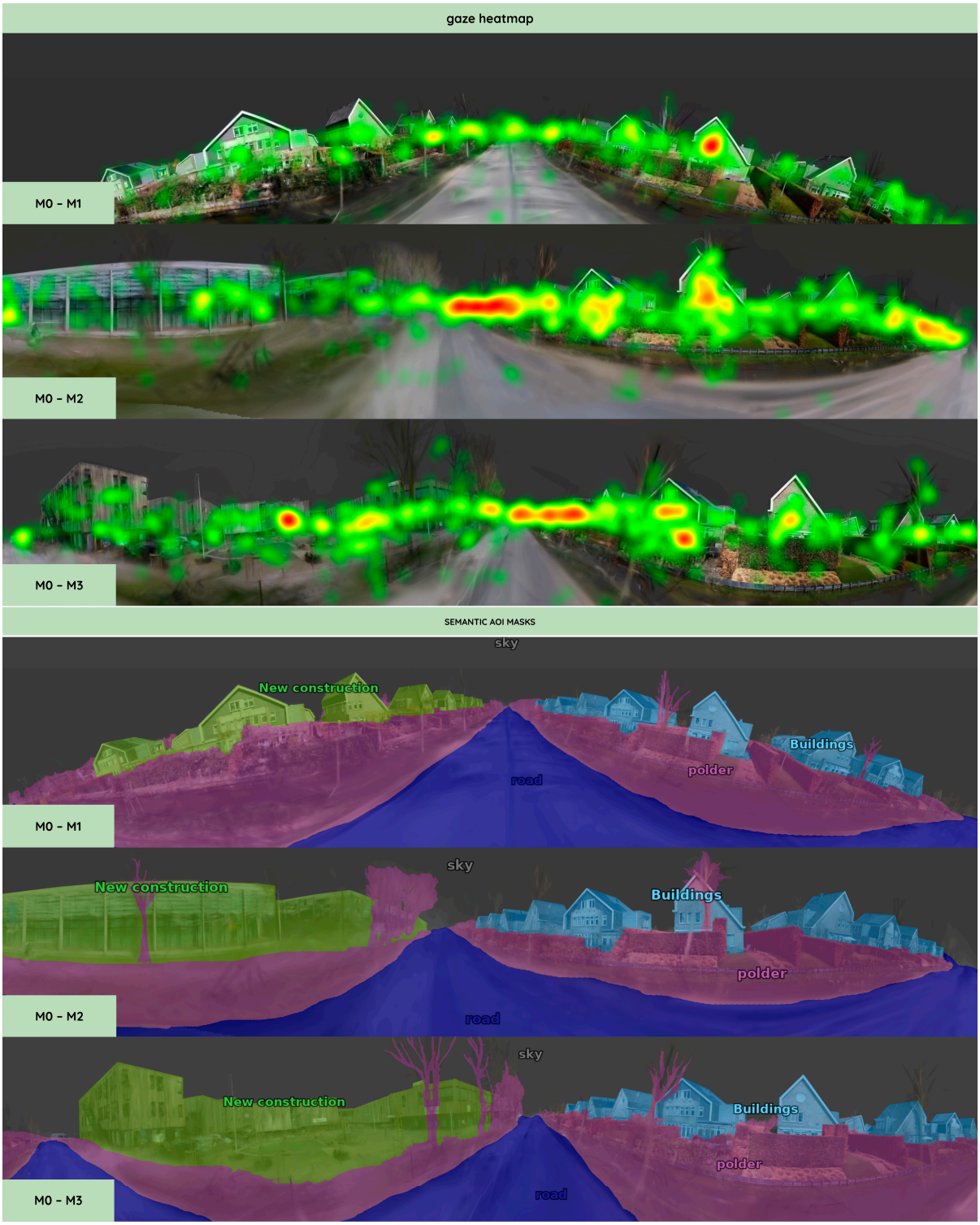

Figure 5. Gaze heatmaps and semantic AOI masks for the proposal scenarios.

https://doi.org/10.3390/drones10060404

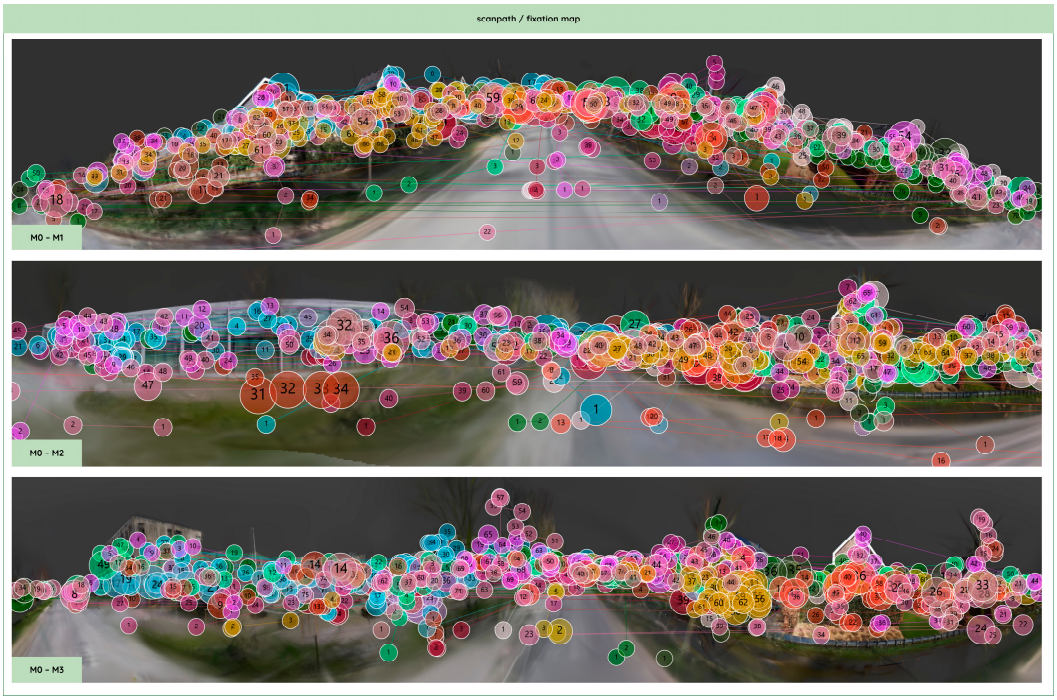

Figure 6. Scanpath and fixation maps for the three proposal scenarios.

4.3.2. Head-Movement Patterns Across Scenarios

Head-movement behavior provided complementary evidence about exploratory viewing, but it should be interpreted more cautiously than gaze allocation (Figure 7). Head-tracking dispersion describes the breadth of embodied scanning during each viewing interval; lower dispersion indicates more constrained exploration, not necessarily higher or lower preference.

Descriptively, all proposal scenarios reduced yaw dispersion relative to M0. M0 showed the broadest horizontal exploration, whereas M3 showed the lowest yaw dispersion, with M1 and M2 occupying intermediate positions. This suggests that the insertion of proposal massing tended to narrow horizontal scanning, especially in M3, where the stronger vertical and centralized configuration appeared to concentrate viewing around the intervention.

Pitch and roll showed more axis-specific patterns. Pitch dispersion was lower in M1 and M2 than in M0, whereas M3 remained closer to the present condition, suggesting that its vertical expression may have induced additional upward and downward adjustment. Roll dispersion was lower in M1 and M3 than in M0, while M2 remained closer to the present condition. These differences indicate that head movement did not follow a fully monotonic scenario order across all axes (Table 10; Supplementary Table S1).

The LMM contrasts support this cautious interpretation (Table 11). M1 and M2 showed significantly lower pitch dispersion than M0, while M3 was not clearly different from the present condition. For roll dispersion, M1 and M3 were significantly lower than M0, whereas M2 showed only a weak difference. For yaw dispersion, M1 and M3 were significantly lower than M0, while M2 showed a non-significant intermediate tendency. Thus, yaw provided the clearest evidence of reduced horizontal exploration, while pitch and roll reflected axis-specific responses to scenario form.

Overall, the head-tracking results indicate that proposal scenarios tended to reduce the breadth of scene exploration, particularly along the yaw axis, but they should not be treated as standalone indicators of scenario compatibility. Their value lies in complementing gaze allocation by showing how different intervention forms changed embodied viewing behavior. The combined behavioral evidence suggests that M1 produced the least

proposal-focused gaze response among the intervention scenarios, M2 produced a stronger but mixed behavioral response, and M3 combined high building-related gaze allocation with the clearest reduction in yaw and roll dispersion. The interpretation of these behavioral patterns in relation to subjective acceptability is developed in Section 4.4.

Table 10. Descriptive summary of head-tracking metrics (The complete results are provided in Supplementary Table S1).

Metric	M0	M1	M2	M3	Explanation
pitch_std	43.75	36.80	36.38	40.76	Intervention scenarios generally reduced vertical exploration, although M3 may have triggered slightly more vertical adjustment.
roll_std	82.52	64.82	77.44	64.50	M0 showed the highest roll dispersion; M1 and M3 were lower, while M2 was intermediate.
yaw_std	93.80	72.59	82.20	63.07	M0 showed the broadest horizontal exploration; M3 was the most constrained, while M1 and M2 were intermediate.

Table 11. Pairwise contrasts from LMMs for head-tracking metrics.

Metric	Contrast	Estimate	SE	t/z	*p*	Interpretation
pitch_std	M1 vs M0	−6.95	2.80	−2.48	0.018	Lower pitch dispersion
	M2 vs M0	−7.37	2.90	−2.54	0.016	Lower pitch dispersion
	M3 vs M0	−2.99	2.60	−1.15	0.260	Not clearly different
roll_std	M1 vs M0	−17.70	6.50	−2.72	0.011	Lower roll dispersion
	M2 vs M0	−5.08	6.20	−0.82	0.420	Weak difference
	M3 vs M0	−18.02	6.70	−2.69	0.012	Lower roll dispersion
yaw_std	M1 vs M0	−21.21	7.50	−2.83	0.008	Lower yaw dispersion
	M2 vs M0	−11.60	7.20	−1.61	0.115	Intermediate
	M3 vs M0	−30.73	8.00	−3.84	<0.001	Clearly lower yaw dispersion

4.4. Subjective Ranking and Comparative Interpretation

After completing the VR session, participants ranked the four conditions from most acceptable to least acceptable in terms of their perceived negative impact on the heritage setting. The resulting ordering was clear and internally consistent: M0 > M1 >> M2 ≈ M3. Thus, the present condition was judged most acceptable, M1 emerged as the most acceptable intervention scenario, and M2 and M3 were consistently perceived as the least acceptable and broadly comparable in impact.

The ranking provides the primary evaluative outcome, whereas the VR-based behavioral measures provide an explanatory layer for interpreting how this outcome was formed during scene viewing. In gaze allocation, M2 and M3 drew stronger building-related attention than M1. In head tracking, M3 in particular showed reduced yaw and roll dispersion, while M2 remained more intermediate and axis-specific. These patterns suggest that the lower-ranked alternatives more strongly reorganized participants' visual engagement around the intervention, whereas M1 produced a weaker proposal-focused behavioral response.

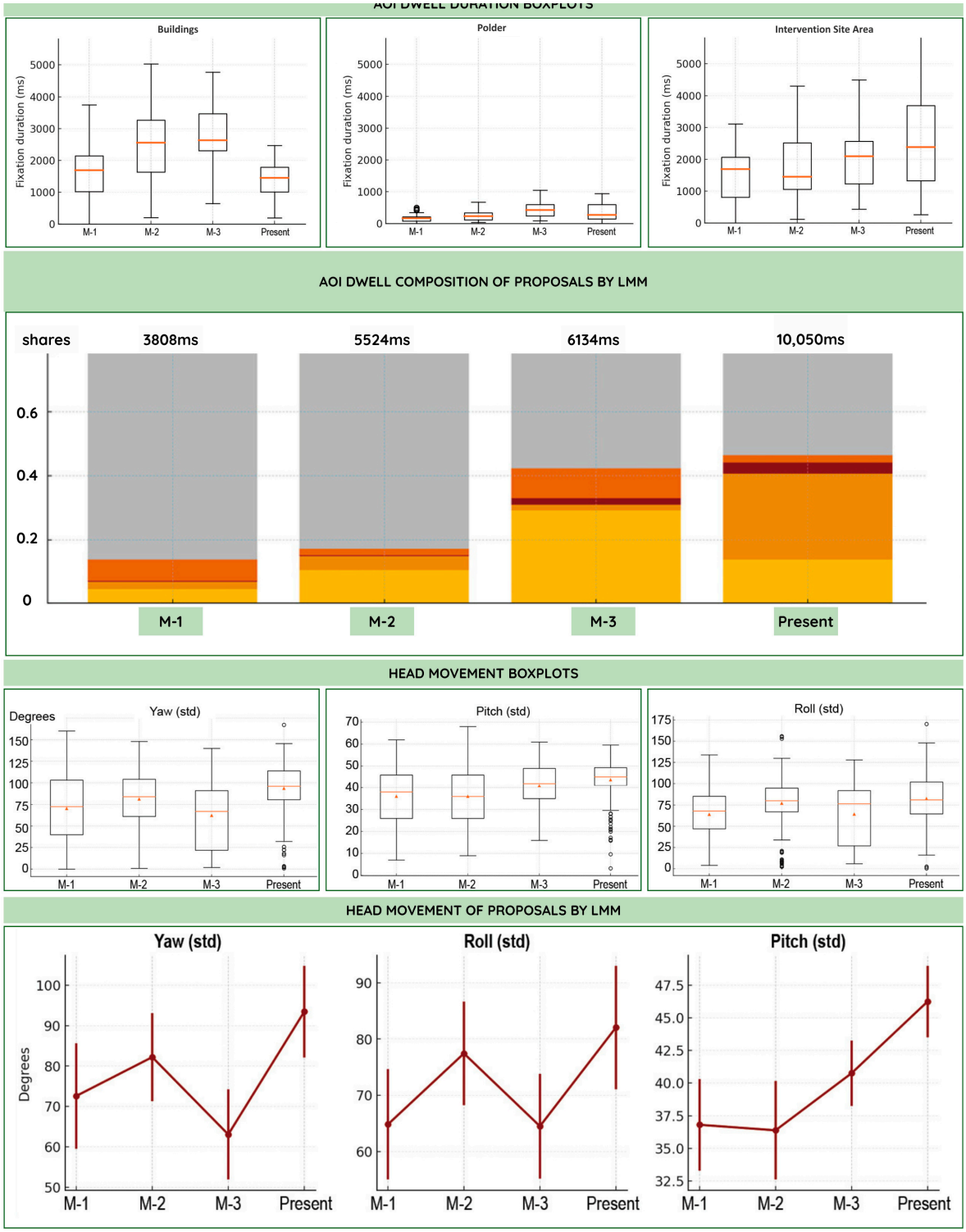

Figure 7. Quantitative summaries of gaze allocation and head-movement behavior across scenarios.

This convergence between explicit ranking and behavioral response strengthens the interpretability of the results. It suggests that the lower acceptance of M2 and M3 was not only expressed at the level of stated preference, but was also accompanied by measurable differences in attention allocation and exploratory viewing during immersive review. The behavioral evidence should therefore be understood as supplementary rather than determinative: it does not replace subjective judgment, but helps explain why certain alternatives were experienced as more visually intrusive within the reviewed heritage landscape.

From the perspective of comparative review, the combined evidence suggests a scenario ordering within this case. Among the three intervention options, M1 appeared to be

the most visually compatible alternative for the Beemster context, whereas M2 and M3 formed a higher-impact, lower-acceptance group. This suggests that, in a visually fragile heritage landscape defined by openness, long sightlines, and low skyline tolerance, dispersed low-rise development may be more compatible with the existing landscape character than centralized and more visually assertive building configurations.

5. Discussions

The results show that the proposed workflow can support both immersive scenario construction and comparative evaluation in a heritage-sensitive planning context. Across the tested scenarios, the technical, behavioral, and ranking results point in the same direction: the workflow produced visually comparable VR scenes, distinguished differences in attention and exploration across alternatives, and supported a consistent ordering of scenario acceptability. The following discussion focuses on what these findings imply for the role of UAV-based capture in heritage-sensitive review, for the appraisal of open rural heritage landscapes, and for the practical use of the workflow in planning and design processes.

5.1. From Documenting Heritage to Supporting Comparative Review

A central implication of this study is that UAV-based heritage recording can extend beyond documentation and function as part of a comparative review process. In many heritage and built-environment applications, UAV workflows still primarily produce orthomosaics, point clouds, or static 3D models. These outputs are valuable for archival, survey, and communication purposes, but they do not by themselves support structured comparison between alternative future interventions. The present results suggest that, when linked with 3DGS and VR, UAV-derived data can also serve as the basis for evaluating possible development scenarios under controlled conditions.

What distinguishes the proposed workflow is not only that it generates an immersive environment, but that it enables repeated comparison within a common scene rather than across separately produced visualizations. This makes scenario differences easier to attribute and supports a more systematic form of review. From this perspective, the contribution of the workflow lies less in producing a single realistic representation than in providing a repeatable process for comparing alternatives under consistent spatial and visual conditions. Its methodological novelty lies in the integration of a reusable UAV-derived immersive baseline, compositional scenario layering, and matched VR comparison within a single workflow.

This also has practical implications. The final output is not a single visualization, but a structured review package that combines the immersive scenario environment, the behavioral evidence recorded during viewing, and the ranking-based judgment outcome. Taken together, these components suggest that the workflow may be useful not only for academic perception studies, but also for planning review, design iteration, and heritage-impact discussions where alternatives need to be compared quickly and consistently.

5.2. Implications for Open Rural Heritage Landscapes

The findings are particularly relevant to open rural heritage landscapes such as those found in the Netherlands, where long sightlines, low skylines, and legible spatial organization make new interventions especially visible [50,51]. In such settings, compatibility appears to depend not only on the amount of new construction, but also on the extent to which development follows the inherited spatial order of the landscape. This includes the existing settlement grain, the distribution of built volumes, and the broader relationship between buildings and the surrounding open polder structure.

From this perspective, the relative preference for M1 should not be read simply as a general preference for low-rise or dispersed development. Rather, the results suggest that proposals more closely aligned with the pre-existing spatial logic of the site are more likely to be perceived as compatible. By contrast, the lower acceptance of M2 and M3 indicates that more centralized and visually assertive configurations may disrupt the established relationship between settlement and landscape, even when their scale remains within a plausible development range.

In this sense, the case suggests that heritage-sensitive appraisal in open rural landscapes benefits from comparing how alternative spatial organizations relate to the existing landscape order. For open Dutch rural heritage environments, the present findings indicate that future development may be better received when it works with, rather than against, the inherited spatial structure of the place

5.3. Behavioral Evidence as a Supplementary Interpretive Layer

The behavioral results provide an additional layer for interpreting the ranking outcome. The ranking indicates which scenarios were considered more or less acceptable, whereas the eye-tracking and head-movement data help clarify how those judgments were formed during immersive viewing. Across the tested scenarios, lower-ranked alternatives were also associated with stronger attention to built elements and more concentrated exploration patterns. By contrast, M1 remained closer to M0 in both gaze allocation and head movement.

This convergence suggests that the lower acceptance of M2 and M3 was not only expressed at the level of explicit preference, but was also accompanied by measurable differences in how participants visually engaged with the scenes. At the same time, the behavioral layer should not be understood as a replacement for subjective judgment. Its value lies in supplementing ranking results with additional evidence about attention allocation and viewing behavior during review.

This is particularly relevant in design and planning contexts where decisions often rely on a combination of visual interpretation, stakeholder judgment, and contextual reasoning [52]. In such settings, behavioral measures may help strengthen the interpretability of comparative review by showing whether explicit preferences are accompanied by corresponding differences in immersive scene engagement [53].

5.4. Limitations and Future Directions

Several limitations should be acknowledged. First, the study is based on a single case in a visually distinctive World Heritage landscape, and the scenario set was intentionally compact. This makes the case well suited for testing the sensitivity of the workflow, but it also limits the generalizability of the findings. The present results show that the workflow can detect meaningful differences in a heritage-sensitive rural context; they do not yet establish that the same perceptual ordering would hold in denser urban settings, enclosed landscapes, or other forms of development pressure. The convenience sample and the modest sample size support a bounded pilot evaluation of the workflow, but they do not support broad claims about stakeholder populations or planning outcomes.

Second, the human-subject evaluation prioritized comparability and control over ecological breadth. Participants viewed all scenarios from the same predefined locations, and the subjective task was intentionally kept lightweight through forced ranking. These choices were appropriate for isolating proposal-driven differences, but they also reduced the variety of movement, interpretation, and reasoning that may occur in less structured public review situations.

Third, the present workflow supports comparative judgment rather than universal thresholds. It allows one scenario to be judged as more or less compatible than another,

but it does not yet define transferable acceptability limits that could be applied across sites or projects. Future work should therefore test the workflow across multiple cases, broader participant groups, and more varied scenario sets in order to determine how robust the observed patterns are and whether more general comparative evaluation criteria can be derived.

Taken together, these limitations do not diminish the central contribution of the study. Rather, they clarify the stage at which the workflow currently operates: it is already capable of linking UAV-derived reality capture, immersive comparison, behavioral evidence, and explicit judgment into a coherent review process, but further work is needed to extend its validity across contexts. Even at this stage, however, the present case demonstrates a meaningful step from heritage recording toward evidence-informed comparative review of future construction scenarios [54].

6. Conclusions

This paper presented a UAV-enabled workflow for constructing scenario-comparable immersive environments for built-environment review. By combining multi-angle UAV imagery with point-cloud-based geometric anchoring, the workflow transformed drone-captured reality into a reusable scene that could support multiple alternative construction scenarios within a shared spatial reference. In this process, 3DGS and VR functioned as downstream technologies that extended the value of UAV data from documentation toward interactive comparison and evaluation.

The Beemster case demonstrated that UAV-derived data can support both visually reliable immersive scene construction and controlled comparison of alternative interventions. The results further showed that behavioral and subjective evidence can be integrated within the same review framework to distinguish between more and less compatible construction scenarios. The main contribution of this study therefore lies in the workflow itself: a practical UAV-based pipeline that links reality capture, immersive scenario comparison, and behavior-informed evaluation. This provides a basis for scenario-based planning and heritage-sensitive design review in contexts where alternatives must be assessed under consistent conditions.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/drones10060404/s1. Table S1: The complete results of the Table 10.

Author Contributions: Conceptualization, X.L. and Y.P.; methodology, X.L. and W.S.; software, W.S.; validation, X.L., W.S. and Y.P.; formal analysis, X.L.; investigation, X.L. and W.S.; resources, Y.P.; data curation, W.S.; writing—original draft preparation, X.L. and W.S.; writing—review and editing, Y.P. and X.L.; visualization, W.S.; supervision, Y.T., Y.P. and Y.Y.; project administration, Y.P. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Doctoral Development Fund of Xi'an University of Architecture and Technology Huaqing College (Grant No. 2026DDF-NS-M07).

Institutional Review Board Statement: This study involved human participants and was approved by the relevant institutional Human Research Ethics Committee prior to the commencement of the experiment.

Informed Consent Statement: All participants provided informed consent prior to participation.

Data Availability Statement: The UAV imagery and derived 3D assets supporting the findings of this study contain location-sensitive cultural heritage information and participant-related data. De-identified evaluation outputs, including aggregated eye-tracking and locomotion metrics, are

available from the corresponding author, Yuyang Peng, upon reasonable request. Access to raw site data may be restricted due to heritage protection requirements and site-access agreements.

Acknowledgments: The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Horvat, N.; Kunnen, S.; Štorga, M.; Nagarajah, A.; Škec, S. Immersive virtual reality applications for design reviews: Systematic literature review and classification scheme for functionalities. *Adv. Eng. Inform.* **2022**, *54*, 101760. https://doi.org/10.1016/j.aei.2022.101760.
2. Jiang, S.; Jiang, W.; Wang, L. Unmanned Aerial Vehicle-Based Photogrammetric 3D Mapping: A survey of techniques, applications, and challenges. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 135–171. https://doi.org/10.1109/MGRS.2021.3122248.
3. Zhang, C.; Lin, G.; Peng, Y.; Yu, Y. A Workflow for Urban Heritage Digitization: From UAV Photogrammetry to Immersive VR Interaction with Multi-Layer Evaluation. *Drones* **2025**, *9*, 716. https://doi.org/10.3390/drones9100716.
4. Jarrin, F.; Koga, Y.; Thomas, D.; Kawasaki, H. Virtual reality-based site layout planning for building design. *Autom. Constr.* **2024**, *167*, 105690. https://doi.org/10.1016/j.autcon.2024.105690.
5. Elghaish, F.; Matarneh, S.; Talebi, S.; Kagioglou, M.; Hosseini, M.R.; Abrishami, S. Toward digitalization in the construction industry with immersive and drones technologies: A critical literature review. *Smart Sustain. Built Environ.* **2020**, *10*, 345–363. https://doi.org/10.1108/sasbe-06-2020-0077.
6. Savini, F.; Cordisco, A.; Fabbrocino, G.; Giallonardo, M.; Trizio, I.; Marra, A. From Drone-Based 3D Model to a Web-Based VR Solution Supporting Cultural Heritage Accessibility. *Drones* **2025**, *9*, 775. https://doi.org/10.3390/drones9110775.
7. Stanga, C.; Banfi, F.; Roascio, S. Enhancing Building Archaeology: Drawing, UAV Photogrammetry and Scan-to-BIM-to-VR Process of Ancient Roman Ruins. *Drones* **2023**, *7*, 521. https://doi.org/10.3390/drones7080521.
8. Ma, J.H.; Erdogmus, E.; Kangisser, S.; Yang, E. A comparative analysis of the effectiveness of immersive virtual reality on end-user design review. *Build. Environ.* **2025**, *267*, 112237. https://doi.org/10.1016/j.buildenv.2024.112237.
9. Peng, Y.; Nijhuis, S.; Wang, Z.; Yu, Y.; Verbree, E.; van Oosterom, P. When to stop? A visual impact assessment framework for incremental urban and community expansion in rural heritage landscapes. *Sustain. Cities Soc.* **2026**, *143*, 107364. https://doi.org/10.1016/j.scs.2026.107364.
10. Colomina, I.; Molina, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. https://doi.org/10.1016/j.isprsjprs.2014.02.013.
11. Siebert, S.; Teizer, J. Mobile 3D mapping for surveying earthwork projects using an Unmanned Aerial Vehicle (UAV) system. *Autom. Constr.* **2014**, *41*, 1–14. https://doi.org/10.1016/j.autcon.2014.01.004.
12. Mohsan, S.A.H.; Khan, M.A.; Noor, F.; Ullah, I.; Alsharif, M.H. Towards the Unmanned Aerial Vehicles (UAVs): A Comprehensive Review. *Drones* **2022**, *6*, 147. https://doi.org/10.3390/drones6060147.
13. Telli, K.; Kraa, O.; Himeur, Y.; Ouamane, A.; Boumehraz, M.; Atalla, S.; Mansoor, W. A Comprehensive Review of Recent Research Trends on Unmanned Aerial Vehicles (UAVs). *Systems* **2023**, *11*, 400. https://doi.org/10.3390/systems11080400.
14. Hou, J.; Zhang, X.; Li, H.; Cui, S. GeoRefGS: Towards Georeferenced 3D Gaussian Splatting from Unmanned Aerial Vehicle Platforms. *Drones* **2026**, *10*, 195. https://doi.org/10.3390/drones10030195.
15. Martinez, J.G.; Alarcon, L.F.; Wandahl, S. Unmanned Aerial Systems (UAS)-Derived 3D Models for Digital Twin Construction Applications. In *Applications of Point Cloud Technology*; Şahin, C., Ed.; IntechOpen: London, UK, 2024. https://doi.org/10.5772/intechopen.1004746.
16. Yan, L.; Fei, L.; Chen, C.; Ye, Z.; Zhu, R. A Multi-View Dense Image Matching Method for High-Resolution Aerial Imagery Based on a Graph Network. *Remote Sens.* **2016**, *8*, 799. https://doi.org/10.3390/rs8100799.
17. Angás, J.; Bea, M.; Valladares, C.; Iranzo, C.; Ruiz, G.; Fatás, P.; Heras, C.d.L.; Sánchez-Carro, M.Á.; Bruschi, V.; Prada, A.; et al. Cave of Altamira (Spain): UAV-Based SLAM Mapping, Digital Twin and Segmentation-Driven Crack Detection for Preventive Conservation in Paleolithic Rock-Art Environments. *Drones* **2026**, *10*, 73. https://doi.org/10.3390/drones10010073.
18. Akgul, M.; Yurtseven, H.; Gulci, S.; Akay, A.E. Evaluation of UAV- and GNSS-Based DEMs for Earthwork Volume. *Arab. J. Sci. Eng.* **2018**, *43*, 1893–1909. https://doi.org/10.1007/s13369-017-2811-9.

19. Chengoden, R.; Victor, N.; Huynh-The, T.; Yenduri, G.; Jhaveri, R.H.; Alazab, M.; Bhattacharya, S.; Hegde, P.; Maddikunta, P.K.R.; Gadekallu, T.R. Metaverse for Healthcare: A Survey on Potential Applications, Challenges and Future Directions. *IEEE Access* **2023**, *11*, 12765–12795. https://doi.org/10.1109/ACCESS.2023.3241628.
20. Calisi, D.; Botta, S.; Cannata, A. Integrated Surveying, from Laser Scanning to UAV Systems, for Detailed Documentation of Architectural and Archeological Heritage. *Drones* **2023**, *7*, 568. https://doi.org/10.3390/drones7090568.
21. Dadrass Javan, F.; Samadzadegan, F.; Toosi, A.; van der Meijde, M. Unmanned Aerial Geophysical Remote Sensing: A Systematic Review. *Remote Sens.* **2025**, *17*, 110. https://doi.org/10.3390/rs17010110.
22. Pádua, L.; Adão, T.; Hruška, J.; Marques, P.; Sousa, A.; Morais, R.; Lourenço, J.M.; Sousa, J.J.; Peres, E. UAS-based photogrammetry of cultural heritage sites: A case study addressing Chapel of Espírito Santo and photogrammetric software comparison. In *ICGDA '18: Proceedings of the International Conference on Geoinformatics and Data Analysis*; Association for Computing Machinery: New York, NY, USA, 2018; pp. 72–76. https://doi.org/10.1145/3220228.3220243.
23. Casisirano, J.D.D.; Claridades, A.R.C. Comparative Analysis of UAV Photogrammetry and Terrestrial Laser Scanning for 3D Building Reconstruction. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2026**, *10*, 143–150. https://doi.org/10.5194/isprs-annals-X-5-W4-2025-143-2026.
24. Wang, F.; Zou, Y.; del Rey Castillo, E.; Ding, Y.; Xu, Z.; Zhao, H.; Lim, J.B. Automated UAV path-planning for high-quality photogrammetric 3D bridge reconstruction. *Struct. Infrastruct. Eng.* **2024**, *20*, 1595–1614. https://doi.org/10.1080/15732479.2022.2152840.
25. Kerbl, B.; Kopanas, G.; Leimkühler, T.; Drettakis, G. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.* **2023**, *42*, 139. https://doi.org/10.1145/3592433.
26. Yu, Y.; Verbree, E.; van Oosterom, P.; Pottgiesser, U.; Peng, Y.; Poux, F. From Comparison to Integration: A Workflow Evaluation of 3D Gaussian Splatting and LiDAR Point Cloud for Modern Architectural Heritage. *Autom. Constr.* **2025**, *180*, 106509. https://doi.org/10.1016/j.autcon.2025.106509.
27. Liu, S.; Yang, M.; Xing, T.; Yang, R. A Survey of 3D Reconstruction: The Evolution from Multi-View Geometry to NeRF and 3DGS. *Sensors* **2025**, *25*, 5748. https://doi.org/10.3390/s25185748.
28. Fei, B.; Xu, J.; Zhang, R.; Zhou, Q.; Yang, W.; He, Y. 3D Gaussian Splatting as a New Era: A Survey. *IEEE Trans. Vis. Comput. Graph.* **2025**, *31*, 4429–4449. https://doi.org/10.1109/TVCG.2024.3397828.
29. Aryal, A.; Giri, S.; Panday, S.P.; Sharma, S.; Dawadi, B.R.; Chalise, S. Efficient 3D Scene Reconstruction From Multi-View RGB Images Using Optimized Gaussian Splatting. *IEEE Access* **2026**, *14*, 1269–1286. https://doi.org/10.1109/ACCESS.2025.3648171.
30. Kim, H.; Lee, I.K. Is 3DGSUseful?: Comparing the Effectiveness of Recent Reconstruction Methods in VR. In Proceedings of the 2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Bellevue, WA, USA, 21–25 October 2024. https://doi.org/10.1109/ISMAR62088.2024.00021.
31. Wang, H. Design of environmental perception and computer virtual reality technology in urban landscape design. *J. Comput. Methods Sci. Eng.* **2025**, *25*, 615–629. https://doi.org/10.1177/14727978251322051.
32. Kwon, O.; Yu, J. Realistic and Interactive Virtual Museum Representation Using 3D Gaussian Splatting. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2025**, *10*, 185–192. https://doi.org/10.5194/isprs-annals-X-M-2-2025-185-2025.
33. Wang, Y.; Xu, H.; Li, Y.; Tian, F. High-Fidelity Virtual Reality With 3D Gaussian Splatting: User Experience Pathways and System Performance Evaluation. *IEEE Access* **2026**, *14*, 8238–8249. https://doi.org/10.1109/ACCESS.2026.3653618.
34. Yu, Y.; Verbree, E.; van Oosterom, P.; Pottgiesser, U. 3D Gaussian Splatting for Modern Architectural Heritage: Integrating UAV-Based Data Acquisition and Advanced Photorealistic 3D Techniques. *AGILE GISci. Ser.* **2025**, *6*, 51. https://doi.org/10.5194/agile-giss-6-51-2025.
35. Zachos, A.; Anagnostopoulos, C.-N. Using TLS, UAV, and MR Methodologies for 3D Modelling and Historical Recreation of Religious Heritage Monuments. *J. Comput. Cult. Herit.* **2024**, *17*, 56. https://doi.org/10.1145/3679021.
36. Pepe, M.; Alfio, V.S.; Costantino, D. UAV Platforms and the SfM-MVS Approach in the 3D Surveys and Modelling: A Review in the Cultural Heritage Field. *Appl. Sci.* **2022**, *12*, 12886. https://doi.org/10.3390/app122412886.
37. Seghetto, I.; Lopes, R.; Lima, F.; Boffi, M.; Rainisio, N.; Stancato, G.; Piga, B.E.A. Virtual Reality as a Tool for Enhancing Understanding of Tactical Urbanism. *Architecture* **2025**, *5*, 26. https://doi.org/10.3390/architecture5020026.
38. Rubio-Tamayo, J.L.; Gertrudix Barrio, M.; García García, F. Immersive Environments and Virtual Reality: Systematic Review and Advances in Communication, Interaction and Simulation. *Multimodal Technol. Interact.* **2017**, *1*, 21. https://doi.org/10.3390/mti1040021.

39. Lucchi, E. Chapter 25—Digital twins, artificial intelligence and immersive technologies for heritage preservation and cultural tourism in smart cities. In *Digital Twin, Blockchain, and Sensor Networks in the Healthy and Mobile City*; Nguyen, T.A., Ed.; Elsevier: Amsterdam, The Netherlands, 2025; pp. 507–520. https://doi.org/10.1016/B978-0-443-34174-8.00026-0.
40. Tsifodimou, Z.E.; Skondras, A.; Stamou, A.; Skalidi, I.; Tavantzis, I.; Stylianidis, E. Monitoring the Impact of Urban Development on Archaeological Heritage Using UAV Mapping: A Framework for Preservation and Urban Growth Management. *Drones* **2025**, *9*, 669. https://doi.org/10.3390/drones9100669.
41. Ali, A.S.; Afifi, S.; Khateeb, S.E.; El Fayoumi, M.A.; ElHusseiny, O.M. Virtual reality simulation in urban design processes: Comparative workflow assessment and implementation outcomes. *Ain Shams Eng. J.* **2025**, *16*, 103671. https://doi.org/10.1016/j.asej.2025.103671.
42. Piga, B.; Boffi, M.; Rainisio, N.; Stancato, G.; Ceravolo, P.; Tavares, G. An interdisciplinary approach to urban planning via Augmented/Virtual Reality and Process Mining. In *C3places Using ICT for Co-Creation of Inclusive Public Space*; Edições Universitárias Lusófonas: Porto, Portugal, 2021; pp. 53–54. https://doi.org/10.24140/2020-sct-vol.4-2.1.
43. Bishop, I.D.; Rohrmann, B. Subjective responses to simulated and real environments: A comparison. *Landsc. Urban Plan.* **2003**, *65*, 261–277. https://doi.org/10.1016/S0169-2046(03)00070-7.
44. Wang, K.; Wang, M.; Zhou, J. Integrating heritage conservation into urban building energy modelling for retrofit decision-making in historic districts. *npj Herit. Sci.* **2026**, *14*, 1. https://doi.org/10.1038/s40494-025-02186-9.
45. Moreno-Arjonilla, J.; López-Ruiz, A.; Jiménez-Pérez, J.R.; Callejas-Aguilera, J.E.; Jurado, J.M. Eye-tracking on virtual reality: A survey. *Virtual Real.* **2024**, *28*, 38. https://doi.org/10.1007/s10055-023-00903-y.
46. Adhanom, I.B.; MacNeilage, P.; Folmer, E. Eye Tracking in Virtual Reality: A Broad Review of Applications and Challenges. *Virtual Real.* **2023**, *27*, 1481–1505. https://doi.org/10.1007/s10055-022-00738-z.
47. Yan, G.; Chen, Y. The Application of Virtual Reality Technology on Intelligent Traffic Construction and Decision Support in Smart Cities. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 3833562. https://doi.org/10.1155/2021/3833562.
48. Nijhuis, S.; Sun, Y.; Cannatella, D.; Xie, G. A Landscape-Based Regional Design Approach for Sustainable Urban Development in the Pearl River Delta, China. In *Adaptive Urban Transformation: Urban Landscape Dynamics, Regional Design and Territorial Governance in the Pearl River Delta, China*; Nijhuis, S., Sun, Y., Lange, E., Eds.; Springer International Publishing: Cham, Switzerland, 2023; pp. 81–114. https://doi.org/10.1007/978-3-030-89828-1_5.
49. James, M.R.; Robson, S.; Smith, M.W. 3-D uncertainty-based topographic change detection with structure-from-motion photogrammetry: Precision maps for ground control and directly georeferenced surveys. *Earth Surf. Process. Landf.* **2017**, *42*, 1769–1788. https://doi.org/10.1002/esp.4125.
50. Janssen, J.; Knippenberg, L. The Heritage of the Productive Landscape: Landscape Design for Rural Areas in the Netherlands, 1954–1985. *Landsc. Res.* **2008**, *33*, 1–28. https://doi.org/10.1080/01426390701773748.
51. Peng, Y.; Nijhuis, S. A GIS-based algorithm for visual exposure computation: the west lake in Hangzhou (China) as example. *J. Digit. Landsc. Archit.* **2021**, 6, 424–435. https://doi.org/10.14627/537705038.
52. Cepero, T.; Montané-Jiménez, L.G.; Toledo-Toledo, G. Visualization Technologies to Support Decision-Making in City Management. *Program. Comput. Softw.* **2021**, *47*, 803–816. https://doi.org/10.1134/S0361768821080107.
53. Spittle, B.; Frutos-Pascual, M.; Creed, C.; Williams, I. A Review of Interaction Techniques for Immersive Environments. *IEEE Trans. Vis. Comput. Graph.* **2023**, *29*, 3900–3921. https://doi.org/10.1109/TVCG.2022.3174805.
54. Yaremych, H.E.; Persky, S. Tracing physical behavior in virtual reality: A narrative review of applications to social psychology. *J. Exp. Soc. Psychol.* **2019**, *85*, 103845. https://doi.org/10.1016/j.jesp.2019.103845.