Digital Scene Reconstruction-Based Immersive VR Aesthetic Education Classroom Model

Aihua Yin, Longbin Song*, Guoyu Liu, Qingzhi Ma

Harbin Institute of Information Technology, Harbin 150431, China

*Corresponding author: hxcisonglongbin@126.com

Abstract. Immersive VR has strong potential for aesthetic education, yet many classroom applications still separate scene display from aesthetic interpretation, creative transfer, and assessment. This revised paper reconstructs the teaching model by clarifying coding definitions, the logic of model derivation, and the assessment design. The study uses a coded public-source corpus of seven policy, platform, and scholarly texts published between 2018 and 2025. Three coding dimensions are employed: scene-reconstruction function, classroom teaching link, and learning-evidence orientation. The last dimension is operationalized as the dominant type of observable student trace linked to four core aesthetic objectives: perceptual attention, interpretive judgment, creative translation, and reflective articulation. The corpus yields 28 scene-function mentions, 24 teaching-link mentions, and 18 learning-evidence mentions. The results show a concentration in reconstruction, roaming, manipulation, observation, and interpretation, while creation and reflection remain relatively weak. On this basis, the paper establishes an explicit chain from coded diagnosis to scenario modules, teaching tasks, and a four-stage classroom model. It further specifies four assessment indicators with concrete tools, four-level scoring criteria, and script-level feasibility checks.

Keywords: digital scene reconstruction, immersive virtual reality, aesthetic education, classroom teaching model, content coding, learning evidence.

1 Introduction

Recent framework guidance and arts-education studies show that immersive VR/XR classroom design is moving toward deeper curricular integration, stronger scaffolding, and clearer scene-to-task links [5-7]. At the same time, immersive virtual reality (VR) and digital scene reconstruction are increasingly used to simulate exhibition environments, heritage spaces, and visual-creation contexts that cannot be fully reproduced in ordinary classrooms [1,4].

The supporting digital ecosystem is also expanding. Recent framework guidance and XR/design studies indicate that accessible resources, creative environments, and immersive platforms are increasingly available for classroom use [5-7].

R. Zur et al. (eds.), *Proceedings of the 2026 5th International Conference on Art Design and Digital Technology (ADDT 2026)*, Advances in Computer Science Research 131,
https://doi.org/10.2991/978-94-6239-737-8_85

Yet many VR-based art or appreciation classes still show three persistent problems: technology-centered demonstration, weak linkage between observation and creation, and insufficient reflective assessment. The present paper therefore asks how digital scene reconstruction can be translated into a teachable VR aesthetic education model rather than a one-off visual spectacle.

2 Methodological Rationale and Framework

2.1 Why Digital Scene Reconstruction Suits Immersive VR Aesthetic Education

Digital scene reconstruction suits aesthetic education because it does not merely present images; it rebuilds cultural or artistic context through spatial structure, object relations, light, viewpoint, and embodied navigation. Prior research shows that immersive VR can strengthen presence and motivation, but learning effects depend on how immersive experience is pedagogically structured [1,2]. In aesthetic education, this structure matters because learners must move from seeing to judging, interpreting, and creating.

This rationale is reinforced by cultural-heritage VR studies, which show that reconstructed digital environments are effective when they support contextual understanding rather than isolated novelty [4]. Accordingly, the key design problem is not whether VR is immersive, but how reconstructed scenes are embedded into a classroom sequence with explicit learning evidence.

2.2 Operational Meaning of the Coding Dimensions

To respond to the review concern regarding coding specificity, the revised study uses three coding dimensions with explicit operational definitions. In particular, learning-evidence orientation refers to the dominant type of observable student trace that a VR-supported activity is expected to produce. It was coded only when a source explicitly linked scene use to one or more core objectives of aesthetic education: perceptual attention to form, interpretive judgment, creative translation, or reflective articulation. The coding dimensions are summarized in Table 1.

Table 1. Coding dimensions, operational definitions, and links to aesthetic education objectives

Dimension	Operational definition	Core objective	Typical textual trigger
Scene-reconstruction function	Explicit technical or spatial affordance offered by a reconstructed scene	Support contextual entry and guided perception	"roaming," "viewpoint switching," "object hotspot"
Classroom teaching link	Explicit teaching action or task sequence attached to VR use	Organize priming, observation, creation, and review	"pre-class task," "guided observation," "peer critique"

Learning-evidence orientation	Explicit student trace expected from the activity output	Judge perception, interpretation, creation, and reflection	"annotated screen-shot," "storyboard," "reflection log"

2.3 Research Framework

The revised framework includes six linked steps: policy and theory scanning, public-source survey, coded diagnosis, intermediate instructional inference, scene-module translation, and evidence-based assessment design. The output is a reusable teaching model that connects scene reconstruction technology with aesthetic learning tasks and observable classroom evidence. The overall framework is illustrated in Figure 1.

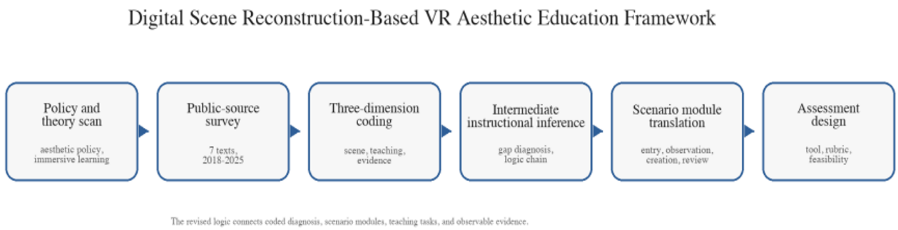

Fig. 1. Research framework for digital scene reconstruction-based immersive VR aesthetic education classroom design.

3 Public-Source Survey and Pedagogical Diagnosis

3.1 Data Sources and Coded Patterns

The corpus contains seven public-source texts published between 2018 and 2025, including four scholarly studies on immersive VR, arts education, and cultural-heritage learning and three official Chinese education documents on aesthetic education and digital resources [1-7]. The coded results yield 28 scene-function mentions and 24 teaching-link mentions. The coded frequencies are summarized in Figure 2.

A further pass on learning-evidence orientation identifies 18 explicit evidence mentions: perceptual attention appears 6 times, interpretive judgment 5 times, creative translation 4 times, and reflective articulation 3 times. This means that current public discourse is comparatively strong in contextual entry, navigation, and guided observation, but less explicit about how students demonstrate creative and reflective achievement.

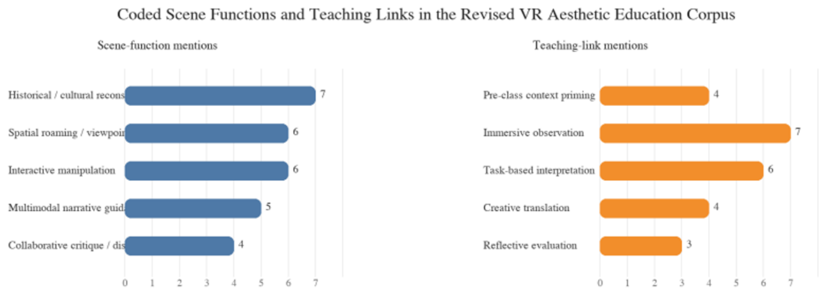

Fig. 2. Frequency of coded scene-function and teaching-link mentions in the VR aesthetic education public-source corpus.

3.2 Intermediate Logic from Coding Diagnosis to Teaching Model

The fourth-order teaching model was not derived by directly renaming the coding categories. Instead, three intermediate inferences were used. First, scene-function coding shows that reconstructed spaces mainly support contextual entry, viewpoint movement, and object-focused observation. Second, teaching-link coding shows that current classroom discourse is concentrated in observation and interpretation, with weaker links to creation and reflective review. Third, learning-evidence coding shows that observable outputs become thinner after the observation stage. Together, these findings justify a model that deliberately extends the sequence from scene entry to observation, creation, and reflection. The empirical chain from diagnosis to model design is summarized in Table 2.

Table 2. Empirical chain from coded pattern to scenario module, teaching task, and model stage

Coded pattern	Scenario module	Teaching task	Model stage
Reconstruction, roaming, and hotspot mentions dominate scene coding	Context-entry and observation modules	Spatial orientation, hotspot tracing, form comparison	Stage 1-2
Observation and interpretation dominate teaching-link coding	Interpretation support module	Prompt cards, annotated screenshots, guided discussion	Stage 2
Manipulation and recomposition appear, but less frequently	Creation module	Storyboard, sketch, recomposition, narrative transfer	Stage 3
Reflective evaluation is least frequent in the corpus	Reflection module	Peer feedback, replay commentary, reflection log	Stage 4

The diagnosis can therefore be stated clearly: the main challenge of immersive VR aesthetic education is not the lack of technology, but the incomplete instructional chain linking reconstructed scene, guided perception, creative task, and assessment evidence.

4 Teaching Model Construction and Assessment Design

4.1 Four-Stage Immersive VR Aesthetic Teaching Model

Based on the logic above, the revised paper proposes a four-stage classroom model: context priming, immersive observation, interpretive creation, and reflective evaluation. Context priming activates historical or artistic background before headset use; immersive observation focuses on viewpoint movement, scale comparison, and hotspot analysis; interpretive creation converts perception into sketching, narrative boards, or recomposition tasks; and reflective evaluation uses replay, peer critique, and rubric evidence to consolidate aesthetic judgment.

This sequence is consistent with the view that immersion alone is not sufficient and must be mediated by cognitive and affective design [2]. It also responds to arts-education research calling for VR to support creativity, contextual understanding, and not only sensory novelty [3].

4.2 Assessment Tools, Scoring Criteria, and Feasibility

Table 3. Assessment indicators, concrete tools, and four-level scoring anchors

Indicator	Measurement tool	Observable evidence	4-level scoring anchor
Presence quality	Teacher checklist + hotspot completion record	Navigation path and task focus	1 off-task; 2 reaches targets; 3 stable attention; 4 active comparison and sustained engagement
Interpretation accuracy	Question sheet + annotated screenshot	Evidence-based visual explanation	1 naming only; 2 partial description; 3 accurate interpretation; 4 accurate and evidence-backed interpretation
Creative translation	Storyboard or sketch rubric	Recomposed visual output	1 copying; 2 minor change; 3 meaningful transformation; 4 transformation with clear aesthetic intent
Reflective articulation	Reflection log + peer feedback card	Comparative judgment and revision statement	1 preference only; 2 simple comment; 3 reasoned comparison; 4 reasoned comparison with revision insight

Feasibility was checked at the script level rather than through unsupported claims of large-sample validation. Each indicator was retained only when it had a fixed collection point in the lesson, at least one observable student artifact, and a four-level rubric that could be completed within one class period. This design makes the four indicators operational rather than merely conceptual. The assessment indicators and scoring anchors are summarized in Table 3.

Under the current Chinese digital-education context, this path is practicable because shared platform resources and recent national smart education initiatives already provide a broad technical base [6,7]. The urgent task is therefore pedagogical scripting and evidence collection rather than hardware novelty.

5 Conclusion

Using a coded corpus of seven public-source texts together with official Chinese education data, this revised paper shows that current immersive VR aesthetic education emphasizes scene presentation, roaming, observation, and interpretation more than creative translation and reflective articulation. The core problem is not technical scarcity but the incomplete connection between reconstructed digital scene, teaching task, and learning evidence.

The revised contribution is threefold. First, it clarifies the operational meaning of learning-evidence orientation and relates it directly to the core objectives of aesthetic education. Second, it reconstructs the explicit chain from coded diagnosis to scenario module, teaching task, and four-stage teaching model. Third, it converts the assessment design into a usable instrument set with concrete tools, four-level scoring criteria, and script-level feasibility checks.

References

1. Radianti, J., Majchrzak, T.A., Fromm, J., Wohlgenannt, I.: A systematic review of immersive virtual reality applications for higher education. Comput. Educ. 147, 103778 (2020). DOI: 10.1016/j.compedu.2019.103778. https://doi.org/10.1016/j.compedu.2019.103778
2. Makransky, G., Petersen, G.B.: The cognitive affective model of immersive learning (CAMIL). Educ. Psychol. Rev. 33, 937-958 (2021). DOI: 10.1007/s10648-020-09586-2. https://doi.org/10.1007/s10648-020-09586-2
3. Gonzalez-Zamar, M.-D., Abad-Segura, E.: Implications of virtual reality in arts education. Educ. Sci. 10(9), 225 (2020). DOI: 10.3390/educsci10090225. https://doi.org/10.3390/educsci10090225
4. Barbara, J.: Re-Live History: An immersive virtual reality learning experience of prehistoric intangible cultural heritage. Front. Educ. 7, 1032108 (2022). DOI: 10.3389/feduc.2022.1032108. https://doi.org/10.3389/feduc.2022.1032108
5. UNESCO: UNESCO Framework for Culture and Arts Education: Implementation guidance. DOI: https://doi.org/10.54675/APCF2495.
6. Liu, X., Shen, Y., Nagai, Y., and Kato, H.: Use of a mixed-reality creative environment in design education. Computers & Education: X Reality 4, 100055 (2024). DOI: https://doi.org/10.1016/j.cexr.2024.100055.
7. Huralna, S., et al.: Application of Audiovisual Technologies in Art Education. 2025. DOI: https://doi.org/10.70389/PJS.100108.