東京大学　関本研究室 / Sekimoto Lab. IIS, the University of Tokyo.

An appearance-contingent generation solution with an automated creation of 3D building digital cousins

Lingfeng Liao, Yoshiki Ogawa, Chenbo Zhao, Yoshihide Sekimoto

Background

As a significant component of urban scenes, researchers have been seeking an effective modeling approach with satisfactory performance to automatically create the building architecture of a digital twin city (DTC). However, most existing methodologies based on reconstructing three-dimensional (3D) urban building models suffered from data retrieval limitations. Under these circumstances, this research raises building digital cousins (BDC), trying to resolve the data difficulties in previous DTC creations by incorporating generative artificial intelligence (AI) approaches, which use open-source building footprint datasets and parameters to regularize the building appearance.

Objectives

- Apply the decoder-only Large Language Model (LLM) theory based on Transformer to create 3D building models in the level of details (LoD) 2 in a natural language-like approach
- Employ additional cross-attention modules parallel to the decoder for embedding the appearance parameters

Methodology

The building models in the standard $< v, f >$ format are interpreted into tokens by a specific discretization approach shown resembling natural languages. We train the decoder-only generator resembling the pattern of LLMs such as generative pre-trained transformer (GPT). As shown in the workflow of Figure 1, parallel cross-attention layers are incorporated in order to embed the appearance-contingent parameters such as roof types and heights. We keep the footprint tokens constantly visible to the model for reducing unnecessary errors.

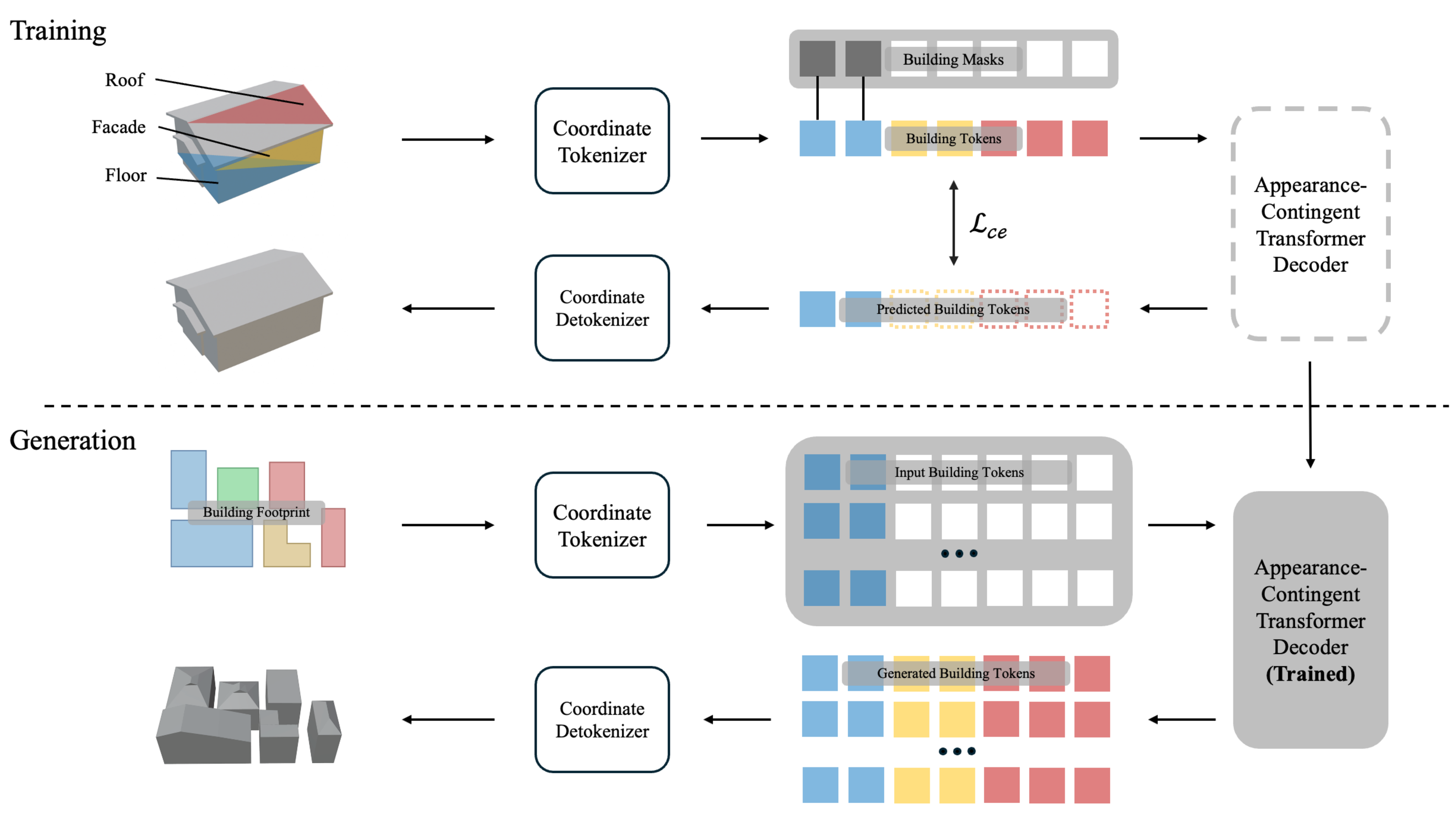

Figure 1: Overall Training-Inference framework of the proposed methodology.

Data sources

3D building dataset from the Project PLATEAU held and managed by the Ministry of Land, Infrastructure, Transportation, and Tourism (MLIT), Japan. Building data in LoD 2 is used for model training, and footprints (LoD 0) are for conditional references.

Results

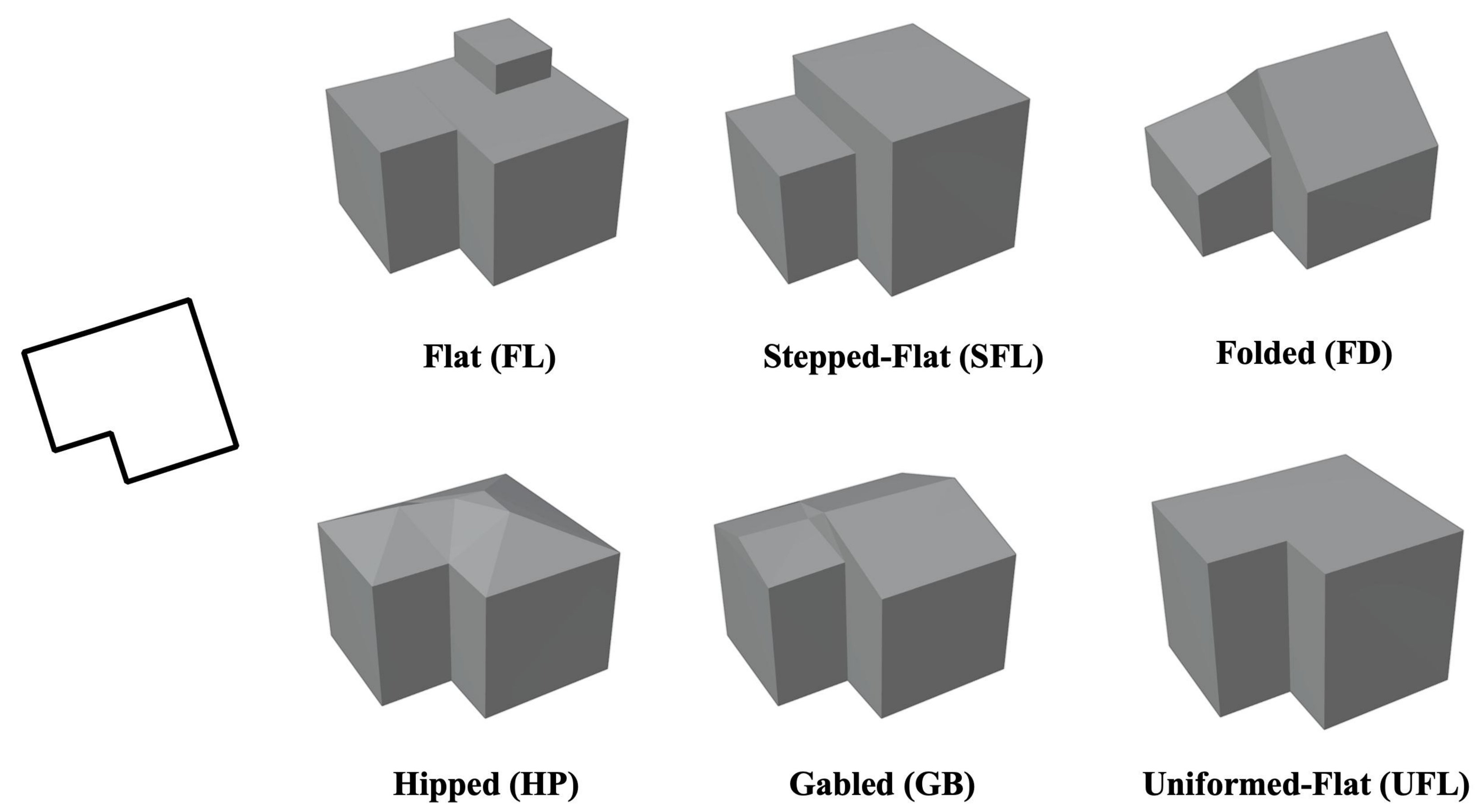

Figure 2: Evaluation of singular BDC creation, using different configurations of roof type parameters.

Figure 3: Evaluation of group-based BDC creations in three selected test areas in Kashiwa City, Chiba.

	Area (a)	Area (b)	Area (c)	Average
RMSE (m)	0.19	0.20	0.17	0.19
BMQI (%)	97.3	98.7	98.2	98.1

Table 1: Quantitative Evaluation of group-based BDC creations in three selected test areas using Root Mean Square Error (RMSE) and Building Mesh Quality Index (BMQI).

Conclusion

We proposed a novel generative framework for mapping wide-range BDC instances based on the prevailing LLM theory. The framework reveals the potential of replacing full-level reconstruction of the real-world urban building constructions in creating DTC representations by involving parameter configuration. Future works may incorporate more detailed information, such as roof orientations and openings, as controls for a more precise building mapping.