

SPATIAL-DISTRIBUTION-AWARE DATA ORGANISATION FOR LARGE-SCALE URBAN 3D MODELS IN SMART-CITY VISUALISATION AND URBAN DIGITAL-TWIN APPLICATIONS

Jinhui Lan
Liu wenbing

Large-scale urban 3D models are now central to smart-city visualisation, digital-twin platforms, urban planning review, and spatial decision support. Their effective deployment, however, remains constrained by the computational burden of real-time rendering and by the semantic fragmentation introduced by conventional tile-based level-of-detail (LOD) scheduling. This paper presents a spatial-distribution-aware data organisation framework for large-scale urban 3D scenes that integrates a statically constructed three-level R-tree with a dynamically constructed adaptive quadtree. The method first classifies buildings at macro-, meso-, and micro-scales according to administrative boundaries, planning blocks, and inter-building similarity, and then records these relationships in an R-tree. A 3D tiled LOD model is subsequently organised through an adaptive quadtree, while the R-tree is used as a semantic constraint to guide tile scheduling and preserve the integrity of the user's area of interest. The framework is demonstrated using the Kowloon Peninsula of Hong Kong, China, covering approximately 39.028 km² with a 90 GB 3D model dataset containing about 26,000 buildings. The model is processed into 21 LOD levels and evaluated against a conventional loading strategy. The results show that the proposed approach reduces visible building fragmentation at macro-, meso-, and micro-scales, preserves semantic coherence within 1.5 s of loading, improves cross-scale loading speed, and increases real-time rendering performance by approximately 10 frames per second. These findings position the method as a practically relevant data-organisation strategy for smart-city systems that require both rendering fluency and cognitively coherent urban scene presentation.

Index Terms — urban 3D model; smart city visualisation; urban digital twin; data organisation; adaptive quadtree; R-tree; semantic integrity; scene rendering

© The author(s) 2025. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license (<http://creativecommons.org/licenses/by/4.0/>).

INTRODUCTION

Large-scale urban 3D models play an increasingly important role in contemporary urban development and smart-city systems because they provide a digital environment in which complex urban information can be visualized, interpreted, and operationalized. Such models support a wide range of planning and management functions, including design review, infrastructure monitoring, environmental assessment, emergency-response preparation, and public-facing digital-twin platforms [1, 6, 10]. Their growing importance reflects a broader transformation in urban governance, where cities increasingly rely on digital spatial representations to improve decision-making, communication, and coordination across departments and stakeholders. In this context, 3D city models are no longer merely illustrative tools; they are becoming core components of digital urban infrastructure.

The practical value of these systems, however, depends not only on geometric realism or visual richness, but also on rendering efficiency, data organization, and cognitive coherence during interaction. Urban planners, managers, and analysts do not simply need visually detailed models; they need models that remain responsive, preserve spatial meaning, and present relevant parts of the city in a way that supports understanding rather than visual overload. Real-time rendering research has long emphasized that performance and perceptual usefulness must be considered together, especially in large environments where the amount of geometry and scene complexity exceeds what can be displayed efficiently through brute-force rendering alone [2]. In smart-city settings, this requirement becomes even more critical because digital urban interfaces are often expected to support interactive exploration rather than static visualization.

The central technical difficulty lies in scale. Urban 3D scenes typically cover very large geographic extents, include dense and uneven building distributions, and contain highly heterogeneous objects with different geometric complexity, semantic roles, and spatial relationships. These characteristics generate very large datasets whose direct rendering can easily exceed the performance limits of conventional pipelines. They also produce organizational challenges because different parts of the city may require different update frequencies, different visual resolutions, and different interaction priorities depending on the user's purpose [1, 5]. As a result, efficient scene management is just as important as accurate model reconstruction.

For this reason, level-of-detail (LOD) modelling remains one of the dominant strategies for large-scale scene visualization. By representing the same urban object at multiple levels of geometric complexity, LOD makes it possible to balance detail and performance according to viewpoint, distance, and user interaction. The formalization of LOD in 3D city modelling has been a major topic in the literature, especially because different levels serve different analytical and visual needs [9]. In practical rendering systems, these ideas are often combined with hierarchical simplification methods such as hierarchical levels of detail (HLOD), which improve display speed by grouping and simplifying large environments while preserving visually meaningful structure [7]. In other words, LOD is not only a graphics technique; it is a foundational organizational principle for city-scale visualization.

In modern web-based and platform-oriented urban visualization systems, LOD strategies are frequently coupled with tile-based streaming mechanisms. Cesium has become one of the most widely used environments for interactive 3D geospatial visualization, and the 3D Tiles specification has emerged as an important standard for streaming, managing, and rendering massive heterogeneous 3D geospatial datasets [3, 4]. These technologies make it possible to load only the visible or potentially relevant parts of a scene at runtime, reducing memory pressure and improving responsiveness. For smart-city applications, this is highly advantageous because city models often need to be accessed over networks, integrated into web services, and displayed on hardware with varying performance capabilities.

Although tiling greatly improves runtime performance, it also introduces a significant drawback: the physical

subdivision of a model does not necessarily respect semantic or perceptual coherence. Buildings that belong to the same urban block, street wall, or visually recognizable district may be split across tile boundaries simply because the streaming structure follows spatial partition rules rather than meaningful urban groupings. This creates the familiar problem of semantic fragmentation, in which only part of a logically coherent area is rendered at sufficient fidelity while adjacent parts remain simplified or delayed. Studies of viewpoint management in urban planning have shown that meaningful navigation and interpretation depend heavily on how users perceive and engage with coherent urban scenes rather than disconnected fragments [6]. When tile boundaries disrupt that coherence, comprehension becomes weaker and interaction becomes less effective.

This problem is especially important in operational urban interfaces. A planner examining a redevelopment district, a mobility analyst reviewing a corridor, or a public user exploring a local neighborhood typically thinks in terms of meaningful urban units rather than in terms of arbitrary data partitions. If only a portion of that semantically meaningful unit is rendered in high quality while the remainder appears later or at lower fidelity, visual continuity is degraded and the scene becomes harder to interpret. The issue is therefore not merely one of visual annoyance; it affects the usability of 3D city models for planning review, policy communication, and analytical exploration. Research on efficient visualization of heterogeneous 3D city models has similarly shown that performance-oriented organization must be balanced against the need to preserve meaningful structure in the rendered city scene [10].

To address this problem, this paper develops a data-organization framework that explicitly considers the spatial distribution of buildings. The core idea is to combine two complementary index structures, each responsible for a different aspect of scene organization and delivery. The first is a three-level R-tree, constructed during preprocessing, which stores building-group classifications at macro-, meso-, and micro-scales. R-tree-based indexing has long been recognized as an efficient way to organize three-dimensional spatial data for virtual geographic environments, especially when rapid retrieval and hierarchical query performance are required [8]. In the present context, however, the R-tree is used not only as a spatial index, but also as a structure for preserving semantically meaningful building groupings derived from spatial distribution patterns.

The second component is an adaptive quadtree, constructed during rendering, which organizes and schedules tiled LOD data according to runtime viewing conditions. Quadtrees are well suited to dynamic view-dependent rendering because they provide a hierarchical decomposition of the scene that can be refined or simplified according to the current area of interest. On their own, however, quadtrees tend to prioritize rendering efficiency rather than semantic continuity. The framework proposed here therefore uses the precomputed R-tree as a semantic constraint on quadtree-driven tile scheduling. In effect, the quadtree manages performance, while the R-tree protects the integrity of meaningful building clusters. This coupling allows the rendering pipeline to remain efficient without sacrificing the spatial coherence of the user's focus area.

This approach is closely related to prior work on organizing complex urban scene data according to visual perception and user-centered scheduling logic. Research in this direction has shown that urban scene organization should not be based solely on raw geometric partitioning, but also on how users perceive districts, clusters, and meaningful visual units within the city [5]. The framework developed in this paper advances that idea by embedding perceptually and spatially meaningful grouping into the indexing structure itself, thereby improving not only rendering speed but also the interpretability of scene transitions during navigation.

The proposed framing is well aligned with the scope of urban development and smart-city research because it addresses an enabling layer of digital urban infrastructure: the reliable organization, streaming, and visual delivery of large-scale city models used in operational platforms and planning systems. A city-scale model is only as useful as its ability to present relevant spatial information clearly, quickly, and coherently to the people who rely on it. For this reason, the contribution of the study is not merely a graphics optimization. It is a smart-city data-management strategy that improves the responsiveness, interpretability, and practical

usability of 3D urban interfaces by connecting spatial indexing, hierarchical rendering, semantic coherence, and user-centered visualization within a single framework [1, 2, 9, 6].

BACKGROUND AND RESEARCH CONTEXT

The construction and interactive visualization of large-scale 3D city scenes typically relies on level-of-detail (LOD) modelling, whereby the same urban object is represented at multiple levels of geometric complexity according to viewing distance, scale, and rendering demand. This principle has become fundamental in 3D city modelling because it allows systems to preserve detail where it is needed while reducing computational load where it is not [9, 2]. In urban environments, where building density, geometric heterogeneity, and spatial extent are all high, such adaptation is essential for maintaining interactive performance. Hierarchical approaches related to LOD, including HLOD concepts, have likewise shown that large static and dynamic environments can be rendered more efficiently when scene complexity is organized into multi-resolution structures rather than processed uniformly [7].

In practical smart-city platforms, LOD is commonly implemented through tiled scene structures that support progressive loading and view-dependent streaming. This approach is especially attractive in web-based and operational 3D urban systems because only the visible or potentially relevant portion of the city scene needs to be transmitted and rendered at a given moment. Frameworks such as Cesium and standards such as 3D Tiles have helped establish this tiled, hierarchical logic as a mainstream solution for large heterogeneous geospatial datasets [3, 4]. Tiled LOD structures are therefore not simply a graphics convenience; they are a core mechanism that makes large-scale urban visualization feasible across different devices, networks, and application contexts.

However, the logic of tile partitioning is fundamentally geometric rather than semantic. A tile is, in essence, a physical data container defined by storage, transmission, and rendering requirements, not by the perceptual or functional structure of the city itself. As a result, buildings and urban blocks that belong to the same visually recognizable or analytically meaningful group may be divided across tile boundaries, particularly in dense, irregular, or topographically varied urban areas [1]. This creates a mismatch between how the scene is stored and how users actually interpret urban space. From the standpoint of planning, urban review, or digital-twin interaction, this is a significant limitation because users typically focus on coherent districts, corridors, or building clusters rather than arbitrary geometric partitions.

The consequence of this mismatch is a weakening of semantic integrity during progressive rendering. When only part of a visually or functionally related building group is refined while adjacent parts remain coarse or load later, the user experiences asynchronous scene formation. This interrupts visual continuity and can reduce comprehension of the area of interest, especially during navigation or detailed inspection. Research on viewpoint management in urban planning has shown that meaningful exploration of 3D city models depends not only on viewpoint control, but also on the coherent presentation of urban space as users move through it [6]. Similarly, studies on efficient visualization of heterogeneous 3D city models have indicated that rendering performance alone is insufficient if the delivered scene lacks structural or perceptual consistency [10]. In this sense, semantic fragmentation is both a technical and a usability problem.

The challenge becomes more complex because 3D city models are rarely homogeneous datasets. They are typically assembled from multiple sources, acquisition campaigns, modelling standards, and semantic layers, resulting in data that are heterogeneous in geometry, scale, quality, and organizational logic. Under such conditions, a single spatial index is often insufficient to support both efficient retrieval and semantically coherent rendering. Traditional indexing methods may perform well for geometric access, but they do not necessarily preserve meaningful building relationships or visual grouping. This limitation has motivated

increasing interest in hybrid organizational strategies that combine multiple index logics rather than relying on one structure alone [8, 5].

Hybrid indexing offers a stronger foundation for large urban scenes because it can separate distinct but interrelated tasks: one structure can preserve meaningful spatial or semantic organization, while another can manage runtime scheduling and refinement. Yet many existing hybrid strategies suffer from two common weaknesses. Some introduce excessive computational overhead by requiring multiple complex structures to evolve simultaneously during rendering, which can reduce efficiency in real-time systems. Others remain too weak in semantic control, meaning that they still privilege geometric partitioning over coherent building-group representation. As a result, they improve access speed without fully addressing the perceptual fragmentation that users encounter in tiled urban scenes [1, 10].

In contrast, the framework presented here assigns complementary rather than competing roles to two different index structures. The R-tree is used to capture semantically meaningful building relationships at three scales, thereby encoding macro-, meso-, and micro-level groupings before runtime begins. This use of an R-tree builds on its established strengths as an efficient spatial indexing method for large 3D geographic environments, but extends its role beyond simple retrieval toward semantically informed organization [8]. The adaptive quadtree, by contrast, is responsible for physical tile scheduling during rendering, organizing which portions of the scene should be loaded, refined, or simplified in response to current view conditions. In this design, the quadtree remains the active runtime mechanism, while the R-tree acts as a static semantic guide that constrains and improves scheduling decisions.

This separation of responsibilities provides an important operational advantage. Because the R-tree is constructed statically during preprocessing and the quadtree is generated dynamically at runtime, only one actively evolving index must be managed during rendering. This reduces the computational burden that would arise if multiple complex structures had to be updated simultaneously during interaction. At the same time, the presence of the precomputed R-tree ensures that runtime efficiency is not achieved at the expense of semantic coherence. The result is a more balanced framework in which performance-oriented scheduling and meaning-oriented organization support one another rather than conflict with one another.

From a research perspective, this background situates the present study at the intersection of real-time rendering, spatial indexing, visual perception, and smart-city data organization. The problem is not simply how to render large urban scenes faster, but how to render them in a way that preserves the integrity of meaningful urban units while remaining computationally tractable. This is particularly important for smart-city applications, where 3D models increasingly serve as operational interfaces for planning, monitoring, design review, and public communication. In such contexts, semantic continuity is not a secondary visual preference; it is part of the functional quality of the digital urban system itself.

METHODOLOGY

Overall framework

The proposed framework consists of four sequential components:

1. construction of a tiled 21-level LOD model from the original 3D city model;
2. multi-scale building classification at macro-, meso-, and micro-scales;
3. storage of classification results in a three-level R-tree;

4. adaptive quadtree scheduling of tiled data under semantic constraints provided by the R-tree.

Figure 1 summarises the workflow.

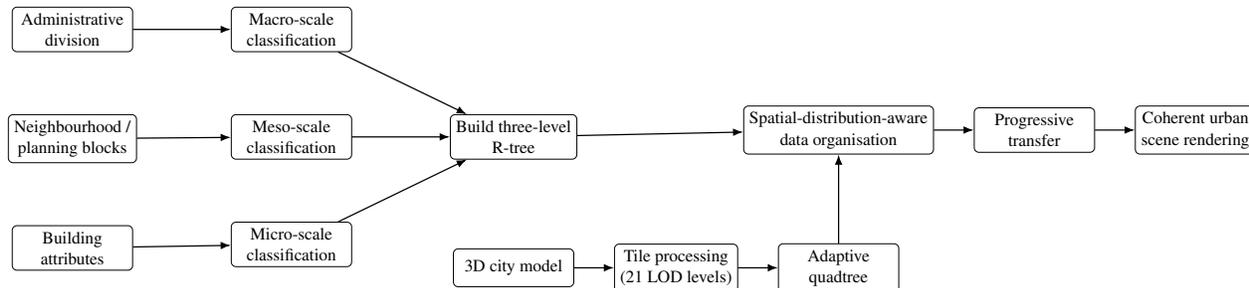


Figure 1: Conceptual workflow of the spatial-distribution-aware data-organisation framework.

Adaptive quadtree construction

The adaptive quadtree is used to organise tiled model data during rendering. The motivation is straightforward: a quadtree recursively divides space, aggregates model data by region, and supports efficient scheduling at runtime. In the present framework, quadtree partitioning is not purely geometric. Instead, it is adjusted according to the spatial distribution of buildings.

The procedure begins with a standard quadtree subdivision of the scene. For each candidate partition, the system determines whether buildings and tiles intersect the dividing line in a way that would split a semantically meaningful group. If such an intersection occurs, the partition range is enlarged. The process is repeated until each subregion contains all tiles belonging to a given classification group. This adaptive expansion reduces the likelihood that a semantically coherent building set will be fragmented by tile boundaries.

R-tree construction for building distribution

The R-tree is used to record building-group relationships so that the scheduler can retrieve semantically coherent building sets efficiently during scene rendering. The index is organised in three levels corresponding to macro-, meso-, and micro-scale classification.

At the macro- and meso-scales, buildings are classified using administrative boundaries and planning-block information. At the micro-scale, a finer grouping is performed using similarity between buildings. The source method describes this as a graph-based classification process in which building relationships are modelled and grouped through a minimum spanning tree constructed after similarity analysis.

The similarity score used for micro-scale grouping can be expressed as:

$$\text{Similarity} = \frac{1}{n} \sum_{i=1}^n (2 - C_i) p_i, \quad (1)$$

where C_i denotes the similarity of the i -th feature between two buildings (with values in $[0, 1]$), p_i denotes the weight assigned to that feature, and n is the number of evaluated building features. In the source framework,

relevant features include colour, height, area, and direction. The resulting similarity structure is then used to build a minimum spanning tree and derive fine-grained building groups.

An important structural detail is that the root of the R-tree does not represent an individual building. Rather, it represents the classification result obtained at the micro-scale and aggregates upward through the three-level hierarchy.

Scale division and LOD mapping

The framework defines three operating scales by reference to CityGML. These scale definitions are then mapped to the 21-level LOD hierarchy used for tile processing.

Table 1: Scale division adopted in the data-organisation framework.

Scale	CityGML reference	LOD levels in the tiled model
Macro-scale	LOD0	Levels 1–5
Meso-scale	LOD1–LOD2	Levels 6–15
Micro-scale	LOD3–LOD4	Levels 16–21

At the macro-scale, users primarily require the overall appearance of buildings, so lower-detail representations are sufficient. At the meso-scale, the general appearance remains important, but a finer level of detail is also needed. At the micro-scale, detailed building representation becomes necessary, and the most refined LODs are used.

Association between the R-tree and the adaptive quadtree

The central innovation of the method lies in how the R-tree and adaptive quadtree are linked through the tiled model.

First, the 3D city model is tiled and stored as a 21-level LOD structure. The system then traverses tiles from coarse to fine levels and queries the corresponding R-tree nodes using each tile's level and spatial attributes. The building classifications returned by the R-tree are written back to the tile level as identifiers, thereby labelling tiles according to semantic grouping.

Once tiles are labelled, the adaptive quadtree is constructed. If the quadtree region does not yet contain all tiles associated with a shared identifier, the region is expanded until the full relevant group is enclosed. In this way, the quadtree inherits semantic constraints from the R-tree and can schedule rendering in a manner that prioritises the integrity of the user's area of interest.

Algorithm 1 Spatial-distribution-aware organisation and scheduling

Require: Original 3D city model, administrative data, planning-block data, building attributes

Ensure: Semantically coherent progressive rendering of the urban scene

- 1: Tile the 3D city model into a 21-level LOD structure
 - 2: Classify buildings at the macro-scale using administrative boundaries
 - 3: Classify buildings at the meso-scale using planning-block divisions
 - 4: Compute inter-building similarity and derive micro-scale groups using a graph-based minimum spanning tree
 - 5: Store macro-, meso-, and micro-scale groupings in a three-level R-tree
 - 6: **for** each LOD level from coarse to fine **do**
 - 7: **for** each tile at the current level **do**
 - 8: Query the R-tree using the tile's spatial extent and scale
 - 9: Assign the corresponding classification identifier(s) to the tile
 - 10: Construct an adaptive quadtree for the tiled scene
 - 11: **for** each quadtree region **do**
 - 12: **if** the region splits tiles sharing the same classification identifier **then**
 - 13: Expand the quadtree region until the relevant group is fully enclosed
 - 14: During rendering, schedule tiles so that tiles with the same classification are prioritised together
-

STUDY AREA AND DATA

The empirical demonstration uses the Kowloon Peninsula in Hong Kong, China. This site is appropriate for evaluating the method because it combines dense urban morphology, multiple administrative districts, and numerous small planning blocks—all of which make building fragmentation more likely under conventional tile scheduling and therefore create a demanding test case for semantically constrained data organisation.

The study area covers approximately 39.028 km². The full 3D city model is approximately 90 GB in size and contains about 26,000 buildings. The dataset is derived from official open data sources in Hong Kong. The research area includes five administrative districts:

1. Kowloon City District,
2. Kwun Tong District,
3. Sham Shui Po District,
4. Wong Tai Sin District,
5. Yau Tsim Mong District.

The original model is processed into a 21-level tiled LOD structure. Level 1 corresponds to the coarsest representation, whereas Level 21 provides the most detailed representation.

RESULTS AND DISCUSSION

Tile processing and multi-scale classification

The first stage of the empirical workflow is the creation of the 21-level tiled LOD model. This establishes the rendering hierarchy required for progressive loading and adaptive scheduling. The coarse-to-fine structure then supports semantic labelling through the R-tree.

Buildings are classified at macro- and meso-scales using the administrative and planning data of the Kowloon Peninsula. At the micro-scale, each planning-area subset is further classified using the building-similarity procedure described above. The resulting classifications are then mapped to tile identifiers so that tile scheduling can be constrained by spatially meaningful building groups rather than geometric partitions alone.

Table 2 reproduces the representative tile-to-class mapping reported for the Sham Shui Po District.

Table 2: Representative tile labelling results reported for Sham Shui Po District.

<i>Adaptive quadtree (tile)</i>	<i>R-tree classification</i>	<i>Scale</i>
Tile_+021_+018.b3dm	Sham Shui Po District	Macro
Tile_+021_+018_L12_0.b3dm	Area1 in Sham Shui Po District	Meso
Tile_+021_+018_L12_1.b3dm	Area1 in Sham Shui Po District	Meso
Tile_+021_+018_L18_0000020t1.b3dm	Class 1 and Class 2 in Area1	Micro
Tile_+021_+018_L18_0000020t2.b3dm	Class 1 and Class 2 in Area1	Micro
Tile_+021_+018_L18_0000020t3.b3dm	Class 1 and Class 3 in Area1	Micro
Tile_+021_+018_L18_0000020t4.b3dm	Class 1, Class 2 and Class 3 in Area1	Micro

Note: Tile labels are assigned by querying the R-tree with the spatial attributes of each tile and mapping the corresponding classification result to the quadtree node, thereby facilitating semantically coherent rendering.

This mapping is especially important at classification boundaries. Where a tile intersects more than one building group, the tile is marked with multiple classifications. This preserves semantic information that would otherwise be lost in a purely geometric tiling scheme.

Rendering performance at macro-, meso-, and micro-scales

The rendering tests compare the proposed method with a traditional 3D scene loading strategy. The source paper evaluates the visual state of the scene at 1.5 s after loading at macro-, meso-, and micro-scales.

The qualitative comparison is consistent across scales:

- *Macro-scale:* under traditional loading, the rendered model at 1.5 s does not yet satisfy human cognitive requirements for the current viewing resolution; under the proposed method, the current-resolution requirement is already met.
- *Meso-scale:* under traditional loading, building fragmentation is visually obvious at 1.5 s; under the proposed method, buildings in the same area are rendered first.
- *Micro-scale:* under traditional loading, visible fragmentation remains evident; under the proposed method, the buildings in the area of immediate user concern retain integrity at 1.5 s.

These observations demonstrate that the method improves not merely technical loading order, but the perceptual coherence of progressive rendering. This is especially relevant to smart-city interfaces in which users inspect districts, planning units, and local building clusters as meaningful urban entities rather than as arbitrary sets of mesh fragments.

Quantitative comparative findings

The source paper also reports a quantitative comparison between the proposed hybrid-index method and a traditional loading approach. Two outcomes are highlighted:

1. the real-time rendering frame rate increases by approximately 10 fps;
2. scene loading becomes faster when the view switches between scales.

Because the reported quantitative figure is presented as a comparative line-chart analysis rather than as a full numeric table, the most defensible source-grounded summary is given in Table 3.

Table 3: Source-grounded comparative findings reported for the proposed method.

<i>Assessment dimension</i>	<i>Traditional loading</i>	<i>Proposed method</i>
Macro-scale rendering at 1.5 s	Current resolution not yet sufficient for user cognition	Current resolution requirement satisfied
Meso-scale rendering at 1.5 s	Visible building fragmentation	Buildings in the same area rendered first
Micro-scale rendering at 1.5 s	Visible building fragmentation in user focus area	Semantic integrity retained in user focus area
Real-time frame rate	Lower baseline performance	Approximately 10 fps higher
Loading during scale transitions	Slower switching response	Faster scene loading across LOD transitions

The combination of qualitative and quantitative evidence supports the central claim of the framework: organising tiled urban data according to building distribution materially improves both rendering fluency and semantic coherence.

Implications for urban development and smart-city systems

From the perspective of urban development and smart-city applications, the significance of this method extends beyond graphics optimisation. Planning support systems, public digital twins, and city information platforms depend on the ability to present urban space in a way that is both fast and cognitively legible. Fragmented rendering can disrupt the interpretation of neighbourhood form, block structure, and local building relationships—all of which matter in planning analysis, stakeholder communication, and operational oversight.

By ensuring that semantically related buildings are rendered together, the proposed framework supports more reliable visual reasoning in urban contexts. The method is therefore particularly relevant to city-scale platforms that must balance large data volumes with real-time interactivity.

CONCLUSION

This paper presents a smart-city-oriented data-organisation framework for large-scale urban 3D models that explicitly incorporates the spatial distribution of buildings. The method combines a statically constructed three-level R-tree and a dynamically constructed adaptive quadtree to organise tiled LOD data in a way that preserves semantic integrity during rendering.

The empirical study on the Kowloon Peninsula of Hong Kong demonstrates the practical value of the approach. Using a 90 GB dataset covering approximately 39.028 km² and about 26,000 buildings, the framework organises the model into 21 LOD levels, classifies buildings at macro-, meso-, and micro-scales, and uses those classifications to guide runtime scheduling. The reported results show reduced visible fragmentation, stronger integrity of the user's area of interest within 1.5 s of loading, faster scale-switch loading, and an improvement of approximately 10 fps in real-time rendering performance relative to a conventional method.

The method is especially suitable for large-scale urban scenes in which planners, analysts, and platform users require coherent, progressive visual access to city data. Its relevance to urban development and smart cities lies in its ability to make city-scale digital environments more interpretable, responsive, and operationally useful.

DATA AVAILABILITY

The data supporting the findings of this study were derived from public-domain sources identified in the source study, including the Hong Kong Common Spatial Data Infrastructure portal and the Hong Kong Planning Department 3D model download service.

CONFLICT OF INTEREST

The author declares no conflict of interest.

REFERENCES

- [1] Yang, X., Huo, L., Shen, T., Wang, X., Yuan, S., Liu, X.: A large-scale urban 3D model organisation method considering spatial distribution of buildings. *IET Smart Cities* 6(1), 54–64 (2024). <https://doi.org/10.1049/smc2.12070>
- [2] Akenine-Möller, T., Haines, E., Hoffman, N.: *Real-Time Rendering*, 4th edn. CRC Press, Boca Raton (2019)
- [3] Cesium: Available online: <https://cesium.com/about>. Accessed 6 June 2023
- [4] 3D Tiles Specification: Available online: <https://github.com/AnalyticalGraphicsInc/3d-tiles/tree/master/specification#tileset-json>. Accessed 6 June 2023
- [5] Chen, X.: Organization and Scheduling Method of Complex 3D Urban Scene Data Driven by Visual Perception. Master's Thesis, Southwest Jiaotong University, Sichuan, China (2018)
- [6] Neuville, R., et al.: 3D viewpoint management and navigation in urban planning: application to the exploratory phase. *Remote Sensing* 11(3), 236 (2019). <https://doi.org/10.3390/rs11030236>

- [7] Erikson, C., Manocha, D., Baxter, W.V.: HLODs for faster display of large static and dynamic environments. In: *Proceedings of the 2001 Symposium on Interactive 3D Graphics*, pp. 111–120 (2001). <https://doi.org/10.1145/364338.364376>
- [8] Zhu, Q., Gong, J., Zhang, Y.: An efficient 3D R-tree spatial index method for virtual geographic environments. *ISPRS Journal of Photogrammetry and Remote Sensing* 62(3), 217–224 (2007)
- [9] Biljecki, F., Ledoux, H., Stoter, J.: Formalisation of the level of detail in 3D city modelling. *Computers, Environment and Urban Systems* 48, 1–15 (2014)
- [10] Buyukdemircioglu, M., Kocaman, S.: Reconstruction and efficient visualization of heterogeneous 3D city models. *Remote Sensing* 12(13), 2128 (2020)

Jinhui Lan, The school of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China

Liu wenbing, The school of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China

Manuscript Published; 15 December 2025.