DOI: 10.54254/2753-7064/2026.32721

Visualization as Dialogic Medium: Generative AI, Positional Repositioning, and the Apparatus of Architectural Representation

Xuanrun Yi

University of California, Los Angeles, USA
yixuanrun0326@g.ucla.edu

Abstract. This paper is a theoretical and literature-based critical analysis examining how generative AI repositions architectural visualization within the design process. It argues that the positional repositioning of visualization, its migration from the terminal stage into the operative middle of the design process, amplifies rather than neutralizes the constitutive agency of the medium: by embedding a high-fidelity apparatus at the site of open-ended conceptual exploration, generative AI subjects the most creatively consequential phase of design to an aesthetic logic the designer did not author and cannot inspect. Drawing on Lawson and Schön's reflection-in-action framework, the paper analyzes the prompt-generation-evaluation cycle as a new form of design conversation. Through Flusser's apparatus theory and Irrgang's extensions, it reveals a structural epistemic asymmetry: the architect controls input and output, but the model's internal aesthetic logic, encoded in a latent space conditioned by historically biased training data, remains opaque. Del Campo's analysis of the architectural latent space identifies this corpus as a site of inherited canonical bias. A design experiment testing a culturally specific vernacular façade vocabulary against generative model outputs demonstrates this asymmetry in practice. The paper concludes that AI visualization is simultaneously expansive and constraining, functioning as a medium with its own constitutive agency rather than a neutral accelerant.

Keywords: generative AI, architectural visualization, dialogic medium, latent space, apparatus theory

1. Introduction

The proliferation of generative AI tools capable of producing architectural imagery from natural language prompts and rough sketches has altered the conditions under which design visualization occurs. What was once a time-consuming and labor-intensive production process that had to be physically rendered and demanded a large amount of time, is now a process that can be accomplished almost instantly through the process of iteration. The feedback connection between conceptual intention and visual expression can be collapsed in one conversation because now architects can make, consider, and discard visual expressions of spatial conceptualizations. It is not

DOI: 10.54254/2753-7064/2026.32721

only technological change, but structural change, the change in the cognitive and procedural situation of design, as such.

This change does not take place evenly. Generative AI has shown its most significant implications in the design middle making phases, as concept exploration, multi-schemes comparison, and atmospheric testing, through a near-instant feedback response to speed up the concept-to-refinement loop [1]. Generative AI output is probabilistic and can not replace conventional rendering pipelines in the terminal stages when a high level of technical precision, regulations, and dimensional accuracy is needed. Current literature has reported the potential of AI to help improve exploration at an early stage [2,3] and proven that the visual representations generated actively affect design judgment but do not reflect it [4]. However, the literature has not sufficiently answered the question of the location of such repositioning, the reasons behind its success during intermediate phases and failure in the final stage, or whether or not the dialogic process between the designer and AI tool has structural biases because such models were created thus way.

It is stated in this paper that the positional repositioning of visualization offered by generative AI, as it shifts to the terminal phase of the design process and is posited in the middle of the process, increases, rather than nullifies, the constitutive agency of this very medium. The integration of a high-fidelity device in the location of open-ended conceptual discovery exposes the creatively most consequential stage of designing to an aesthetical logic that the designer has not just written but has no capability of scrutinising. The paper utilizes the theory of the apparatus developed by Flusser and the current elaborations by Irrgang to argue that it is the realization of such a causal relationship between the point at which the visualization currently functions and why that point is important that is the way to use these tools with a critical consciousness as opposed to naive fluency.

2. Research approach

This paper is a theoretical and literature-based critical analysis. It is carried through three registers. The former is empirical and is an academic review of peer-reviewed studies on generative AI in architectural visualization that determine the point of entry and usage of these tools in the design process [1-3]. The second one is theoretical, the use of Flusser apparatus theory as developed by Irrgang to prompt-response setup of AI image tools in addition to the analysis of the latent space of architectural instruments presented by del Campo. The third is analytical where a close reading of a design experiment where culturally specific vernacular prompts were tested as opposed to the output of generative model was used to provide a concrete example of the asymmetry of the apparatus that theoretical framework suggests. No empirical research was done; the experiment is an analytic case but not a source of data. This method concurs with the history of critical theory in design research where theoretical frameworks are evaluated and developed in relation to particular design context instead of controlled experimentation.

3. Visualization in architectural design: before generative AI

3.1. Representation as constitutive act

Architectural visualization has long been understood as a post-design instrument deployed after fundamental design decisions have been made. In this traditional framing, the rendering conveys a complete scheme to client and regulators but never takes part in designing. This realization has never been full. Architectural representations, as Christenson puts it, are process representations that

DOI: 10.54254/2753-7064/2026.32721

are constitutive but not descriptive in nature [5]. In each kind of rendering, there is a choice of viewpoint, state of light, material resolution, and quality of atmosphere. The study of language of architectural rendering shows that these choices actively form an argument concerning what a design is, rather than a clear account of it [6]. A language of persuasion is and never a language of transcription.

When there is constitutive as opposed to transparent representation, the implication of cognition is obvious. Studies on information visualization and its impact on judgment and decision making prove that how information is presented visually affects the quality of decision and its direction, where cognitive load is overall altered, and evaluative criteria are reformulated to suit the user and he or she is often unaware of it [4]. A high-fidelity facade visualisation illuminated by warm afternoon light does not simply depict one of the design possibilities. It renders such an option both emotionally and cognitively present in a manner inaccessible to an unrendered section. The judgment that the visualization seems to be merely recording is also partly generated by the visualization, a fact which then achieves structural significance when the very mechanism that creates the visualization itself has its own aesthetics.

3.2. Design as dialogue and the uneven distribution of fidelity

Lawson, drawing on Schön's concept of reflection-in-action, establishes that design proceeds through a dialogic encounter between designer and material [7,8]. The designer has to intervene, the material has to react, and interpretation of such reaction provokes a new action. According to Goldschmidt [9], this is the dialectics of sketching where the designer shifts back and forth between reflective criticism and analogical reinterpretation. The drawing talks back. This dialogic role has always been fulfilled in the middle phases of design by sketchings and vague drawings. They are quick, inexpensive and epistemically philanthropic; their ambiguity actively seeks to be redecoded not shut down.

On the other hand, high-fidelity visualization has traditionally been restricted to such late phase services due to its prohibitive cost of production at an earlier stage. The architect could not do away with attempting several competing schemes just to experiment which was right. This produced an inequality in the dialogic ability of visualization. Instead, low-fidelity talk was offered all the way across, but high-fidelity visual discourse, that kind that uses atmospheric judgment, material intuition and experience anticipation, was to be economized all the way to the terminal. It is the state wherein generative AI interferes that is depicted in Figure 1 [2,3,7].

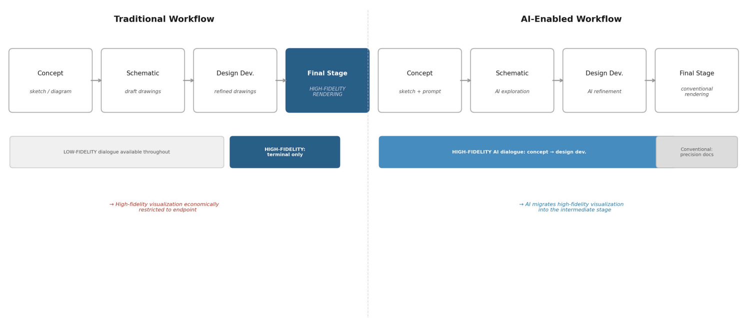

Figure 1. Positional repositioning of visualization in the design process: traditional workflow (left) vs. AI-enabled workflow (right) (picture credit: original)

DOI: 10.54254/2753-7064/2026.32721

4. Generative AI and the positional repositioning of visualization

4.1. Language as a new design input medium

Text-to-image and sketch-to-image models have placed a historically new state of affairs. Architectural form can now be generated directly using natural language as a medium. The research of Attentional Generative Adversarial Networks (AttnGAN) that was presented by Del Campo at the Taubman College reveals that a language-based model allows designers to question a space of architecture through the use of programmatic text, which directs an input of written design intent into a resulting visual form of space with no initial drawn form [10]. In his Harvard GSD thesis project Guida further extends it to multimodal diffusion models, which allows designers to wander through what he terms as a multimodal feedback loop of information generated by language and facilitated by AI tools, where language is itself a generative medium [11]. A study by Carnegie Mellon also shows that design sketching as a generative AI input opens up opportunities to ideation as well as presents novel issues regarding representational control and authorship [12].

Quick access interaction then the reflection-in-action pattern that Lawson and Schoen use to characterize feedback is recaptured behaviorally by a prompt-based interaction, except that now it is a structure of high-fidelity visual output, and not pencil drawing [7,8]. The architect is able to argue with atmospheric images, experimenting with material senses and space arrangements and commit to a scheme. Li et al. prove this with a concrete example of Tsinghua University workflow which has been shown to produce conceptual floorplans and 3D models out of simple sketches to facilitate rapid ideation and generate architectural output under controlled context based on textual description [13]. Figure 2 illustrates the cycle [7,8,14-16].

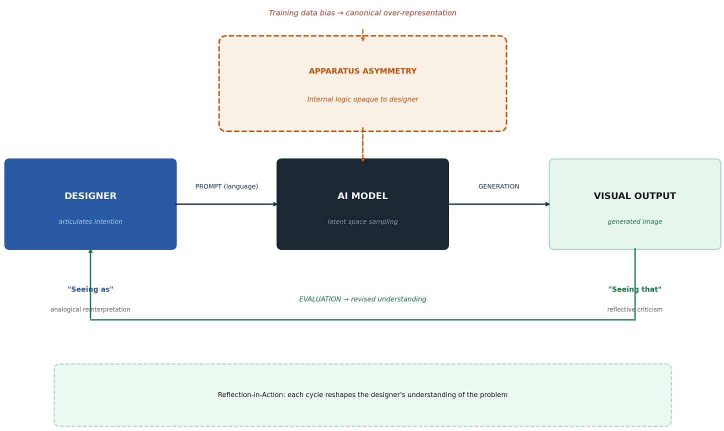

Figure 2. The prompt-response cycle as reflection-in-action, showing the epistemic asymmetry introduced by the apparatus (picture credit: original)

4.2. Visualization enters the design core, and its limits

Lee et al. show that text-to-image technology may be effectively used in exploratory mode since the earliest stage of design, and in the case of their study, more than 10,000 style-trained images were generated with significantly better visualization efficiency at the earliest stages of the project. Shi et al. go even further by integrating the intent to design the text with sketch-based input using ControlNet and LoRA fine-tuning and, as a result, allows comparing multiple schemes in real time in the conceptual phase. Table 1 presents a summary of the functional difference during the design

DOI: 10.54254/2753-7064/2026.32721

stages. This positional change should however be qualified. Ko et al. note that although AI tools can speed up the early concept development, architecture is presented with a situation quite different to the engineering sector where the performance criteria can be accurately measured. Each generated image is a probabilistic realization and not a deterministic description in such a way that generative AI is structurally unreliable at the end-point in which regulatory adherence necessitates rigidity. This difference is exactly the idea behind the positional repositioning concept. The implementation of AI has also shifted the point of control of high-fidelity visual dialogue to the middle of the process, but not to the end.

Table 1. Conventional rendering vs. AI-generated visualization across design stages

Design stage	Conventional rendering	AI-Generated visualization	Source
Early / conceptual	Low-fidelity sketches only; high-fidelity not produced	Text-to-image from initial phase; 10,000+ style-trained images; sketch-to-image via ControlNet	[2,12,13]
Intermediate / schematic	High cost prohibits multi-scheme comparison at full fidelity	Compound input (sketch + textual intent); real-time atmospheric comparison; co-creation workflows	[3,17,18]
Terminal / documentation	Deterministic, dimensionally accurate; primary mode for regulatory submission	Probabilistic output; cannot guarantee dimensions or compliance; structurally unreliable	[1,18,19]

Note: All sources verified in the project reference library.

5. The apparatus and the asymmetry of the dialogue

5.1. The latent space as black box interior

Del Campo and Manninger describe the latent space as a high-dimensional abstraction in which data points are mapped according to features learned during training, constituting a multidimensional domain of learned data patterns and hidden relationships [20]. In areas that require a generative model to create an image, it samples a probability distribution instead of creating an image based on acknowledged geometric principles. The preceding language-based GAN studies by Del Campo also exemplify that this process results in that models differentially form their ownally specific sensibility, which cannot be entirely controlled by the designer [10]. In a strict theoretical perspective, the interior of a black box consists of the latent space, which is visible at the output and input face, but opaque in between.

5.2. Flusser's apparatus theory and the case of vernacular prompting

Flusser is of the view that the camera is a device whose internal software generates symbolic constructs based on possibilities that the user can trigger but still not fully understand [14]. The photographer has dominion over what is input and has the choice of the output, but the transformative process within is obscure. According to Flusser, the control of a game is performed by the functionaries who do not possess competence over it. This is further extended by Irgang. The technical image is a projective apparatus that organises human imagination without the conscious understanding of the user, which is especially true in the relationship with or in AI generative systems the latent space of which codifies aesthetical possibilities that are not selected by the designer [15]. The HCI position by Weisz et al. attests to the fact that the opaceness of the generative

DOI: 10.54254/2753-7064/2026.32721

AI systems is inherently a challenge to the issue of user control, as the user cannot access the inner workings that govern the way the model will or will not generate something [21].

The case analysis is made practical with regards to the architectural implications as presented below. When a designer created a responsive facade to a culturally dense urban residential area, a location in the heritage of Chicano urbanism based around improvised material accretion, overlaying street-level activism, and the spatial aesthetic of rasquachismo, a Stable Diffusion model was created using words selected directly out of that vocabulary, such as layered surfaces, improvised assembly, mural-activated street wall, and informal canopy. Images obtained by the model were coherent and resolved in the atmosphere. All outputs however reverted to the formal grammar of the polished historic districts of Europe, with regularized masonry coursing, the rhythm of fenestration being regular, atmosphere lighting being cleansed and not in actual state of the site. It did not appear in the outputs of the site because the prompt used to capture the site was not able to convey it or because the model had not been trained on non-canonical architectural publications and databases of images and therefore had never discovered it. What the model is not able to notice in its training set, it also fails to produce in response to prompt. The system of dialogue between apparatus and designer was real, but the conditions upon which this dialogue was practiced had been preconstituted by the prejudices of that body of training.

5.3. Training data and historical bias

Del Campo identifies the theoretical ground of this failure [16]. Categorization of architectural form had always been a power exercise. This political aspect is shared by machine learning models. The training data is a historical conditioned selection that expresses the prejudices of archives and institutional databases that it has been compiled. There is overrepresentation of canonical buildings and underrepresentation or absence of vernacular, native and non-western buildings. Recognition is determined by classification and generation is determined by the recognition. This trend is supported by the strength of the systematic review of generative AI in architectural design by Yang et al. who note that, although AI systems approach substantially help to decrease the preliminary design time, their performance is predisposed by the representational biases of training corpora that largely indicates Western canonical backgrounds [18]. The architect who is experimenting with AI-generated alternatives is thus not examining the entire architectural space but a historically filtered version of it that has already been constrained in some way that is not perceivable at the interface level.

6. Conclusion

This paper has argued that the positional repositioning of visualization, its migration from the terminal stage into the operative middle of the design process, amplifies rather than neutralizes the constitutive agency of the medium. The reflection-in-action structural reflection is echoed in the prompt-generation-evaluation cycle, but via an apparatus that possesses an inner aesthetic logic the logic of which is the result of a latent space historically conditioned but opaque to the designer. This is concrete in the vernacular prompting case. Installing a high-fidelity device at location on open-ended conceptual probe does not only expedite the intents of the designer; it concerns them through a field of visual capability preset with the prejudice of the training corpus. This is not something that future model advancements will merely remove, it is an inherent property of the way that generative models encode and decode visual knowledge.

DOI: 10.54254/2753-7064/2026.32721

The awareness of the dual nature of AI visualization, broad scope, limited field is the condition of a critical fruitful application of such tools in an architectural practice. Further studies ought to explore how designers might create what Flusser has introduced as the capacity to play with the apparatus that goes against the aesthetic logic encoded in generative models without losing the power of exploration of generative models. The problem, as del Campo puts it, is to develop a form of intelligence that avoids the totalizing tendency of identification. To meet it, the structural, positional, and dialogic conditions, as well as the asymmetric ones, have to be mapped first in which the designer now works.

References

[1] Ko, J., Ajibefun, J. and Yan, W. (2025). Generative AI-Powered Parametric Modeling and BIM for Architectural Design and Visualization. Proceedings of the Design Society 5 (ICED25).

[2] Lee, J.-K., Yoo, Y. and Cha, S. H. (2024). Generative Early Architectural Visualizations: Incorporating Architect's Style-Trained Models. Journal of Computational Design and Engineering 11. DOI: 10.1093/jcde/qwae065.

[3] Shi, M., Seo, J., Cha, S. H., Xiao, B. and Chi, H.-L. (2024). Generative AI-Powered Architectural Exterior Conceptual Design Based on the Design Intent. Journal of Computational Design and Engineering 11. DOI: 10.1093/jcde/qwae077.

[4] Eberhard, K. (2023). The Effects of Visualization on Judgment and Decision-Making: A Systematic Literature Review. Management Review Quarterly 73, 167–216. DOI: 10.1007/s11301-021-00235-8.

[5] Christenson, M. (2019). Theories and Practices of Architectural Representation. Routledge.

[6] Bernardello, R. A. and Momolo, A. (2021). The Language of Rendering in Architectural Visualisations. In Mediati, D. et al. (eds.), Connettere: 42nd International Conference of Representation Disciplines Teachers. UID.

[7] Lawson, B. (2005). Design as Conversation and Perception. In How Designers Think (4th ed.). Architectural Press.

[8] Schön, D. A. (1983). The Reflective Practitioner. Basic Books.

[9] Goldschmidt, G. (1991). The Dialectics of Sketching. Creativity Research Journal 4.2, 123–143.

[10] Del Campo, M. (n.d.). Architecture, Language and AI: AttnGAN and Architecture Design. Taubman College of Architecture, University of Michigan.

[11] Guida, G. (2022). Multimodal Architecture: Applications of Language in a Machine Learning Aided Design Process (MArch II Thesis). Harvard GSD.

[12] Zhang, C., Wang, W., Pangaro, P., Martelaro, N. and Byrne, D. (2023). Generative Image AI Using Design Sketches as Input: Opportunities and Challenges. C&C '23. ACM.

[13] Li, P., Li, B. and Li, Z. (2023). Sketch-to-Architecture: Generative AI-Aided Architectural Design. In Chaine, R., Deng, Z. and Kim, M. H. (eds.), Pacific Graphics 2023.

[14] Flusser, V. (2000). Towards a Philosophy of Photography. Reaktion Books.

[15] Irrgang, D. (2023). Projective Imagination: Flusser's Concept of the Technical Image. Theory, Culture & Society. DOI: 10.1177/02632764231165244.

[16] Del Campo, M. (2023). Between Architectural Canon and Latent Space: The Problem of Classification in AI Models. Architectural Intelligence 2.1, 1–18.

[17] Verheijden, M. P. and Funk, M. (2023). Collaborative Diffusion: Boosting Designerly Co-Creation with Generative AI. CHI EA '23. ACM.

[18] Yang, Y., Li, Y., Bai, X., Zhang, W. and Chen, S. (2026). Research Progress and Frontier Trends in Generative AI in Architectural Design. Buildings (MDPI).

[19] Muller, M., et al. (2022). GenAICHI: Generative AI and HCI. CHI EA '22. ACM.

[20] Del Campo, M. and Manninger, S. (2024). Tectonics of the Latent Space. ACADIA 2024: Small Designing Change, 167–175.

[21] Weisz, J. D., He, J., Muller, M., Hoefer, G., Miles, R. and Geyer, W. (2024). Design Principles for Generative AI Applications. CHI '24. ACM.