Enhancing Point Cloud Segmentation of Chinese Historical Buildings with Synthetic Data

Jiangping Ma[1,2,3,a], Zhiyuan Guo[1,2,3,b], Wenpeng Li[1,2,3,c], Weiya Chen[1,2,3,d*]

[1]School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan, 430074, China
[2]National Center of Technology Innovation for Digital Construction, Wuhan, 430074, China
[3]International Joint Research Laboratory of Smart Construction, Wuhan 430074 China

[a]m202271554@hust.edu.cn, [b]d202180541@hust.edu.cn, [c]m202371912@hust.edu.cn, [d]weiya_chen@hust.edu.cn
*Corresponding author

Abstract. Considering the challenges of segmenting architectural components in point cloud data, particularly for Chinese historical buildings, we develop an efficient method for constructing datasets to enhance deep learning techniques for precise segmentation. Surface sampling is integrated with advanced virtual laser scanning technology in our approach. Initially, labeled point cloud data is captured through surface sampling. Subsequently, the HELIOS++ simulation platform mimics real-world scanning to generate unlabeled data resembling actual point clouds. Precise alignment and label transfer between these two types of data result in an annotated dataset that preserves authentic scanning characteristics. Additionally, we introduce a symmetry-axis-based point cloud completion technique to address data loss during scanning, leveraging the inherent symmetry found in Chinese historical buildings. To validate the effectiveness of our dataset, two state-of-the-art deep learning models are selected for comprehensive evaluation. Experimental results demonstrate that our dataset supports efficient and stable model training, exhibits strong generalization capabilities, and provides a robust foundation for semantic segmentation of historical buildings.

Keywords: Historical buildings; Point cloud semantic segmentation; Synthetic dataset

1 Introduction

Historical buildings, as carriers of traditional culture and craftsmanship, exhibit distinctive styles and profound cultural significance. Historic Building Information Modeling (HBIM) is recognized as an effective method for managing data of historical buildings in both geometric and informational contexts. A significant challenge in HBIM lies in constructing building objects. Unlike modern buildings that often have detailed drawings and specifications facilitating 3D modeling, comprehensive records are typically lacking for historical buildings, posing hurdles to accurate modeling.

© The Author(s) 2025
Y. Qiu et al. (eds.), *Proceedings of the 2024 7th International Conference on Civil Architecture, Hydropower and Engineering Management (CAHEM 2024)*, Advances in Engineering Research 256,
https://doi.org/10.2991/978-94-6463-650-5_15

Segmentation of point clouds into distinct architectural elements is crucial for transitioning from point clouds to 3D models. In recent years, various new methods for point cloud segmentation have been introduced, with deep learning (DL) methods demonstrating promising performance. However, their application in historical building segmentation remains limited due to the scarcity of widely accessible annotated point clouds.

Haznedar et al. [1] introduced the HBIM dataset containing 3D point cloud data from historical buildings in Gaziantep, Turkey. However, the dataset's building data is insufficient for training deep learning models and is limited to training models on individual rooms. The ArCH (Architectural Cultural Heritage) [2] includes various types of cultural heritage buildings, but the lack of similar design and structural features among buildings makes it difficult for deep learning models to learn and understand. Given challenges in acquiring large real-world datasets, there is growing interest in synthetic data for point cloud segmentation. For instance, studies [3,4] have introduced methods for generating synthetic point cloud datasets using virtual laser scanners with variable distances, simulating diverse point cloud scenarios. Y. Ji et al. [5] sampled surfaces of 3D MAX models to obtain annotated point cloud data of historical buildings but only attempted segmentation on roofs.

For Chinese historical buildings, existing datasets are inadequate for deep learning model training due to mismatched segmentation content, limited data scale, and insufficient classification detail. To address this, a high-quality dataset has been developed, combining real data collection and synthetic data generation. We leverage the inherent axial symmetry of Chinese historical buildings and propose a symmetry axis-based point cloud completion algorithm, ensuring dataset integrity and accuracy. Additionally, we innovatively combine surface sampling and virtual laser scanning techniques, producing synthetic data that retains real scan characteristics while achieving higher annotation accuracy.

2 Methodology

2.1 Symmetry Axis Identification and Point Cloud Void Completion

The dimensions of Chinese historical buildings are often substantial, and the scanning environment is complex, resulting in significant data voids in scanned results due to numerous obstacles. Leveraging the inherent symmetry found in traditional Chinese historical buildings, a novel approach is proposed that utilizes symmetry planes to facilitate the reconstruction of incomplete sections in point clouds of Chinese historical buildings. The mathematical modeling of reflection symmetry in architectural analysis, particularly in historical buildings characterized by bilateral and longitudinal symmetries, can be expressed as follows:

$$\mathrm{T}(p) = \left(p_x - 2\left(p_x \cos\theta + p_y \sin\theta - \rho\right)\cos\theta, p_y - 2\left(p_x \cos\theta + p_y \sin\theta - \rho\right)\sin\theta\right) \quad (1)$$

Here, p_x and p_y represent the Cartesian coordinates of point p, while θ and ρ define a line with the polar equation $x\cos\theta + y\sin\theta = \rho$. To determine the parameters of either Equation 1, we propose the following optimization problem:

$$\arg min \sum_{i=1}^{n} \parallel t_i - T(x_i) \parallel \tag{2}$$

In this equation, t_i denotes the closest point to x_i after being mapped through a symmetry plane. Particle Swarm Optimization (PSO) is employed to efficiently determine the parameters of the symmetry transformation. PSO is a global optimization algorithm that can find the optimal solution within a large search space. It demonstrates strong robustness and stability, especially when handling point clouds with noise and incomplete data. After identifying the symmetry axis, the point cloud is mirrored along this axis to replenish missing points. The procedure begins with constructing a KD-tree for efficient nearest neighbor searches within the point cloud. For each point in the original cloud, its mirrored counterpart is computed based on the symmetry axis's ρ and θ values. Once mirrored points are computed, we use k-nearest neighbor searches to verify if these points adequately distance themselves from existing points on the opposite side of the symmetry axis in the point cloud. If a mirrored point meets the distance criteria, it is considered a missing point and subsequently added to the cloud.

2.2 Synthetic Point Cloud Dataset Generation Pipeline

Point clouds sampled directly from Building Information Modeling (BIM) models often show discrepancies in point distribution and density compared to real world point clouds. This disparity underscores the significant differences between synthetic and real data. To overcome this limitation, we propose a novel method for generating annotated dataset, which is depicted in Fig.1. This innovative workflow produces labeled synthetic point clouds that closely resemble features observed in real LiDAR scans. Moreover, this approach substantially reduces the time and effort typically associated with manual labeling.

Initially, we select 3D models of Chinese historical buildings from multiple open-source repositories. Using Navisworks, a professional BIM software tool, we conduct comprehensive collision detection analysis on all models to ensure their spatial non-conflict and metric accuracy. To ensure uniformity and interoperability in data processing, the output format of all 3D models is standardized to OBJ files. During the point cloud data acquisition phase, we innovatively combine two approaches: surface sampling and Virtual laser scanning. Firstly, SketchUp software is utilized to precisely classify the 3D models by structural components (such as beams, columns, doors, etc.) and export them into CloudCompare software for surface sampling, which generates segmented point cloud data based on components. This approach allows us to swiftly annotate entire components, avoiding the complexity of segmenting entire buildings. Secondly, we employ virtual laser scanning technology to simulate the scanning process in a real-world environment. The advanced open-source virtual laser scanning tool HELIOS++ [6] is used to conduct virtual laser scans on the undivided OBJ models. HELIOS++ is based on ray tracing technology and excels in simulating complex light reflections. It accepts scene models in OBJ format and supports laser scanning simulations across various platforms, including airborne, drone-based, and ground mobile systems. For this study's specific requirements, HELIOS++ is configured to simulate a RIEGL VZ400 laser scanner, with a scanning range of 120 meters and precision of 8 millimeters.

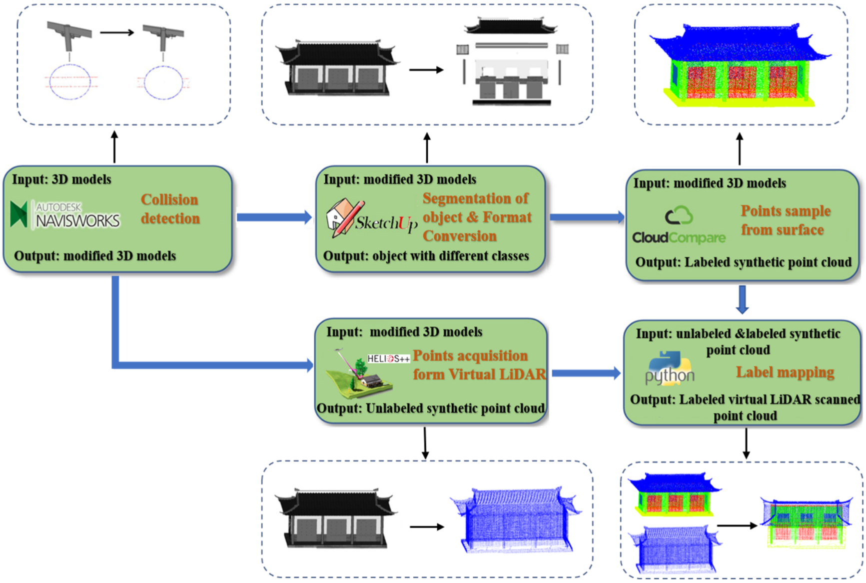

Fig. 1. The workflow of generating synthetic point clouds

After acquiring both the labeled point cloud obtained from surface sampling and the unlabeled point cloud scanned by the virtual laser scanner, label transfer is performed. Specifically, the Iterative Closest Point (ICP) algorithm is utilized to align the two types of point cloud data, and the K-D tree algorithm is applied to each point in the unlabeled point cloud to locate its nearest neighbor in the labeled point cloud, thereby transferring the label to the unlabeled point cloud.

3 Chinese Historical Buildings Point Cloud Segmentation Dataset

We conduct a comprehensive survey of existing ancient buildings and antique buildings in Wuhan to select suitable collection objects, and finally obtain 56 real point cloud data including temples, pavilions, and more.

During the collection of real point cloud data, many non-building component points are present in the acquired data due to the surrounding environment and the transformation of the building itself, such as some decorative items, pedestrians, trees, etc. These non-building points interfere with the accurate collection of building structure data, leading to incomplete datasets. Consequently, data cleaning is performed by manually removing non-building components. For missing parts of the point cloud data, we employ the method outlined in Section 2.1 to fill in the gaps, as illustrated in Fig.2. This approach effectively utilizes the symmetrical characteristics of Chinese historical buildings to partially restore missing point cloud data. However, the effectiveness of the completion is limited, as mirroring can only reconstruct areas proximate to the symmetry plane on both sides.

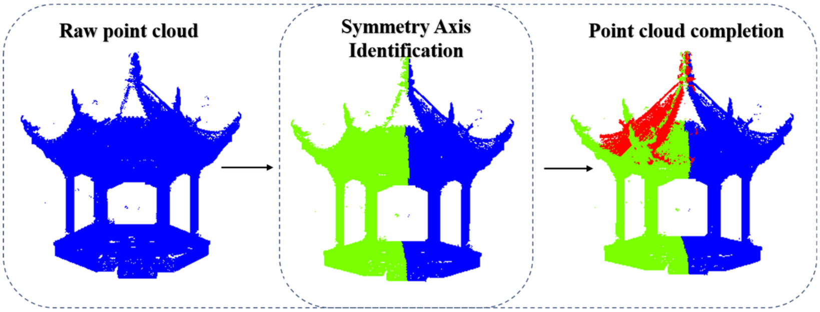

Fig. 2. Schematic of point cloud data complementation based on symmetry axis (The red points are complementary points)

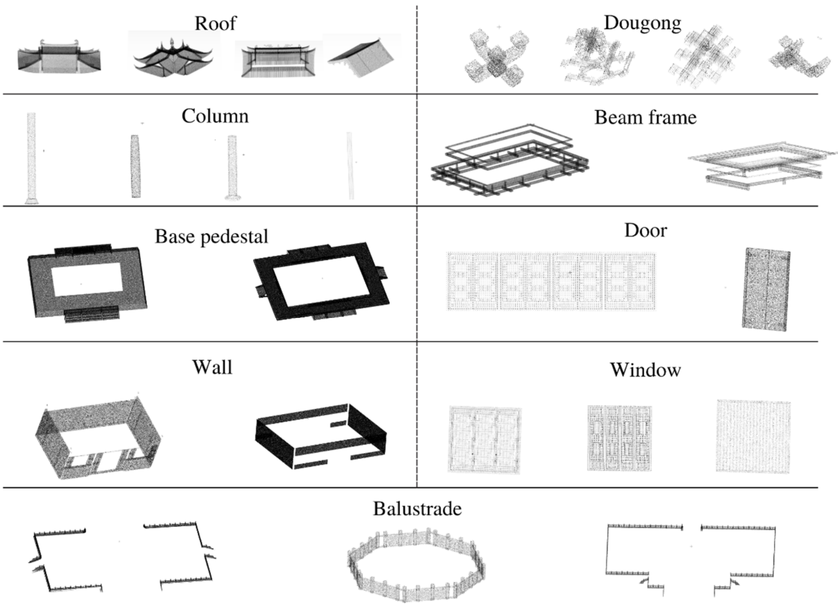

Fig. 3. Categories of components in Chinese historical buildings

Meanwhile, leveraging the method described in Section 2.2 for generating high-quality, accurately labeled synthetic point cloud data, we select 22 3D models of Chinese historical building from various public model libraries to create synthetic point cloud data. Finally, a well-annotated dataset of Chinese historical buildings is built to facilitate data-driven algorithms. The dataset consists of 56 real and 22 synthetic point clouds. Moreover, to facilitate the process of reconstructing 3D geometries for HBIM models, the point cloud is divided into 9 categories: roof, window, dougong, wall, column, base pedestal, balustrade, beam frame, and door, as illustrated in Fig.3.

4 Evaluation

To comprehensively evaluate the dataset's potential and applicability in complex scenarios, two state-of-the-art deep learning algorithms, Stratified Transformer [7] and EPCL [8], are adopted for assessing the semantic segmentation performance enabled by our dataset.

4.1 Semantic Segmentation Results

Our experiments are implemented in PyTorch, and we use the Adam optimizer with an initial learning rate of 10-2. Each model is trained for 100 epochs, with the learning rate dropped by 5% after each epoch. The Cross-Entropy loss is used for training. We evaluate the model only on real point clouds to confirm the performance of the model trained on synthetic data that represent realistic scenarios. To accurately measure the semantic segmentation model, three general evaluation metrics (overall accuracy (OA), mean Intersection over Union (mIoU), mean class Accuracy (mAcc)) are employed. The results are shown in Table 1.

Table 1. Comparison of two models on Chinese historical building dataset

Models	OA (%)	mACC (%)	mIoU (%)
Stratified Transformer	80.6	68.9	56.3
EPCL	81.1	69.6	57.3

Evidently, both models exhibit an OA exceeding 80%, indicating that the dataset comprises rich and representative samples encompassing various aspects and details of historical architecture, thereby supporting highly generalized model training. However, the relatively low mAcc and mIoU suggest the presence of certain categories or complex scenarios that are challenging to distinguish within the dataset. For instance, windows and doors in Chinese historical buildings exhibit similar features, making differentiation difficult. The performance of the EPCL model on some test data is visually demonstrated in Fig.4, recognition errors in specific scenarios are revealed by magnification of details.

It is evident that the model performs well in identifying components such as roofs, base pedestals, and balustrades. However, the boundaries of doors, windows, and walls remain contentious, making accurate recognition difficult. This indicates that our dataset is both reasonable and challenging. The limitations of current point cloud segmentation algorithms when applied to Chinese historical buildings are also highlighted. To enhance segmentation performance, targeted adjustments based on the dataset's characteristics are necessary.

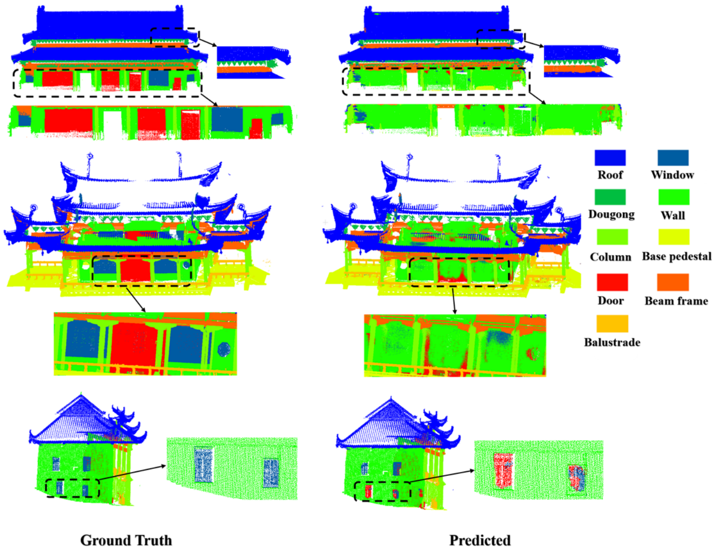

Fig. 4. Segmentation results of EPCL model on Chinese historical building dataset

5 Conclusion and Future Work

In recent years, numerous applications in the digitalization of historical architecture necessitate segmenting architectural components from point cloud data for tasks such as building information modeling and digital archiving. Given the existing historical building datasets are not suitable for the semantic segmentation of Chinese historical buildings, this study has constructed a dataset of Chinese historical buildings through both real data acquisition and synthetic data generation. The dataset includes 56 real and 22 synthetic point cloud, effectively facilitating the application of deep learning techniques in the segmentation of Chinese historical buildings.

Nonetheless, several limitations have been identified in this study that need to be addressed in future work, including issues of data class imbalance and similarity. Furthermore, for the segmentation of point clouds of Chinese historical buildings, we intend to explore new deep learning methods tailored to the specific characteristics of historical building data to improve segmentation performance promoting detailed reconstruction.

Acknowledgement

This study was supported by the National Natural Science Foundation of China (Grant No. 72001086).

References

1. Haznedar, B., Bayraktar, R., Ozturk, A. E., & Arayici, Y. (2023). Implementing PointNet for point cloud segmentation in the heritage context. Heritage Science, 11(1):2-19. https://doi.org/10.1186/s40494-022-00844-w.
2. Matrone, F., Lingua, A., Pierdicca, R., Malinverni, E. S., Paolanti, M., Grilli, E., Remondino, F., Murtiyoso, A., & Landes, T. (2020). A benchmark for large-scale heritage point cloud semantic segmentation. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.Paris.pp.1419–1426. https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-1419-2020.
3. Pierdicca, R., Mameli, M., Malinverni, E. S., Paolanti, M., & Frontoni, E. (2019). Automatic generation of point cloud synthetic dataset for historical building representation. In: Augmented Reality, Virtual Reality, and Computer Graphics: 6th International Conference, AVR 2019. Santa Maria al Bagno. pp.203-219. https://doi.org/10.1007/978-3-030-25965-5_16
4. Morbidoni, C., Pierdicca, R., Paolanti, M., Quattrini, R., & Mammoli, R. (2020). Learning from synthetic point cloud data for historical buildings semantic segmentation. Journal on Computing and Cultural Heritage (JOCCH).13(4):1-16. https://doi.org/10.1145/3409262.
5. Ji, Y., Dong, Y., Hou, M., Qi, Y., & Li, A. (2021). An extraction method for roof point cloud of ancient building using deep learning framework. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Barcelona. pp.321–327. https://doi.org/10.5194/isprs-archives-XLVI-M-1-2021-321-2021.
6. Winiwarter, L., Esmorís Pena, A. M., Weiser, H., Anders, K., Martínez Sánchez, J., Searle, M., & Höfle, B. (2022). Virtual laser scanning with HELIOS++: A novel take on ray tracing-based simulation of topographic full-waveform 3D laser scanning. Remote Sensing of Environment. 269: 112772. https://doi.org/10.1016/j.rse.2021.112772.
7. Lai, X., Liu, J., Jiang, L., Wang, L., Zhao, H., Liu, S., Qi, X., & Jia, J. (2022). Stratified transformer for 3D point cloud segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans.pp.8500–8509. https://doi.org/10.48550/arXiv.2203.14508
8. Huang, X., Huang, Z., Li, S., Qu, W., He, T., Hou, Y., Zuo, Y., & Ouyang, W. (2024). Frozen CLIP transformer is an efficient point cloud encoder. In: Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI-24). Washington, D.C. pp. 2382–2390. https://doi.org/10.48550/arXiv.2212.04098

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.