

# INCREMENTAL MAP REFINEMENT OF BUILDING INFORMATION USING LIDAR POINT CLOUDS

Q. Zou<sup>1</sup>, M. Sester<sup>1</sup>

<sup>1</sup> Institute of Cartography and Geoinformatics, Leibniz University Hannover, Germany – {Qianqian.Zou, Monika.Sester}@ikg.uni-hannover.de

Commission II, WG II/3

**KEY WORDS:** Mobile Mapping, 3D Quality Map, Building Modelling, LiDAR, Incremental Refinement, Point Cloud

## ABSTRACT:

For autonomous systems, an accurate and precise map of the environment is of importance. Mapping from LiDAR point clouds is one of the promising ways to generate 3D environment models. However, there are many problems caused by inaccurate data, missing areas, low density of points and sensor noise. Also, it is often not possible or accurate enough to generate a map from only one measurement campaign. In this paper, we propose a method to incrementally refine the map by several measurements from different campaigns and represent the map in a hierarchical way with a measure indicating uncertainty and the level of detail for objects. The idea is thus to store all captured information with a tentative semantics and uncertainty – even when it is not yet complete. Hence, occluded areas are presented as well, which can be possibly improved by the supplemental observation from the next measurement campaign. The proposed 3D environment model framework and the incremental update method are evaluated using LiDAR scans obtained from Riegl Mobile Mapping System.

## 1. INTRODUCTION

A precise 3D map is of great importance to many applications, such like autonomous systems. In outdoor environment, stereo images and light detection and ranging (LiDAR) are the popular technology to obtain data. Mobile LiDAR is a powerful way to measure dense point clouds along roads. However, due to the constraints of scanning distance, field of view, object self-occlusion and occlusion by others, it is difficult to obtain a complete and sufficient sampling of all the building surface in one measurement campaign (Wang et al., 2018). Many researchers come up with methods to reconstruct the urban scene, during which many problems exist such as inaccurate data, missing areas, low density of points and sensor noise. For stationary scanning method, multiple scans at different locations are applied to get complete datasets. Likewise, for mobile mapping system, point clouds from several measurement campaigns can be used to improve the map. In this paper, we propose a method to refine the map by several measurements from different campaigns incrementally: the map is represented in a hierarchical way with a measure indicating uncertainty and the level of detail for objects. This way, high-level objects like the façade structures are automatically modelled in an incremental and hierarchical way, as more data become available.

There are a variety of 3D urban model researches, among which building modelling is particularly addressed. The grammar-based reconstruction (Alegre et al. 2004, Ripperda and Brenner 2009) was developed to extract building structures, where statistical approaches, e.g. MCMC, rjMCMC, are utilized to generate the models. In some other researches, more semantical context of building structures are used as prior knowledge. (Xiong et al., 2013) use contextual relationships to label patches and (Malihi et al., 2016) pre-cluster the building point clouds and model the planes, edges and vertices later with the geometrical constraints from plane intersection.

The incomplete point cloud and the occlusion in the measurement is a challenge for urban scene modelling. It is difficult to measure the back and the roof of the building from the Mobile Mapping System mounted on a vehicle (Wang et al., 2018). Some existing methods deal with the problem with prior knowledge such as the repetitive and symmetric geometry of building structures (Ripperda and Brenner 2009, Malihi et al. 2016). Some methods like (Xiong et al., 2013) utilize learning-based algorithms to model the structures and fill the holes (Stutz and Geiger, 2020). (Tutzauer and Haala, 2015) fuse geometric features with color information to better model windows. In (Li and Wu, 2020), topological relations are applied as constraints to reconstruct complex buildings from incomplete point clouds. All the approaches try to identify complete semantic objects (building components). However, they do not model partial or even completely unknown information. This is where our approach comes in: it also tries to capture the partial knowledge. For example, when a laser beam only hits a few points on the façade, it is not possible to reconstruct the façade; however, it provides some information about its tentative position and orientation.

This paper proposes a method to automatically generate and update the map incrementally from more than one measurement campaign and potentially with different sensors. In the beginning, the initial urban scene might be mapped incompletely and partially inaccurate, where the modelling does not require the complete datasets for every side of the building. The map is refined by more measurements obtained later. Moreover, the existing methods often just fill the unknown occluded region with their models, but none of them stores this information to indicate the uncertainty there. This information is provided as a part of the integrity measure in our map and serves for the next refinement. In this way, a model stores not only the information it knows, but also the information it does not know. In addition, general knowledge about the object (here building and its parts) and the sensors will be used for the integration of subsequent measurements in a hierarchical fashion. E.g. if initially a façade is instantiated with high accuracy, however containing missing

parts due to occlusion, later the missing parts may only be filled with sub-objects of a façade, and taking the geometric constraints of the original façade into account. Thus, each new measurement will contribute to the completion of the model and/or to an improvement of its accuracy and integrity.

## 2. METHOD

### 2.1 Overview

The proposed 3D model has a hierarchical structure. An object has its own sub-elements and corresponding integrity measure. For example, a building consists of façades while a façade consists of planes, windows and doors, with an integrity measure showing the point density or quality of the plane parameters, as well as occluded regions. The workflow goes from the building to each façade, to planar patches, and to more detailed elements.

Figure 1 shows the framework of our method. To begin with, the point cloud labelled as “building” is segmented into individual building instances. In this step, ground plans from cadastral maps are applied as constraints to separate them.

For an individual building, the RANSAC shape detection (Schnabel et al., 2007) is used to extract the planes and pre-cluster the points to different façades of the building. Then more detailed components are estimated for each façade. The points are projected to 2.5D, where windows, doorways and occlusions are detected and estimated. For vertical planes, the depths of different planes are estimated by a Gaussian Mixture Model. With the assumptions that the façade elements are repetitive and symmetric, and the boundary of the façade intersects with ground planes, some missing area can be preliminarily completed. One option is to fill the missing information by learning-based method (it is not discussed in detail in this paper). In all these cases, the original occluded areas remain in the object and its uncertainty is indicated.

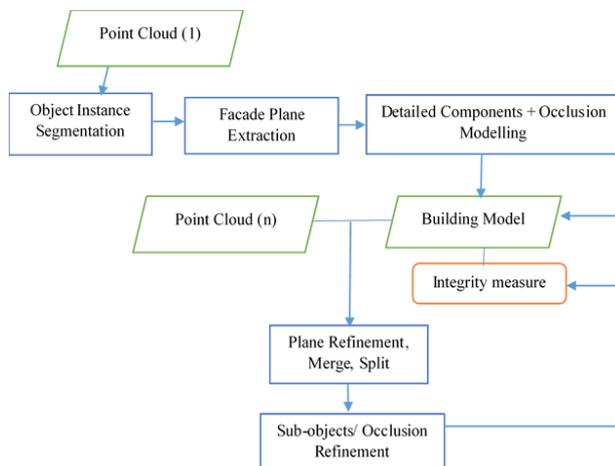


Figure 1. Overview of the workflow

The last – iterative – step is the refinement from the next measurement dataset. The occluded region is likely to be measured and corrected with the new measurements. The problematic facade side with only few measured points and difficulty to detect the elements can get enriched information. Some objects can be merged and completed. The parameters of the planar patches are improved as well.

The inference process is iteratively and hierarchically instantiating façades and the elements on a façade, which is

controlled by the hierarchical building model given in Figure 2. On a coarse level, the representation is a bounding box of the façade, together with the accuracy of the depth. When more information is available, this façade can be split into several façade parts, which, in turn, can be specified into different (semantic) parts, such as windows, doors and structural elements on the façade, including balconies, small extrusions and decorative elements, as illustrated in Figure 2. Methods are devised, allowing for the extraction and delineation of these objects, which are described as polygons. In the following, the methods for extracting façades, simple windows, as well as occluded areas are described.

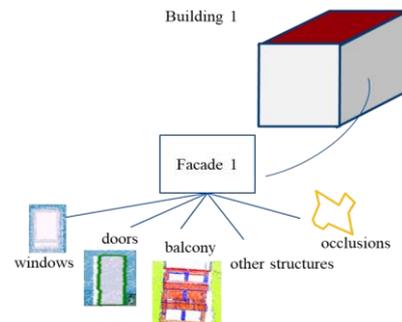


Figure 2. Hierarchical model

### 2.2 Pre-processing and assumptions

To obtain the “building” point clouds, raw point clouds are pre-classified by deep learning-based method (Peters and Brenner, 2019). In this approach, labelled scan strips are generated by projecting 2D prediction to 3D point clouds using fully calibrated mobile mapping data. Since this leads to various types of label noise due to occlusions, the point cloud must then be corrected, which was done in this case using a supervised neural network.

The ground plans cannot accurately separate these points directly, as they are not the true footprints of buildings. The protrusions of different buildings have various distances to the main walls, which are mostly generalized in the cadastral map. In addition, there are other objects in urban scenes, e.g. vegetation and cars, close to buildings. It would be improper to extract all building points by a buffer of ground plan polygons. This is why the learning-based classification is performed first.

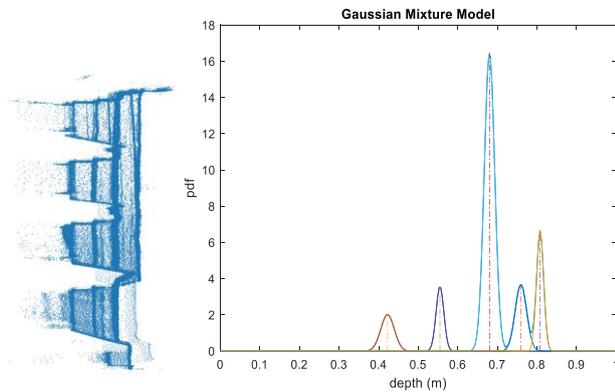
In cadastral maps, the footprints of buildings are represented as polygons, in WGS 1984 UTM coordinates. The point clouds are measured in the same coordinate system. With the assumption that most of the classified points are buildings, these points are connected to the nearest building instance in the cadastral map, involving a spatial join. This way, the points are segmented to individual building instances, as shown in Figure 3. To deal with little remaining label noise in the deep learning classification, points with a distance of more than two meters are ignored.



Figure 3. Segmentation of the buildings.

### 2.3 Points segmentation on a façade

For an individual building, each façade is processed separately, by projecting the 3D points to 2.5D planes in regular grids. The normal vector of the façade and the preliminary centre are estimated by RANSAC shape detection (Schnabel et al., 2007). Points on the same façade are segmented to the planes with different depths by Gaussian Mixture Model (GMM), as shown in Figure 4. The Gaussian component with the highest weight is considered as the main plane, which is the third one for the façade in Figure 4 b). In our assumption, the main plane of a façade is a vertical plane, representing the orientation or normal of the façade. Therefore, the points belonging to this plane are extracted to calculate a more accurate normal vector for the façade. In a Gaussian distribution  $X \sim \mathcal{N}(\mu, \sigma)$ , about 99.6% of values are supposed to distribute within three standard deviation ( $\sigma$ ) away from the mean ( $\mu$ ). Thus, points distributed within this interval are extracted. Then Principle Component Analysis (PCA) is performed for these points and the resulting third dimension is the normal direction of the façade. Points are transformed to a new 2.5D space, with the new normal and a new Gaussian Mixture Model is estimated again to segment them to planes.



a) Profile of the façade      b) Gaussian Mixture Model  
**Figure 4.** Segmentation of points based on depth

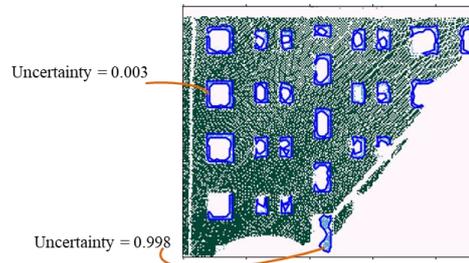
The accuracy of a plane is estimated with the standard deviation  $\sigma$ . On the same plane layer, different objects are then segmented by iterative region growing. The shape of an object is a polygon calculated using alpha-shape algorithm. The uncertainty is relevant to the confidence whether the object belongs to a plane, as computed in equation (1):

$$\alpha = 1 - (cdf(\mu + |\mu - \bar{d}|) - cdf(\mu - |\mu - \bar{d}|)), \quad (1)$$

where  $\alpha$  is the confidence level,  $\mu$  is the mean of the corresponding Gaussian component,  $\bar{d}$  is the average distance to  $\mu$ ,  $cdf$  is the cumulative density function of the distribution.

Figure 5 shows an example for the extracted window casings. To make the uncertainty measure grows when errors increase, in practice, the uncertainty measure is obtained as equation (2). In Figure 5, the object with quite large uncertainty 0.998 indicates that it is probably not on this plane, but since there is no sufficient information available for it yet, it is stored in this way.

$$\epsilon = 1 - \alpha \quad (2)$$

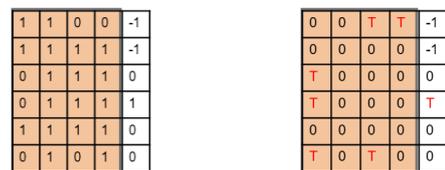


**Figure 5.** Shape and uncertainty of the window casing

### 2.4 Windows and occlusions

Windows and occlusions are both holes in the point cloud. To distinguish them, a ray-tracing algorithm is applied. The points measured in front of the main façade plane in the neighbour region are all considered for computation. The area that has been either self-occluded or occluded by foreground objects like vehicles, vegetation or a road sign, are marked as occlusion mask.

The remaining holes are considered as the candidates for windows and doorways. In the projected plane, if the ratio of pixels in a hole to pixels in the total bounding box of the hole is larger than the threshold value  $Ta$ , the hole is selected as a candidate for windows. After edge detection of the hole, vertical and horizontal lines are extracted by Hough Transform, and then compose candidate rectangles. The rectangle with the best estimation of a hole is stored as the shape of the window. The criteria is to minimize the uncertainty, which is the number of different rectangular polygon and original hole area, as shown in Figure 6 and equation (3), where “1” denotes the hole, “0” is the occluded area, “-1” indicates the area with non-window objects detected, and the orange polygon is a candidate window rectangle. The presented whole grid is the bounding box of the hole. If a candidate rectangle contains pixels with value not equal to one, or there are “1” pixels not in the orange polygon, the error will increase, as denoted as red “T” in Figure 4 b).



a) Shape estimation of the hole      b) Different pixels

**Figure 6.** Window shape estimation

$$\epsilon_w = \frac{N_{diff} + N_{obj}}{N_{hole}}, \quad (3)$$

where  $\epsilon_w$  is the uncertainty of the window,  $N_{hole}$  is the total number of hole pixels,  $N_{diff}$  is the number of different pixels with respect to the window shape,  $N_{obj}$  is the number of non-window object pixels included in the shape.

After window detection and modelling, the irregular holes not selected for windows are labelled as unknown, together with the occlusion mask marked in the beginning in Figure 7 a). The uncertainty of these holes are marked as infinite, and they are represented by polygon shape as well. Occlusions and self-occlusions are both presented as the polygons with white boundaries in Figure 7 b).

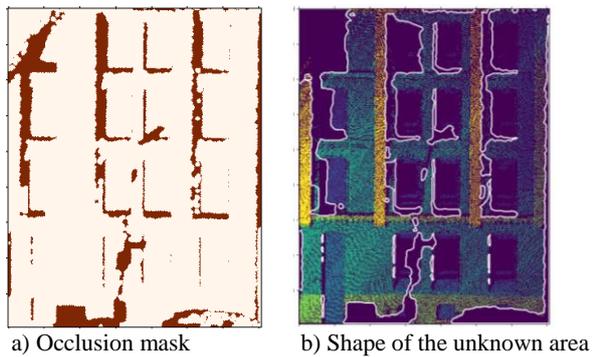


Figure 7. Shape of occlusions

## 2.5 Refinement

The refinement of the existing model with new coming data is based on object-level. The detailed façade elements are refined with new estimated objects from new measurements. The individual building is segmented from the new measurement by the same step as above. For a building, points are classified to the existing façades according to their distances to the main plane. If the distance is larger than two meters, it is marked as unclassified and if possible, these points will be computed to find a new façade of the building using RANSAC algorithm.

The points of the same façade are applied to new normal vector calculation and refinement with PCA, which is then combined to the normal vector from the existing model with the weight  $w$  computed by the point size of the new cloud. Gaussian Mixture Model estimation and points segmentation are performed for the new measurement data. As the new data and existing data share the same original depth values, two statistical models can be combined with Bayes principle. The new distribution multiplies the existing GMM, as an analogue of Kalman Filter. The following equations demonstrate the multiplication of each Gaussian component.

$$\mathcal{N}(\mu_0, \sigma_0) \cdot \mathcal{N}(\mu_1, \sigma_1) = \mathcal{N}(\mu', \sigma') \quad (4)$$

$$k = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2} \quad (5)$$

$$\mu' = \mu_0 + k(\mu_1 - \mu_0) \quad (6)$$

$$\sigma'^2 = \sigma_0^2 - k \cdot \sigma_0^2 \quad (7)$$

where  $(\mu_0, \sigma_0)$  are the mean and standard deviation of Gaussian component for the existing model,  $(\mu_1, \sigma_1)$  are the parameters of the new Gaussian component,  $\mathcal{N}(\mu', \sigma')$  is the refined model.

The façade elements and occlusions are modelled as same as the approach in 2.3 and 2.4. The modelled objects are compared with the existing objects. If they belong to the same plane class and the shape polygons overlap with a significant area (e.g. IoU), two objects can be likely merged. For example, in Figure 8 a), there are two windows from the existing model (red) and new model (orange), respectively. The original hole area is recorded in the model, with the pixel value as the number indicating how many times a pixel is detected as “window” in every single measurement campaign, which will be treated as weights for the best window shape estimation. The region detected as other objects instead of the occlusion would count as error with a negative weight; i.e. if these pixels are contained in the rectangle, the error will increase. Vertical and horizontal lines from the existing window shape and new shape compose a new set of

candidate rectangles. The best shape is found by minimizing different weighted pixel numbers between original hole area and shape polygon. The ratio of weighted number of difference to the total weighted number indicates the uncertainty of the object. In Figure 8, two windows before merging have the uncertainty measures as 0.1353 and 0.0867, while the uncertainty of the merged window shape is 0.0762.

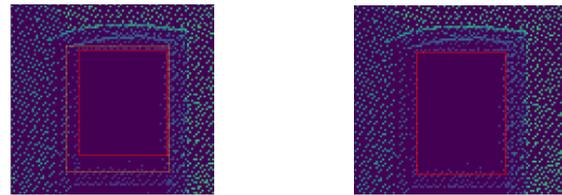


Figure 8. Window shape refinement

Other objects, which have irregular shapes derived by the alpha-shape algorithm, are merged with similar criteria but slightly more complicated. If the polygon to be merged has occupied the pixels detected as objects on other plane layer, their uncertainty will be compared, and the conflicting pixels will be allocated to the one with better accuracy.

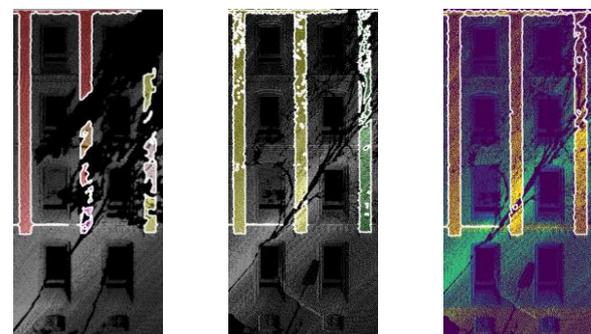


Figure 9. Object shape refinement: small vertical extrusions

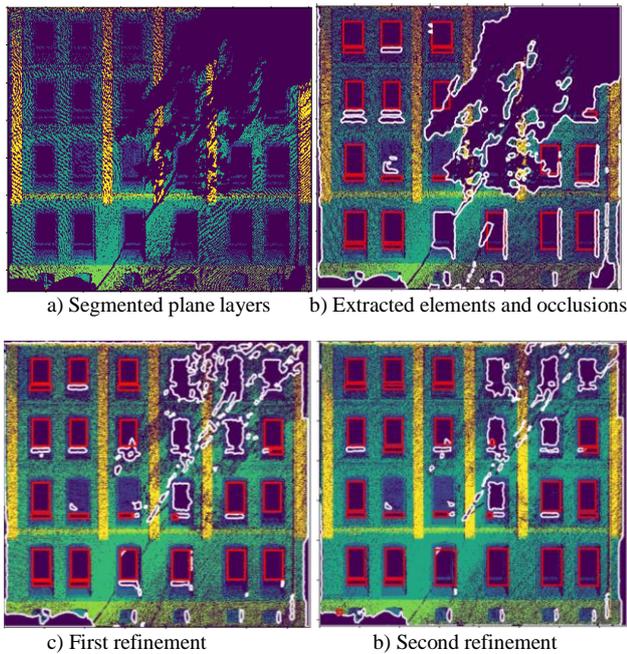
For different measurement campaigns, the occluded area typically is not exactly the same, which helps to merge objects and complete the holes in the existing model. It can happen that one façade structure is separated by occlusions and modelled as several objects in the earlier model, as presented in Figure 9. In the refinement, they can be merged to one object and the occluded part can be completed.

## 3. DATA AND EXPERIMENTS

In the experiments, the 3D Lidar datasets were observed from the Riegl Mobile Mapping System VMX 250 in Hannover, Germany. The measurement quality of the scanner system is specified as 5-mm precision, 10-mm accuracy. As the GPS positioning inaccuracy is larger than the laser scanner noise, the measurements from different campaigns were aligned with the strip adjustment method proposed by (Brenner, 2016), with a standard deviation lower than 2-cm. The point clouds are classified with a deep learning approach proposed by (Peters and Brenner, 2019) and the ones labelled as “building” are investigated in the experiments.

The individual building instances are segmented with the ground plan constraints. Different façades of the building are then extracted. The clustered points of a façade are projected to 2.5D as a depth image, as shown in the Figure 10. It can be partially

segmented to plane layers based on the depth value. Figure 10 a) shows a general picture of the different elements on a façade. Thus, the façade model after the analysis of the first measurement consists of main façade (green), structures on the facade (yellow), as well as the ground level façade (light green). All of these objects are represented as 2D-polygons, including their depth, semantics and uncertainty values.

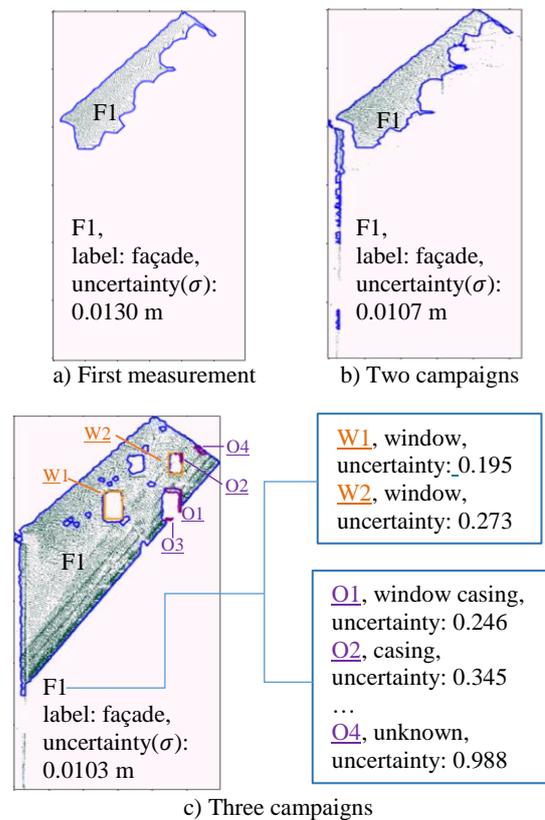


**Figure 10.** Segmented depth image

To eliminate the impact of the low sampling resolution of the sensor in some area, the point density is analyzed to better detect the holes. Afterwards, an iterative region growing is performed to extract windows and occlusions. The shapes and locations of windows are estimated and the presence of the unknown area are stored. The estimated façade elements in Figure 10 b) are presented with modelled shapes and the uncertainty measures for windows (red polygons), as well as the unknown occluded patches (white polygons). If the façade elements are predicted or completed in the unknown area afterwards, the initial infinite uncertainty can be improved.

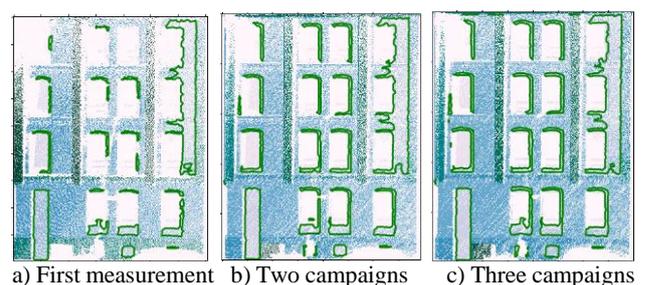
When the next point cloud measurements are obtained for this region, the model will be refined. Figure 10 c) and d) are the refinement results. The occluded region is significantly eliminated. The detection of windows is more complete.

Figure 11 shows a façade that is not well measured with only one measurement campaign but the information is enriched incrementally, from a coarsely estimated irregular plane in Figure 11 a) and b) to a plane with few windows (yellow polygons) in Figure 11 c). The accuracy of the façade plane improves with more data and refinement. The hierarchical model of this façade is illustrated in Figure 11 c) as well, where the purple polygons are the captured detailed elements. Most of the objects are window casings while O4 with large uncertainty measure is likely not on the same plane as others and could be better recognized with more new measurements.



**Figure 11.** Enriched façade information

Figure 12 presents the results of the refinement of object shapes (mainly window casings). The polygons with green boundaries are supposed to be located on the same plane with similar depth. The object shapes are more complete from the first measurement to two campaigns. With three campaigns, the boundaries of the objects are smoother, however, the improvement is not that remarkable as before. The reason could be that the mobile mapping system has the same route for measurements in this case, and these two measurements have been pretty close in time. The occlusions close to the ground are still there.



**Figure 12.** Completed object shapes

The façade in Figure 12 is not on the street where the mobile mapping system drives but perpendicular to it. The street is on the left side of the facade, and thus, the range is longer for the right part, which leads to a worse accuracy. This explains the presence of the large polygon on the right side, resulting from the inaccuracy of depth. It should be a part of the main façade plane but it is not corrected by new measurements here as the mobile mapping system measured similar data with the same route.

#### 4. CONCLUSION

In this paper, a workflow for a hierarchical building modelling and the refinement of models with new coming measurements is

demonstrated. In particular, the representation of the façade models is described, taking different semantic objects, as well as unknown objects with their respective qualities into account. The resulting façade elements are represented with polygon shapes and uncertainty measures. The experimental results show that the fusion and refinement from more measurements improve the accuracy of the existing map.

The model can be used for different applications, especially in the automotive context, where the benefits are obvious: whereas traditional 3D city models and HD maps rely on complete information, which is difficult, expensive and time consuming to acquire, this model delivers partial information as well. E.g. the façade element F1 in Figure 11: an autonomous vehicle whose laser beam hits this element can safely use this information for self-localization.

The paper serves as a proof of the concept, which will be elaborated more in the future. In this work, the occluded parts are currently not particularly modelled. The hole completion methods for the occluded area before the actual data measured need to be further studied as well. The inference process has been demonstrated for the detection of facades and windows; this will be extended to other (semantic) objects.

Although most of the geo-objects can be estimated better with more data, there are some areas that can never be measured due to the limitation of MLS location, e.g. the roof with quite flat angle. In the future, more sensors can be fused to compensate it, like aerial sensors. Thus, the refinement process should further consider the data from other sensors, with different characteristics, e.g. drone data. The workflow presented in this paper allows this in principle.

In general, it is hard to tell how many campaigns are sufficient for a precise map, as it depends on how often the measurement runs, if some objects, e.g. parked cars, have moved, and if the route of the mobile mapping system has passed the backside of the building. More investigation experiments will be done to have a better insight into that. The assumption is the map will be continuously refined by more new measurements – which will be captured by potentially all future vehicles with more and more sensors incorporated. Thus, another aspect of future research will be to simulate automotive grade sensors and check the update quality of those sensors. In addition, as the environment change may occur, the prediction and detection of the temporal change could be studied to collaborate with it.

#### ACKNOWLEDGEMENTS

This work was funded by the German Research Foundation (DFG) as a part of the Research Training Group GRK2159, “Integrity and collaboration in dynamic sensor networks” (i.c.sens).

#### REFERENCES

Alegre, O and Dellaert, F., 2004. A probabilistic approach to the semantic interpretation of building facades, In *Int. Workshop on Vision Techniques Applied*, pp. 1-12.

Brenner, C., 2016. Scalable estimation of precision maps in a mapreduce framework. In: M. Ali and S. Newsam (eds), *SIGSPACIAL '16: Proceedings of the 24th ACM SIGSPACIAL International Conference on Advances in Geographic Information Systems*, ACM Press, New York, New York, USA.

Li, Y. and Wu, B., 2020. Automatic 3D Reconstruction of Complex Buildings from Incomplete Point Clouds with Topological-Relation Constraints. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 5, pp.85-92.

Malihi, S., Zoj, M.V., Hahn, M., Mokhtarzade, M. and Arefi, H., 2016. 3D building reconstruction using dense photogrammetric point cloud. *Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3, pp.71-74.

Peters, T. and Brenner, C., 2019. Automatic Generation of Large Point Cloud Training Datasets Using Label Transfer. In: 39th Annual Scientific and Technical Conference of the DGPF eV, in Vienna, Austria.

Ripperda, N., Brenner, C., 2009. Application of a Formal Grammar to Facade Reconstruction in Semiautomatic and Automatic Environments. In: *Proceedings of 12th AGILE Conference on GIScience*, Hannover, Germany.

Schnabel, R., Wahl, R., Klein, R., 2007. Efficient RANSAC for point cloud shape detection, *Computer Graphics Forum*. Wiley Online Library, pp. 214-226.

Stutz, D., and Geiger, A., 2020. Learning 3d shape completion under weak supervision. *International Journal of Computer Vision*, 128(5), 1162-1181.

Tutzauer, P., and Haala, N. 2015. FAÇADE Reconstruction Using Geometric and Radiometric Point Cloud Information. In: *IAPRS*, Munich, XL-3/W2, pp. 247-252.

Xiong, X., Adan, A., Akinci, B. and Huber, D., 2013. Automatic creation of semantically rich 3D building models from laser scanner data. *Automation in construction*, 31, pp.325-337.