

# Symmetry-aware Texture Refinement for 3D Building Models via Massing Decomposition and Generative AI

Fan Xue<sup>1,\*</sup>, Yijie Wu<sup>2</sup>, Maosu Li<sup>3</sup>

<sup>1</sup> The University of Hong Kong, Pokfulam, Hong Kong (China SAR) - xuef@hku.hk

<sup>2</sup> Hong Kong Polytechnic University, Hunghom, Hong Kong (China SAR) - yijie.wu@polyu.edu.hk

<sup>3</sup> Hong Kong University of Science and Technology (Guangzhou), Guangdong, China – maosuli@hkust-gz.edu.cn

**Keywords:** Mesh texture, Massing decomposition, 3D building model, Generative AI, Symmetry, vision language model.

## Abstract

Three-dimensional (3D) building models with accurate geometry and realistic textures remain essential for city information modeling and digital twin applications. However, photogrammetric reconstructions consistently suffer from severe texture defects caused by occlusions, shadows, distortions, and projection errors. Existing approaches either rely on rigorous photometric optimization that demands topological correctness and multi-view imagery, or employ flexible AI-driven generation that leverages semantics but often lacks geometric constraints. This paper presents a novel hybrid framework that exploits architectural regularities—specifically massing decomposition and partial symmetries—to guide high-fidelity texture refinement. We first decompose building meshes into mass-aligned convex volumes using MorphCut. Textures are then reprojected onto these volumes, followed by Building Section Skeletons to pair symmetric facades and establish precise geometric correspondences. Finally, generative AI is applied using symmetry-aware constraints to achieve contextually accurate inpainting and correction. Pilot studies on three Hong Kong buildings demonstrate robust decomposition, faithful texture transfer, and effective defect mitigation, while revealing current limitations of zero-shot vision language models (VLMs) in preserving floor counts and structural regularity. The proposed symmetry-guided pipeline notably advances the reliable and semantically coherent reconstruction of textures for complex urban buildings.

## 1. Introduction

### 1.1 Background

Three-dimensional (3D) building models with accurate geometry, realistic texture, and rich semantics are fundamental to city information modeling (CIM) (Kamra et al., 2022; Xue et al., 2021). The 3D building models and CIM subsequently support modern urban planning, digital twin development, and immersive simulation environments (Biljecki et al., 2015; Li et al., 2026; Meng et al., 2026). Modern photogrammetry and LiDAR methods have made notable advancements in this domain (Jiang et al., 2021; Xue et al., 2021).

However, a primary challenge in all methods is the unsatisfactory texture quality of buildings, particularly for those with complex geometry. Such texture defects can arise from data acquisition, including occlusions, heavy shadows, distortions, and missing images. They can also result from reconstruction pipelines, such as geometric mesh errors, vertex misalignments, clutter disruptions, and projection errors across multiple images.

### 1.2 Existing methods

Texture refinement of 3D building models has been addressed in many studies. Broadly, there are two categories: photometric optimization and AI-driven generation. Photometric optimization mainly relies on classical photogrammetric principles. These methods iteratively adjust the input mesh, including its geometry, to minimize photometric errors between the geometry and the source data. One example is Rothermel et al. (2020),

who adapted the texture consistency of photographs from high-resolution multi-view satellite imagery to refine geometric details. Iwaszczuk and Stilla (2017) focused on line-based matching and camera pose refinement to process thermal infrared images for uncertain 3D models with image sequences.

The second category is AI-driven generation, which applies modern deep learning, particularly diffusion models and Generative Adversarial Networks (GANs), to synthesize, repair, and enhance textures (Buyukdemircioglu et al., 2022). Diffusion models learn to reverse a gradual noising process and have proven effective for upsampling low-quality textures and inpainting missing regions for 3D texture refinement. One example is Elevate3D, which balanced quality and fidelity through variable noise handling (Ryu et al., 2025). Another example is TwinTex, which used a fine-tuned diffusion model to inpaint missing regions on abstracted building models (Xiong et al., 2023). GANs can translate semantic label mappings into photorealistic images. For instance, Shang et al. (2023) repaired texture defects using facade semantics such as “window”, “wall”, and “door”, along with architectural style images.

In general, photometric optimization methods are stringent about data inputs, including a coarse mesh model and multiple views of imagery, and rely less on semantics. The topological correctness of the coarse mesh enables iterative geometric and texture optimization. In contrast, AI-driven generation methods are more flexible in terms of data inputs but closely relate to building semantics, whether explicitly or implicitly. Some studies hybridized processes from both categories to take advantage of both topology (and geometry) and semantics for 3D building models. For example, Texture2LoD3 integrated ray-casting, image rectification, semantic segmentation, and data fusion to combine low-detail semantic building models (e.g., LoD1 or LoD2 CityGML) with geo-referenced panoramic street-level im-

\* Corresponding author

agery (Tang et al., 2025). The AI components were particularly effective in isolating facade regions and removing cars, pedestrians, and vegetation from textures.

### 1.3 Motivation

Architectural regularities include lines, planes, volumes, symmetries, and their combinations. Such regularities serve as ubiquitous and natural guides in texture refinement, particularly within photometric optimization. Examples are the line matching for uncertain 3D models in Iwaszczuk and Stilla (2017) and planar consistency for texture defragmentation in Liu et al. (2024). A recent advance for regular architectural volumes is the Building Section Skeletons (BSS) method, which pairs internal and partial symmetric facades (Wu et al., 2024). Based on the BSS, MorphCut decomposes convex volumes of building massing to align with architectural structures (Wu et al., 2025). These developments in processing regular lines and planes, particularly symmetric massing volumes, motivate this study and serve as its foundational preprocessing.

Another motive arises from emerging applications of generative AI (GenAI). GenAI can synthesize novel, photorealistic texture data to overcome data defects. It can inpaint missing regions, upscale blurry images to high fidelity, and generate contextually accurate details (e.g., windows and doors) from semantic prompts and guides. Even plausible “hallucinations” of entirely missing textures can supplement information creatively to produce complete, detailed, and visually consistent 3D building models.

## 2. Research Methods

The methods presented in this paper aim to leverage building symmetry and massing decomposition, as well as GenAI, for texture inpainting and correction of 3D building models. There are three steps, with one more preprocessing step. Section 2.1 emphasizes the preparatory step of decomposing building models into mass-aligned volumetric components. Section 2.2 details how to reproject photogrammetric textures onto the volumes. Section 2.3 pairs the symmetric skeletons and the projected textures. The final refinement step in Section 2.4 utilizes GenAI for high-fidelity, context-aware error mitigation.

### 2.1 Preprocessing of 3D building model decomposition by MorphCut

The preprocessing step involves segmenting the 3D building model into mass-aligned parts that capture the volumetric and semantic structure of architectural designs. Building mass is defined as the aggregate form composed of distinct elements, such as podiums and towers. The building mass serves as a critical prior for texture analysis, as it correlates with material and stylistic consistency within components while allowing for variation across them.

Then, we employ MorphCut in (Wu et al., 2025) to decompose an input building into mass-aligned parts. The input is a 3D building model represented as a triangle mesh. MorphCut first processes the model by reconstructing its geometry, eliminating topological errors, and ensuring a watertight manifold surface.

The core of MorphCut’s approach is a best-first branch-and-bound (BBnB) procedure (Land and Doig, 2009), enhanced with key morphological characteristics of buildings. The BBnB

algorithm decomposes the building model into as few convex or near-convex parts as possible, ensuring that the resulting segments align closely with the building’s underlying massing structure through the integration of architectural morphology.

### 2.2 Step 1: Texture reprojection

After the decomposition, textures from the original photogrammetric mesh are reprojected onto the external facets of the decomposed parts. The resulting grouped textures serve as priors for subsequent texture refinement using GenAI.

We then implement this via a straightforward reprojection procedure. First, the decomposed model and the original photogrammetric mesh are aligned at their centroids and base elevations. Their orientations remain consistent, as MorphCut’s reconstruction preserves the original model’s orientation.

Next, rays are emitted from the external faces of the decomposed model along their respective face normals to intersect the photogrammetric mesh. Upon intersection, each ray retrieves color information from the corresponding location on the photogrammetric surface and transfers it back to the originating face. A higher sampling density on the external faces results in higher-resolution textures but typically incurs greater computational cost.

### 2.3 Step 2: Pairing BSS construction and texture

The final phase involves constructing BSS within each decomposed part and pairing symmetric facades to quantify texture inconsistencies, thus establishing geometric constraints for GenAI-based refinement. BSS, conceptualized as medial sheets approximating midpoints between parallel surfaces, extends traditional skeletonization techniques to architectural contexts, capturing partial symmetries reflecting common design regularities.

For each decomposed part, we use ODAS (Xue et al., 2019) to detect its primary vertical symmetry plane. ODAS formulates symmetry detection as a derivative-free optimization problem. With 1,000 iterations, it reliably produces precise symmetry planes with low root-mean-square deviation (RMSD). This detected plane serves as the primary symmetry plane for the part. Additionally, we compute a secondary vertical symmetry plane orthogonal to the primary one, as many convex building components exhibit cuboid-like geometry, possessing two orthogonal pairs of symmetric facades.

Once the primary and secondary symmetry planes are determined for each part, the faces parallel to these planes are projected orthogonally onto the respective symmetry planes to delineate their vertical extents. Since all parts are convex or near-convex, the resulting projection outlines are typically simple. For computational efficiency and simplicity, we approximate these outlines as rectangles. This allows us to pair the textures on opposing faces parallel to each symmetry plane. Consequently, each symmetry plane is associated with a corresponding pair of textures, which guide symmetry-aware texture refinement in subsequent GenAI processing.

### 2.4 Step 3: Texture refinement

The texture refinement task includes a shadow removal to correct anomalously dark areas and a subsequent cross-image enhancement to improve illumination consistency between the detected symmetric facade pairs.

The proposed vision-language model (VLM) refinement applies one standard prompt to achieve the task: “The uploaded image contains two opposite facade textures: the left half is one facade of a building, and the right half is the opposite facade. Remove shadows on the two facade textures, harmonize the luminance, structure, and color between the two halves. Output the results.” Note that the pair of textures is temporarily merged into one image for maximized compatibility of most VLMs.

A baseline process is employed using two manipulation algorithms in Fig. 1 in the CIE LAB (Lightness, a and b) color space (ISO/CIE 11664-4:2019). The baseline shadow removal applies a per-row multiplicative gain to the L channel relative to a natural brightness level, then corrects A and B color casts. Then the baseline unifies the L profiles of two paired texture images to their element-wise maximum and shifts the A and B channels toward their grand mean to correct white-balance discrepancies.

---

#### Algorithm 1: Per-image shadow removal

---

**Input:** Image

```

1 Convert Image to CIE  $\rightarrow L, A, B$ ;
2  $natural_L \leftarrow$  60th percentile of  $smooth(rowMean(L))$ ;
3 foreach  $r$ -th row do
4    $gain[r] \leftarrow clip(scale\ L[r]$  to  $natural_L, 1, 3.5)$ ;
5 end
6  $L' \leftarrow clip(L \times gain, 0, 255)$ ;
7  $A', B' \leftarrow$  correct color cast( $A, B$ );
8 return LAB2BGR( $L', A', B'$ );

```

---



---

#### Algorithm 2: Cross-image harmonization

---

**Input:** List of images

```

1 Convert all images to LAB;
2  $ref_L \leftarrow$  max envelope of all L profiles (smoothed);
3 foreach image do
4    $gain \leftarrow clip(ref_L / image\ L, 1, 4)$ ;
5    $L' \leftarrow clip(L \times gain, 0, 255)$ ;
6    $A', B' \leftarrow$  Harmonize image  $A, B$  to global means;
7   Save LAB2BGR( $L', A', B'$ );
8 end
9 return harmonized images;

```

---

Figure 1. Two-stage shadow removal and harmonization pipeline in CIELAB color space

The visual consistency between each paired texture is quantified by the Structural Similarity Measure (SSIM) (Wang et al., 2004) and the Root Mean Square Deviation (RMSD) of the Gaussian-smoothed  $L$ ,  $a$ , and  $b$  channel maps. The SSIM is calculated to compare local patterns of pixel intensities, normalized for luminance and contrast, to assess structural alignment. The three RMSD metrics, denoted as  $RMSD_L$ ,  $RMSD_a$ , and  $RMSD_b$ , are compared to suppress high-frequency structural content and to isolate the low-frequency illumination field.

### 3. Experimental Results

#### 3.1 Experimental settings

The proposed method was evaluated through pilot studies on three buildings in Hong Kong, as shown in Figure 2. The three test cases represent low-rise, mid-rise, and high-rise building structures from 3D building models and the iB1000 topographical 3D map from the Common Spatial Data Infrastructure of Hong Kong (Available at: <https://portal.csd.gov.hk>).



Figure 2. Three test cases of building models in this paper.

Figure 2 illustrates the 2D footprints and the 3D realistic meshes for these cases. The cases include one school and two residential buildings; The selected buildings exhibit clear massing structures and high symmetry. Notably, Case 2 is situated in a relatively dense environment, where textures may be incomplete or distorted due to limited scanning angles and occlusions. Case 3 is exceptionally tall and surrounded by other high-rise structures, which results in heavy shadowing on the textures.

#### 3.2 Results of Step 1

Figure 3 shows the building convex decomposition, the corresponding BSS, and the reprojected textured parts for the three cases. The MorphCut method successfully decomposed the original meshes into compact, mass-aligned convex parts that preserve the geometric integrity of the buildings. The extracted BSS captured the dominant structural symmetry and directions, particularly the two principal vertical symmetric planes in each case. The results demonstrate the robustness of the extraction process even for complex multi-tower configurations.

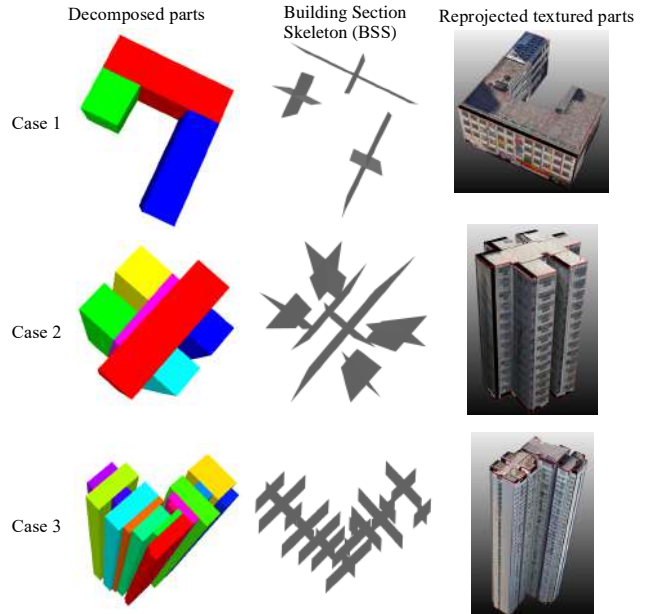


Figure 3. Decomposed parts, BSS, and reprojected textured parts of three test cases.

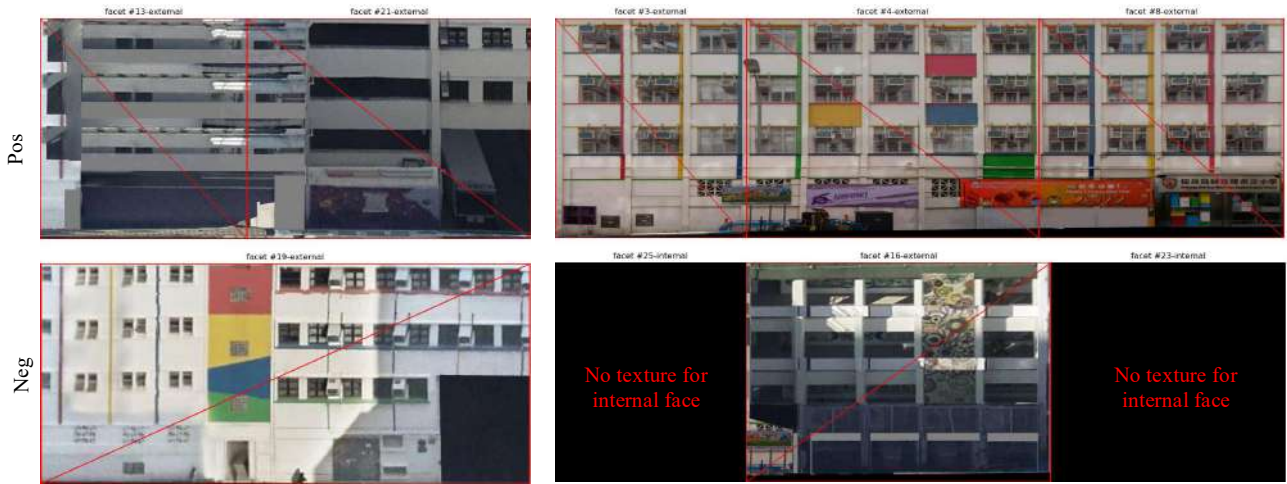


Figure 4. Paired textures for Case 1. Note: ‘Pos’ and ‘Neg’ rows denote the symmetric and asymmetric sides; The left pair indicates the blue part of Case 1 in Figure 3, while the right pair is associated with the red part.

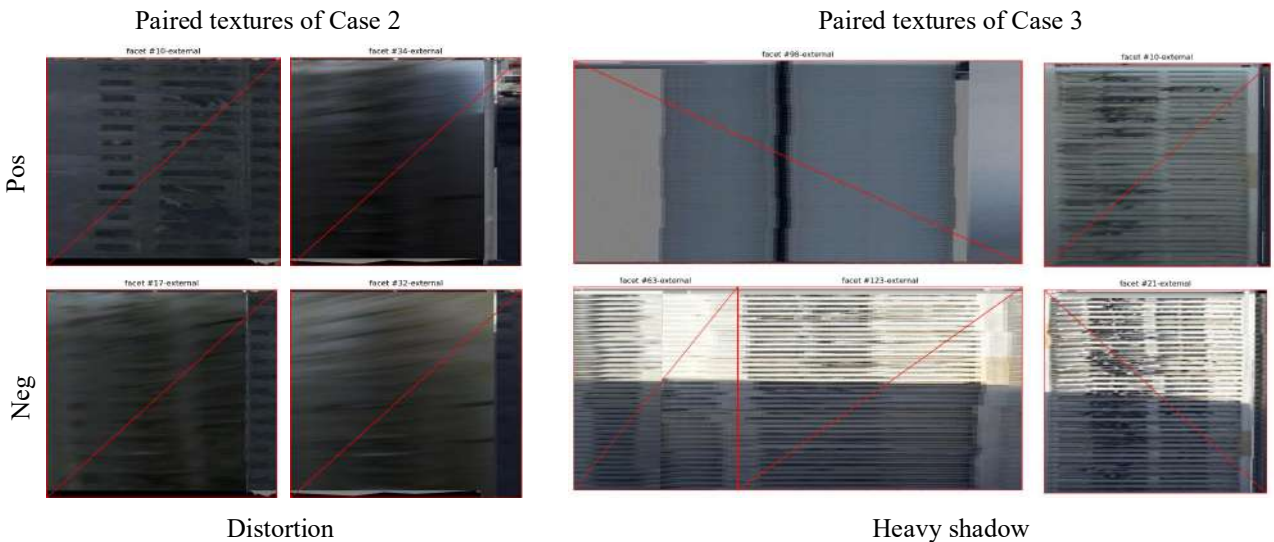


Figure 5. Distortion and heavy shadows observed in the projected textures of 3D volumetric parts in Cases 2 and 3.

The reprojected textured parts in Figure 3 further confirm the accuracy of the decomposition and reprojection steps, as the textures remain well aligned with the facades and roofs of the original photogrammetric meshes, maintaining visual realism and architectural consistency across the reconstructed models.

### 3.3 Results of Step 2

Figure 4 shows paired textures for decomposed parts in Case 1. ‘Pos’ and ‘Neg’ rows denote the positive and negative symmetric sides of each of the three volumetric parts in the first row of Figure 3, respectively. The textures on the left correspond to the blue part of Case 1 in Figure 3; those on the right correspond to the red part. Note that some facets on the symmetric planes of the red part lie inside the building; therefore, no textures were projected onto them in the bottom right texture image.

As the building structure in Case 1 is located in a relatively open environment, the 3D photogrammetry model had high-quality geometry and textures. Subsequently, the textures reprojected in Figure 4 also generally exhibited high quality, with a few

minor defects in shadow effects. Textures on the faces differ noticeably in appearance due to variations in lighting and viewpoint; however, the geometric structures and details, including the number of floors and associated windows, floor heights, remain consistent and accurate. This confirmed the accuracy of the symmetry detection and texture projection, which are required to upgrade the iB1000 polygon model’s Level of Detail (LOD) from 2 to 3.

More pronounced issues appear in the paired textures in Cases 2 and 3 in Figure 5. The primary reason was that the two input models were of lower quality, partly due to their complex geometries and surrounding environments in surveying and photogrammetry. In the left column of Figure 5, there are notable texture distortions for Case 2. We examined the environment and found that the sides of the building are in close proximity to neighboring structures, which obstruct many viewing angles during aerial or ground image acquisition. In Case 3, heavy shadows are cast on the main facades because the tall building is surrounded by other high-rise structures, which resulted in

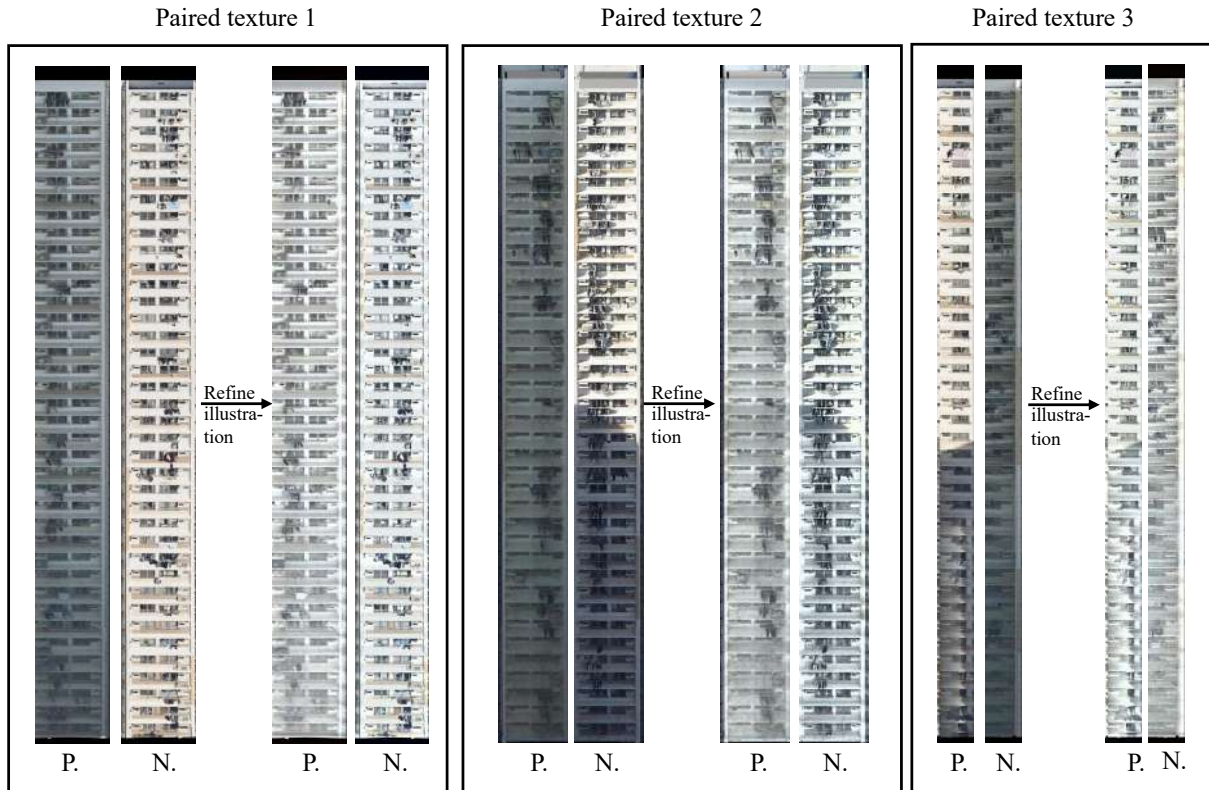


Figure 6. Results of the baseline CIE process for three decomposed parts in Case 3. A reduction in shadow and color-temperature discrepancies between symmetric pairs was observed.

pronounced shadowing and slight texture distortions near the lower levels. These defects call for rectification and refinement in the next step.

### 3.4 Results of Step 3

Case 3 with the heaviest shadow effects, as shown in Figure 5, was selected to demonstrate this step. Table 1 lists the quantitative metrics of visual consistency, i.e.,  $SSIM$ ,  $RMSD_L$ ,  $RMSD_a$ , and  $RMSD_b$ , of the refinement results.

Table 1. Visual consistency of facade textures before and after the illumination refinement procedure. (Note: Best values in bold.  $\uparrow$ : higher is better;  $\downarrow$ : lower is better.)

| Process       | $SSIM \uparrow$ | $RMSD_L \downarrow$ | $RMSD_a \downarrow$ | $RMSD_b \downarrow$ |
|---------------|-----------------|---------------------|---------------------|---------------------|
| Input pair    | 0.251           | 0.277               | 0.023               | 0.051               |
| Baseline CIE  | 0.189           | 0.153               | 0.0102              | 0.025               |
| Gemini2.5-Pro | 0.263           | 0.159               | 0.018               | 0.047               |
| GPT-Image2    | 0.250           | <b>0.096</b>        | <b>0.0097</b>       | <b>0.022</b>        |
| Grok4-Fast    | <b>0.273</b>    | 0.180               | 0.013               | 0.046               |
| QWen3.6-Plus  | 0.248           | 0.260               | 0.0317              | 0.052               |

The  $SSIM$  metric received losses, ranging from  $-0.10$  to  $-0.02$ . This mild drop in  $SSIM$  reflects contrast loss due to higher brightness or clipping at highlights (Wang et al., 2004; Pambrun and Noumeir, 2015). Nevertheless, this mild drop in  $SSIM$  due to brightness is acceptable because removing inconsistent lighting allows the underlying structural similarities between the symmetric facade pairs to be captured more accurately. In comparison, the  $RMSD$  values of the smoothed  $L$ ,  $a$ , and  $b$  consistently decrease across all tested pairs. The reduced  $RMSD$  values confirm that the illumination refinement stage effectively

suppressed heavy shadows and reduced brightness discrepancies between the detected symmetric facades of the high-rise building.

Figure 6 shows baseline CIE process’s refinement results for three pairs of facade textures for Case 3. It can be observed that the refinement process improves brightness and color in general, as well as suppresses shadow effects. Overall, the color-temperature discrepancy was also minimized. These effects echo the considerable decrease in  $RMSD$  of all three  $L$ ,  $a$ , and  $b$  channels as well as the decrease in  $SSIM$ .

Figure 7 shows example results of texture refinement by four VLMs for the same pairs of facades for Case 3. In contrast with the baseline CIE process, all VLMs’ results kept consistent in  $SSIM$ . Two VLMs, i.e., Gemini2.5-Pro and ChatGPT-Image2, achieved 0.263 and 0.273  $SSIM$  values, which were slightly greater than the input’s value at 0.251 but considerably higher than baseline’s 0.189. This indicates the color, contrast and structural consistencies were maintained by VLMs.

Furthermore, results from ChatGPT-Image2 in Figure 7 represent almost no shadows and a much more consistent illumination on all facades. That was why ChatGPT-Image2 returned the lowest  $RMSD$  values in the LAB channels. Light shadows can be observed in the results from Grok4-Fast and Qwen3.6-Plus.

Furthermore, some realistic details, such as clothes hanging outside windows below the 15th floor, were missing from Grok4’s results. This unfavorable moderation may be the main reason for achieving the highest  $SSIM$  value.

In terms of consistency across symmetry-guided parts, both ChatGPT-Image2 and Grok4-Fast achieved satisfactory levels of consist-

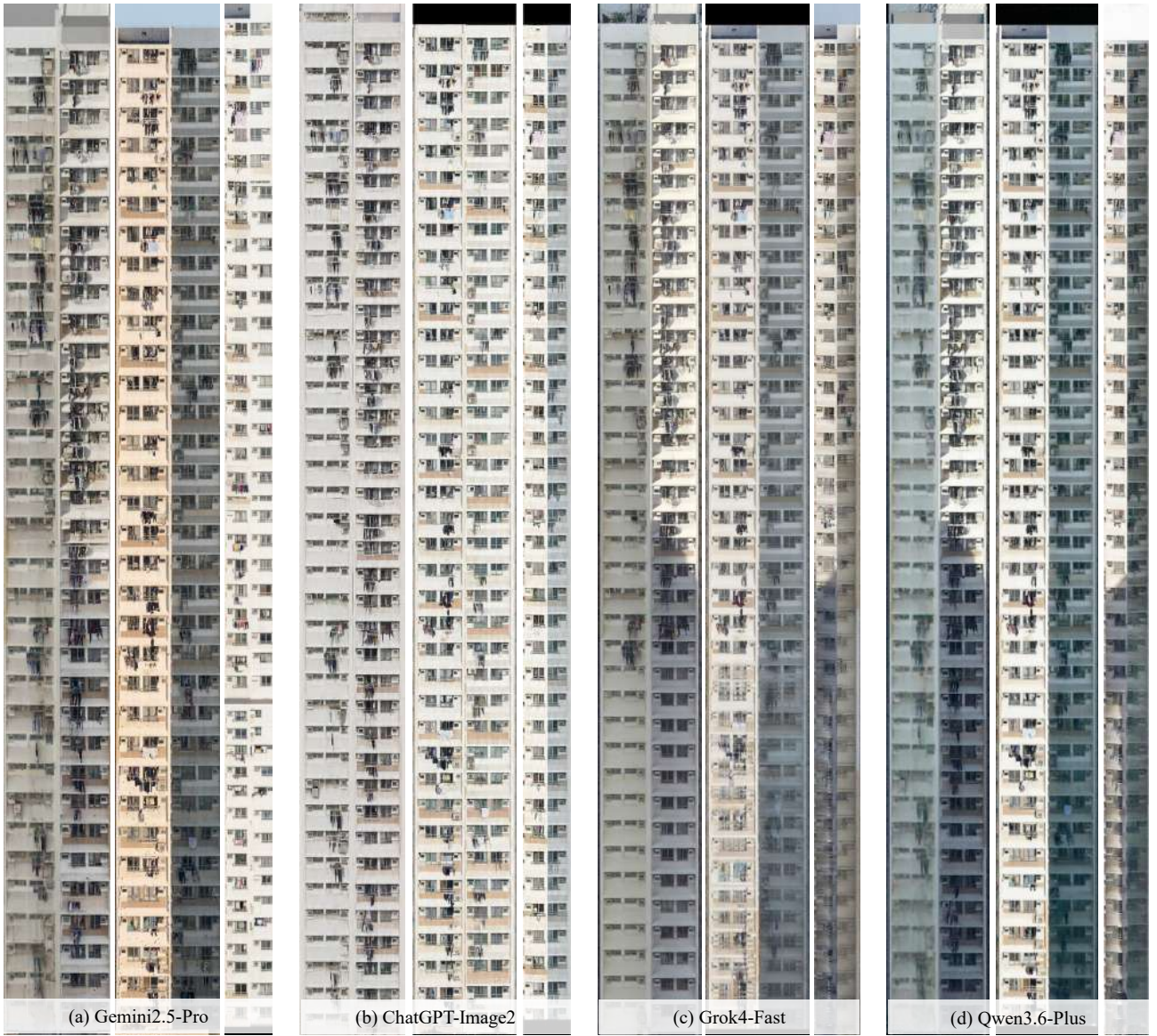


Figure 7. Texture refinement of the three pairs of facades in Figure 6 using VLMs. (a) Gemini2.5-Pro; (b) ChatGPT-Image2 (with least  $RMSD$ ); (c) Grok4-Fast (with highest  $SSIM$ ); (d) Qwen3.6-Plus.

ency in contrast and illuminance. In comparison, the second pair with Gemini2.5-Pro and all pairs with Qwen3.6-Plus show poor consistency.

These results indicate that zero-shot VLMs, even without fine-tuning and careful conditioning, may effectively handle cross-texture refinement and consistency. This highlights the potential of VLMs, with field-specific model adaptation strategies, to achieve perceptually consistent and structurally faithful facade texture refinement across paired inputs.

#### 4. Conclusion and Further Research

This study proposes a symmetry-aware, GenAI-supported framework for refining 3D building textures. The method integrates mass-aligned decomposition, texture reprojection, and the pairing of symmetric facades through Building Section Skeletons (BSS). The experimental results across low-rise, mid-rise, and high-rise buildings demonstrate that the approach preserves geometric integrity, enhances texture alignment, and captures struc-

tural symmetry, effectively removing heavy shadows. The findings confirm that architectural regularities, such as symmetry, when combined with generative capabilities, can improve texture completeness and visual realism. The work also highlights the potential of integrating classical photogrammetric refinement of facade texture models for high-fidelity, context-aware texture refinement.

However, several limitations remain in the paper, despite promising results. First, the framework relies on accurate mass decomposition and reliable symmetry detection, which may be compromised by highly irregular geometries or non-orthogonal photogrammetric meshes, such as heritage structures common in European or Asian historic cores. Secondly, pre-trained general-purpose GenAI models can introduce dimensional and structural inconsistencies, particularly when generating textures for high-rise facades with many repeated patterns. The current prompting strategy is over-simplified and not fully optimized for architectural constraints. Last but not least, there were too few test cases and insufficient benchmarks and metrics in this

preliminary study on integrating building mass symmetry with GenAI.

Accordingly, future research can incorporate explicit geometry-preserving priors, multi-view constraints, and adaptive prompts informed by building semantics. Integrating physics-based rendering cues, photometric-semantic integrations (Wu et al., 2026), and training task-specific generative models may also improve structural fidelity. Expanding the evaluation to more diverse building types, such as those in European cities (Wysocki et al., 2026), and large-scale urban scenes will further validate the robustness and generalizability of the proposed workflow.

### Acknowledgment

The work presented in this paper was supported by the Hong Kong Research Grants Council (RGC) (No. 17201325) and the Department of Science and Technology of Guangdong Province (GDST) (2023A1515010757).

### References

- Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., Çöltekin, A., 2015. Applications of 3D city models: State of the art review. *ISPRS International Journal of Geo-Information*, 4(4), 2842–2889. <https://doi.org/10.3390/ijgi4042842>.
- Buyukdemircioglu, M., Kocaman, S., Kada, M., 2022. Deep learning for 3D building reconstruction: A review. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 359–366. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-359-2022>.
- Iwaszczuk, D., Stilla, U., 2017. Camera pose refinement by matching uncertain 3D building models with thermal infrared image sequences for high quality texture extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 132, 33–47. <https://doi.org/10.1016/j.isprsjprs.2017.08.006>.
- Jiang, S., Jiang, W., Wang, L., 2021. Unmanned Aerial Vehicle-Based Photogrammetric 3D Mapping: A survey of techniques, applications, and challenges. *IEEE Geoscience and Remote Sensing Magazine*, 10(2), 135–171. <https://doi.org/10.1109/MGRS.2021.3122248>.
- Kamra, V., Kudeshia, P., ArabiNaree, S., Chen, D., Akiyama, Y., Peethambaran, J., 2022. Lightweight reconstruction of urban buildings: Data structures, algorithms, and future directions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16, 902–917. <https://doi.org/10.1109/JSTARS.2022.3232758>.
- Land, A. H., Doig, A. G., 2009. An automatic method for solving discrete programming problems. *50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art*, Springer, 105–132.
- Li, M., Guo, S., Duarte, F., Kumar, A., Kobori, N., Xue, F., Zhuang, W., Yeh, A. G., Ratti, C., 2026. Influence of objective and perceived exposures to urban nature on people's happiness. *npj Urban Sustainability*, 6, 6. <https://doi.org/10.1038/s42949-025-00306-9>.
- Liu, B., Liu, W., Lei, Z., Zhang, F., Huang, X., Awwad, T. M., 2024. A Planar Feature-Preserving Texture Defragmentation Method for 3D Urban Building Models. *Remote Sensing*, 16(22), 4154. <https://doi.org/10.3390/rs16224154>.
- Meng, S., Wu, L., Li, M., Yeh, A. G., Xue, F., 2026. Automated window detection for digital twin buildings: A generalized 'sandwich' model and practical guidelines. *Energy and Buildings*, 360, 117384. <https://doi.org/10.1016/j.enbuild.2026.117384>.
- Pambrun, J.-F., Noumeir, R., 2015. Limitations of the SSIM quality metric in the context of diagnostic imaging. *2015 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2960–2963.
- Rothermel, M., Gong, K., Fritsch, D., Schindler, K., Haala, N., 2020. Photometric multi-view mesh refinement for high-resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 166, 52–62. <https://doi.org/10.1016/j.isprsjprs.2020.05.001>.
- Ryu, N., Won, J., Son, J., Gong, M., Lee, J.-H., Cho, S., 2025. Elevating 3D models: High-quality texture and geometry refinement from a low-quality model. *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*, 165, 1–12.
- Shang, Q., Hu, H., Yu, H., Xu, B., Wang, L., Zhu, Q., 2023. Semantic Image Translation for Repairing the Texture Defects of Building Models. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1–20. <https://doi.org/10.1109/TGRS.2023.3338962>.
- Tang, W., Li, W., Liang, X., Wysocki, O., Biljecki, F., Holst, C., Jutzi, B., 2025. Texture2LoD3: Enabling LoD3 building reconstruction with panoramic images. *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2016–2026.
- Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), 600–612. <https://doi.org/10.1109/TIP.2003.819861>.
- Wu, X., Yang, T., Yu, L., Cao, J., Si, H., 2026. Efficient semantic-aware texture optimization for 3D scene reconstruction. *Computers & Graphics*, 134, 104529. <https://doi.org/10.1016/j.cag.2025.104529>.
- Wu, Y., Xue, F., Li, M., Chen, S.-H., 2024. A novel Building Section Skeleton for compact 3D reconstruction from point clouds: A study of high-density urban scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 209, 85–100. <https://doi.org/10.1016/j.isprsjprs.2024.01.020>.
- Wu, Y., Xue, F., Nan, L., Wu, L., Stoter, J., Yeh, A. G., 2025. MorphCut: an efficient convex decomposition method of 3D building models for urban morphological analytics. *International Journal of Geographical Information Science*, 1–23. <https://doi.org/10.1080/13658816.2025.2562251>.
- Wysocki, O., Schwab, B., Biswanath, M. K., Greza, M., Zhang, Q., Zhu, J., Froech, T., Heeramaglore, M., Hijazi, I., Kanna, K. et al., 2026. TUM2TWIN: Introducing the large-scale multimodal urban digital twin benchmark dataset. *ISPRS Journal of Photogrammetry and Remote Sensing*, 232, 810–830. <https://doi.org/10.1016/j.isprsjprs.2025.12.013>.

Xiong, W., Zhang, H., Peng, B., Hu, Z., Wu, Y., Guo, J., Huang, H., 2023. Twintex: Geometry-aware texture generation for abstracted 3D architectural models. *ACM Transactions on Graphics*, 42(6), 1–14. <https://doi.org/10.1145/3618328>.

Xue, F., Lu, W., Webster, C. J., Chen, K., 2019. A derivative-free optimization-based approach for detecting architectural symmetries from 3D point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 148, 32–40. <https://doi.org/10.1016/j.isprsjprs.2018.12.005>.

Xue, F., Wu, L., Lu, W., 2021. Semantic enrichment of building and city information models: A ten-year review. *Advanced Engineering Informatics*, 47, 101245. <https://doi.org/10.1016/j.aei.2020.101245>.