

How machine learning can improve 3D city models

3D building models are important for tasks such as urban planning, flood risk assessments, and daylight studies. However, creating accurate building models is not always straightforward. Roofs can be difficult to represent correctly, which may lead to incomplete models. In this study, machine learning is used to recognize different roof shapes automatically from laser-scanned data. By improving how roofs are represented, the method can contribute to more useful digital models of our cities.

3D building models have a wide range of applications and play an important role in decision-making processes related to urban planning and analysis. Point clouds are among the most common data formats used for generating 3D representations of buildings. Point clouds consist of millions of measured 3D points collected through laser scanning technology and are commonly used to digitally represent buildings and urban environments. However, extracting detailed building geometry from such data remains a challenging task, often resulting in reconstructed models with limited geometric detail and inconsistencies. Applications such as daylight simulations and water flow analysis often require more detailed and consistent 3D models than models reconstructed solely from point cloud data. In addition, gaps, noise, and structural inconsistencies frequently occur in the reconstructed building models, reducing their suitability for simulation and analysis purposes.

This study proposes a method for classifying roof types from point cloud data and reconstructing buildings parametrically based on the predicted roof type. Instead of manually modelling buildings, the proposed workflow combines parametric modelling, synthetic training data and automated roof classification to support automated 3D reconstruction. Parametric modelling is a method where building shapes are generated automatically from predefined rules and characteristics, such as roof type and building height. To achieve this, synthetically generated 3D building models were used to create training data for a machine learning model. Synthetic training data refers to computer-generated examples rather than measurements collected from real buildings. By learning geometric patterns from synthetic roof models, the model can automatically identify roof structures in real-world point cloud data.

The results show that machine learning can successfully classify roof types when tested on synthetic data, although the performance decreases when applied to real-world data due to differences in complexity and noise between the datasets. The study also highlights several challenges related to roof geometry, building footprints, and ambiguities in real roof structures.

Despite these challenges, the method demonstrates potential for future automated 3D city modelling workflows. As open geospatial data becomes increasingly available, methods such as this could contribute to faster and more efficient generation of 3D building models for urban planning, visualization, and environmental analyses, while reducing the need for time-consuming manual modelling of buildings.