UNLOCKING POINT CLOUD POTENTIAL: FUSING MLS POINT CLOUDS WITH SEMANTIC 3D BUILDING MODELS WHILE CONSIDERING UNCERTAINTY

O. Wysocki*, Y. Xu*, U. Stilla*

Photogrammetry and Remote Sensing, Technical University of Munich, 80333 Munich, Germany
(olaf.wysocki, yusheng.xu, stilla)@tum.de

KEY WORDS: MLS Point Cloud, Semantic 3D City Model, Data Fusion, Bayesian Network, Façade Refinement, Uncertainty

ABSTRACT:

Throughout the years, semantic 3D city models have been created to depict 3D spatial phenomenon. Recently, an increasing number of mobile laser scanning (MLS) units yield terrestrial point clouds at an unprecedented level. Both dataset types often depict the same 3D spatial phenomenon differently, thus their fusion should increase the quality of the captured 3D spatial phenomenon. Yet, each dataset has modality-dependent uncertainties that hinder their immediate fusion. Therefore, we present a method for fusing MLS point clouds with semantic 3D building models while considering uncertainty issues. Specifically, we show MLS point clouds coregistration with semantic 3D building models based on expert confidence in evaluated metadata quantified by confidence interval (CI). This step leads to the dynamic adjustment of the CI, which is used to delineate matching bounds for both datasets. Both coregistration and matching steps serve as priors for a Bayesian network (BayNet) that performs application-dependent identity estimation. The BayNet propagates uncertainties and beliefs throughout the process to estimate end probabilities for confirmed, unmodeled, and other city objects. We conducted promising preliminary experiments on urban MLS and CityGML datasets. Our strategy sets up a framework for the fusion of MLS point clouds and semantic 3D building models. This framework aids the challenging parallel usage of such datasets in applications such as façade refinement or change detection. To further support this process, we open-sourced our implementation.

1. INTRODUCTION

MLS point clouds are characterized by high temporal resolution, density and relative 3D point accuracy. These traits make them ideal sources for a myriad of applications. However, MLS point clouds' low absolute 3D point accuracy hinders their usability in tasks such as a refinement of 3D city models wherein overlying datasets' spatial consistency is pivotal. The absolute 3D point accuracy especially varies within urban environments due to factors such as multipath and non-line-of-sight (NLOS) signal phenomena for GNSS receivers (Zhang et al., 2018, Lucks et al., 2021). Furthermore, the point clouds' unstructured nature and lack of semantics stand as challenging tasks for building reconstruction purposes (Xu and Stilla, 2021).

On the contrary, semantic 3D city models represent a structured data of rich semantic information and can reach absolute 3D point accuracy of centimeter levels. Such 3D city models are widely available not only as proprietary or governmental but also as open datasets[1]. The semantic 3D city models have already proved to yield credible features in a wide range of applications (Biljecki et al., 2015). Yet, they are mostly represented by generalized planar geometries. While that kind of representation may be valid for various applications, it may be not sufficient when it comes to tasks requiring very detailed geometries. For example, such representation does not suffice for testing of complex physical sensor effects of automated driving functions (Wysocki et al., 2021, Schwab and Kolbe, 2019) or may limit achievable accuracy of solar potential analysis (Willenborg et al., 2018).

Therefore, it is believed that fusion of MLS point clouds and semantic 3D city models should yield enhanced quality of 3D spatial information, deriving from the definition that the data fusion should result in a maximization of data potential that simultaneously decreases their limitations (Hall and Llinas, 1997). Since *enhanced quality* is a nebulous term, it should be defined
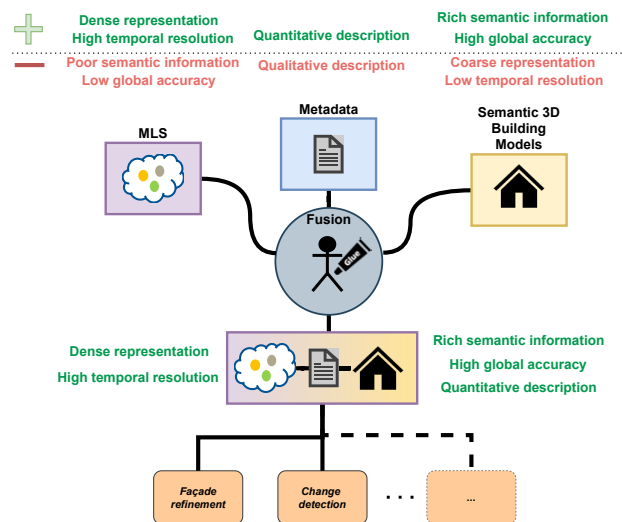
[1] https://github.com/OloOcki/awesome-citygml

Figure 1. The fusion supported by expert knowledge maximizes the potential of semantic 3D building models and MLS point clouds in a myriad of applications.

as application-specific (Schmitt and Zhu, 2016). Additionally, the fusion of datasets is inevitably connected to their inherited uncertainties which, while not appropriately addressed, may impact the interpretation of the final results (Anderson et al., 2017). Such uncertainties are caused by for example global registration accuracy, acquisition technique, or vaguely described metadata. While application-specific requirements and inherited uncertainties are implicitly expressed, they should be addressed by a skilled operator. The operator's knowledge and metadata should serve as a binder for the input datasets. To this end, we present our work with contributions listed as follows:

- We propose the strategy for multimodal datasets, fusion of

MLS point clouds and semantic 3D city models in a challenging urban environment.

- We present the method coregistering datasets by addressing uncertainty issues by incorporation of CIs estimated on a basis of expert knowledge and dataset specification. This increases method flexibility and avoids pitfalls of predecessors using fixed thresholds.

- Matching steps are developed, adapting to measures obtained in the coregistration part to update CIs and set bounds for the matching. Simultaneously, the semantic information is transferred to unstructured MLS point clouds. Whereas, MLS point clouds enhance geometrical representation of building models.

- Attribute estimation is performed in the designed modular BayNet that propagates uncertainties throughout the network to the target and intermediate nodes. This allows for propagation of uncertainties throughout the process and seamless application-specific expansion of the strategy.

Preliminary experimental results on MLS point clouds and CityGML building models in urban environment are given as well, demonstrating promising performance. As such, our approach shows how to explore the potential of simultaneous usage of MLS point clouds and semantic 3D city models addressing the uncertainty of both datasets (as illustrated in Figure 1). The implementation is available within the attached repository[2].

2. RELATED WORK

Generally, a typical data fusion process is composed of several steps. Firstly, heterogeneous datasets need to be coregistered to achieve spatial harmonization. Then, datasets should be matched to create an association among them. These steps are rather generic for most of applications. Yet, the attribute or identity estimation is an application-specific operation. This is a core step of data fusion as it defines what information is distilled in the entire process (Schmitt and Zhu, 2016).

2.1 Coregistration & matching

There are three main strategies of multiple-points sets' registration for building reconstruction purposes, namely: point-based, primitive-based, and global-based (Xu and Stilla, 2021). The point-based strategy aims to find corresponding pairs of points from different point clouds, whereas primitive-based registration searches for geometric primitives shapes within point clouds to which correspondence can be found. The global-based approaches, however, neglect local information and focus on global features for entire point clouds.

One of the well-established examples of point-based strategy is the iterative closest point (ICP) algorithm and its variations point-to-point (Besl and McKay, 1992) and point-to-plane (Rusinkiewicz and Levoy, 2001). Although it is a generic algorithm feasible in many applications, it is challenging to apply it for point clouds depicting an outdoor environment (Dong et al., 2020). The main obstacles involve varying density, low overlap, and occlusions, among others (Dong et al., 2020, Xu and Stilla, 2021). On the contrary, primitive-based solutions seem

to be tailored to the harsh nature of registration in an urban environment. This is implied by many artificial structures localized within urban areas that are often represented by geometrical primitives. So far researchers have investigated possibilities to formulate the registration process on the basis of found edges, crest curves, planes, and other geometric primitives (Xu and Stilla, 2021). However, this strategy encounters another set of challenges such as quality and consistency issues of extracted primitives or high computational time (Xu and Stilla, 2021). While numerous works tackled the challenge of terrestrial laser scanning (TLS) point cloud registration (Dong et al., 2020, Xu and Stilla, 2021), the registration of such point clouds based on semantic 3D city models seems to be obscured.

According to our meticulous research, only a few research groups specifically tackled the coregistration and matching of point clouds and semantic 3D city models (Goebbels and Pohle-Fröhlich, 2018, Goebbels et al., 2018, Goebbels et al., 2019, Lucks et al., 2021). While (Goebbels and Pohle-Fröhlich, 2018, Goebbels et al., 2018) incorporate a Mixed Integer Linear Program to find correspondences between modalities, the latest work uses a modified ICP point-to-plane (Goebbels et al., 2019). Yet, the publications of Goebbels et al. focus on point clouds generated from images using the structure from motion (SfM) algorithm. This implies the usage of radiometric features for prefiltering (Goebbels et al., 2019) that may remove valid building features. (Lucks et al., 2021) incorporate the ICP point-to-plane algorithm for matching of MLS point clouds and semantic 3D city models. To increase the matching accuracy, they introduce random forest to select only point clouds' points depicting façades. The predecessors assume absolute 3D point accuracy of point clouds at a fixed value (Goebbels et al., 2018, Goebbels et al., 2019, Lucks et al., 2021) that is believed to limit the scalability of the developed methods. Moreover, the analysis of the 3D models' coverage by point clouds is not presented (Goebbels et al., 2018, Lucks et al., 2021).

2.2 Identity estimation

Each measurement and subsequent fusion is inevitably related to uncertainty issues. This might be expressed by, for example, acquisition technique, registration accuracy, or expert belief in metadata (Chuprikova, 2019). One of the measures of uncertainty in the observed dataset is the CI (Chuprikova, 2019). Within the scope of 3D building reconstruction, the CI measure is used, for example, to accommodate for a map, roof extensions, and feature extraction inaccuracies (Suveg and Vosselman, 2000). However, this measure is often obscured in the domain of map accuracy assessment, which then results in false full certainty (Anderson et al., 2017).

To tackle the actual fusion step, several approaches have been introduced. The approach of (Hebel et al., 2013) of single sensor measurements fusion of airborne laser scanning (ALS) point cloud is based on Dempster–Shafer theory (DST). In this approach, the scene is partitioned into voxels that are empty or occupied. This is defined on a basis of ray tracing. The degree of ignorance introduces uncertainty to spaces introducing another state - unknown. (Hebel and Stilla, 2012) investigate the registration accuracy for change detection and perform an automatic alignment on the fly and a simultaneous calibration of the laser systems. On the contrary, (Gehrung et al., 2017) favor the Bayesian approach over the DST as the latter may yield non-intuitive results. Their work fuses single sensor measurements of MLS point clouds to remove moving objects. The main idea behind the concept is to create a voting process for solid objects. Firstly, the occupancy probabilities for voxels are accumulated. Subsequent decreasing happens if in the following

---

[2] https://github.com/OloOcki/fusing-mls-with-blds

timestamp these are unoccupied. Their approach is based on the work of (Hornung et al., 2013) where a concept of *OctoMap* had been introduced. Building upon the concept of the space occupation, (Tuttas et al., 2015) introduce validation of BIM elements using photogrammetric point clouds. Also within this work the *OctoMap* (Hornung et al., 2013) framework is utilized. Here multimodal datasets need to be analyzed. The occupancy grid concept combined with additional measures is introduced. However, the BIM model area of possible measurement position is represented by a fixed bounding box, thus, statically addressing BIM model accuracy uncertainties.

Another approach to data fusion is the BayNet, which is an acyclic graph incorporating joint probability distribution (Stritih et al., 2020, Chen and Pollino, 2012). This probabilistic concept allows propagating the uncertainties throughout the fusion process (Stritih et al., 2020). Moreover, it can incorporate uncertainties expressed by operator beliefs as well as quantitative and qualitative measures. The BayNets have already proved that they are suitable for modeling of complex environmental systems within a Geographic Information System (GIS) community (Chen and Pollino, 2012, Stritih et al., 2020, Chuprikova, 2019). However, the modeling of uncertainty propagation throughout the point clouds and semantic 3D models fusion process seems to be under-explored.

3. STRATEGY AND METHODS

Therefore, to address the aforementioned challenges, we present a strategy shown in Figure 2. The strategy aims to fuse MLS point clouds with semantic 3D building models addressing the uncertainty issue. Firstly, the metadata and expert beliefs shall be determined according to input datasets. Then, estimation of priors is conducted (see subsection 3.1). Following the alignment and matching methods the workflow leads to BayNet (see subsection 3.2), where an identity estimation is conducted. Since the fusion shall be steered by an end goal, we present possible applications (see subsection 3.3) and preliminary experiments at the end of the paper (see section 4.).

We use semantic 3D building models in the CityGML standard (Gröger et al., 2012) as a representation of a group of semantic 3D city models. The standard is an application schema of the Geography Markup Language (GML) issued by the Open Geospatial Consortium (OGC) (Gröger et al., 2012). The strategy can be seamlessly adjusted to other city objects and other encoding such as CityJSON.

3.1 Estimation of priors

The first essential step in the data fusion process is to achieve satisfying alignment of heterogeneous datasets. In this case, the associated uncertainties are errors of modalities' absolute 3D registration. For example, in the case of a facade-based registration, the uncertainty can be defined on a basis of footprints' absolute 3D point accuracy, whereas for point clouds this can be the absolute 3D point registration accuracy. Both pieces of information can be extracted from the metadata. However, often the metadata is sparse and the belief of the skilled operator needs to be incorporated. For matching, however, not only coregistration accuracy but also other factors impact uncertainty. For instance, CityGML Level of Detail (LoD) (Gröger et al., 2012) plays an important role in the expert's confidence in a generalization level of a depicted phenomenon. Depending on the situation at hand, this confidence might be intertwined with the coregistration deviations and respective errors should propagate throughout the workflow.
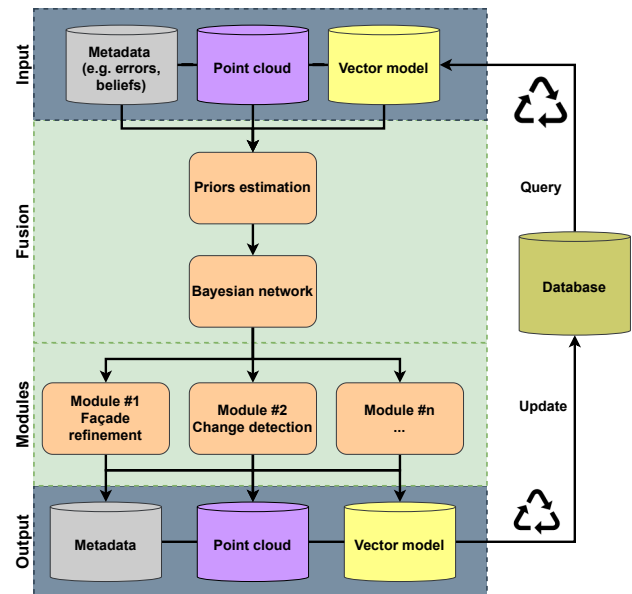

Figure 2. Strategy overview: database (green) query of vector model (yellow), point cloud (purple), and metadata (gray) should trigger the workflow (blueish) and subsequent processes (orange) ultimately leading to the database update.

The strategy also allows for adding other factors of uncertainty, as shown in Figure 3.

To quantitatively measure the uncertainty the CI is used and to estimate it certain set of parameters is required. These depend on the specified uncertainties and input datasets. The most important are: confidence level (CL) with associated z value (z), standard deviation (σ), and mean (μ). Naturally, for spatial data, the estimations are performed via the L2 norm. While this is computationally expensive, here, an approach for estimation in the L1 norm is preferred. This overestimates the CI range while still incorporating desired measures. Moreover, assuming Gaussian distribution, the next tier of computation simplification is added. The estimated upper bounds of confidence are discretized to L1 norm 3D buffers calculated based on the building's polygons positions. Such 3D boxes are used to set bounds for the coregistration process, as shown in Figure 4. Then they are dynamically altered to the size set up by the matching process.

Firstly, the alignment of semantic 3D models and MLS point clouds is performed. This operation consists of several steps. The workflow starts with the estimation of CI based on the introduced parameters and data. Then, the feature coverage analysis is performed to select targets that are eligible for the estimation of the matrix that ultimately rectifies global position accuracy. As a result the position is altered and associated errors are passed to the matching process.

3.1.1 Confidence intervals estimation: For the MLS point clouds and semantic 3D models, several pivotal uncertainties are defined. The expert should incorporate the known error (e_1) of MLS point cloud global registration. The error is expressed as an anticipated maximal deviation from the true value. The certainty of the anticipated deviation is calculated based on the confidence level (CL_1) introduced by an operator. The upper bound and confidence level serve to find the standard deviation for the MLS measurement (σ_1). The same idea is employed to the reference buildings. Also here building error is the assumed maximal discrepancy between the true value and a model (e_2). The belief in the error prior information is expressed by confidence
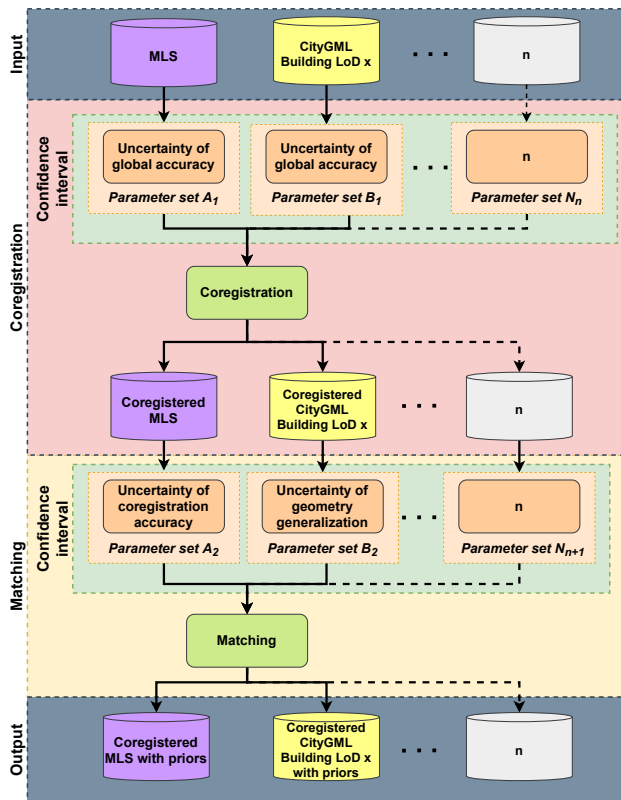
Figure 3. Estimation of priors: point clouds (purple), vector datasets (yellow), and defined uncertainties (orange) are the core elements of the process that is adaptable to new insights (gray).

level (CL_2). The found CI results in the estimation of the standard deviation for the semantic 3D building model (σ_2). The total standard deviation (σ) is found using the formula (Suveg and Vosselman, 2000):

$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2} \qquad (1)$$

The final upper and lower bounds are found on a basis of $[\mu - 2\sigma, \mu + 2\sigma]$. This is valid for the inaccuracies that are assumed to represent Gaussian distribution and overestimate the bounds by operating in the L1 norm. The rationale for the presented estimations derives from the work of (Suveg and Vosselman, 2000). The bounds might be perceived as 3D boxes in the discretized space as shown in Figure 4.

3.1.2 Feature coverage analysis: At this stage, it is feasible to estimate the coverage of the building's features by MLS point clouds. The idea is loosely based on the work of (Wysocki et al., 2021). The main difference is that the current algorithm operates in 3D space conducting more robust estimations whereas the aforementioned algorithms consider density and uniformity estimation solely within 2D space.

The first estimation determining adequately covered building features is a density measure. It calculates a total number of points per building feature within the introduced 3D boxes. Then a threshold rejecting 3D boxes not reaching the 60th percentile value is executed. The second tier measurement is the uniformity estimation introduced to neglect 3D boxes with small and densely populated concentrations of points. This measure is represented by a ratio $r \in [0, 1]$ of the MLS point cloud volume within the 3D box (v_1) to the total volume of the 3D box (v_2).
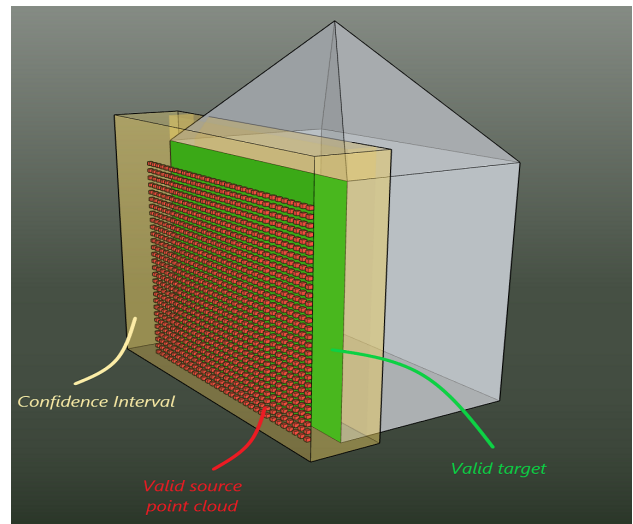
$$r = \frac{v_1}{v_2} \qquad (2)$$

Figure 4. CI (yellow) discretized in L1 norm for a valid LoD 2 building model wall (green) and respective MLS point cloud (red).

The estimated ratio should exceed a value of 0.6 to pass a threshold for further processing of a particular 3D box. Within the 3D box for MLS point clouds, a RANdom SAmple Consensus (RANSAC) algorithm is applied to find a vertical-like plane within the shrunken area to maximize the correct correspondence (Wysocki et al., 2021).

The previous steps yield two point clouds: representing valid 3D building features (target) and respective MLS point cloud (source), as illustrated in Figure 4. Since our solution requires no campaign trajectory, we present an alternative solution to find valid normal directions. The normal values are calculated both for target and source point clouds in a homogeneous local coordinate system originating in the center of a scene. This step reduces large float numbers and computational time and assures consistent normals directions.

3.1.3 Coregistration: The situation in hand is now tailored for the ICP point-to-plane algorithm (Rusinkiewicz and Levoy, 2001). It is because the *plane* is a target set, which in this case, is represented by (sampled) planar surfaces of semantic 3D building models. On the other hand, *point* is an MLS point cloud filtered to represent planar elements of a 3D building model. Additionally, the even voxelization at rate 0.1 [m] for both target and source point cloud diminishes the effect of uneven point cloud distribution. It also adapts to changing sizes of input 3D boxes. As the terrain depicted by point clouds is uneven on the contrary to bottom edges of 3D models geometries, this might cause false Z-coordinate rectification. To avoid that, the height rectification is performed by estimation of mode height within a 3D box of the introduced CI.

Since the point clouds are assumed to be coarsely registered, the initialization matrix is represented by the identity matrix to not alter the initial position of the point cloud. The maximal corresponding distance corresponds to the CI (see subsubsection 3.1.1). The convergence criteria are met if the fitness score or relative root mean square error (RMSE) reaches $1 * 10^{-6}$ and performs maximal 30 iterations. The result of registration is a matrix applied back to the whole object in question and associated error. Therefore, even slight global position enhancements are encompassed in the subsequent matching step.

3.1.4 Matching: The matching extensively uses advances of the coregistration process. The matching adapts CIs and discretizes them again to the 3D boxes (as depicted in Figure 4). Since the raw ultimate misalignment is curbed, the inaccuracies are lower also and the confidence level is higher too. The quantitative measure, RMSE, of the error is inherited from the coregistration process and can be directly incorporated. New borders (the 3D boxes) delineate matching bounds. The CIs serve as transmitters of the semantics and geometry between 3D building models and respective point clouds. Such 3D boxes of rich semantic, geometry, and incorporated uncertainty serve as priors for the BayNet.

3.2 Bayesian network

The data fusion method that considers uncertainties needs to quantify them throughout the process. In this work, the BayNet that explicitly maps uncertainties is incorporated into the workflow. The designed BayNet consists of 5 nodes with 2 target ones for estimation of occupied spaces of 3D building models. The network is visualized in Figure 5. The nodes have 2 mutually-exclusive states *occupied* and *unoccupied*. The fusion should couple datasets confirming a certain phenomenon. Thus, the unknown state is neglected assuming full visibility. For 2 states a joint probability distribution is calculated $P(X, Y)$. This is expressed by conditional probability table (CPT) where the probability distribution of each node for each combination of its parent nodes' states is prescribed. This defines causal relationships between the nodes and results in belief estimation. The relationship is visualized by edges in Figure 5. To calculate a probability for a selected state the marginalization process is used (Kjaerulff and Madsen, 2008, Stritih et al., 2020).

For such compiled networks, the obtained priors in the matching step serve as a basis for the BayNet. Such data is referred to as soft evidence as all of the priors contain explicit uncertainties expressed by CIs. However, the presented strategy also allows for qualitative input or hard evidence that has underlying 100% certainty. This makes the network flexible for new insights and pieces of evidence. The input uncertainty is propagated through the inference process (Stritih et al., 2020). The process estimates a posterior probability distribution (PPD) for each node in the network. This yields the expected state and related uncertainty for target nodes.

As shown in Figure 5, the first designed node in the network is *Occupied spaces for building envelope*. This node takes as input MLS point cloud and CityGML dataset at given global registration error with associated confidence and generalization range of 3D models, respectively. This node propagates towards the possible position and associated elements of a building. Roof and façade can have unmodeled surfaces such as dorms for roof and balconies for façade elements. Thus, further separation of building elements based on semantics leads to split for elements belonging to a wall (*Occupied spaces for wall and its elements*) or a roof (*Occupied spaces for roof and its elements*). This second tier of nodes narrows the possibilities of false associations. Both nodes lead to separate *Occupied spaces for building walls* and *Occupied spaces for building roofs* target nodes of the BayNet. To the target nodes, the additional raw geometries of the CityGML model are added as soft evidence. This time each of the evidence has only global registration error (without generalization) and belief to check whether the explicitly modeled geometries are confirmed in the reality. If there is no confirmation then there is a high chance that an MLS observation is missing or there is no building feature present (P_{low}). On the contrary, if there is a high probability of the occupation, the elements are confirmed and probably do not require refinements
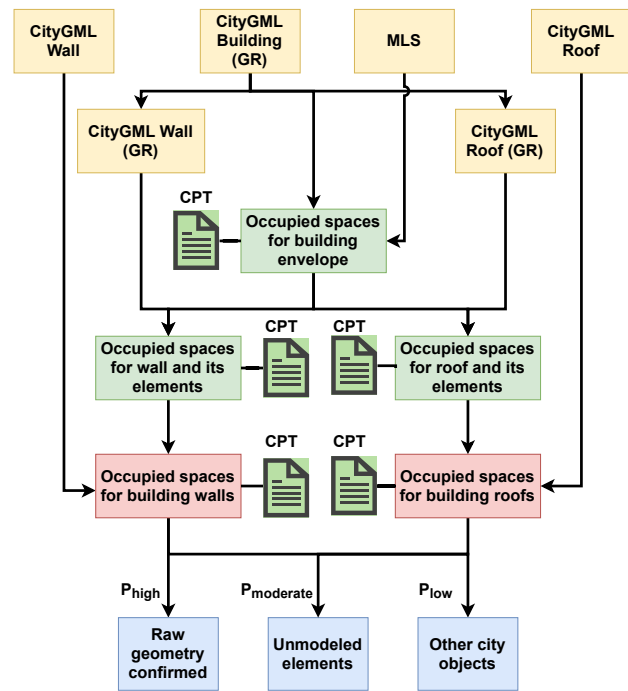

Figure 5. The BayNet: Target nodes (red), soft evidence (yellow), and nodes (green) with CPT. GR stands for generalized range of city model with associated uncertainty.

and shall be fused (P_{high}). The in-between probability measure indicates discrepancies between a model and MLS measurements that should trigger further application-specific investigations ($P_{moderate}$). Therefore, based on the evaluated end probabilities, the distinction to confirmed, unmodeled, and other city objects is made. The BayNet enables to seamlessly rebuild the structure, add new components, and obtain results at the intermediate steps as well.

3.3 Fusion

The threshold probabilities in Figure 5 are generic since they should depend on an end application. The fusion itself aims to couple confirming object parts. Thereby, only the output of *Raw geometry confirmed* in Figure 5 should be used for such purposes. Also, P_{high} and $P_{moderate}$ thresholds may be used for façade reconstruction purposes, whereas this might not be the case for change detection tasks. For example, the fusion process assumes that 3D building models exist and seek confirmation of the structure in MLS measurements. Thereby, for an outdated city models a further application-specific module has to be designed. Primarily, P_{low} should reject areas of less priority. As such, other city objects (e.g., road) and building parts unconfirmed by MLS observations are considered of low importance. The rationale for this is that occlusions are blocking spatial information acquisition. This, however, limits 3D processing capabilities within these areas and as such should be rejected.

Moreover, the presented publication introduces CIs discretized to 3D boxes. As shown in Figure 6, these might be used directly as soft evidence or further discretized to 2D patches, voxels, or octrees. This process, however, shall avoid element-wise discretization. Also, point clouds are often not parallel to walls (as for simplicity shown in Figure 6). In such cases the additional intersection conflicts have to be solved. To avoid this, a scene-wise partition is designed. Furthermore, the discretization method should be application-specific. The same applies to a discretization unit size. For example, detection of discrepancies due
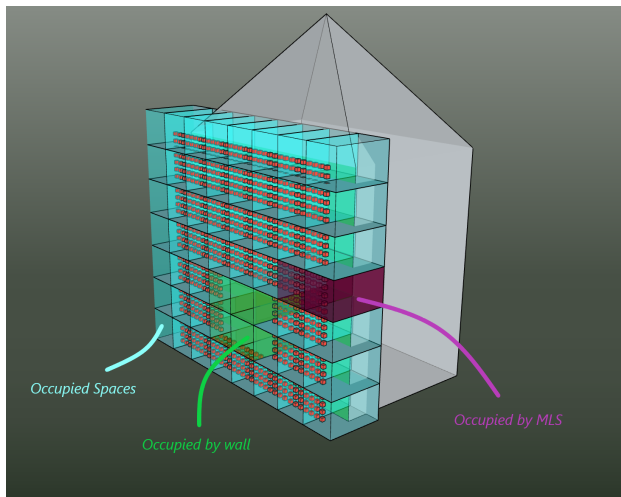
Figure 6. Visualization of voxelized CI space: Confirmed raw geometries with P_{high} (blue), unmodeled elements with $P_{moderate}$ (purple), and P_{low} for other city objects (green).

to wall openings will not yield plausible results when using 2D patches of 5.0 [m] x 5.0 [m] size which should not be the case for 3D voxels of 0.01 [m] x 0.01 [m].

4. EXPERIMENTS AND PRELIMINARY RESULTS

As mentioned previously, each data fusion workflow should be steered by an application. Each application puts different requirements on the expected quality improvements implying a definition of parameter values, necessary features, and required processing steps. Therefore, the tests are conducted on a selected application. Namely, for façade refinement purposes where the detection of deviations between raw semantic 3D wall models and acquired MLS point cloud is the core step.

4.1 Experimental datasets

The tests are performed within the city center of Munich, Germany. The testing dataset consists of CityGML LoD2 building models and MLS point clouds. The CityGML models represent governmental data based on cadastre and ALS point clouds[3]. The TUM-MLS-2016 is the MLS point cloud used for testing (Zhu et al., 2020). Both datasets are coarsely coregistered. The defined uncertainty issues for these semantic 3D models are: footprint absolute 3D point accuracy deviations and implied generalization of walls at the LoD2 (Gröger et al., 2012). For the MLS modality the main uncertainty is the absolute 3D point accuracy due to GNSS/IMU and loop-closures issues. All the errors are associated with expert beliefs regarding the attached metadata.

4.2 Estimation of priors evaluation

The set of parameters was selected to represent $e_1 = 1.00$ [m], $CL_1 = 80\%$ for MLS point cloud and $e_2 = 0.03$ [m], $CL_2 = 95\%$ for CityGML building model. The CI for buildings in question is estimated at 1.10 [m].

4.2.1 Feature coverage analysis: This allowed for creation of 3D boxes within which density and uniformity estimation were performed. Since the building in question had no major extruded parts these tests were performed without vertical-like filtering implementation. Also, since the MLS acquisition geometry implies

capturing of outer walls, the roof elements were assumed uncovered. This test rejected 15 wall segments as not adequately covered and 6 passed both tiers of measurement. The visual assessment of the results confirmed the validity of the outcome.

4.2.2 Coregistration: Based on the obtained matrices, the ICP point-to-plane algorithm achieved rectification of error of around 0.5 [m] to 0.04 [m].

4.2.3 Matching results: This translated to new CI at rates $e_1 = 0.04$ [m], $CL_1 = 90\%$ for MLS point cloud and $e_2 = 2.00$ [m], $CL_2 = 95\%$ for generalization of CityGML building model. This estimated CI at 1.44 [m] for the possible range of building elements, this is depicted in Figure 7 as GR. Since the BayNet aims to find confirmation of the raw CityGML wall models too, the CI for these geometries was obtained. Thereby, $e_1 = 0.04$ [m], $CL_1 = 90\%$ for MLS point cloud and $e_2 = 0.03$ [m], $CL_2 = 95\%$ for CityGML wall objects was estimated at 0.03 [m] CI.

4.3 Bayesian network performance

This BayNet was tailored to detect confirmation between raw CityGML wall geometry and MLS point clouds. As such, the utilized BayNet was a distilled version of the general concept shown in Figure 5. The BayNet was designed in GeNIe Modeler[4]. The previously obtained discretized CIs were the basis for soft
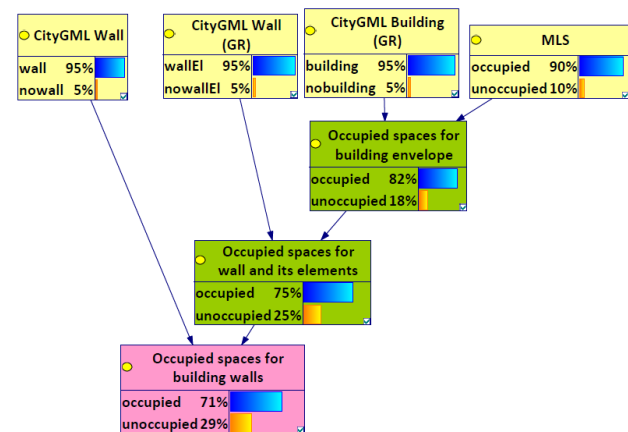

Figure 7. The BayNet with soft evidences (yellow), nodes (green), and target node (pink) and associated inference posterior probability distribution (PPD) scores. GR stands for generalized range of city model with associated uncertainty.

spatial evidence incorporation. In the case of this test, these 3D boxes were further discretized to 3D vertical patches. Each patch had the size of 0.05 x 0.05 [m] and height derived from the respective CI 3D box. As shown in Figure 7, the network consisted of 4 evidence, 2 nodes, and 1 target node. While respective CPT are not shown in Figure 7, they are available under the attached repository[2]. The soft evidence explicitly incorporated probabilities derived from the CI calculation. This was propagated throughout the network to the target node. The network inference was performed in R bnspatial package[5]. The probability for target node state *Occupied* (see Figure 7) was calculated using *bnspatial* function. For the designed BayNet, probability thresholds were set accordingly: $P_{high} > 0.7$, $0.7 <= P_{moderate} => 0.3$, $P_{low} < 0.3$.

---

[3] Landesamt für Digitalisierung, Breitband und Vermessung (https://www.ldbv.bayern.de/produkte/3dprodukte/3d.html)

[4] BayesFusion, LLC (http://www.bayesfusion.com/)

[5] https://cran.r-project.org/web/packages/bnspatial/vignettes/bnspatial.html

4.4 Fusion results

Each spatial output split by probability thresholds might trigger various further investigations. The coarse fusion of both datasets
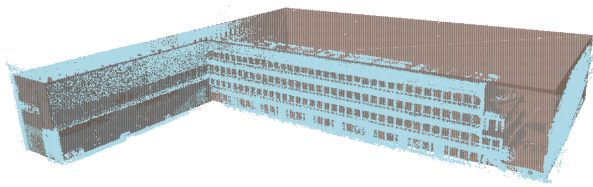

Figure 8. The raw CityGML façade geometry (gray) and coregistered MLS point cloud (blue) within CI for walls.

might be plausible for some applications (see Figure 8). The core fusion task, however, was performed on the basis of the high confirmation rate. Thus, spatial objects with P_{high} were used for coupling of MLS point clouds and CityGML models of façades. Each of the blocks linked to a specific part of the raw model and point cloud. Additionally, the blocks inherited calculated probabilities that can serve as metadata for further processing steps or a database update. As depicted in Figure 9 and in Figure 10, the
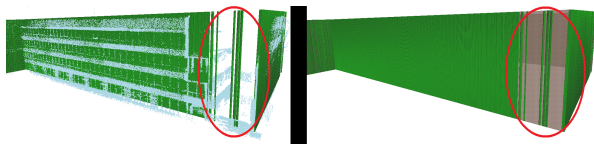

Figure 9. P_{high} blocks (green) indicate confirmed parts of the raw façade geometry (gray) by MLS point clouds (blue).

confirmation of closely and densely aligned point clouds with raw 3D models was achieved. The vertical 3D patches were suitable for coarse wall confirmation but should be finer for other purposes. For instance, 3D vertical patches did not consider building openings. However, this could be mitigated by for example utilization of grid voxels or octree structures of small size. The method
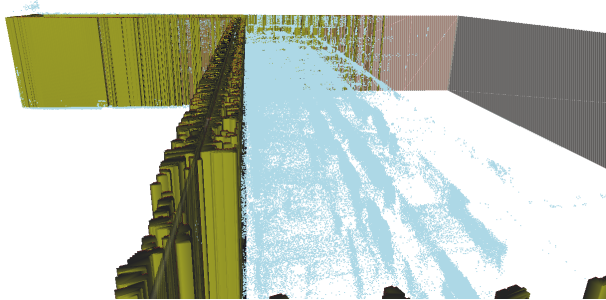

Figure 10. $P_{moderate}$ as an indicator of possible deviations between raw model (gray) and MLS observations (blue). Such deviations indicate possibility of unmodeled surfaces.

achieved satisfactory results also regarding $P_{moderate}$ and P_{low} thresholds. As depicted in Figure 10, $P_{moderate}$ together with P_{high} might serve as a trigger for the façade refinement purposes. For instance, P_{high} patches were the areas not desired for the refinement. Whereas $P_{moderate}$ marked possibly unmodeled elements or strong deviations w.r.t. the raw model. Moreover, the inner points were marked as P_{low} probability to belong to the façade. This might mitigate their negative impact on the façade refinement process. The one-sided Hausdorff distance (Cignoni et al., 1998) was used to quantitatively measure fusion performance. The tests were conducted on a raw situation before apply-

Fusion set	MLS vs. LoD2 façade		
	max	μ	RMS
P_{high}	0.5	0.04	0.06
Raw	0.5	0.20	0.23

Table 1. Comparison of raw and P_{high} fusion set deviations - all metrics given in meters.

ing the strategy and an end fusion using P_{high} areas. To compare deviations under constant conditions, the maximal deviation was set to 0.5 [m]. The results are shown in Table 1 and in Figure 11.
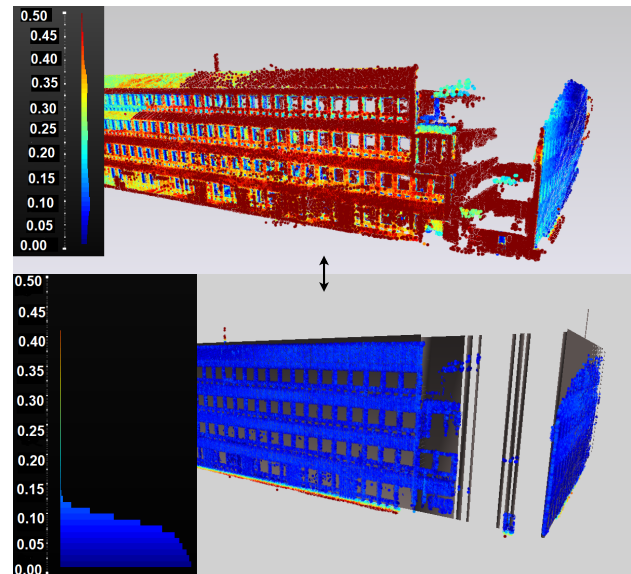

Figure 11. Raw (above) and P_{high} (below) fusion sets with deviations obtained by Hausdorff metric given in meters.

5. CONCLUSION

This work presents a strategy for the fusion of MLS point clouds and semantic 3D building models while considering uncertainty issues. The strategy avoids the pitfalls of fixed bounding boxes through the usage of dynamic CIs. Such an advantage shall pave the way for the following works and increase the scalability of data fusion solutions. Moreover, the method to align and match multimodal datasets illustrates how the traits of structured and unstructured datasets can be coupled while addressing the associated uncertainties. Thereby, enabling modeling of the spatial phenomena in a closer-to-reality manner as multimodal datasets are inevitably depicting an object with a dataset-specific degree of accuracy. Furthermore, the matching outcome propagates throughout the BayNet's identity estimation step to the application-specific goal, thus, increasing datasets' potential in applications such as façade refinement. The designed BayNet is modular and, as such, is extendable beyond the fusion of tested 3D building models and should be applicable for other city objects too. The preliminary experiment presents promising results for the strategy. In the next steps, the concept should be tested on various building types and other city objects. Moreover, the strategy implies that not only the fusion algorithms but also the data model standardization (Beil et al., 2021) will be cornerstones of point clouds and semantic 3D city models integration.

ACKNOWLEDGMENTS

REFERENCES

Anderson, L., Cheek, D., Aragao, L., Andere, L., Duarte, B., Salazar, N., Lima, A., Duarte, V. and Arai, E., 2017. Development of a point-based method for map validation and confidence interval estimation: a case study of burned areas in Amazonia. *J. Remote Sens. GIS* 6(1), pp. 193.

Beil, C., Kutzner, T., Schwab, B., Willenborg, L., Gawronski, A. and Kolbe, T. H., 2021. Integration of 3D point clouds with semantic 3D city models – Providing semantic information beyond classification. *ISPRS - Int. Ann. Photogramm. Remote Sens. Spatial Inf. Sci. (accepted)*.

Besl, P. and McKay, N. D., 1992. A method for registration of 3D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* 14(2), pp. 239–256.

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S. and Çöltekin, A., 2015. Applications of 3D City Models: State of the Art Review. *ISPRS Int. J. of Geo-Inf.* 4(4), pp. 2842–2889.

Chen, S. H. and Pollino, C. A., 2012. Good practice in bayesian network modelling. *Environ. Model. Softw.* 37, pp. 134–145.

Chuprikova, E., 2019. Visualizing Uncertainty in Reasoning. Dissertation, Technische Universität München, München.

Cignoni, P., Rocchini, C. and Scopigno, R., 1998. Metro: measuring error on simplified surfaces. In: *Computer Graphics Forum*, Vol. 17, Blackwell Publishers, pp. 167–174.

Dong, Z., Liang, F., Yang, B., Xu, Y., Zang, Y., Li, J., Wang, Y., Dai, W., Fan, H., Hyyppä, J. et al., 2020. Registration of large-scale terrestrial laser scanner point clouds: A review and benchmark. *ISPRS J. Photogramm. Remote Sens.* 163, pp. 327–342.

Gehrung, J., Hebel, M., Arens, M. and Stilla, U., 2017. An approach to extract moving objects from MLS data using a volumetric background representation. *ISPRS Int. Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* IV-1/W1, pp. 107–114.

Goebbels, S. and Pohle-Fröhlich, R., 2018. Line-based registration of photogrammetric point clouds with 3d city models by means of mixed integer linear programming. In: *VISIGRAPP (4: VISAPP)*, pp. 299–306.

Goebbels, S., Pohle-Fröhlich, R. and Kant, P., 2018. A linear program for matching photogrammetric point clouds with CityGML building models. In: *Operations Research Proceedings 2017*, Springer, pp. 129–134.

Goebbels, S., Pohle-Fröhlich, R. and Pricken, P., 2019. Iterative closest point algorithm for accurate registration of coarsely registered point clouds with CityGML models. *ISPRS - Int. Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* IV-2/W5, pp. 201–208.

Gröger, G., Kolbe, T. H., Nagel, C. and Häfele, K.-H., 2012. OGC City Geography Markup Language CityGML Encoding Standard. Open Geospatial Consortium: Wayland, MA, USA, 2012.

Hall, D. and Llinas, J., 1997. An introduction to multisensor data fusion. *Proceedings of the IEEE* 85(1), pp. 6–23.

Hebel, M. and Stilla, U., 2012. Simultaneous calibration of ALS systems and alignment of multiview LiDAR scans of urban areas. *IEEE Trans. Geosci. Remote Sens.* 50(6), pp. 2364–2379.

Hebel, M., Arens, M. and Stilla, U., 2013. Change detection in urban areas by object-based analysis and on-the-fly comparison of multi-view ALS data. *ISPRS J. Photogramm. Remote Sens.* 86, pp. 52–64.

Hornung, A., Wurm, K. M., Bennewitz, M., Stachniss, C. and Burgard, W., 2013. OctoMap: An efficient probabilistic 3D mapping framework based on octrees. *Autonomous robots* 34(3), pp. 189–206.

Kjaerulff, U. B. and Madsen, A. L., 2008. Bayesian networks and influence diagrams. *Springer Science+ Business Media* 200, pp. 114.

Lucks, L., Klingbeil, L., Plümer, L. and Dehbi, Y., 2021. Improving trajectory estimation using 3D city models and kinematic point clouds. *Transactions in GIS* 25(1), pp. 238–260.

Rusinkiewicz, S. and Levoy, M., 2001. Efficient variants of the icp algorithm. In: *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pp. 145–152.

Schmitt, M. and Zhu, X. X., 2016. Data fusion and remote sensing: An ever-growing relationship. *IEEE Geoscience and Remote Sensing Magazine* 4(4), pp. 6–23.

Schwab, B. and Kolbe, T. H., 2019. Requirement Analysis of 3D Road Space Models for Automated Driving. *ISPRS - Int. Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* IV-4/W8, pp. 99–106.

Stritih, A., Rabe, S.-E., Robaina, O., Grêt-Regamey, A. and Celio, E., 2020. An online platform for spatial and iterative modelling with bayesian networks. *Environ. Model. Softw.* 127, pp. 104658.

Suveg, I. and Vosselman, G., 2000. 3D reconstruction of building models. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XXXIII, Part B2, pp. 538–545.

Tuttas, S., Stilla, U., Braun, A. and Borrmann, A., 2015. Validation of BIM components by photogrammetric point clouds for construction site monitoring. *ISPRS - Int. Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II-3/W4, pp. 231–237.

Willenborg, B., Pültz, M. and Kolbe, T. H., 2018. Integration of Semantic 3D City Models and 3D Mesh Models for Accuracy Improvements of Solar Potential Analyses. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-4/W10, pp. 223–230.

Wysocki, O., Schwab, B., Hoegner, L., Kolbe, T. H. and Stilla, U., 2021. Plastic surgery for 3D city models: A pipeline for automatic geometry refinement and semantic enrichment. *ISPRS - Int. Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* V-4, pp. 17–24.

Xu, Y. and Stilla, U., 2021. Towards building and civil infrastructure reconstruction from point clouds: A review on data and key techniques. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 14, pp. 2857–2885.

Zhang, S., Lo, S., Chen, Y.-H., Walter, T. and Enge, P., 2018. GNSS multipath detection in urban environment using 3D building model. In: *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, IEEE, pp. 1053–1058.

Zhu, J., Gehrung, J., Huang, R., Borgmann, B., Sun, Z., Hoegner, L., Hebel, M., Xu, Y. and Stilla, U., 2020. TUM-MLS-2016: An annotated mobile lidar dataset of the TUM city campus for semantic point cloud interpretation in urban areas. *Remote Sensing* 12(11), pp. 1875.