OPEN ACCESS PRESS (OAP)

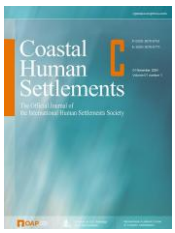
OpenAccessPress.com/Journal/CHS

Coastal Human Settlements

ISJN: CHS

Volume 2 Issue 2 - 2025

ISSN: 3078-8773 (P) / 3078-8765 (E)



Research Article | Open Access | DOI: Registering

Mapping Solar Radiation to Architectural Surface Textures: A Generative Workflow Using Stable Diffusion and ControlNet

Haoyi Chen *
School of Architecture and Art Design, Inner Mongolia University of Science & Technology
Haoyi.chen7@gmail.com

Ao Huang
Academy of Arts & Design, Tsinghua University
ha25@mails.tsinghua.edu.cn

Yiqun Wang
Laboratory of Landscape Development (LAND), School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
yiqunwang24@gmail.com, yiqun.wang@epfl.ch

*** Correspondence:** yiqunwang24@gmail.com, yiqun.wang@epfl.ch

ABSTRACT: Current computational design methodologies face a distinct dichotomy: parametric tools offer quantitative precision but lack morphological complexity, while generative AI models excel in aesthetic synthesis yet struggle with spatial fidelity. To address this gap, this study proposes a *Conditioned Morphogenesis* workflow that integrates a multimodal diffusion model, Stable Diffusion, with a spatial control network, ControlNet, to translate solar radiation heatmaps into functional architectural surface textures. Tree bark morphology is employed as a stress test for self-shading performance, and a Human-in-the-Loop methodology is used to calibrate the interaction between data granularity and generative parameters. Quantitative evaluation using Shannon Entropy, Gray-Level Co-occurrence Matrix (GLCM) Contrast, and Pearson correlation reveals a non-linear optimization relationship, indicating that a grid resolution of 0.6 m achieves optimal morphological stability, with the lowest entropy value of 4.89 bits while effectively balancing texture definition and artifact suppression. Furthermore, a multidimensional performance matrix identifies an optimal convergence state in which moderate data correlation (Pearson r = 0.31) coincides with maximum architectural depth (GLCM Contrast = 236.3), thereby avoiding the rigidity associated with over-controlled data mapping. This framework establishes a validated approach for data-driven texture generation and reframes the architect's role from direct form-making to parameter curation within generative design systems.

KEYWORDS: Generative AI; Conditioned Morphogenesis; Stable Diffusion; ControlNet; Solar Radiation Mapping

RECEIVED: 15 Oct 2025
REVIEWED: 26 Dec 2025
ACCEPTED: 28 Dec 2025
PUBLISHED: 22 Jan 2026

ACADEMIC EDITOR(S): Eugene Oks
OPERATING EDITOR(S): Kumar Shrestha
REVIEWER(S): Eugene Oks
CITATION: Registering
DOI: https://doi.org/10.65736/chs02021

1. Introduction

1.1. The Technical Divergence in Computational Design

Contemporary architectural design is characterized by a schism between analytical precision and morphological complexity. Traditional parametric tools enable strict geometric definitions based on environmental data but often fail to generate the organic intricacy required for advanced biological emulation [1]. Conversely, the recent proliferation of latent diffusion models (LDMs) has democratized high-fidelity image synthesis, yet these stochastic models historically lack the spatial rigor required for engineering applications [2]. When design intent demands strict adherence to site-specific conditions, such as solar radiation mapping, conventional text-to-image models fail to provide sufficient geometric accuracy. This disconnect constitutes a fundamental barrier in sustainable design: the difficulty of simultaneously harnessing the generative richness of AI and the data fidelity of parametric engineering.

1.2. The Evolution of Biomimetic Envelopes: From Analogue to Parametric

Biomimicry in architecture has evolved from stylistic formalism to functional emulation. Early biomimetic explorations, exemplified by the works of Gaudí and Frei Otto, relied on analogue form-finding methods to mimic natural structural logic. These approaches, while aesthetically compelling, lacked quantitative environmental feedback. The advent of parametric design tools (e.g., Grasshopper/Rhino) marked a significant advance by allowing architects to directly link environmental data, such as solar radiation or wind pressure, to geometric parameters [3]. However, parametric biomimicry faces inherent limitations in modeling organic complexity. Generating intricate, non-repetitive textures, such as tree bark or coral, requires explicitly defined mathematical functions (e.g., Voronoi tessellations or reaction–diffusion equations). As complexity increases, computational overhead becomes prohibitive, and the resulting geometries often appear rigid or overly rationalized, failing to capture the stochastic nuance of biological systems [4]. Consequently, a critical gap persists: how to generate biologically complex, data-driven morphologies without the bottleneck of explicit parametric coding.

1.3. Generative AI in Architecture: GANs vs. Diffusion Models

The integration of artificial intelligence into architectural workflows has offered new avenues for overcoming these parametric limitations. Over the past five years, Generative Adversarial Networks (GANs), such as Pix2Pix and CycleGAN, have been the dominant tools for style transfer and floor plan generation [5]. Although capable of learning stylistic distributions, GANs notoriously suffer from mode collapse and training instability, often producing blurry or structurally incoherent results

when constrained by strict engineering data [6]. Since 2022, Denoising Diffusion Probabilistic Models (DDPMs), represented by Midjourney and Stable Diffusion, have largely superseded GANs due to their superior image fidelity and output diversity. Recent studies have demonstrated the potential of these models in early-stage ideation and material synthesis [7]. However, a critical deficiency remains in the existing literature: most AI-aided design workflows prioritize visual aesthetics over the integration of performance data. As a result, AI is frequently employed as a "sketching tool" rather than a "simulation driver," leading to generative hallucinations that disregard physical laws or site-specific constraints [8]. This study addresses this gap by proposing a workflow that enforces strict environmental conditioning within the diffusion process.

1.4. Bridging the Gap with Controllable Generations

To resolve this technical divergence, this research introduces a workflow that integrates Stable Diffusion with ControlNet. Unlike purely text-based generation, this approach anchors the stochastic nature of AI to deterministic environmental constraints [8]. By treating a solar radiation heatmap not merely as a visual reference but as a strict spatial condition, the proposed method transforms the generative process from random aesthetic exploration into a controlled mechanism for performance-based texture mapping. This approach aligns with the principles of Performance-Based Design (PBD), in which quantitative targets guide qualitative form generation [3].

1.5. Justification of the Case Study: Bark Texture as a Stress Test

This study selects tree bark texture as the primary case study for specific methodological reasons rather than purely aesthetic biomimicry. In façade design, bark represents a morphology characterized by complex depth variation and significant self-shading potential [9]. It features intricate ridges and fissures that are functionally analogous to high-performance building skins designed for solar regulation [10]. Consequently, generating bark texture serves as an ideal stress test for the proposed workflow. If the AI can accurately translate the grayscale gradients of a radiation map into corresponding variations in depth and density within a bark-like structure, this demonstrates the system's capacity to handle complex, non-linear performance data, thereby shifting biomimicry from visual imitation toward functional simulation.

2. Materials and Methods

2.1. Methodological Framework Overview

The research proposes a systematic, data-driven workflow (Figure 1) designed to translate environmental performance data into two-dimensional architectural surface textures. The framework operates through three distinct stages to bridge the gap between analytical simulation and morphological generation. First, high-fidelity solar radiation data are simulated and structured into a pixel-based format. Second, these data are processed through a conditional generative AI pipeline to synthesize biomimetic morphologies. Third, the output is subjected to a quantitative evaluation framework to ensure statistical alignment between the environmental input and the generated texture.

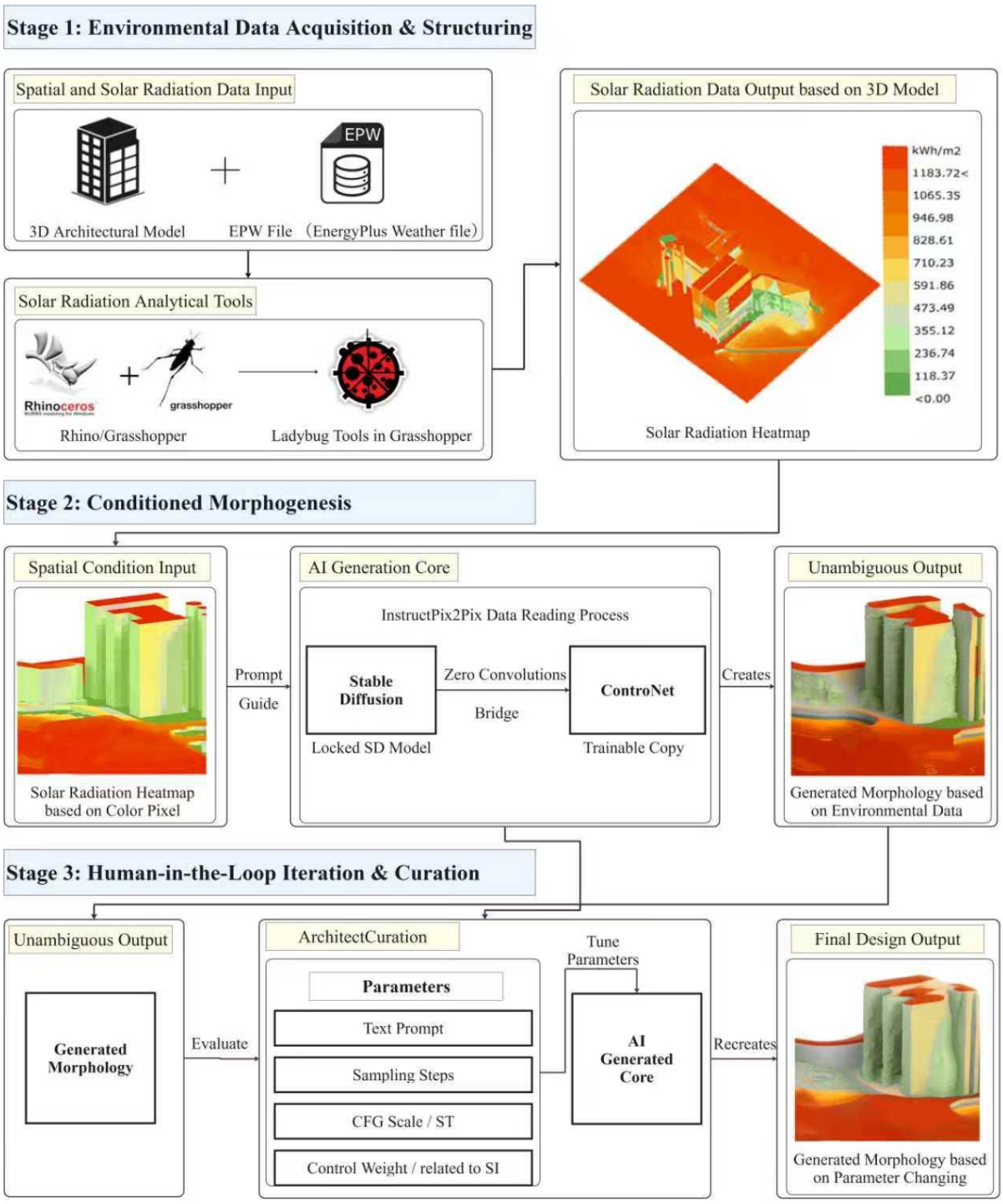


Figure 1. The Data-Driven Morphogenesis Workflow.

2.2. Stage One: Environmental Data Simulation

To ensure the physical validity of the input data, this study employs Ladybug Tools within the Rhino–Grasshopper environment. This platform was selected because it integrates validated simulation engines, specifically EnergyPlus and Radiance, thereby ensuring scientific accuracy in environmental analysis. The data acquisition process begins with the import of an EnergyPlus Weather (EPW) file for the target site, which provides hourly meteorological data, including temperature and solar irradiance. Subsequently, a geometric site model is constructed to serve as the spatial framework for analysis. The simulation calculates cumulative solar radiation intensity across the building façade over an annual cycle. These radiation values are then

normalized and converted into a pixel-based heatmap, in which color gradients represent varying levels of solar exposure. This image serves as the primary spatial condition for the subsequent AI-based generation phase.

2.3. Stage Two: Conditioned Morphogenesis

The core generative engine utilizes the Stable Diffusion v1.5 architecture coupled with the ControlNet extension. Unlike unconstrained text-to-image models that rely solely on semantic prompts, this configuration enables precise spatial conditioning [8].

The technical innovation of ControlNet lies in its ability to inject spatial conditions without corrupting the pre-trained diffusion model. It employs a "locked copy" of the large-scale diffusion model to preserve its semantic knowledge and a "trainable copy" to learn the specific condition. These two branches are connected via Zero Convolution layers, namely 1×1 convolutions initialized with both weight and bias set to zero. Mathematically, for a feature map x, the output of a zero convolution layer $Z(x; \Theta)$ is initially zero. This ensures that, at the beginning of training, ControlNet does not introduce any perturbation to the Stable Diffusion backbone, effectively causing the model to behave identically to the original architecture. As training progresses ($\Theta \neq 0$), the zero convolution layers gradually learn to transmit spatial features from the solar radiation heatmap into the latent space of the diffusion model. This mechanism is critical for the present study, as it ensures that the generated architectural textures are structurally guided by environmental data while retaining the high-fidelity textural details learned by the base model.

Table 1. Specification of Computational Environment and Generation Parameters

Parameter Category	Specific Setting / Value	Technical Justification
Base Model	Stable Diffusion v1.5 (Checkpoint: v1-5-pruned-emaonly.ckpt)	Selected for its proven stability in texture synthesis and compatibility with ControlNet v1.1.
Control Network	ControlNet v1.1 (Model: control_v11p_sd15_inpaint)	Chosen to strictly enforce spatial correspondence between the input radiation heatmap and the generated texture.
Sampling Method	Euler a (Euler Ancestral)	Utilized for its rapid convergence speed, allowing for efficient iterative testing during the prompt tuning phase.
Sampling Steps	Variable (Tested range: 20–110)	Adjusted to investigate the correlation between denoising iterations and the emergence of high-frequency texture details.
Guidance Scale	CFG Scale = 7.0	Maintained at a neutral value to balance prompt adherence with image coherence.
Input Resolution	Variable (0.3 m, 0.6 m, 1.2 m grid equivalents)	Varied to analyze the trade-off between data granularity and the occurrence of visual artifacts.
Seed	Fixed (e.g., 42) for comparative runs	Fixed (e.g., 42) for comparative runs

2.4. Stage Three: Human-in-the-Loop as a Filtering Mechanism

The workflow incorporates a Human-in-the-Loop (HITL) approach that functions as a critical quality control mechanism [11]. Rather than acting merely as artistic curators, human designers serve to filter out generative artifacts and logical inconsistencies. Raw outputs from the AI model often contain hallucinations or geometric distortions that do not align with architectural logic. Through iterative calibration, the designer adjusts key parameters, specifically sampling steps and the prompt guidance ratio (ST), to ensure that the emerging texture maintains structural integrity.

2.5. Quantitative Evaluation Framework

To transition from subjective aesthetic judgment to objective performance validation, this study establishes a multi-dimensional evaluation framework. Given the generative nature of the workflow, a single metric is insufficient to capture the nuanced trade-offs between data adherence and morphological richness. Therefore, we employ a methodological triangulation approach, utilizing three distinct computer vision metrics to quantify the Information Complexity, Architectural Depth, and Data Fidelity of the generated surfaces.

2.5.1. Shannon Entropy for Information Density

To detect potential noise artifacts and evaluate image complexity, we utilize Shannon Entropy (H) [12]. This measures the randomness of the pixel intensity distribution and helps identify over-processed images:

$$H = -\sum_{\{k\}p_k} \backslash log_{2(p_k)}$$

where p_k is the probability of intensity level k. While adequate entropy signifies rich detail, excessively high entropy values, particularly at high resolutions, often suggest the presence of visual noise or fragmentation rather than useful architectural detail.

2.5.2. GLCM Contrast for Texture Depth

To quantify the morphological definition of the generated surfaces, this study employs the Gray-Level Co-occurrence Matrix (GLCM) contrast [13]. In digital image processing, GLCM calculates the probability of adjacent pixel pairs exhibiting specific gray-level values. The contrast statistic specifically measures local variations within the gray-level co-occurrence matrix. In the context of this architectural study, GLCM contrast serves as a direct computational proxy for physical texture depth. High contrast values indicate sharp transitions between light and dark pixels, which, when interpreted as a displacement map, correspond to deep ridges and steep fissures in the bark texture. Conversely, low contrast values imply a smooth or relatively flat surface. Therefore, maximizing GLCM contrast, within stable limits, is equivalent to maximizing the self-shading potential of the façade, as deeper textures provide greater solar obstruction at oblique angles. It is calculated as follows:

$$\text{Contrast} = \sum_{\{i,j\}|i-j|}^{2} p(i,j)$$

where $\boldsymbol{p(i,j)}$ is the probability of adjacent pixel values $\boldsymbol{i}$ and $\boldsymbol{j}$. A higher contrast value indicates deeper ridges and clearer texture definition, representing a more developed morphological structure.

2.5.3. Pearson Correlation for Data Fidelity

To validate the accuracy of data mapping, the Pearson correlation coefficient (r) is calculated between the input radiation heatmap (x) and the generated texture density (y). This metric is critical for establishing the data-driven nature of the proposed workflow:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \overline{y})}{\sqrt{(\sum(x_i - \bar{x})^2)\sqrt{(\sum(y_i - \overline{y})^2)}}}$$

3. Case Study Analysis: Generating Form from Solar Radiation Data

To validate the proposed workflow, a typical high-rise office building in Baotou, China (Cold Climate Zone II), was selected as the experimental subject. The south-facing façade, characterized by significant variation in solar exposure, served as the primary testbed. Unlike controlled laboratory conditions, this site exhibits a complex gradient of radiation intensities, ranging from deep shading (0 kWh/m²) to intense exposure (>1100 kWh/m²), thereby providing a rigorous context for testing the AI's capacity to differentiate and map textures responsively. The following subsections present a three-phase experimental analysis focusing on data granularity, temporal generation steps, and parameter calibration.

3.1. The Impact of Data Granularity on Morphological Precision

3.1.1. Rationale for Resolution Selection

It is critical to establish the architectural logic underlying the selected grid resolutions. Accordingly, the study determines resolution values based on standard tectonic scales in façade construction. A low resolution of 1.2 m × 1.2 m corresponds to the standard module of a unitized curtain wall panel and is used to evaluate whether the AI can generate macro-patterns suitable for large-scale fabrication. A medium resolution of 0.6 m × 0.6 m represents the typical subdivision of façade sub-frames or louvers, balancing visual continuity with manufacturing feasibility. A high resolution of 0.3 m × 0.3 m approaches the scale of detailed material texture; the working hypothesis is that higher resolution would lead to superior detail, an assumption that this experiment is designed to critically examine.

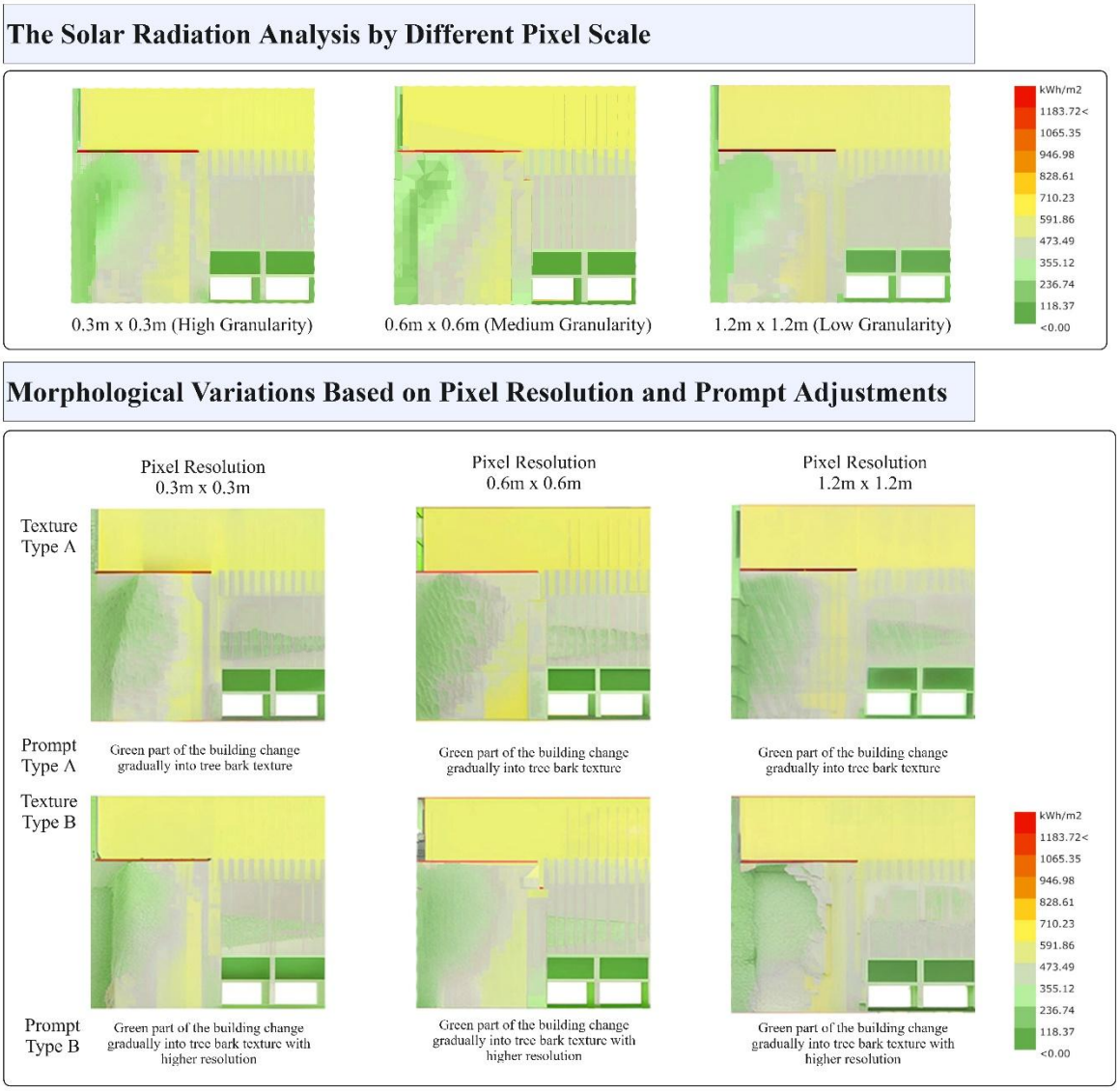


Figure 2. Comparative Analysis of Morphological Outputs Across Varying Data Granularity and Text Prompts**.**

3.1.2. Morphological and Quantitative Analysis

The systematic comparison of these resolutions (Figure 2) reveals a complex, non-linear relationship between data input and morphological output. Visual inspection further indicates a trade-off between textural detail and structural coherence. The high-resolution input of 0.3 m produces rich detail but also introduces significant visual artifacts and fragmentation. Conversely, the low-resolution input of 1.2 m lacks definition, resulting in blurred boundaries and pixelated blocks that fail to convey the

intended biomimetic depth. By contrast, the 0.6 m resolution yields the most structurally coherent morphology, effectively balancing texture definition with surface continuity.

To rigorously validate these qualitative observations, Shannon entropy (H) was calculated for each output. As defined in Section 2.5.2, entropy serves as a measure of information disorder; accordingly, a lower entropy value within a complex system typically indicates higher structural order and stability. As illustrated in the quantitative analysis chart (Figure 3), the results exhibit a distinctive U-shaped optimization curve rather than a linear progression.

The analysis yields several specific findings. The 0.3 m resolution produces the highest entropy value, at 5.5069 bits, quantitatively confirming the presence of excessive high-frequency noise, in which pixel-level grain is interpreted by the generative model as disjointed morphological features. The 1.2 m resolution also exhibits an elevated entropy of 5.2479 bits, indicating that the coarse input introduces aliasing artifacts and sharp pixelation boundaries that are interpreted as disorder rather than coherent geometry. Crucially, the 0.6 m resolution achieves the lowest entropy value of 4.8965 bits, forming an "entropy valley" that mathematically demonstrates this resolution as the state of maximum morphological stability.

Consequently, the quantitative data support the conclusion that 0.6 m represents the optimal granularity. This resolution effectively minimizes both the stochastic noise observed at higher resolutions and the aliasing artifacts associated with lower resolutions, ensuring that the generated texture maintains a high degree of architectural coherence.

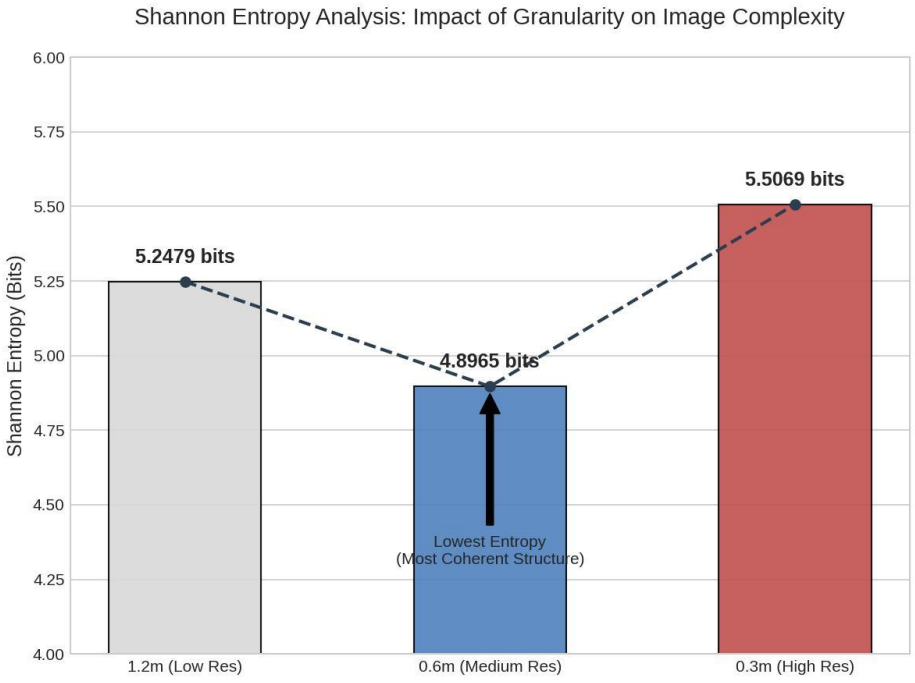


Figure 3. Quantitative Analysis of Image Complexity using Shannon Entropy

3.2. Iterative Exploration Through Sampling Steps

3.2.1 The Logic of Denoising Schedules

The sampling steps parameter governs the discretization of the reverse diffusion process. In standard text-to-image tasks, 20–30 steps are typically sufficient for convergence. However, for architectural morphogenesis that requires volumetric depth and spatial articulation, the optimal number of steps remains undefined. Accordingly, this experiment establishes a testing range from 10 to 110 steps to probe model behavior. The working hypothesis is that while early steps primarily establish global composition (layout), extended steps are necessary to develop the volumetric texture required for effective self-shading.

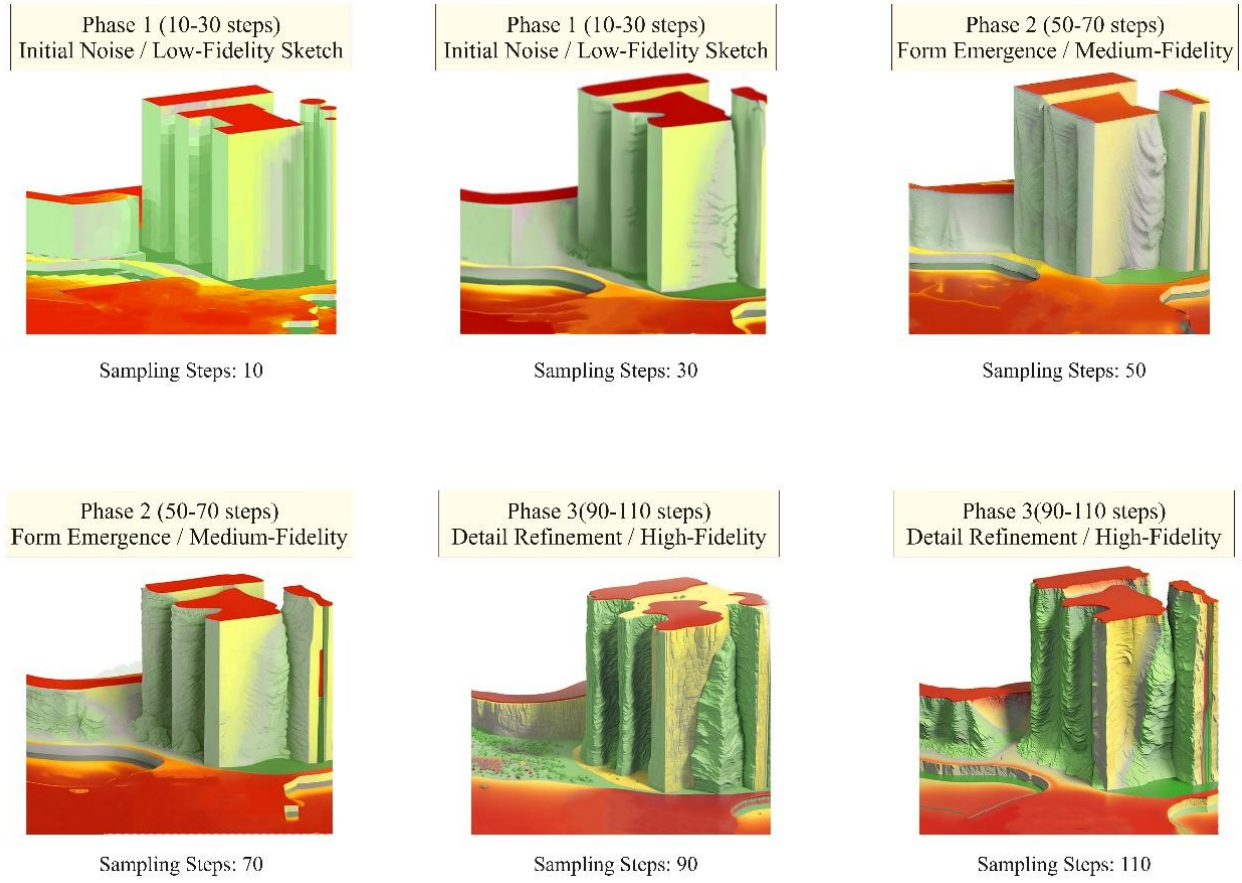


Figure 4. Morphological Evolution Guided by Sampling Steps.

3.2.2. Analyzing Morphological Convergence

The visual progression shown in Figure 4 demonstrates a clear transition from abstract noise to defined geometry. At lower iteration levels, between 10 and 30 steps, the morphology remains in a nascent stage. The resulting surfaces appear relatively smooth and lack the characteristic roughness of the intended biomimetic reference. As the number of steps increases to the 50–70 range, the form becomes significantly more articulated. Deep ridges and fissures begin to emerge, gradually resembling the self-shading properties observed in natural bark.

To objectively quantify this evolution, GLCM contrast was calculated for each stage. As described in Section 2.5.1, this metric measures local intensity variations and serves as a computational proxy for texture depth. As shown in the quantitative trend analysis (Figure 5), the contrast values follow a distinct multi-stage growth curve.

The analysis identifies three critical phases in the generative process. During the latent phase (10–30 steps), contrast values remain low, ranging between 22.2 and 19.0, quantitatively indicating that insufficient iterations result in a lack of textural definition. In the emergence phase (50–70 steps), the metric exhibits a marked increase, rising from 65.6 to 93.5, which reflects the rapid accumulation of morphological detail and the establishment of the texture's geometric structure. In the intensification phase (90–110 steps), rather than reaching a plateau, the contrast value continues to accelerate, reaching 139.6 and peaking at 236.3, demonstrating that higher sampling steps actively deepen the texture. This statistical evidence clarifies the convergence behavior: while the overall geometric form may stabilize at approximately 70 steps, textural depth continues to intensify up to 110 steps. Consequently, higher sampling steps are methodologically justified when the objective is to achieve the maximum biomimetic depth required for effective self-shading.

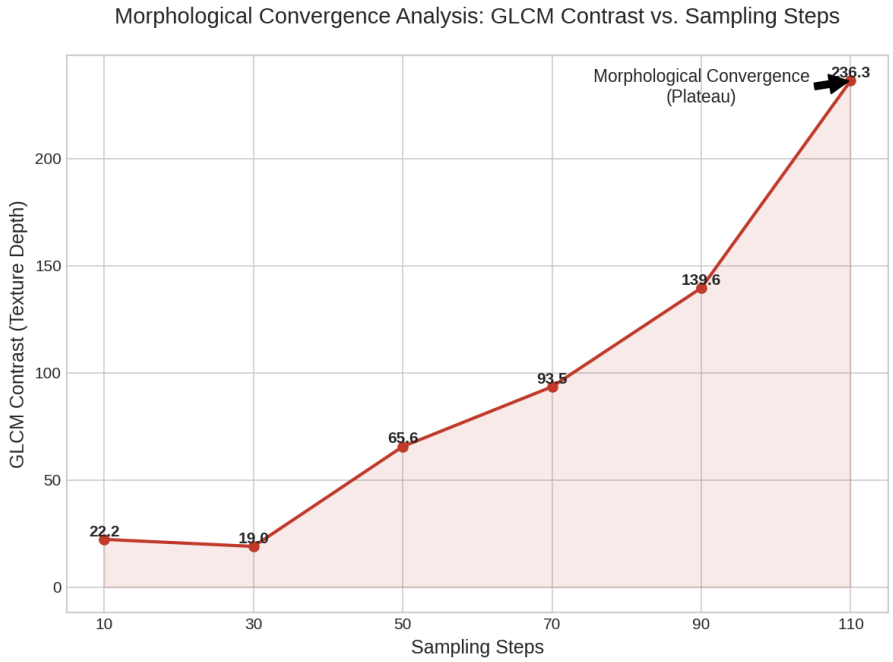


Figure 5. Quantitative Analysis of Morphological Convergence

3.3. Calibration of Generative Control and Verification of Data Fidelity

3.3.1. Defining the Parameter Search Space

The final stage of the analysis focuses on the precise calibration of generative parameters to ensure alignment between environmental inputs and architectural outputs. The primary challenge lies in balancing the creative interpretation driven by the text prompt with the rigid spatial constraints imposed by the solar radiation heatmap. This balance is governed by two key parameters: the prompt guidance ratio (ST) and the control strength (SI).

The visual matrix in Figure 6 systematically illustrates the morphological impact of these variables. The prompt guidance ratio (ST) determines how strictly the model adheres to the semantic description. Experimental results indicate that an ST value of approximately 8.0 provides an effective equilibrium. When ST is set excessively high, such as 20.0, the model tends to overfit the text description, producing abstract forms that disregard the input geometry. Conversely, a low ST value yields generic outputs that lack the specific biomimetic characteristics defined in the prompt. Similarly, the control strength (SI) governs the degree of influence exerted by the input heatmap. The results show that a value range between 1.4 and 1.6 effectively constrains generation to the heatmap's zoning while preserving sufficient freedom for texture synthesis.

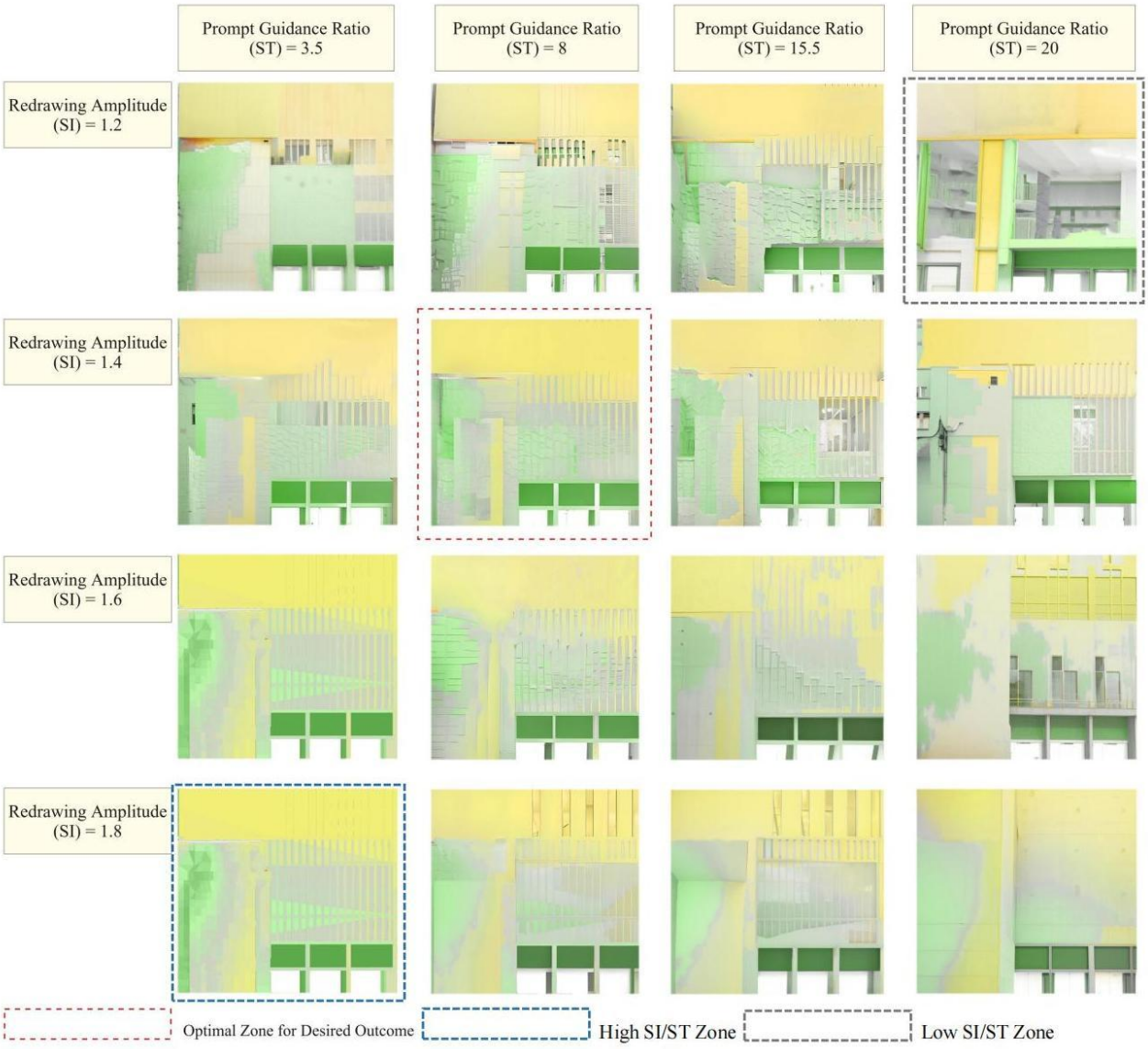


Figure 6. Parameter Calibration Matrix Showing the Impact of ST and SI on Morphological Generation.

3.3.2. Quantitative Verification: The Correlation Paradox

To rigorously validate the selection of the optimal parameter set, a comparative statistical analysis was conducted across three distinct generative scenarios: under-controlled (low SI/ST), optimal (red box), and over-controlled (high SI/ST). Figure 7 presents the corresponding scatter plot distributions and Pearson correlation coefficients (r) for each case.

The comparative analysis yields several critical findings across the three generative scenarios. In the over-controlled case (high SI/ST), the model achieves the highest correlation ($r = 0.50$), which mathematically indicates the strictest adherence to the input data; however, visual inspection shows that this results in a direct cloning of the heatmap gradients. The excessive linearity suppresses high-frequency textural details, producing a flattened surface that fails to achieve the intended biomimetic depth. In contrast, the under-controlled case (low SI/ST) exhibits a lower correlation ($r = 0.27$), with a more diffuse scatter plot distribution, indicating that texture generation is driven primarily by stochastic noise rather than environmental guidance, leading to reduced data fidelity. By comparison, the optimal balance achieved with the selected parameter set yields a moderate correlation ($r = 0.31$), which represents an effective "sweet spot." As illustrated in the central plot of Figure 7, this configuration maintains a positive linear trend that respects macro-level environmental zoning while preserving the degree of variance necessary for meaningful morphological expression.

The results of this calibration, validated through Pearson correlation analysis (Figure 7), provide a critical insight. A comparison between the over-controlled case and the optimal balance shows that enforcing a high correlation (r = 0.50, high SI) compromises architectural quality by reducing GLCM contrast and flattening geometric expression. This finding supports the premise that rigid data mapping conflicts with generative synthesis. The selected parameter set (ST ≈ 8.0, SI ≈ 1.4–1.6) therefore represents a negotiated equilibrium, achieving a moderate correlation (r = 0.31) that is sufficiently data-driven while retaining the variance necessary for tectonic architectural expression.

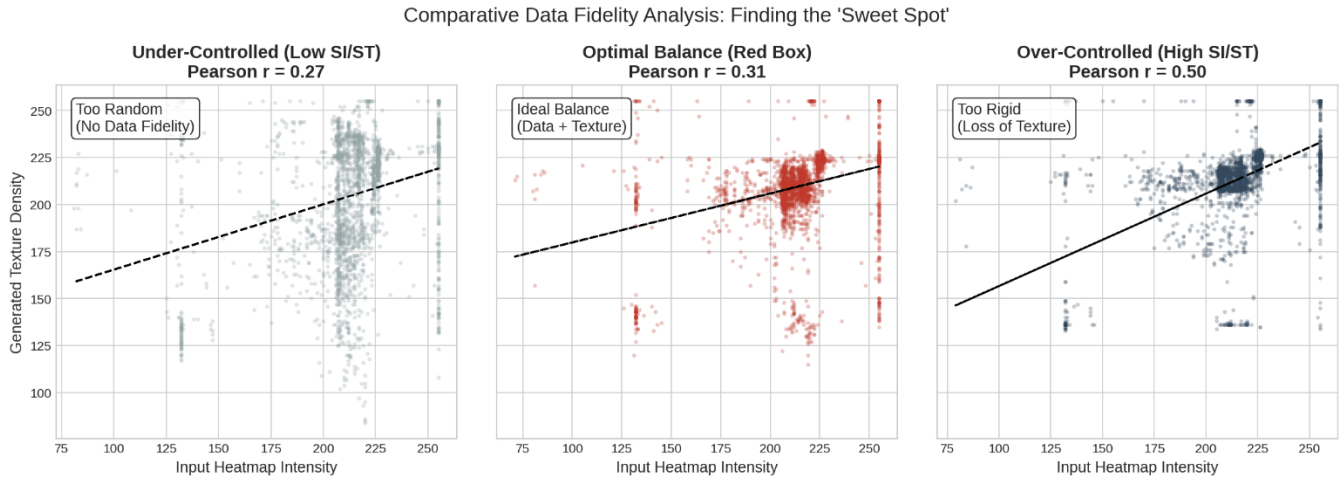


Figure 7. Comparative Data Fidelity Analysis.

4. Discussion

4.1. The Generative Control Trilemma: A Systemic Conflict

The core theoretical contribution of this study lies in identifying a structural trade-off inherent in AI-driven morphogenesis, which we term the Generative Control Trilemma. As illustrated in the Multi-Dimensional Performance Optimization Matrix, the generative process is constrained by three orthogonal dimensions: strict adherence to input data (fidelity), the richness of generated geometry (depth), and the structural coherence of the resulting pattern (stability).

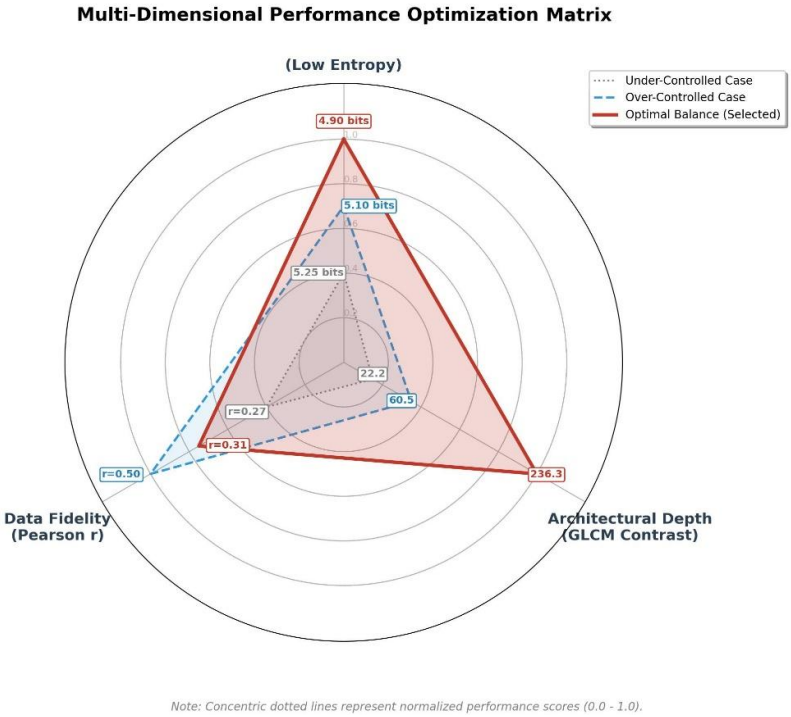


Figure 8. The Multi-Dimensional Performance Optimization Matrix. The radar chart reveals the structural trade-off between Data Fidelity (Pearson r), Architectural Depth (GLCM Contrast), and Morphological Stability (Entropy). The "Optimal Balance" (Red Triangle) maximizes the functional area, demonstrating that moderate correlation yields the highest morphological quality.

The radar chart demonstrates that these three objectives cannot be simultaneously maximized. A distinct inverse relationship exists between data fidelity and architectural depth. As the system prioritizes pixel-perfect alignment with the solar radiation map, it inevitably sacrifices the volumetric complexity required for effective biomimetic shading. This finding fundamentally challenges the linear logic of traditional parametric design, in which higher precision is invariably equated with better performance. In the context of diffusion models, optimization is therefore not a trajectory toward a single maximum, but rather a negotiation for the largest functional area within this triangular equilibrium.

4.2. The Paradox of Fidelity: Why r =0.50 Fails

The over-controlled case, represented by the blue dashed trajectory in Figure 8, illustrates the pitfalls of algorithmic rigidity. Although this scenario achieves the highest statistical correlation (Pearson r = 0.50), it paradoxically yields the poorest architectural performance. The GLCM contrast decreases markedly to 60.5, indicating a flattening of the surface texture. This outcome occurs because excessive control forces the Stable Diffusion model to function as a literal mapping mechanism rather than as a generative synthesizer. As a result, the model suppresses the generation of three-dimensional ridges and fissures, since such deviations would mathematically reduce the correlation score. Consequently, the output effectively becomes a two-dimensional grayscale reproduction of the heatmap—accurate but architecturally inert. This finding offers a critical warning for computational designers: within AI-driven workflows, statistical perfection, reflected by a high correlation value, can operate as a restrictive constraint that suppresses morphological emergence.

4.3. Functional Fidelity: The Optimization Equilibrium

The optimal balance strategy, represented by the red solid triangle, demonstrates the necessity of controlled stochasticity. By allowing the correlation coefficient to relax to r = 0.31, the system achieves a substantial increase in architectural depth, with

GLCM contrast rising to 236.3—nearly a fourfold improvement compared to the over-controlled case. We define this state as functional fidelity. Unlike geometric fidelity, which reproduces the formal appearance of data, functional fidelity captures its underlying intent. In this configuration, the AI respects the macro-level zoning of the solar radiation map by assigning denser textures to high-radiation areas, while retaining sufficient micro-level freedom to generate complex, self-shading geometries. Furthermore, this balance point coincides with the lowest entropy value (4.90 bits), suggesting that this specific parameter configuration aligns most naturally with the latent space's internal logic of bark texture. The red triangle therefore represents the only viable design strategy: it sacrifices a marginal degree of statistical linearity to achieve a substantial advancement in tectonic capability.

4.4. Scope Delineation: The Strategic Necessity of 2D Displacement Mapping

A critical distinction must be made regarding the dimensionality of the generated output. Although contemporary research increasingly explores direct three-dimensional generation, such as voxel- or point-cloud–based models, this study deliberately limits its scope to two-dimensional depth mapping for specific morphological reasons.

First, current direct-to-3D generative models often lack the resolution required to synthesize high-frequency surface details, such as the micro-fissures found in tree bark. By focusing on two-dimensional latent diffusion, this study leverages the model's superior capacity for texture synthesis to produce high-fidelity grayscale displacement maps. Within standard computational design workflows, these maps are not merely images but function as topographic data carriers that can subsequently be extruded into three-dimensional geometry.

Second, the two-dimensional–first approach aligns with the computational efficiency required in early-stage design. Generating and evaluating three-dimensional meshes incurs substantial computational overhead, thereby hindering rapid iteration. By decoupling texture generation (2D) from geometric extrusion (3D), the proposed workflow enables architects to explore hundreds of morphological variants in near real time before committing to computationally expensive three-dimensional reconstruction. Accordingly, the two-dimensional outputs of this study should be understood as a necessary intermediate data layer—a bridge between environmental analysis and final volumetric fabrication**.**

5. Conclusions

This study establishes a validated computational framework for Conditioned Morphogenesis, effectively bridging the persistent gap between quantitative environmental analysis and qualitative AI-driven generation. By systematically calibrating the interaction between Stable Diffusion and ControlNet, the research moves beyond subjective visual assessment to provide a statistically grounded evaluation of generative design.

First, the quantitative analysis identifies input resolution as a critical determinant of morphological stability. The results indicate that a grid resolution of 0.6 m corresponds to the optimal "entropy valley" (4.89 bits). This threshold effectively balances the suppression of aliasing artifacts inherent in coarse input data (1.2 m) with the mitigation of high-frequency stochastic noise observed at finer resolutions (0.3 m), thereby validating the need to align data granularity with the semantic scale of the latent space.

Second, the study elucidates a fundamental trade-off between data fidelity and architectural depth. Contrary to the prevailing assumption in parametric design that higher precision necessarily leads to superior performance, maximizing the Pearson correlation coefficient ($r = 0.50$) paradoxically results in a flattening of surface morphology, as indicated by a reduced GLCM contrast value of 60.5. The identified optimal balance strategy ($r \approx 0.31$) demonstrates that a moderate relaxation of strict data adherence is required. This state of functional fidelity enables the generative model to synthesize deep, self-shading structures

(GLCM contrast = 236.3) that capture the intent of the environmental data without being constrained to an isomorphic replication of its form.

Finally, this research provides a practical calibration framework for early-stage design by recommending specific hyperparameter settings—namely, sampling steps above 90 and control strength values between 1.4 and 1.6—to achieve structurally coherent results. Nevertheless, an important limitation must be acknowledged: the current outputs are restricted to two-dimensional depth maps. Future research will therefore focus on the volumetric reconstruction of these textures through displacement mapping onto three-dimensional mesh vertices. Subsequent work will further integrate computational fluid dynamics (CFD) simulations, using solvers such as OpenFOAM, to calculate convective heat transfer coefficients, thereby closing the loop between generative form-finding and rigorous thermodynamic verification.

Author Contributions: Conceptualization, Haoyi Chen; methodology, Haoyi Chen; software, Haoyi Chen and Ao Huang; validation, Haoyi Chen and Yiqun Wang; formal analysis, Haoyi Chen and Ao Huang; investigation, Haoyi Chen and Yiqun Wang; resources, Yiqun Wang; data curation, Haoyi Chen; writing—original draft preparation, Haoyi Chen; writing—review and editing, Yiqun Wang; visualization, Haoyichen; supervision, Yiqun Wang; project administration, Yiqun Wang; funding acquisition, Yiqun Wang. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data is unavailable due to privacy restrictions.

Acknowledgments: During the preparation of this work the author(s) used GPT-4o and Google Gemini to correct grammatical errors and optimize the language. After using these tools/services, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication. Authors gratefully acknowledge Bo Liu for his valuable support in the preparation of the diagrams presented in this paper.

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

References

1. Menges, A. Material computation: Higher integration in morphogenetic design. *Archit. Des.* **2012**, *82*(2), 14–21. https://doi.org/10.1002/ad.1374
2. Zhang, L.; Rao, A.; Agrawala, M. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, France, 2–6 October 2023; pp. 3836–3847. https://doi.org/10.1109/ICCV51070.2023.00355
3. Oxman, R. Performance-based design: Current practices and research issues. *Int. J. Archit. Comput.* **2008**, *6*(1), 1–17. https://doi.org/10.1260/147807708784640090
4. Carpo, M. *The Second Digital Turn: Design Beyond Intelligence*; MIT Press: Cambridge, MA, USA, 2017.
5. Chaillou, S. ArchiGAN: A generative stack for apartment building design. *NVIDIA Technical Blog*, 17 July 2019. Available online: https://developer.nvidia.com/blog/archigan-generative-stack-apartment-building-design/ (accessed on 12 December 2025).
6. Saxena, D.; Cao, J. Generative adversarial networks (GANs): Challenges, solutions, and future directions. *ACM Comput. Surv.* **2021**, *54*(3), 1–42. https://doi.org/10.1145/3446374
7. del Campo, M. Ontology of diffusion models: Tools, language and architecture design. In *Diffusions in Architecture: Artificial Intelligence and Image Generators*; del Campo, M., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2024; pp. 44–54. https://doi.org/10.1002/9781394191802
8. Zhang, L.; Rao, A.; Agrawala, M. Adding conditional control to text-to-image diffusion models. *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)* **2023**, 3836–3847. https://doi.org/10.1109/ICCV51070.2023.00355
9. Badarnah, L. Form follows environment: Biomimetic approaches to building envelope design for environmental adaptation. *Buildings* **2017**, *7*(2), 40. https://doi.org/10.3390/buildings7020040
10. Al-Obaidi, K.M.; Ismail, M.A.; Hussein, H.; Rahman, A.M.A. Biomimetic building skins: An adaptive approach. *Renew. Sustain. Energy Rev.* **2017**, *79*, 1472–1491. https://doi.org/10.1016/j.rser.2017.05.028
11. Leach, N. *Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects*; Bloomsbury Visual Arts: London, UK, 2022. https://doi.org/10.5040/9781350165557
12. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
13. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*(6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314