

SEMANTICALLY ENRICHED HIGH RESOLUTION LOD 3 BUILDING MODEL GENERATION

A. Gruen¹, S. Schubiger¹, R. Qin¹, G. Schrotter¹, B. Xiong¹, J. Li¹, X. Ling¹, C. Xiao¹, S. Yao¹, F. Nuesch¹

¹ETH Zurich, Future Cities Laboratory, Singapore-ETH Centre, Singapore - agruen@geod.baug.ethz.ch, simon.schubiger@fhnw.ch, qin.324@osu.edu, gerhard.schrotter@zuerich.ch, b.xiong@whut.edu.cn (jiaqiang.li, ling.xiao, xiao, yao, fiona.nuesch)@arch.ethz.ch

KEY WORDS: Semi-automation, LoD3 Building Reconstruction, Procedural Modelling, Image and Point Cloud Processing, Semantic Labelling, Deep Learning, User-Interface

ABSTRACT:

This paper reports about an effort to generate LoD3 models of buildings semi-automatically, with the highest possible level of automation. It is work in progress. We use multi-sensor data like aerial images from a 5-head camera with a GSD of 10cm, UAV images, and aerial and mobile LiDAR point clouds. We distinguish two cases: LoD2 models are available and they are not. We apply Multi-Photo Geometrically Constrained Least Squares Matching for different kind of point measurements. The regularity of many building façades in Singapore leads us to the idea to generalize the measurement procedure towards using measurement macros (geometrical primitives, i.e. windows, doors, etc.) and combine reality-based with procedural modelling. In parallel we try to model these façade elements from LiDAR point cloud data. In another research line we do building detection by a novel approach to land-cover classification, incorporating features of the façades to improve the classification accuracy. To generate the semantic labels of the façades, we developed a spatially unrelated mean-shift clustering method to yield structurally confined segments. It is the characteristic of automated and even semi-automated procedures that the results need some amount of editing. We therefore work on interactive post-editing approaches on CityGML building models containing semantic information of each surface. Maintaining the semantic information throughout the editing process is essential but often lack the support from current tools. Accordingly, we implement a method to synchronize CityGML models. Overall this project consists of a great number of different algorithmic components, which can only be coarsely explained in this paper.

1. INTRODUCTION

3D computer-based city models have not only become an indispensable basis for a variety of tasks related to issues of urban sustainability, but also for real estate business, virtual tourism, microclimate studies, among others. Today, they are essential to the concept and realization of Smart Cities. However, a key problem is to generate such models in a way that they are inexpensive, accurate, complete, easy to update, and useable as the basis for spatial information systems.

Based on previous work at the Simulation Platform of the Future Cities Laboratory, we will present the R&D work in 3D/4D city modelling that is currently performed under the Virtual Singapore project, with the goal to model the geometry at level of detail 3 (LoD3). The NUS (National University of Singapore) campus is used as the test area for the algorithms and procedures.

Over the years many fully automated approaches to building extraction have evolved, but only very few were designed to be semi-automated from the very beginning. Among those CyberCity Modeler (Gruen, Wang, 1998) is the most prominent one. Very often, procedures are declared automatic but require so much post-editing that their status as automatic methods becomes questionable.

Object extraction in general consists of three steps: detection, reconstruction (geometric modelling) and attribution (semantic modelling). This sequence even defines a processing strategy, but also represents a path towards increased complexity. At the detection level cues like colour and Digital Surface Model (DSM) data have proven to be particularly valuable.

They are first used to separate a man-made object from vegetation and other natural features; then to distinguish buildings from other man-made objects. Good success has been reported with isolated houses. Complex urban structures still widely resist this approach.

In reconstruction one encounters a great variety of methods depending on the type of building, level of required detail, number of images, kind of image cues and image primitives used, and utilized external and a priori information, level of automation and operator interference. Recently, there has been a clear trend towards the use of multiple (more than two) images, colour cues, early transition to 3D processing, and geometrical constraints. Image cues may involve texture, colour, shadows, and reflection properties. Image primitives include points, double- and triple-legged vertices, linear elements, and homogeneous regions.

There are basically two fundamental approaches which may lead to progress in automation: (a) advancement in image understanding and (b) multi-sensor, multi-data techniques. While (a) is a very hard problem to solve with a very long-term perspective, (b) can be realized more readily. We are combining both approaches in a project that is funded by the NRF (National Research Foundation) Singapore and part of the larger effort towards creating a high-resolution 3D city model of Singapore (Virtual Singapore, Call for Proposals: Towards Automatic Acquisition Of 3D City Models For Virtual Singapore)

For this purpose we are looking into the problems of

1. Geometric modelling by image analysis and point cloud processing

2. Landcover classification and semantic modelling
3. Generic modelling and UI (User Interface) issues

We target a system of LoD3 model generation with operability potential by integrating three necessary sub-components: 1) semi-automatic 3D roof and façade geometry and topology reconstruction, 2) façade elements semantic modelling and procedural modelling of façades and 3) interactive 3D editing based on polygon editing and procedural generation. Our proposed research combines the reality-based and procedural modelling, adopts semantic image understanding and interactive editing to accelerate the overall workflow of LoD3 model generation and enables various semantic attributes.

This project presents work in progress. Here we report about the state of our work.

We have defined 3 work packages (WPs), whose interaction is shown in Figure 1 and which will be explained in detail in chapter 2.

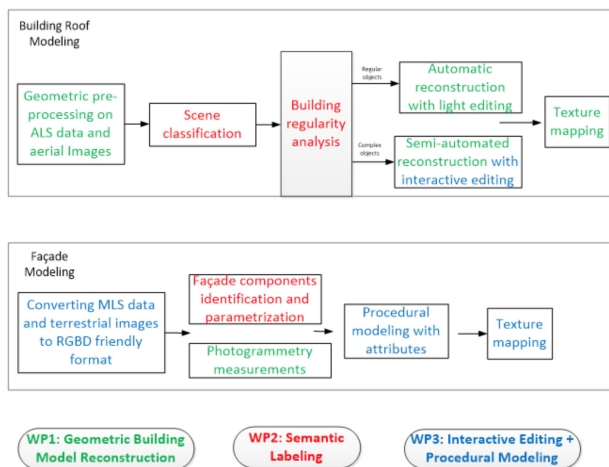


Figure 1. Interaction of work packages

The overall workflow is demonstrated in Figure 2.

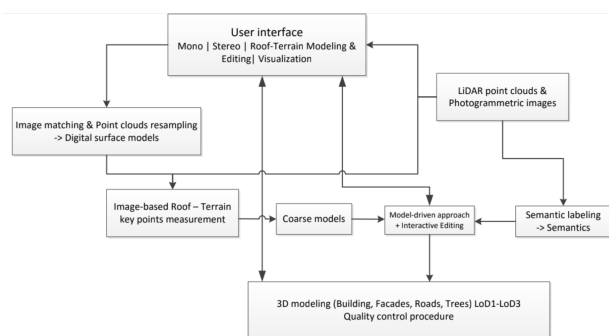


Figure 2. Overall workflow of the project functions

2. DEFINITION OF WORKPACKAGES

2.1 WP1: Automatic/Semi-automatic Building Model Reconstruction By Use Of Images And Point Clouds

There are three major procedures in the LoD3 building reconstruction:

1. Roof reconstruction. This procedure is to reconstruct the roof's boundary, roof planes and roof parts above the roof planes.
2. Ground surface reconstruction: The ground surface can either be taken from existing Digital Terrain Models (DTMs) or automatically/semi-automatically reconstructed.
3. Wall reconstruction. This procedure is to reconstruct the walls and façade elements. Here we concentrate on reconstructing three typical façade elements: windows, doors and balcony.

For our reconstruction tasks we do have available: Aerial images from a 5-head camera with a GSD of 10cm, selected images from UAV flights, aerial LiDAR point clouds and terrestrial point clouds from a mobile mapping system (MMS).

2.1.1 Reconstruction From Images Figure 3 shows the workflow for semi-automated geometric building model reconstruction from images.

The basic module “point measurement” is to accurately measure the point's location both in image space and 3D space. It is applied to all procedures throughout the LoD3 reconstruction. We use an advanced image matcher (Geometrically Constrained Multi-Image LS Matching) to determine the geometry. This approach solves in one system the image-based matching parameters of all images and the object space coordinates X,Y,Z of the point in question simultaneously. The covariance matrix computation allows for an advanced quality analysis (Gruen, 1985, Gruen, Baltsavias, 1986).

We consider three principles when designing our algorithmic approach:

1. minimize the amount of human interaction.
2. maximize the location accuracy of vertices of the building models.
3. optimize the reliability of reconstruction.

Two concepts are used throughout the workflow to achieve these goals:

1. develop measurement macros.
2. apply least squares matching.

The operator will determine a key point in the façade for which matching has to be done. Given a surface model of the area under consideration (which can be derived from the aerial images) we can compute the object space coordinates of the keypoint. Since LSM is a non-linear procedure we need in all images approximate positions for the corresponding image points. Those positions are for instance obtained by back-projection of the keypoint's object space coordinates into the image spaces.

2.1.2 Measurement Macros The measurement macros are combinations of several basic 2D/3D shapes, e.g. rectangles, circles, cubes, cones etc., to represent complicated building elements, such as balconies, windows etc. In such case we go through the following operator-executed steps:

Selection of area containing repetitive patterns → Selection of several key points to create a measurement macro → Entering of the vertical and horizontal element numbers; the measurement macro will be applied to the whole area. The basic idea is illustrated in Figure 6.

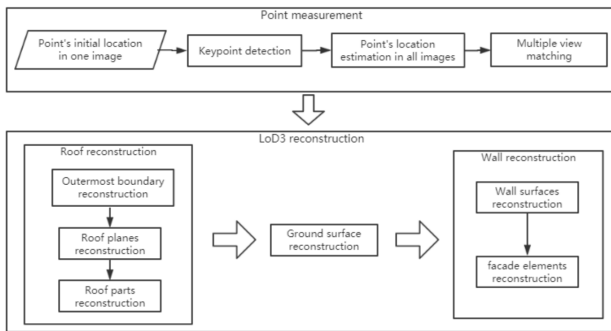


Figure 3. Workflow of the image-based semi-automated geometric building model reconstruction

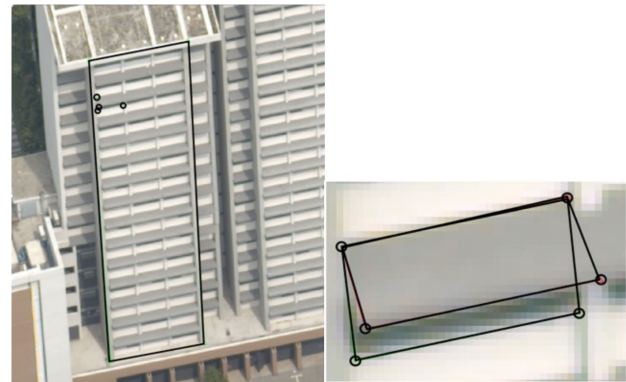


Figure 7. The four corners of a window are selected by the operator, then the window is represented by two polygons.



Figure 4. Left is the master image with the keypoint marked. Center and right are slave images with the initial locations of the keypoint

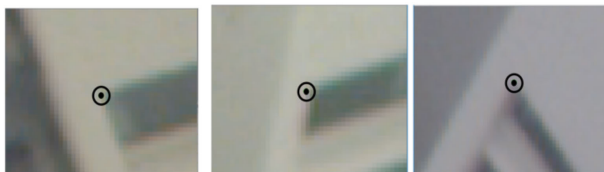


Figure 5. Final matching result after a number of LSM iterations

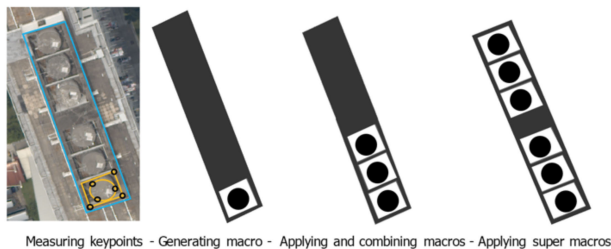


Figure 6. Basic idea of measurement macros

The measurement macros can greatly reduce the workload for operators when dealing with repeated patterns. For example: A façade filled with windows.

1. The operator selects an area containing repetitive patterns
2. He/she selects several keypoints to create a measurement macro (Figures 7 and 8)
3. He/she enters the vertical and horizontal element numbers; the measurement macro will be applied to the whole area (Figure 8)

2.1.3 Opening Detection In Point Clouds Façade elements as windows and doors are sometimes substituted by the term of openings. Therefore, the opening detection and reconstruction contribute to the upgrade of models from the LoD2.

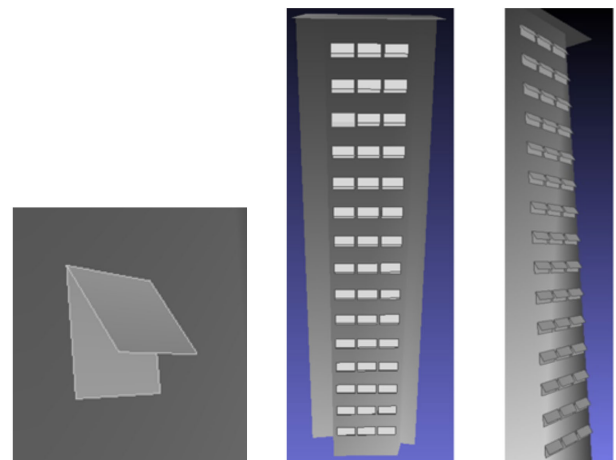


Figure 8. Left: 3D model of the window after its four corners have been measured. Right: automatically generated windows after the operator enters the vertical and horizontal window numbers

In particular, the terrestrial LiDAR data is a favourable data set supplying façade information for the detection of openings. In general, hole-based (Pu, Vosselman, 2009, Zolanvari, Laefer, 2016), local and global feature based (Pauly et al., 2008) and rule-based (Becker, 2009, Nan et al., 2010) are the frequently used approaches in detecting openings from point clouds.

In order to extract feature points of openings (e.g. window corners) and to mitigate the noise effect from point clouds, we use a sliding window method (Li et al., 2018). Firstly, the façade is inspected by the predefined horizontal and vertical sliding windows. During the inspection, we record the point density of the search window coupled with its current location. Followed by a sequenced slide, we observe that the set of features presents a certain regularity because of the different point density of the façade elements. Therefore, we follow the law in the feature domain to extract the borders of openings. Intersecting the border lines, corner points are eventually generated. This approach conducts a recovery of points in partly missing data areas by inferred symmetric rules. The extracted points and rules can be integrated into measurements on 2D images (for example see Figure 9 right) and opening reconstruction.

Although feature points can be extracted by the mentioned approach, there still remains challenges in the opening detection

from point clouds. As a data-driven method, the precision and completeness of detections are to rely on the data quality. It is subject to outliers, noisy points, occlusions and inhomogeneous point density which cause impaired detection results. Moreover, many of the works assume that the openings are rectangular and repetitive with neat alignment.

For ease of the above discussion, we adopt a hole-based algorithm which uses the geometry information from point clouds to reconstruct the configuration of openings (see Figure 9 left). From the observations by our experiments, we assume that the distribution of current detected data more or less follows certain rules. In other words, there would be a generic rule describing the relationship between openings on a façade. Thus, our future work is generating a generic rule for openings reconstruction and façade modelling.

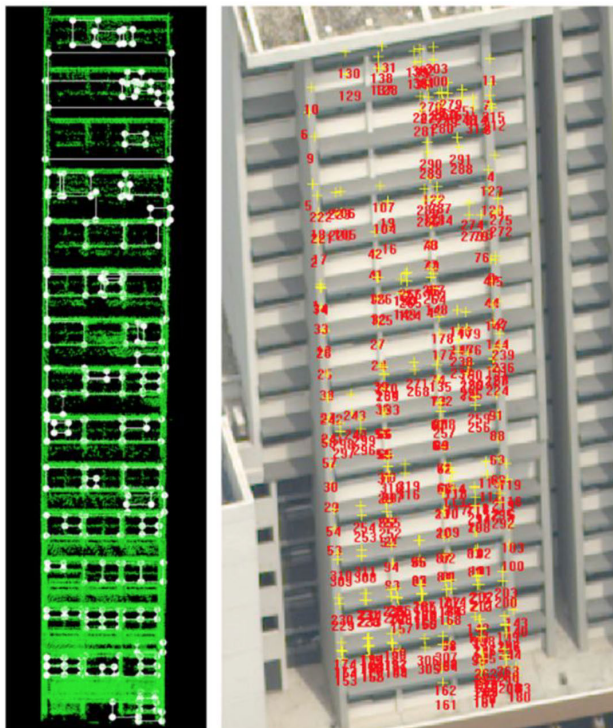


Figure 9. Reprojection of extracted feature points from the 3D point cloud to a 2D oblique Image; Left: The green colour represents the inhomogeneous mobile LiDAR data of a façade. White dots are the extracted feature points; Right: yellow points are the reprojected points from the white dots. Numbers show the topology of the points

2.2 WP2: Landcover Classification, Semantic Enrichment And Texture Mapping

Assigning attributes to the components of the models is an important issue for LoD3 modelling. We aim to employ image analysis and machine learning techniques to enhance the semantic content of the 3D models. In line with image analysis, we also implement texture mapping to assign photorealistic appearance to the models.

The major tasks in WP2 include the land-cover classification of the top-view images and façade semantic labelling using the oblique images. The semantic labels for the top-view and façade objects serve for two purposes: first of all, the land-cover maps provides ROIs (regions of interest) such that DSM-guided

measurement can be performed directly to locations with appropriate zoom-in levels where building roofs or other objects such as trees. Secondly, the façade labels on one hand can serve as ROI for façade object measurement, and on the other hand function as important contextual information for assigning attributes of façade objects for LoD3 city modelling. Under such a context, our solutions for these two tasks aim to employ the images and their intermediate product (e.g. DSM) through machine learning algorithms to achieve accurate land-cover classification and semantic labelling of façades. These tasks are highlighted in the following two sub-sections (a) Land-cover classification using façade information from oblique images, and (b) a novel pattern-based labelling method for labelling via segmentation and convolutional neural network (CNN).

2.2.1 Land-Cover Classification With Façade Information

Land-cover classification is a well-studied yet challenging topic, especially for very-high-resolution data (2m GSD (ground sampling distance) or less). Spectral and textural features extracted from top-view images as conventional approaches are standard and well-practiced, while its capacity can be limited due to the high intra-class variability and inter-class similarity. Such ambiguities may be readily addressed by using oblique-view images since they can provide useful information to differentiate objects with a similar appearance in top-view. Therefore, this contribution utilizes the oblique aerial images by extracting respective features and incorporates them with the top-view features in a novel way, with the aim to push the boundary of land-cover classification accuracies.

To assign appropriate façades to each of the top-view segment, we hence propose the following procedure: Firstly, we assume only exist on above-ground objects and thus we segment these objects from the DSM (digital surface model) using a morphology-based method (Qin, Fang, 2014). Secondly, we locate the nearest façades of the object by assuming it as a planar face, therefore oblique images can be projected to the corresponding face for feature extraction. The procedure is illuminated in Figure 10.

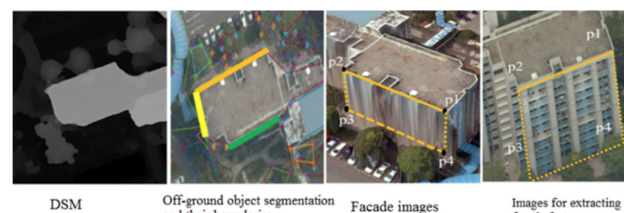


Figure 10. The façade texture extraction procedure from oblique images.

In this work, the average color within each segment and their standard deviation in R, G, B channel, as well as the Haar-like features (Crow, 1984, Viola et al., 2001) that are defined as the difference of the sum of pixel values inside different areas are stacked as feature vectors. Features are extracted for each planar face of the top-view object and thus we eventually obtain aggregated features in each side of the object.

To analyze the improvements of the classification with the façade information, we test the classifications at five sites near and in the National University of Singapore (NUS) campus. For each test site, the 1% of the marked labels is selected as train samples for a random forest classifier (Breiman, 2001). In our experiment, dimension of the feature vector is 28 in which 12

are the façade features and the others include the principal components (PCA) of the color in the top-view orthophoto, DMTHP features of the brightness image, darkness image, and the DSM as described in (Zhang et al., 2015). The final results can be found in Table 1. We can observe that even with very limited training samples, the façade features have improved the average overall accuracy around 4.9%, especially the building accuracy has been improved by 10%, more details can be found in (Xiao et al., 2019).

Site	Facade	Ground	grassland	car	Rain-shield	Tree	building	Overall accuracy	kappa
A	Yes	38.48	97.45	95.27	64.28	60.51	95.04	81.94	73.07
	No	36.11	97.55	95.20	62.49	74.26	58.91	65.94	53.79
B	Yes	68.13	86.32	68.75	82.10	74.07	97.75	82.59	77.31
	No	70.23	82.95	65.59	80.81	88.85	84.50	80.36	74.44
C	Yes	62.65	87.23	92.06	79.57	97.80	97.69	92.00	87.99
	No	65.04	85.26	92.06	80.69	89.03	97.85	88.86	83.23
D	Yes	58.26	58.42	47.80	99.97	71.90	97.15	67.76	60.76
	No	57.69	57.01	39.70	99.97	67.82	97.15	65.87	58.39
E	Yes	80.72	80.39	78.77	97.33	82.04	93.52	86.86	81.29
	No	76.46	86.78	77.21	94.35	82.33	93.23	85.65	79.62
Avg.	Yes	61.65	81.96	76.53	84.65	77.26	96.23	82.23	76.08
	No	61.11	81.91	73.95	83.66	80.46	86.33	77.34	69.89

Table 1. The land-cover classification with/without façade information.

2.2.2 Façade Semantic Labelling Using Image Segmentation And CNN Semantic labels of the façade provide contextual information for LoD3 modelling. Existing deep learning-based methods do not consider the regularity and façade repetitive patterns, thus generating noisy and distorted labeling results. In this work, we developed a spatially unrelated mean-shift clustering method to yield structurally confined segments, which is based on the hypothesis that most of the façade elements are regular and can be seen as repetitive or partially repetitive, thus to regularize the predicted façade labels to be visually consistent with the façade image patterns.

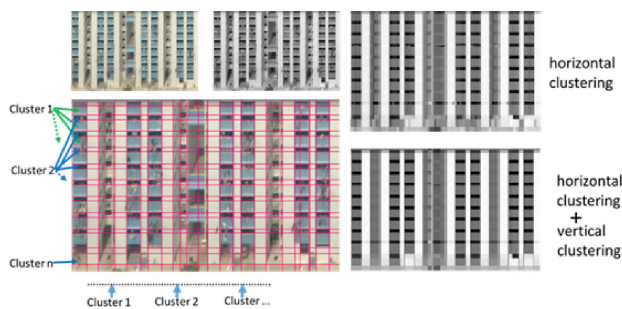


Figure 11. Spatially unrelated image orthographic clustering and segmentation.

Since most of the façade elements for medium or high-rise buildings follow a grid pattern, we firstly rectify the façade images (Figure 11), and we then treat each row or column as a feature vector and cluster all of them into different groups, as a spatially unrelated mean-shift clustering. By applying this clustering, the entire façade image can be segmented into different grids as displayed in Figure 11.

We use a U-net (Ronneberger et al., 2015) to classify the façade with different labels. Since U-net predicts per-pixel labels the yielded segmentations can be blurry in its element boundaries (Figure 12, lower centered image), the segmentations performed using our spatially irrelevant mean-shift segmentation can further regularize the predicted labels (Figure 12, the right

image), thus producing a much clearer semantically labeled image.

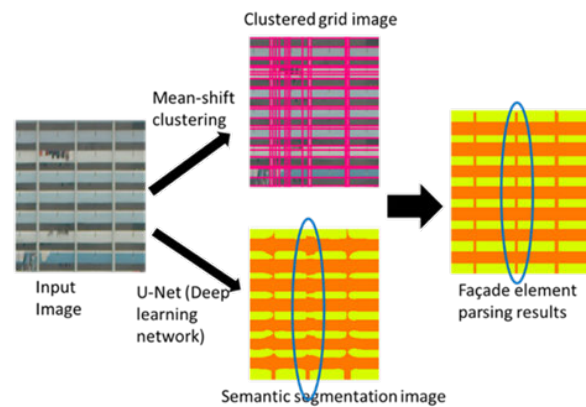


Figure 12. Façade semantic labeling using segmentation and CNN

2.3 WP3: Interactive Post-Editing And Procedural Modelling Of The Building Façade

Since the quality of the models generated by WP1 and WP2 can only be as good as the input data, further refinement of the models may be necessary for certain type of LoD3 applications. An important correction tasks is e.g. the extension of façade patterns into occluded areas. Due to vertical greenery, which is frequently found on façade in Singapore, underlying façade elements may be covered and therefore not reconstructed correctly by WP1 and WP2. This is why WP3 implements a custom 3D editor that offers interactive editing and procedural modelling techniques to enhance the models.



Figure 13. Typical procedural façade templates reality-based on Singapore's HDB buildings.

A set of façade templates (see Figure 13 for some examples) is developed which allows efficient modelling and correction of models. Data from WP1 and WP 2 is used to parameterize the templates according to the measurements.

The editor allows changing the parametrization either through explicit entering of measured values, interactive adjustment via the mouse cursor as well as snapping to the point cloud data (Kelly et al., 2015). Visually overlaying point cloud data with the 3D models in the same editor view enables the operator to adjust the model as well as verify it by comparing it to the actual LiDAR and photogrammetric measurements.

In addition to template-based façade modelling, state of the art interactive 3D editing based on push/pull (Lipp et al., 2014) add classic polygon editing capabilities and geometric primitives which let the operator handle irregular cases that cannot efficiently be represented with rule-based approaches. Interactive changes are automatically classified and matched to the CityGML data model of the editor through model synchronization. The automatic matching and classification significantly reduce the need for manual labelling of interactively constructed façade elements.

2.3.1 Interactive Editing While template-based façade modelling is very effective for modern high-rise, regular architectural styles, varying façades such as those found in Singapore’s Chinatown (see Figure 14) often require considerable manual rework and effective tool support is crucial.

Interactive push/pull geometry modifications are state of the art and implemented in most commercial 3D modelling suites (e.g. Trimble SketchUp, Autodesk Fusion 360). They intuitively transform a polygonal mesh locally by interactive vertex, edge or face drag operations. PushPull++ (Lipp et al., 2014) adds novel methods for adaptive face insertion, adjacent face updates, edge collapse handling, and an intuitive user interface that automatically proposes useful drag directions. PushPull++ has shown to reduce the complexity for common modelling tasks by up to an order of magnitude when compared to existing tools and is therefore well suited for an efficient LoD3 modelling process. We use a PushPull++ implementation derived from Esri CityEngine for interactive geometry modifications.



Figure 14. Traditional shop houses in Singapore’s Chinatown

2.3.2 CityGML Model Synchronization CityGML is the hierarchical data model used throughout the project. CityGML’s building model is a structured geometry with semantic labels for each surface (roof, wall, ground, window, door, etc.). The semantic information of the building model is important for applications beyond visualisation such as governing, planning, simulation, measuring and prediction.

Maintaining that semantic information throughout the editing process is essential for an efficient workflow. Current imple-

mentations of interactive editing, geometry optimisation, and clean-up tools often lack the support for semantic information and only preserve UVs and materials across geometry manipulations. In order to provide a flexible integration of existing as well as future editing tools, an automatic matching method (see Figure 15) between the modified geometry and the data model had to be found.

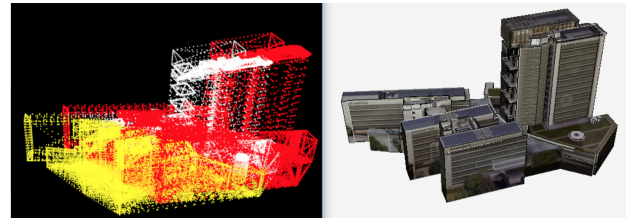


Figure 15. Automatic matching of 3D model changes to CityGML data model

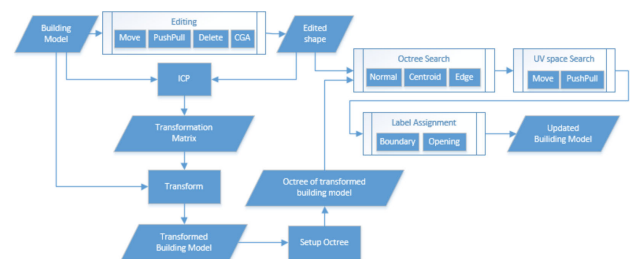


Figure 16. Flow chart of automatic CityGML model synchronisation

Figure 16 gives an overview of our model synchronisation method. Applying an interactive tool results in an edited shape which lacks semantic information because we do not expect this information to be preserved by tools. The original building model is the source shape for the ICP algorithm whereas the edited shape is its target shape. The output of ICP is a transformation matrix which is applied to the original building model yielding a transformed building model. An octree is constructed which includes information (face normal, centroid, length of edges) from the original building model. An octree search yields matching rigid transformed faces and their original semantic labels. Removed faces are handled in a similar way by an additional UV space search.

Newly added faces can be classified as boundary surfaces and openings. For boundary surfaces, we distinguish roof, wall, or ground faces based on their normal and centroid position. For openings, we use again the octree and UV space to find the position in the CityGML data model, finally resulting in an updated building model which preserves as much as possible the original hierarchy information and semantic labels.

2.3.3 Façade Templates Reconstruction based on purely photogrammetric methods faces multiple challenges in an urban environment:

1. Occlusions (e.g. greenery, pedestrians, cars, etc.)
2. Insufficient data due to privacy
3. Duplicate measurement for the same elements

In order to fill these gaps, we are building a reality-based template library which can be parametrized by measurements from

WP1 and WP2. Figure 14 shows a few examples of typical façade patterns found on Singapore HDB buildings. Esri's procedural runtime (Esri, 2013) is used to generate geometry from these templates which are then integrated into the final LoD3 building model. The procedural runtime is a library implementing the "Computer Graphics Architecture" (CGA) shape grammar introduced in (Parish, Müller, 2001), which defines a set of operations for transforming shapes in space, with a special focus on patterns found in architecture.

3. CONCLUSIONS

Our paper reports about the algorithmic approaches that we have chosen/developed in order to generate an operable procedure for the production of LoD3 buildings in a 3D city model. We have explored the use of different kind of raw data: Aerial images from a 5-head camera, aerial LiDAR point clouds, terrestrial mobile mapping system point clouds, and if, required, UAV images. We have considered two cases: (a) LoD2 models of buildings are already given (which is the case in Singapore) and (b) no a priori information on buildings is given. In order to be able to handle case (b) we also have developed a novel method for land cover classification, including beside the vertical images also 3D information from oblique images. This added information improves the results of land cover classification significantly.

Our reconstruction from images uses Geometrically Constrained Least Squares Matching for accurate point measurements. In many cases Singaporean buildings have façades with a quite regular arrangement of features (windows, balconies, doors, etc). Therefore we have developed a macro-based measurement procedure which uses higher level geometrical elements "macros") for the effective measurement of repetitive features.

Thus we replace point measurements by macro-based measurements. At the same time this opens us the possibility to combine reality-based with generic modelling.

In a parallel path we investigate how LiDAR-based point clouds can be used for reconstruction. We use a technique that detects openings in point clouds under the assumption that these are windows in a façade. Obviously aerial LiDAR data cannot be used for this approach because it will not give enough information on façades.

A crucial aspect of our work is the semantic modelling of façade elements. For this purpose we achieve façade semantic labelling using image segmentation and CNN. In this work we developed a spatially unrelative mean-shift clustering method to yield structurally confined segments, which is based on the hypothesis that most of the façade elements are regular.

We are aiming at semi-automated procedures. This is why we also need an advanced interactive editing procedure. We use a PushPull++ implementation derived from Esri CityEngine for interactive geometry modifications. Maintaining the semantic information throughout the editing process is essential for an efficient workflow. Current implementations of interactive editing, geometry optimisation, and clean-up tools often lack the support for semantic information. Therefore we have implemented a method that allows us to synchronize CityGML models when going through the editing procedure.

Singapore is a city very rich on trees. What is pleasant for people is a problem for the modeler, because of substantial occlusions, this is why we implemented a custom 3D editor that offers interactive editing and procedural modelling techniques to enhance the models in areas where there is no substantial information.

This report shows work in progress. As such we are aware that we do not present a closed solution to the problem, but rather building blocks which still have to be evaluated and integrated into an operable system.

ACKNOWLEDGEMENTS

This material is based on research/work supported by the National Research Foundation under Virtual Singapore Award No.NRF2015VSG-AA3DCM001-024. We are grateful to Singapore Land Authority (SLA) for their great support on multi-sensor data.

REFERENCES

- Becker, S., 2009. Generation and application of rules for quality dependent façade reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6), 640–653.
- Breiman, L., 2001. Random forests. *Machine learning*, 45(1), 5–32.
- Crow, F. C., 1984. Summed-area tables for texture mapping. *ACM SIGGRAPH computer graphics*, 18, ACM, 207–212.
- Esri, 2013. Procedural Runtime Whitepaper.
- Gruen, A., 1985. Adaptive least squares correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography*, 14(3), 175–187.
- Gruen, A. W., Baltasavias, E. P., 1986. Adaptive least squares correlation with geometrical constraints. *Computer Vision for Robots*, 595, International Society for Optics and Photonics, 72–82.
- Gruen, A., Wang, X., 1998. CC-Modeler: a topology generator for 3-D city models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(5), 286–295.
- Kelly, T., Wonka, P., Müller, P., 2015. Interactive dimensioning of parametric models. *Computer Graphics Forum*, 34, Wiley Online Library, 117–129.
- Li, J., Xiong, B., Biljecki, F., Schrotter, G., 2018. A SLIDING WINDOW METHOD FOR DETECTING CORNERS OF OPENINGS FROM TERRESTRIAL LIDAR DATA. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W10, 97–103.
- Lipp, M., Wonka, P., Müller, P., 2014. PushPull++. *ACM Transactions on Graphics (TOG)*, 33(4), 130.
- Nan, L., Sharf, A., Zhang, H., Cohen-Or, D., Chen, B., 2010. Smartboxes for interactive urban reconstruction. *ACM Transactions on Graphics (TOG)*, 29, ACM, 93.

Parish, Y. I., Müller, P., 2001. Procedural modeling of cities. *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, ACM, 301–308.

Pauly, M., Mitra, N. J., Wallner, J., Pottmann, H., Guibas, L. J., 2008. Discovering structural regularity in 3d geometry. *ACM transactions on graphics (TOG)*, 27, ACM, 43.

Pu, S., Vosselman, G., 2009. Knowledge based reconstruction of building models from terrestrial laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6), 575–584.

Qin, R., Fang, W., 2014. A hierarchical building detection method for very high resolution remotely sensed images combined with DSM using graph cut optimization. *Photogrammetric Engineering & Remote Sensing*, 80(9), 873–883.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, Springer, 234–241.

Viola, P., Jones, M. et al., 2001. Rapid object detection using a boosted cascade of simple features. *CVPR (1)*, 1(511–518), 3.

Xiao, C., Qin, R., Ling, X., Yuan, H., 2019. URBAN LAND-COVER CLASSIFICATION WITH FAADE FEATURE FROM OBLIQUE IMAGES. *2019 IEEE International Geoscience and Remote Sensing Symposium IGARSS 2019*.

Zhang, Q., Qin, R., Huang, X., Fang, Y., Liu, L., 2015. Classification of ultra-high resolution orthophotos combined with DSM using a dual morphological top hat profile. *Remote Sensing*, 7(12), 16422–16440.

Zolanvari, S. I., Laefer, D. F., 2016. Slicing Method for curved façade and window extraction from point clouds. *ISPRS journal of photogrammetry and remote sensing*, 119, 334–346.