Robust and Detailed 3D Building Model Generation from Airborne Laser Scanning Point Clouds via Piecewise Affine-Linear Mumford-Shah Segmentation

Julian Rasch[1,2], Jonas Luft[1], Ali Razavian[1], Jannik Reinke[1]

[1] syte GmbH, Münster, Germany - jonas.luft@syte.ms
[2] Department of Engineering Physics, Münster University of Applied Sciences, Steinfurt, Germany - julian.rasch@fh-muenster.de

Keywords: 3D Building Reconstruction, Mumford-Shah Segmentation, LoD3, Roof Topology, Airborne Laser Scanning.

Abstract

The increasing availability of high-resolution Airborne Laser Scanning (ALS) data has expanded the potential for high-fidelity 3D city modeling. At the same time, more and more use cases have surfaced which crucially require these high-fidelity 3D models to derive geometrical building properties in a more straightforward manner. However, standard automatically generated Level of Detail 2 (LoD2) models often fail to accurately capture complex building and roof topologies due to their reliance on overly simplistic geometric primitives, while creating more detailed models typically requires tedious manual editing. To address this gap, we propose a fully automated, data-driven pipeline to generate detailed, LoD3-like roof models utilizing only publicly available LiDAR point clouds and cadastral footprints. Our methodology introduces a robust piecewise affine-linear Mumford-Shah functional for initial point cloud segmentation, followed by discrete graph-cut polygon regularization and a global roof plane estimation step that mathematically encourages watertight, closed seams. Experimental evaluation on the ISPRS Vaihingen benchmark yields a highly competitive planimetric RMSE of 0.27 m. Furthermore, evaluation on a dataset of 109 buildings in Münster (Westf.), Germany, demonstrates that our method reduces the mean point-to-surface RMSE by up to 82% in complex urban areas compared to state-provided LoD2 baselines, achieving a superior geometric fit in more than 98% of the evaluated structures and produces visually appealing buildings. Relying on only three fixed global parameters, the proposed deterministic framework offers a highly scalable, mathematically sound path toward detailed, nationwide urban reconstruction.

1. Introduction

Accurate three-dimensional building models are a strict requirement for a wide range of urban applications: estimating living space in the building stock (Dochev et al., 2020), modeling energy demand (Luft et al., 2025), and assessing solar potential (Li et al., 2021; Ni et al., 2024). High-fidelity geometry is also essential for texture mapping and facade- and roof-level image analysis such as window counting or material classification (Meixner and Leberl, 2010; Forth et al., 2024), as oversimplified shapes cause visible distortion and inaccuracies. However, most large-scale 3D building inventories use Level of Detail 2 (LoD2), representing roofs through a small set of geometric primitives (Kolbe et al., 2005). While automatically generated and widely available, LoD2 models frequently fail to capture complex roof topologies: multi-component gable roofs, dormers, and chimneys are reduced to overly simplistic shapes with significant geometric error. More detailed models approaching LoD3 traditionally require tedious manual effort, restricting their availability to small, project-specific datasets.

This lack in accuracy coincides with a window of opportunity: German state surveying agencies and their counterparts in other countries increasingly publish open geospatial data, including airborne laser scanning point clouds. The respective resolution has improved with every acquisition campaign, now reaching densities that make detailed roof reconstruction feasible at scale. Numerous approaches have been proposed to bridge the gap between demanding applications and the lack of detail in LoD2. Direct meshing from point clouds (Bauchet and Lafarge, 2019) can achieve high precision but produces irregular surfaces without preserving architectural composition. As a consequence, ridges, dormers, and roof shapes dissolve into unstructured triangles. Modern methods employing neural networks (Agoub et al., 2019; Liu et al., 2024) offer scalability but currently lack the geometric accuracy required for detailed applications. Terrestrial scanning (Hanke et al., 2025) achieves very high fidelity but does not scale to large areas. Image-based approaches (Zebedin et al., 2006, 2008; Schuegraf et al., 2023; Mei et al., 2025) can provide complementary line features but are susceptible to shading and contrast artifacts. More structured reconstruction strategies incorporate architectural priors. Model-driven methods fit predefined primitives to observed data (Vosselman and Dijkman, 2001; Verma et al., 2006), but struggle with complex multi-component buildings. Segmentation-based methods partition point clouds and reconstruct geometry from these: Huang et al. (2022) proposed City3D based on PolyFit (Nan and Wonka, 2017), while Jung et al. (2017) used a minimum description length principle for regularization. This is conceptually related to the boundary-length-weighted Potts model we employ in this work, which has been introduced by Zebedin et al. (2008) and used by Peters et al. (2022) to generate LoD2.x models for the Netherlands.

We propose a method for automatically generating detailed 3D roof models from publicly available ALS point clouds and cadastral footprints. The per-building pipeline comprises four stages: (1) Point cloud segmentation into piecewise-planar surfaces via a second-order Mumford–Shah model, (2) roof polygonization using intersection lines, step edges, and bounding boxes to partition the footprint into candidate polygons, (3) polygon regularization through a boundary-length-weighted Potts model solved via graph cuts, controlling the trade-off between LoD2-like and LoD3-like detail, (4) 3D extrusion through globally constrained plane fitting that encourages closed seams along shared edges.

https://doi.org/10.5194/isprs-annals-XII-4-W1-2026-285-2026

A central contribution is the formulation of these stages as linked optimization problems with very few free parameters, yielding high robustness across diverse building shapes and point cloud resolutions without per-dataset tuning. More precisely, we use a total of only three user-facing scalar parameters, one related to each of the steps segmentation, polygon regularization and plane fitting, which we do not change throughout the different experiments. For this work, we restrict input data to LiDAR and cadastral footprints; experiments with orthophoto-derived line features showed no quality gain that outweighed the additional artifacts caused by noise and shadows.

We evaluate our work on the ISPRS Vaihingen 3D Reconstruction Benchmark (Rottensteiner et al., 2012) and a custom data-set of 109 buildings around Münster (Westf.), Germany, covering complex urban and simpler suburban roofs. Results show a better base geometry fit than LoD2 for over 95 % of buildings, with an average improvement in RMSE of 67–82 % for complex geometries. In the ISPRS benchmark dataset of Vaihingen, we achieve a competitive accuracy of 0.27 m in planimetric RMSE and 0.37 m RMSE in height.

2. Methods

Our goal is to construct three-dimensional building models, represented as closed set of planar polygon shapes, from a LiDAR point cloud (alternatively, a digital surface model). To do so, we need to separate the point cloud into its different segments, i. e. roof planes, dormers, chimneys, etc. We then need to find appropriate boundary lines around these segments to obtain simple polygons that cover the footprint in a watertight and disjunct fashion, which can eventually be extruded to their original height. Additionally, we seek for a way to control the scale of details of the roof, i. e. manipulate how much details are visible on the roof; less details resembling a classic LoD2-type reconstruction with large primary roof shapes only, more details allowing a path towards LoD3-type reconstruction including secondary roof parts such as dormers and chimneys.

To achieve the above, we follow a similar approach as Zebedin et al. (2008), Jung et al. (2017), and Peters et al. (2022), and employ multiple subsequent steps to transition from point cloud to polygons while controlling the scale: (1) point cloud segmentation, (2) roof line detection and polygonization, (3) polygon regularization and scale selection, (4) (global) roof plane estimation and roof extrusion.

Before we further outline the different steps, let us fix the notation. Let $\mathcal{P}$ be the complete set of 3D LiDAR points, with elements $p = (x_p, y_p, z_p) \in \mathcal{P}$, denoting easting, northing and elevation. Subsets of the point cloud are denoted as $\mathcal{P}(\cdot)$, for example, $\mathcal{P}(l)$ denotes the subsets of points with a certain label l. For this work, we only consider building LiDAR points that are located within a building's cadastral footprint, and translate the data to be centered around its centroid for better algorithm stability.

We are generally concerned with estimating a plane approximation $z = Ax+Bx+C$ for given sets of LiDAR points. To intuitively map the plane parameters $\theta = \{A, B, C\}$ to their respective entities, we use the notation $\theta[\cdot]$, for example, $\theta[l] = \theta[\mathcal{P}(l)]$ denotes the plane associated with all points with label l. Treating the plane as a function, its height at a 2D coordinate $\mathbf{v} = (x, y)$ is evaluated directly as $\theta(\mathbf{v}) = \theta(x, y) = Ax + By + C$.

2.1 Piecewise affine-linear LiDAR segmentation

Our initial goal is to find a segmentation of our LiDAR point cloud into planes. Treating the point cloud as a digital surface model (filtering the raw cloud the to the topmost points if necessary), we can consider the data a surface function $z(\mathbf{v})$ such that $z(x_p, y_p) = z_p$ for all $p \in \mathcal{P}$. The classic continuous second-order Mumford-Shah functional (Mumford and Shah, 1989) seeks an optimal piecewise affine-linear surface approximation on partitions $\mathcal{S}$ and a set of partition boundaries $\mathcal{B}$ that fit the true surface data $z(\mathbf{v})$ while penalizing boundary length. In our notation, the discrete piecewise affine-linear Mumford-Shah problem reads

$$\arg\min_{\mathcal{S},\theta} \sum_{s\in\mathcal{S}} \sum_{p\in\mathcal{P}(s)} |\theta[s_k](x_p, y_p) - z_p|^2 + \frac{\gamma}{2} \sum_{b\in\mathcal{B}} |b|, \quad (1)$$

where $\mathcal{S}$ is a partitioning of the building footprint, and $\mathcal{B}$ the set of shared 2D boundary segments, with $b = (s, s') \in \mathcal{B}$ representing the boundary of length $|b|$ between adjacent partitions $s, s' \in \mathcal{S}$. Note that the partitioning here refers to a partitioning of the point cloud such that $\bigcup_{s\in\mathcal{S}} \mathcal{P}(s) = \mathcal{P}$ and $\mathcal{P}(s) \cap \mathcal{P}(s') = \emptyset$ for all $s \neq s' \in \mathcal{S}$. We later aim to "convert" this partition of LiDAR points into a partitioning of the building footprint using polygons. As demonstrated by Kiefer et al. (2020) on images, the two-dimensional problem (1) can be solved efficiently by decomposing it into one-dimensional scan lines L along principal directions across the pixel grid. Treating the LiDAR data as a surface, we can adapt the approach to LiDAR data and choose a similar set of directions aligned with the building's principal directions and diagonals. Since LiDAR resolution varies, the points along L are evaluated using a dynamic step size, searching for the closest next point along L. Along any such scan line, the two-dimensional plane $\theta[s_k]$ simplifies to a one-dimensional Taylor jet (a linear segment $z(t) = m \cdot t + c$), allowing the functional to be minimized efficiently. We adapted the implementation and solver provided by Kiefer et al. (2020)[1].

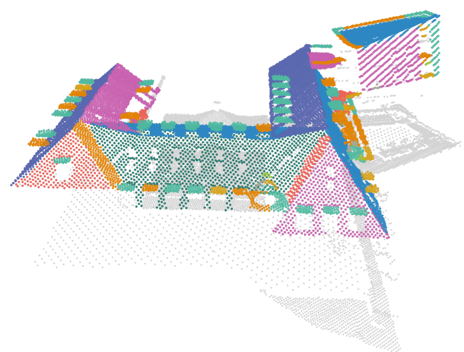


Figure 1. LiDAR segmentation of a complex building roof into planes using a piecewise affine-linear Mumford-Shah model.

The output of the LiDAR segmentation is a set of labels $\mathcal{L} = \{l_1, l_2, \ldots, l_M\}$ with associated points $\mathcal{P}(l_i)$ and planes $\theta[l_i]$ (see Figure 1). The parameter $\gamma > 0$ controls the aggressiveness of the boundary penalty; the larger γ, the fewer segments

[1] https://github.com/lu-kie/PALMS_ImagePartitioning

created. For practical purposes, we found that with correct data normalization and centering, the same γ could be used without adjusting throughout the datasets we investigated. When in doubt, we can afford having "too many" segments here, as the subsequent steps are able to merge them again. The single segmentation model drawback is that it inherently approximates planes, so roof structures such as domes may be poorly represented depending on their size (see also Section 4.2).

2.2 Roof line detection and polygonization

From the initial planar LiDAR segments $\mathcal{P}(l_i)_{i=1}^{M}$ and the associated planes $\theta[l_i]_{i=1}^{M}$ we now estimate a set of roof lines that separate the different segments as best as possible. To find candidate lines, we follow a similar approach to Jung et al. (2017) and Peters et al. (2022), and use (1) the footprint boundary, (2) plane intersection lines (for sloped segments), (3) step edge lines (height difference between adjacent segments), (4) bounding box lines (for horizontal segments).

The footprint boundary can typically be obtained from cadastral data and serves as the natural boundary shape for our polygons and building model. To identify plane intersection lines and step lines, we construct a spatial tree of neighboring roof segments and analyze each pair and their associated planes separately. If at least one of the planes is sloped, we determine their intersection line and constrain its length so that it traces the outline of at least one of the two segments by checking whether the distance of the orthogonal projection of any point onto the line is sufficiently close. Additionally, we determine any large height discontinuity at the boundaries between two neighboring segments and use the resulting lines as candidate lines. In order to include chimneys and vertically offset dormers properly, we assign a footprint-oriented bounding box to each horizontal plane (which does not intersect with any other planes) and add the associated line segments as candidate lines.

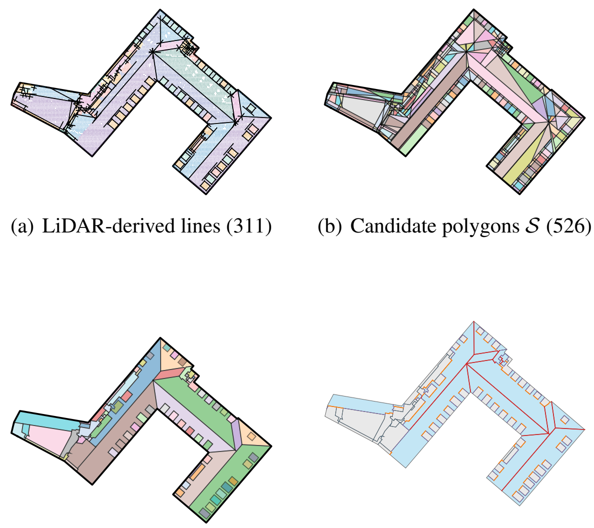
The result of this step is a set $\mathcal{S} = \{s_1, s_2, \ldots, s_K\}$ of initial 2D polygons generated by cutting the building footprint along the above lines. An example of lines and initial candidate polygons is shown in Figures 2(a) and 2(b), respectively. It is worth noting that the quality of the set of lines is crucial for the eventual building model, as the shape and structure of the resulting polygons cannot change after this (see also Section 4.2).

2.3 Polygon regularization and scale selection

The set of polygons $\mathcal{S}$ created in the previous step is generally too granular, and a single roof plane is often represented by a larger number of polygons instead of a single one. We now aim to reassign polygon labels (which correspond to planes here) and eventually merge them to represent each roof plane with constant slope by a single polygon if possible. During this regularization step, we can control how many details of the original roof segmentation we want to keep in the final result, which we can use to decide whether to include dormers, etc. into the roof reconstruction or keep only larger surfaces. This problem can be formulated as a discrete label assignment problem, and the idea has been taken from (Zebedin et al., 2008), but modified to include normalization and weighting. The optimal discrete label assignment can be computed as follows:

$$\min_{\mathbf{l}} E(\mathbf{l}) = E_{\text{data}}(\mathbf{l}) + \lambda E_{\text{reg}}(\mathbf{l}). \tag{2}$$

The state variable is the set of polygon labels $\mathbf{l} = [l_1, \ldots, l_K]^T$, where $l_k \in \mathcal{L}$ is the discrete label (respectively plane) assigned



(a) LiDAR-derived lines (311) (b) Candidate polygons $\mathcal{S}$ (526)

(c) Final polygons $\mathcal{F}$ (42) (d) Plane estimation constraints

Figure 2. From line detection to constrained roof plane estimation. Numbers in parentheses denote the respective number of lines and polygons. Sloped constraints (25) are marked in red, boundary constraints (66) in yellow. Slanted roof parts are colored in blue, horizontal parts in grey.

to polygon s_k, i.e., $\theta[s_k] = \theta[l_k]$. The data fidelity term E_{data} evaluates how well the LiDAR points physically located within each polygon s_k fit its assigned candidate plane $\theta[s_k]$:

$$E_{\text{data}}(\mathbf{l}) = \sum_{s_k \in \mathcal{S}} \frac{|s_k|}{|\mathcal{P}(s_k)|} \sum_{p \in \mathcal{P}(s_k)} \text{dist}_{\perp}(p, \theta[s_k]) \tag{3}$$

scaling by area $|s_k|$ ensures that larger polygons exert a proportionally larger influence on the total data cost, preventing small high-error polygons from dominating. The regularization term E_{reg} penalizes the assignment of different planes to adjacent polygons along shared boundaries, utilizing a boundary-length-weighted Potts model (piecewise constant Mumford-Shah):

$$E_{\text{reg}}(\mathbf{l}) = \sum_{b=(s,s') \in \mathcal{B}} |b| \cdot \chi(\theta[s] \neq \theta[s']) \tag{4}$$

The indicator function χ evaluates to 1 if the assigned planes of adjacent polygons differ and 0 if they are identical. The energy (2) can be efficiently minimized using a multi-label graph cut with alpha expansions.[2] After label (re)assignment, adjacent polygons sharing the same label are merged to form the finalized topological faces $\mathcal{F} = \{f_1, \ldots, f_N\}$, with an associated set of edges $\mathcal{E}$, with $e = (f, f') \in \mathcal{E}$ representing an edge between adjacent faces f and f'.

The regularization parameter $\lambda > 0$ plays a crucial role in determining the eventual scale or roof structure and its details (cf. also (Zebedin et al., 2008)). A high λ results in few final polygons and hence only the large roof surfaces represented in the final result. The smaller λ, the more details remain in the polygons, and hence we can find dormers and chimneys in the final reconstruction. Setting λ to different values allows us to create reconstructions of similar scale as LoD2, or go beyond it and transition towards LoD3. In practice, the choice of λ poses no

[2] We use pygco https://github.com/Borda/pyGCO.git.

https://doi.org/10.5194/isprs-annals-XII-4-W1-2026-285-2026 | © Author(s) 2026. CC BY 4.0 License.

particular challenges: we used the same λ for all shown reconstructions and evaluations (see also Section 3).

Before we estimate their plane parameters, we transform the set of finalized two-dimensional polygons $\mathcal{F}$ into a planar half-edge graph. To align with CityGML semantics, we explicitly support both exterior boundaries and interior rings (holes) within the graph faces. We merge vertices within a defined spatial tolerance to establish topological nodes and link twin half-edges to encode the adjacency between neighboring roof segments. Finally, we apply a simplification pass to remove topologically redundant collinear vertices without introducing self-intersections. This yields a robust graph structure with faces $\mathcal{F}$, edges $\mathcal{E}$ and associated vertices $\mathcal{V}(e)$ for $e \in \mathcal{E}$ for the subsequent plane optimization (see Figure 2(d)).

2.4 Roof plane estimation

The next step is to extrude the roof surfaces back to their original height in three dimensions. A simple approach is to use RANSAC or singular value decomposition to individually estimate the associated plane $\theta[f_i]$ for each face $f_i \in \mathcal{F}, i = 1, \ldots, N$, with parameters $\theta_i = (A_i, B_i, C_i)$. The drawback is that seam lines between neighboring faces such as ridges do not necessarily have to be closed, creating undesirable jumps across ridges. A better idea is to do a global estimation of plane parameters for each face simultaneously, and constrain the edges between neighboring faces to remain attached under certain conditions. Intuitively, we want to keep sloped faces attached to each other to create closed ridge lines, but let horizontal faces such as dormers or flat parts of the roof "float above" neighboring faces. To achieve this, we start with any

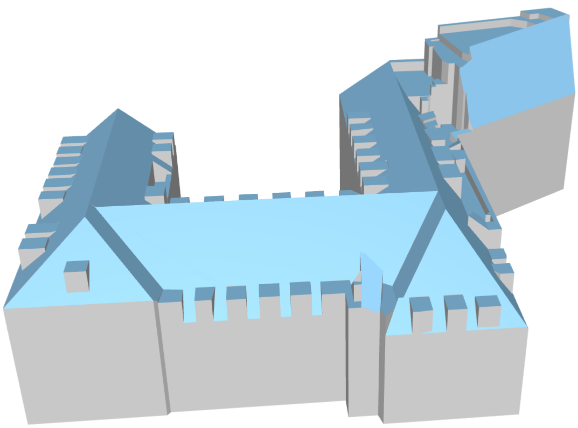


Figure 3. Final building model. Note the artifact for the dome in front due to model violations.

decent approximation of the plane parameters θ_i (again using SVD or RANSAC), such that a first approximation is available. We then define two angles ϕ_{flat} and $\phi_{\max}$ representing the maximum angle for which a polygon is considered flat, and the maximum slope we allow for roofs (before they become a wall). We define the set of flat faces as

$$\mathcal{F}_{\text{flat}} = \left\{ f \in \mathcal{F} \,\middle|\, \sqrt{A^2 + B^2} \leq \tan(\phi_{\text{flat}}) \right\} \qquad (5)$$

and set their plane equation to be perfectly horizontal with the average height of associated LiDAR points,

$$\theta[f](x, y) \equiv \frac{1}{|\mathcal{P}(f)|} \sum_{p \in \mathcal{P}(f)} z_p \quad \text{for } f \in \mathcal{F}_{\text{flat}}. \qquad (6)$$

All remaining faces $\mathcal{F}_{\text{sloped}} = \mathcal{F} \setminus \mathcal{F}_{\text{flat}}$ are considered (potentially) sloped and active for subsequent optimization. Using $t_{\max} = \tan(\phi_{\max})$, we define the set of slope constraints as[3]

$$\Theta = \left\{ \theta = (A, B, C)^T \in \mathbb{R}^3 \,\middle|\, |A| \leq t_{\max}, |B| \leq t_{\max} \right\} \qquad (7)$$

Then we define the set of plane parameters as $\mathbf{x} = \left\{ \theta[f] = (A, B, C) \in \Theta \,|\, f \in \mathcal{F}_{\text{sloped}} \right\}$ and we solve for $\mu > 0$,

$$\arg\min_{\mathbf{x}} E(\mathbf{x}) = E_{\text{data}}(\mathbf{x}) + \mu E_{\text{reg}}(\mathbf{x}). \qquad (8)$$

The data term E_{data} is a standard least-squares estimate of the plane parameters for each face, measuring the fit between the ideal plane $\theta[f_i]$ and the associated points $\mathcal{P}(f_i)$:

$$E_{\text{data}}(\mathbf{x}) = \sum_{f_i \in \mathcal{F}} \frac{1}{|\mathcal{P}(f_i)|} \sum_{p \in \mathcal{P}(f_i)} \left(\theta[f_i](x_p, y_p) - z_p \right)^2 \qquad (9)$$

Using the initial plane estimate, we compute the maximum vertical gap g_e along every edge segment $e \in \mathcal{E}$ to define weights for our edges, $w_e = |e| \cdot \exp\left(-g_e^2 / 2\sigma^2\right)$. The weight w_e is large for gaps smaller than $\sigma > 0$, then decays quickly to zero, putting emphasis on seam lines that are close together, while ignoring segments that are vertically far from each other. We set σ to be a small multiple of the LiDAR resolution to avoid a too aggressive bias towards fusing line seams. Weighting by the edge length $|e|$ puts more weight onto the ridges of large roof planes, which we prioritize to be closed. The regularization term E_{reg} then penalizes the weighted squared vertical difference between the plane height at the vertices for each edge $e \in \mathcal{E}$:

$$E_{\text{reg}}(\mathbf{x}) = \sum_{e = (f, f') \in \mathcal{E}} w_e \sum_{\mathbf{v} \in \mathcal{V}(e)} \left(\theta[f](\mathbf{v}) - \theta[f'](\mathbf{v}) \right)^2 \qquad (10)$$

If the difference vanishes for any pair of edges, they are entirely sealed. This approach resembles the work of (Verma et al., 2006), who create a roof topology graph similar to the one depicted in Figure 2(d) for the ridges and then estimate the roof planes while keeping the ridges closed. The main limitation of their approach is that they restrict themselves to a set of given roof shapes for which they can construct the topology graph; more complex roof shapes cannot be covered. We replace this restriction with a softer constraint, at the cost of possibly allowing ridges that cannot be closed entirely. In particular, since polygons are not required to be convex in our formulation, two faces can have multiple edges between them. In this case, there is not necessarily a solution where all vertex differences can vanish, and small gaps in ridge lines can remain.

It is worth pointing out that for any edge $e = (f, f')$ between a sloped face f and a horizontal face f', $\theta[f'](\mathbf{v})$ is a constant, as we do not estimate the parameters of horizontal planes; the sloped plane is merely encouraged to snap onto the horizontal boundary.[4] Problem (8) can be solved using classic techniques

[3] Note that constraining A and B separately instead of $\sqrt{A^2 + B^2}$ is easier for the optimization.

[4] The method can easily be extended to also estimate the z value of the

such as L-BFGS-B, and can even be solved in parallel if the edge graph $\mathcal{E}$ decouples into multiple disconnected subgraphs.

2.5 Roof extrusion

Following the planar optimization, the two-dimensional topological faces and their associated plane equations $\theta[f]$ are extruded into a three-dimensional building model. To ensure a continuous watertight mesh, roof vertices sharing the same 2D coordinates across adjacent faces must resolve to a unified elevation. We employ a Union-Find clustering algorithm to group identical (x, y) topological vertices. Their local plane-derived elevations are then averaged to establish a single canonical height, effectively eliminating vertical gaps between adjacent roof segments.

Using the harmonized heights, we systematically construct the manifold 3D geometry. Roof surfaces are elevated to 3D, explicitly preserving interior rings (such as dormers or skylights) as topological holes within the parent surface. Vertical discontinuities between adjacent roof segments are sealed with step walls. To maintain manifold constraints, these step wall quads are dynamically split at intermediate heights, preventing degenerate triangles and enforcing exact vertical alignment. Outer graph edges lacking a topological twin are extruded downwards to the ground plane to form exterior boundary walls. Internal building voids, such as courtyards, are extruded the same way to form interior walls. The resulting polygonal 3D models can be exported as CityGML/CityJSON or, after triangulation into meshes, as OBJ files.

3. Evaluation/Results

We evaluate our method on two datasets with different characteristics.

3.1 ISPRS Vaihingen benchmark

The ISPRS benchmark for urban object classification and 3D building reconstruction (Rottensteiner et al., 2012) provides airborne LiDAR data at approximately 4 pts/m^2 per swath, with partially overlapping swaths leading to locally higher but irregular point density. The benchmark covers three areas: Area 1 (“Inner City”, 38 buildings with complex historic roofs), Area 2 (“High Riser”, 15 high-rising residential buildings), and Area 3 (“Residential”, 57 small detached houses). We use the provided ground truth roof outlines as building footprints. Our pipeline produced results for 106 of the 110 reference buildings; four small buildings could not be included due to insufficient point density or excessive noise. We set $\lambda = 0.1$ (high detail) for all experiments.

3.2 Münster (Westf.) dataset.

To compare against state-provided LoD2 models in a practical setting, we selected two representative blocks in and around Münster (Westf.), Germany, and evaluate *all* buildings within each: (1) an urban block of 21 complex “Altbau” buildings with multi-part mixed-gable roofs covering approximately 28 000 m^2 (see Figure 4), and (2) a suburban block with 88 mostly single-family houses (simple gable, flat, and tent roofs) covering approximately 32 000 m^2 (see Figure 5). LiDAR point clouds ($\sim$10 pts/m^2), LoD2 models (latest campaign 2024), and cadastral building footprints (ALKIS) are all publicly provided by Geobasis NRW (Bezirksregierung Köln). Again, we set $\lambda = 0.1$ for all experiments.

horizontal planes, however, we did not find any benefit.

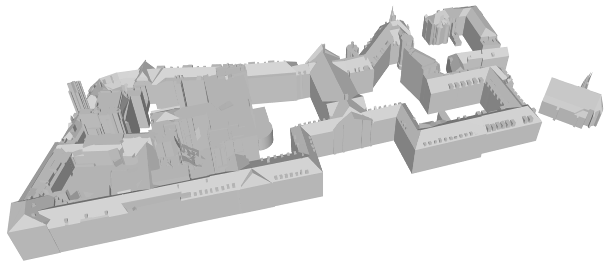


Figure 4. Reconstruction of an urban area in Münster (Westf.).

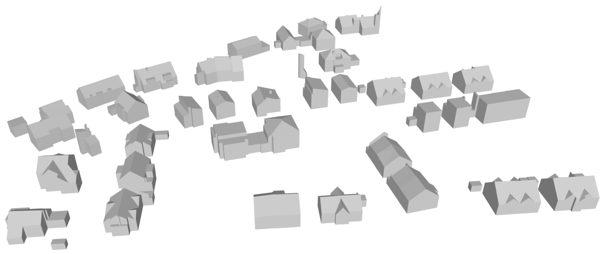


Figure 5. Reconstruction of a residential area in Münster (Westf.).

3.3 Evaluation Metrics

For the ISPRS benchmark, we follow the evaluation protocol of Rottensteiner et al. (2014). Since our method uses the provided ground truth roof outlines as input footprints, the area-based completeness, correctness, and quality metrics primarily reflect footprint coverage rather than roof reconstruction quality and are hence always close to 1, so we omit them. Geometric accuracy is assessed via RMSE in planimetry (RMSE$_{xy}$) and height (RMSE$_z$), both computed as vertex-to-vertex distances with a 3 m outlier threshold.

For the Münster dataset, no manual ground truth is available for the complex roof shapes. We therefore compute the point-to-surface RMSE: the root mean square of Euclidean distances from each LiDAR point to the nearest face of the reconstructed model, using the 95th percentile to mitigate the effect of noise and misclassified points. This metric is applied to both our LoD3-like models and the state LoD2 models. We additionally report the per-building win rate, i.e. the fraction of buildings for which our model achieves equal or lower RMSE than the LoD2 baseline.

3.4 Results on the ISPRS benchmark

Table 1 summarizes our results per-area along with the range of scores reported for other methods on the ISPRS benchmark leaderboard (Rottensteiner et al., 2014). Our RMSE$_{xy}$ is consistently below the range of competing methods, reaching less

Area	#	RMSE$_z$ [m]		RMSE$_{xy}$ [m]	
1	36	0.514	[0.2–0.6]	0.293	[0.8–1.1]
2	14	0.344	[0.1–1.2]	0.141	[0.3–1.2]
3	56	0.288	[0.1–0.6]	0.267	[0.6–1.1]
	106	**0.372**		**0.259**	

Table 1. Results on the ISPRS Vaihingen benchmark. Benchmark ranges in brackets indicate the spread of scores reported by other methods (Rottensteiner et al., 2014).

Area	#	RMSE (Ours) [m]		RMSE LoD2 [m]	
		mean	σ	mean	σ
Urban	21	0.120	0.079	0.678	0.648
Suburban	88	0.140	0.198	0.428	0.678

Table 2. Point-to-surface RMSE (95th percentile) on the Münster dataset. σ denotes the standard deviation within the dataset.

Method	Dataset	#	pts/m^2	RMSE [m]
Huang et al.	AHN3 (NL)	14	~ 8	0.128
Huang et al.	DALES (CA)	8	~ 50	0.184
Huang et al.	Vaihingen (DE)	57	~ 4	0.157
Ours	**Münster urban**	**21**	$\sim$ **10**	**0.120**
Ours	**Münster suburb.**	**88**	$\sim$ **10**	**0.140**

Table 3. Cross-dataset comparison of point-to-surface RMSE with Huang et al. (2022).

than half of the best value previously reported in Area 1. However, this score might be skewed due to our use of building footprints as input, and thus only judging the internal roof plane boundaries. RMSE$_z$ falls within the benchmark range but towards the upper end for Area 1, likely due to the complex historic roofs and low point density for some buildings. Among more recent work, Jung et al. (2017) report a planimetric RMSE of 0.76 m on the 110 full Vaihingen buildings, but do not report RMSE$_z$. Processing the 106 ISPRS Vaihingen buildings takes approximately 1.4 seconds per building on average on a standard MacBook Air M2 2022 without special focus on Python runtime optimization.

3.5 Results on the Münster (Westf.) dataset

Table 2 compares our building models against the state-provided LoD2 models. Our method reduces the mean RMSE by 82 % in the urban and 67 % in the suburban area compared to the state-provided LoD2 models. For more than 98 % of all buildings, our model achieves a lower point-to-surface error than the LoD2 baseline. Since our Münster metric (point-to-surface RMSE) matches the one used by Huang et al. (2022), a rough cross-dataset comparison is possible, though limited by differences in building types and point cloud characteristics (see Table 3). Our point-to-surface RMSE on complex urban roofs (0.120 m) is comparable to or lower than the values reported by Huang et al. (2022) across all their datasets, despite our evaluation covering all buildings in contiguous urban blocks rather than selected examples. We note that the RMSE of 0.157 m reported by Huang et al. (2022) for Vaihingen cannot be directly compared with our vertex-based ISPRS benchmark results (RMSE$_z$ $= 0.371$ m), as the two metrics measure fundamentally different quantities.

4. Conclusion

4.1 Discussion

While our overarching pipeline adapts the concept of discrete optimization and graph-cut frameworks proposed by Zebedin et al. (2008) and later adopted for large-scale modeling by Peters et al. (2022), our methodology introduces critical advancements. A primary limitation of the existing pipelines is the lack of explicit formulations both for the segmentation as well as the 3D roof extrusion phase; for example, the exact mechanics of transitioning from 2D topological partitions to a watertight 3D mesh are left unexplained. To address this ambiguity and improve overall geometric fidelity, we systematically improve upon their foundation through the following contributions.

First, we replaced the segmentation step with a highly robust piecewise affine-linear Mumford-Shah functional, achieving a highly accurate initial plane extraction that is resilient to LiDAR noise and varying point densities. Second, we modified the existing polygon regularization approach with a graph cut by introducing proper normalization and replacing the vertical distance by the orthogonal distance to the plane. Third, to resolve some of the extrusion ambiguity present in previous works, we introduce an explicitly formulated global roof plane estimation step. Rather than estimating the spatial parameters of each roof face independently, our framework optimizes all active, sloped roof faces simultaneously. By penalizing the vertical differences along shared topological edges via a gap-weighted regularization term, we encourage closed seam lines.

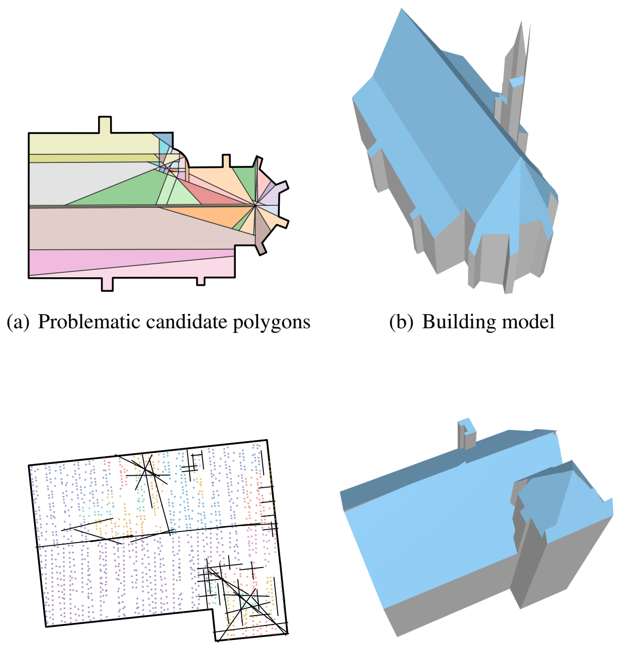
The experimental results demonstrate that the proposed piecewise affine-linear Mumford-Shah segmentation, coupled with discrete graph-cut regularization, provides a highly robust framework for automated 3D roof reconstruction. In particular, the segmentation delivers a highly accurate plane approximation which is crucial for the subsequent line detection. On the ISPRS Vaihingen benchmark, our method achieved a highly competitive planimetric RMSE of 0.27 m, highlighting the effectiveness of the segmentation as well as the subsequent steps. While the overall height RMSE (0.37 m) remains very competitive, the slight elevation in Area 1 suggests that the strict piecewise affine-linear assumption can sometimes struggle to perfectly approximate highly irregular, historic roof structures when point densities are low.

The evaluation on the Münster dataset underscores the practical superiority of our method over standard state-provided LoD2 models. By reducing the mean point-to-surface RMSE by 82 % in complex urban environments and 67 % in suburban areas, our pipeline proves that standard LoD2 primitives frequently fail to capture the geometric reality of multi-part roofs, dormers, and chimneys. Achieving a 98 % win rate against baseline LoD2 models confirms that our data-driven polygonization strategy yields significantly higher architectural fidelity. Furthermore, the ability to control the structural scale and transitioning seamlessly toward LoD3-like detail, using only three fixed, user-facing global parameters (γ, λ, μ) demonstrates robustness without the need for exhaustive per-dataset tuning.

Finally, the computational efficiency of the pipeline, averaging 1.1 seconds per building on standard hardware, proves the viability of this approach for large-scale urban reconstruction. Our deterministic optimization framework inherently enforces closed seams and strictly manifold geometry through globally constrained plane fitting, offering a scalable, reliable, and mathematically sound path toward nationwide LoD3 city modeling.

4.2 Limitations

Although our proposed methodology demonstrates high robustness, it is not without limitations. As discussed in Section 2, the Mumford-Shah segmentation inherently assumes that roof segments are piecewise planar. This assumption prevents the accurate representation of curved architectural features, such as domes, barrel roofs, or conical structures (see Figure 6(a) and 6(b)). Figure 6(a) illustrates the candidate polygons that



(a) Problematic candidate polygons (b) Building model

(c) Problematic line detection (d) Building model

Figure 6. Limitations of the method: Artifacts due to model violations (conic structures and domes), slither polygons due to faulty line detection, and incorrect jump lines due to irregularities and gaps in LiDAR data. Buildings are from Münster urban (above) and Vaihingen area 1 (below).

can arise from a bad segmentation due to steep conical structures and faulty subsequent line generation. A remedy for the incorrect plane assumption in some cases is the integration of surfaces of revolution (e.g. Zebedin et al., 2008), which could be selectively applied where the piecewise affine-linear segmentation shows high planar fitting errors. Furthermore, our approach processes both the LiDAR data and the resulting reconstruction purely as a 2.5D surface model, ignoring explicit wall points. Future iterations could incorporate a preprocessing step to detect and integrate vertical structures directly from the data.

Secondly, the fidelity of the final building reconstruction relies heavily on the robust estimation of 2D lines derived from the initial segmentation. Because these lines and their intersections strictly define the topological roof structure, insufficient or faulty line detection can occasionally lead to undesired geometric artifacts. On the tower structure in Figure 6(a) and 6(b), too many lines are generated for a small area, resulting in tiny polygons with very few LiDAR points, which the regularization in (2) cannot successfully remove. This is not a problem restricted to structures that fail to meet the modeling assumptions of being a plane, but can also arise in other places where too many lines generate tiny slither polygons that can cause ragged artifact lines. Such instances can typically be identified by a correspondingly high residual error against the original LiDAR point cloud. To mitigate this, the pipeline can be iteratively restarted with increased regularization parameters, promoting fewer, larger segments and, consequently, more reliable boundary lines.

In addition, imprecise roof lines can introduce minor topological anomalies during the polygonization and regularization phases. For example, over-extended boundary lines may generate artificially small corner polygons. This can be resolved directly during the construction of the half-edge graph by either isolating these anomalies as separate planar surfaces or dynamically removing the vertices from the affected polygons.

Lastly, the inherent quality, resolution, and density of the input LiDAR data fundamentally constrain the reconstruction detail. Figures 6(c) and 6(d) illustrate the problems that can arise due to irregular LiDAR sampling and gaps in the LiDAR data in the Vaihingen dataset; for large gaps, it is hard to place lines, in particular step lines, at the correct location, which can be observed in the tower area. To ensure adaptability, all geometric distance thresholds within our pipeline, in addition to the core regularization weights (γ, λ, μ), can be dynamically scaled as a multiple of the estimated LiDAR resolution. In practice, we utilize the 95th percentile of nearest-neighbor distances to robustly estimate the local point cloud resolution, effectively accounting for density variations across different roof slopes. However, naturally, as the point cloud becomes coarser, a reliable estimation of topological boundary lines becomes increasingly challenging. A possible remedy here is the use of lines detected from other imagery, as has been demonstrated by Zebedin et al. (2006, 2008) which, however, requires substantially more data and comes with its own challenges. It should be noted that the use of existing cadastral building footprints is a rather small restriction and can be replaced by any estimate of the building footprint (e.g. Jung et al., 2017), to also unlock roof overhangs or possible balconies that extend over the cadastral footprint.

4.3 Future work

Future research will primarily focus on scaling the proposed methodology from regional datasets to a nationwide application in Germany. This expansion requires adapting the pipeline to robustly handle the inherent heterogeneities in ALS point densities and the varying data quality provided by different federal state agencies. To overcome the current algorithmic limitations associated with the piecewise affine-linear assumption, future iterations will explore the integration of generalized non-planar primitives to accurately reconstruct non-planar architectural features, such as domes and barrel roofs. To further enhance the geometric fidelity of the models, roof overhangs can also be included. A practical approach is to buffer the LiDAR points by some margin around the footprint and trace the roof planes that extend over the footprint until they end. The difference between the LoD building presented in this work and the extended planes is a good candidate for the automatic modeling of roof overhangs and eaves. Finally, transitioning our highly detailed roof models and walls into complete Level of Detail 3 representations remains a primary objective. We intend to enrich the generated building models with detailed facade semantics by incorporating automated window detection and wall texturing pipelines. This will be achieved by using complementary data sources such as oblique imagery (e.g. Xia et al., 2025), moving the framework toward the fully automated generation of comprehensive and visually realistic 3D city models.

Acknowledgements

This work is part of the ENERsyte project, funded by the state of North Rhine-Westphalia through the EFRE/JTF program “Grüne Gründungen” with funding code EFRE-20800322 and project number IN-GG-1-007a.

References

Agoub, A., Schmidt, V., Kada, M., 2019. Generating 3D City Models Based on the Semantic Segmentation of Lidar Data Using Convolutional Neural Networks. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-4-W8, 3–10.

Bauchet, J.-P., Lafarge, F., 2019. City Reconstruction from Airborne Lidar: A Computational Geometry Approach. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-4-W8, 19–26.

Dochev, I., Seller, H., Peters, I., 2020. Assigning Energetic Archetypes to a Digital Cadastre and Estimating Building Heat Demand. An Example from Hamburg, Germany. *Scientific Journal of Riga Technical University. Environmental and Climate Technologies*, 24, 233–253.

Forth, K., Noichl, F., Borrmann, A., 2024. LCA Calculation of Retrofitting Scenarios Using Geometric Model Reconstruction and Semantic Enrichment of Point Clouds and Images. *Computing in Civil Engineering 2023*, American Society of Civil Engineers, Corvallis, Oregon, 390–397.

Hanke, F., Bieringer, A., Wysocki, O., Jutzi, B., 2025. CM2LoD3: Reconstructing LoD3 Building Models Using Semantic Conflict Maps. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4-W6-2025, 81–88.

Huang, J., Stoter, J., Peters, R., Nan, L., 2022. City3D: Large-Scale Building Reconstruction from Airborne LiDAR Point Clouds. *Remote Sensing*, 14(9).

Jung, J., Jwa, Y., Sohn, G., 2017. Implicit Regularization for Reconstructing 3D Building Rooftop Models Using Airborne LiDAR Data. *Sensors*, 17(3).

Kiefer, L., Storath, M., Weinmann, A., 2020. PALMS Image Partitioning - A New Parallel Algorithm for the Piecewise Affine-Linear Mumford-Shah Model. *Image Processing On Line*, 10, 124–149.

Kolbe, T. H., Gröger, G., Plümer, L., 2005. CityGML: Interoperable Access to 3D City Models. P. Van Oosterom, S. Zlatanova, E. M. Fendel (eds), *Geo-information for Disaster Management*, Springer Berlin Heidelberg, Berlin, Heidelberg, 883–899.

Li, P., Zhang, H., Guo, Z., Lyu, S., Chen, J., Li, W., Song, X., Shibasaki, R., Yan, J., 2021. Understanding Rooftop PV Panel Semantic Segmentation of Satellite and Aerial Images for Better Using Machine Learning. *Advances in Applied Energy*, 4, 100057.

Liu, Y., Obukhov, A., Wegner, J. D., Schindler, K., 2024. Point2Building: Reconstructing Buildings from Airborne LiDAR Point Clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 215, 351–368.

Luft, J., Bagherinejad, N., Razavian, A., Reinke, J., Sommer, P., Rasch, J., 2025. Heat Planning with a Single Click - Using Open Data for a Scalable Planning Tool. *Tagungsband der 3. Konferenz der Norddeutschen Wärmeforschung*, Bremen, 60–65.

Mei, Q., Bulatov, D., Iwaszczuk, D., 2025. Polygonizing Roof Segments from High-Resolution Aerial Images Using Yolov8-Based Edge Detection. *Proceedings of the 20th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 39–49.

Meixner, P., Leberl, F., 2010. Interpreting Building Façades from Vertical Aerial Images Using the Third Dimension. *Proc. Joint Symposium of ISPRS Technical Commission IV \& AutoCarto*, 15–19.

Mumford, D., Shah, J., 1989. Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems. *Communications on Pure and Applied Mathematics*, 42(5), 577–685.

Nan, L., Wonka, P., 2017. PolyFit: Polygonal Surface Reconstruction from Point Clouds. *2017 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Venice, 2372–2380.

Ni, H., Wang, D., Zhao, W., Jiang, W., Mingze, E., Huang, C., Yao, J., 2024. Enhancing Rooftop Solar Energy Potential Evaluation in High-Density Cities: A Deep Learning and GIS Based Approach. *Energy and Buildings*, 309, 113743.

Peters, R., Dukai, B., Vitalis, S., van Liempt, J., Stoter, J., 2022. Automated 3D Reconstruction of LoD2 and LoD1 Models for All 10 Million Buildings of the Netherlands. *Photogrammetric Engineering & Remote Sensing*, 88(3), 165–170.

Rottensteiner, F., Sohn, G., Gerke, M., Wegner, J., Breitkopf, U., Jung, J., 2014. Results of the ISPRS Benchmark on Urban Object Detection and 3D Building Reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 256–271.

Rottensteiner, F., Sohn, G., Jung, J., Gerke, M., Baillard, C., Benitez, S., Breitkopf, U., 2012. The ISPRS Benchmark on Urban Object Classification and 3D Building Reconstruction. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-3, 293–298.

Schuegraf, P., Fuentes Reyes, M., Xu, Y., Bittner, K., 2023. ROOF3D: A Real and Synthetic Data Collection for Individual Building Roof Plane and Building Sections Detection. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-1-W1-2023, 971–979.

Verma, V., Kumar, R., Hsu, S., 2006. 3D Building Detection and Modeling from Aerial LIDAR Data. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 2, 2213–2220. ISSN: 1063-6919.

Vosselman, G., Dijkman, S., 2001. 3D Building Model Reconstruction from Point Clouds and Ground Plans. *Proceedings of the ISPRS Workshop: land surface mapping and characterization using laser altimetry, 22-24 October 2001, Annapolis, Maryland*, International Society for Photogrammetry and Remote Sensing (ISPRS), 37–43.

Xia, Y., Gao, W., Stoter, J., 2025. Enriching LoD2 Building Models with Façade Openings Using Oblique Imagery. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W6-2025, 225–232.

Zebedin, L., Bauer, J., Karner, K., Bischof, H., 2008. Fusion of Feature- and Area-Based Information for Urban Buildings Modeling from Aerial Imagery. D. Forsyth, P. Torr, A. Zisserman (eds), *Computer Vision – ECCV 2008*, Springer, Berlin, Heidelberg, 873–886.

Zebedin, L., Klaus, A., Gruber-Geymayer, B., Karner, K., 2006. Towards 3D Map Generation from Digital Aerial Images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(6), 413–427.