Quality dependent reconstruction of building façades

Susanne Becker[1], Norbert Haala[1]

[1] Institute for Photogrammetry, Universitaet Stuttgart, 70174 Stuttgart, Germany
forename.lastname@ifp.uni-stuttgart.de

Abstract. The paper describes an approach for the quality dependent reconstruction of building façades using 3D point clouds from mobile terrestrial laser scanning. Due to different look angles, such measurements frequently suffer from different point densities at the respective building façades. In order to support the interpretation at areas, where no or only limited LiDAR measurements are available, a quality dependent processing is implemented. First, façades are reconstructed at areas of sufficient LiDAR point availability. Based on this reconstruction, rules are derived automatically, which together with the respective façade elements constitute a so-called façade grammar. It holds all the information which is necessary to reconstruct façades in the style of the given building. In our quality dependent approach, this grammar is used as knowledge base for the verification of a façade model reconstructed at areas of limited sensor data quality. Additionally, it is applied for the generation of synthetic façades where no LiDAR measurement is available.

Keywords: Architecture, Modelling, Interpretation, Building, Segmentation, 3D point clouds, laser scanning.

1 Introduction

Modelling and visualisation of 3D urban landscapes has been a topic of major interest in the past years. A number of tools for the production of virtual city models were developed, which are usually based on 3D measurements from airborne stereo imagery or LiDAR. In addition to this area covering airborne data collection, which mainly provides the outline and roof shape of buildings, terrestrial laser scanning (TLS) is frequently used, especially if a more accurate and detailed three-dimensional mapping of man-made structures is required. Beside measurements from fixed viewpoints, these scanners are often mounted on moving platforms. Such mobile mapping systems are usually also equipped with multiple video or digital cameras and allow for the rapid and cost effective capturing of 3D data for larger areas. This variety of sensors and platforms results in heterogeneous urban data sets of different accuracy, coverage and amount of detail. These differences in quality have to be considered during further processing.

Existing solutions are available for a purely geometric evaluation. One example is the automatic registration and transformation of the captured data sets to a reference coordinate system by so-called integrated georeferencing. For this purpose, least squares approximations are often used, which can also provide explicit information on the underlying geometric data quality. However, such standard tools are not available if semantic information has to be derived for tasks like object reconstruction. One application in this context is the automatic extraction of façade structures for the generation of highly detailed 3D city models, which is the focus of our work. Generally, the reliability and accuracy of façade models derived from measured data depends on data quality in terms of coverage, resolution and accuracy. Façade parts for which only little or inaccurate 3D information is available cannot be reconstructed at all or require considerable manual pre- or postprocessing. In order to avoid such time-consuming user interaction, automatic algorithms for façade modelling which can cope with data of heterogeneous quality become important.

In our application, a formal grammar is applied for the explicit reconstruction of building façades using point clouds from terrestrial laser scanning. The main problem in this respect is the strong variation of the available point densities at the building façades, which frequently results from different look angles during the scanning process. We combine bottom-up and top-down modelling to handle this inhomogeneous data quality during reconstruction. The bottom-up modelling ensures flexibility to capture the great variety of real-world façade structures whereas the top-down modelling achieves topological correctness and robustness against potentially incomplete data sets of heterogeneous quality. While existing algorithms based on formal grammars still require manual interaction either for rule definition or façade interpretation, our algorithm runs fully automatically during all processing steps. The core of our façade reconstruction approach, thus, is the automatic generation of a formal grammar, which will be part of quality sensitive modelling strategies.

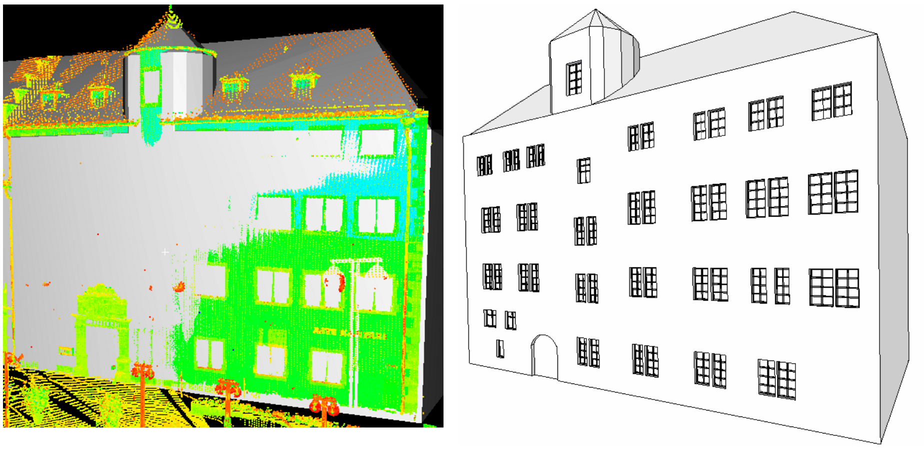
Our algorithm is implemented as follows: Firstly, rules are extracted automatically from observed façade geometries, which are - due to limitations during data acquisition - mostly available only for parts of a building. As discussed in section 2, these rules are represented by a so-called façade grammar. The rules then can be applied to generate façade structure for the remaining parts of the building. As demonstrated in section 3, we take advantage of this in various ways. Top-down predictions are activated and used for the verification and robustification of the reconstruction result that has already been derived from the observed measurements during the bottom-up modelling. Moreover, the façade grammar can be applied to synthesise façades for which no sensor data is available. Concerning the uncertainty of the generated façade models, first considerations on strategies for quality evaluation are presented in section 4. Exemplary 3D reconstruction results are given in section 5.

2 Grammar based modelling of building façades

Our algorithm starts with the bottom-up modelling of façade geometries using terrestrial LiDAR and image data as discussed in section 2.1. After this interpretation step, the resulting reconstructed façade serves as a knowledge base for further processing. Dominant or repetitive features and regularities as well as their hierarchical relationship are detected from the modelled façade elements. At the same time, production rules are automatically inferred. The rules together with the 3D representations of the modelled façade elements constitute a formal grammar which we will call *façade grammar*. It contains all the information which is necessary to reconstruct façades in the style of the respective building during top-down modelling. Section 2.2 will give a short overview of the use of formal grammars for the representation of building architecture. The automatic inference will be discussed in section 2.3.

2.1 Data driven reconstruction

The approach for the data driven façade reconstruction aims at refining an existing coarse building model by adding 3D geometries to the planar façades [1]. Windows, doors and protrusions are modelled from the LiDAR data by searching for holes in the point cloud measured at the building façade. These structures are then refined by integrating further 3D information derived from images of high resolution. The modelling process applies a 3D object representation by cell decomposition. The idea is to segment an existing coarse 3D building object with a flat front face into 3D cells. Each 3D cell represents either a homogeneous part of the façade or a window area. Therefore, they have to be differentiated based on the availability of measured LiDAR points. After this classification step, window cells are eliminated while the remaining façade cells are glued together to generate the refined 3D building model. The difficulty is finding planar delimiters from the LiDAR points that generate a good working set of cells. Since our focus is on the reconstruction of windows, the delimiters have to be derived from 3D points at the window borders by searching for holes in the point cloud. For the exemplary dataset “Alte Kanzlei, Stuttgart”, Fig. 1a depicts the coarse building model with the aligned LiDAR points. Fig. 1b shows the refined façade after reconstruction.



(a) (b)

Fig. 1. Alte Kanzlei, Stuttgart: LiDAR points aligned with coarse 3D building model (a) and refined façade model (b).

2.2 Formal grammars for the modelling of architecture

Usually formal grammars are applied during object reconstruction to ensure the plausibility and the topological correctness of the reconstructed elements. A famous example for formal grammars is given by Lindenmayer-systems (L-systems). Originally used to model the growth processes of plants, L-systems serve as a basis for the development of further grammars appropriate for the modelling of architecture. For instance, [2] produce detailed building shells without any sensor data by means of a shape grammar.

In our application, a formal grammar will be used for the generation of façade structure where only partially or no sensor data is available. In principle, formal grammars provide a vocabulary and a set of production or replacement rules. The vocabulary comprises symbols of various types. The symbols are called non-terminals if they can be replaced by other symbols, and terminals otherwise. The non-terminal symbol which defines the starting point for all replacements is the axiom. The grammar's properties mainly depend on the definition of its production rules. They can be, for example, deterministic or stochastic, parametric and context-sensitive. A common notation for productions which we will refer to in the following sections is given by

$$id: \quad lc < pred > rc : cond \quad \rightarrow \quad succ : prob$$

The production identified by the label *id* specifies the substitution of the predecessor *pred* for the successor *succ*. Since the predecessor considers its left and right context, *lc* and *rc*, the rule is context-sensitive. If the condition *cond* evaluates to true, the replacement is carried out with the probability *prob*. Based on these definitions and notations we develop a façade grammar which allows us to synthesise new façades of various extents and shapes. The axiom refers to the new façade to be modelled and, thus, holds information on the façade polygon. The sets of terminals and non-terminals, as well as the production rules are automatically inferred from the reconstructed façade as obtained by the data driven reconstruction process.

Existing systems for grammar based reconstruction of building models which derive procedural rules from given images or models still resort to semi-automatic methods [3], [4] and [5]. In contrast, we propose an approach for the automatic inference of a façade grammar in the architectural style of the observed building façade.

2.3 Automatic inference of façade grammar

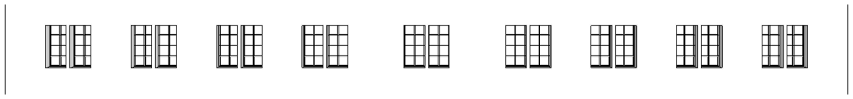
Based on the data driven reconstruction result, the façade grammar is automatically derived by searching for terminals, their interrelationship, and production rules.

Searching for terminals. In order to yield a meaningful set of terminals for the façade grammar, the building façade is broken down into some set of elementary parts, which are regarded as indivisible and therefore serve as terminals. For this purpose, a spatial partitioning process is applied which segments the façade into floors and each floor into tiles. Tiles are created by splitting the floors along the vertical delimiters of *geometries*. A geometry describes a basic object on the façade that has been generated during the data driven reconstruction process (section 2.1). It represents either an indentation like a window or a protrusion like a balcony or an oriel.

Two main types of tiles can be distinguished: wall tiles, which represent blank wall elements, and geometry tiles, which include structures like windows and doors. All these tiles are used as terminals within our façade grammar. In the remaining sections of the paper, wall tiles will be denoted by the symbols W for non-terminals and w_i for terminals. Geometry tiles will be referred to as G and g_i in case of non-terminals and terminals, respectively.

Interrelationship between terminals. Having distinguished elementary parts of the façade we now aim at giving further structure to the perceived basic tiles by grouping them into higher-order structures. This is done fully automatically by identifying hierarchical structures in sequences of discrete symbols. The structural inference reveals hierarchical interrelationships between the symbols in terms of rewrite rules. These rules identify phrases that occur more than once in the string. Thus, redundancy due to repetition can be detected and eliminated. As an example, Fig. 2a shows a modelled floor of the data set "Prinzenbau, Stuttgart". While Fig. 2b depicts the corresponding tile string in its original version, the compressed string and the extracted structures S_i (i=1,2,3) are given in Fig. 2c. The hierarchical relations between the façade elements can be stored in a parse tree illustrated in Fig. 2d.

a)



b) floor 1 → $w_1\ g_1\ w_3\ g_1\ w_1\ g_1\ w_3\ g_1\ w_1\ g_1\ w_3\ g_1\ w_1\ g_1\ w_3\ g_1\ w_2\ g_1\ w_3\ g_1\ w_2\ g_1\ w_3\ g_1\ w_1\ g_1$
$w_3\ g_1\ w_1\ g_1\ w_3\ g_1\ w_1\ g_1\ w_3\ g_1\ w_1$

c)

floor 1 → $w_1\ S_3\ w_2\ S_1\ w_2\ S_3\ w_1$

$S_1 \rightarrow g_1\ w_3\ g_1$
$S_2 \rightarrow S_1\ w_1\ S_1$
$S_3 \rightarrow S_2\ w_1\ S_2$

d)

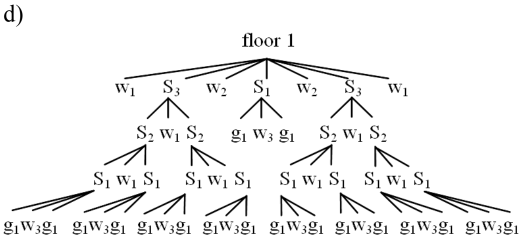


Fig. 2. Modelled floor of the data set "Prinzenbau, Stuttgart" (a), corresponding tile string (b), compressed tile string and extracted structures (c), parse tree (d).

Inference of production rules. Based on the sets of terminals $T=\{w_1, w_2, \ldots, g_1, g_2, \ldots\}$ and non-terminals $N=\{W, G, \ldots, S_1, S_2, \ldots\}$, which have been set up previously, the production rules for our façade grammar can be inferred. Following types of production rules are obtained during the inference process:

p_1: $F \rightarrow W+$
p_2: $W : cond \rightarrow W\ G\ W$
p_3: $G : cond \rightarrow S_i : P(x|p_3)$
p_4: $G : cond \rightarrow g_i : P(x|p_4)$
p_5: $lc < W > rc : cond \rightarrow w_i : P(x|p_5)$

The production rules p_1 and p_2 stem from the spatial partitioning of the façade. p_1 corresponds to the horizontal segmentation of the façade into a set of floors. The vertical partitioning into tiles is reflected in rule p_2. A wall tile, which in the first instance can stand for a whole floor, is replaced by the sequence wall tile, geometry tile, wall tile. Each detected structure gives rise to a particular production rule in the form of p_3. This rule type states the substitution of a geometry tile for a structure S_i. In addition, all terminal symbols generate production rules denoted by p_4 and p_5 in the case of geometry terminals g_i and wall terminals w_i, respectively.

3 Application – Knowledge propagation

Our façade grammar implies information on the architectural configuration of the observed façade concerning its basic façade elements and their interrelationships. This knowledge is applied in three ways. First, the façade model generated during the data driven reconstruction process can be verified and made more robust against inaccuracies and false reconstructions due to imperfect data (section 3.1). Second, façades which are only partially covered by sensor data are completed (section 3.2). Third, totally unobserved façades are synthesised by a production process (section 3.3).

3.1 Verification

The result of the data driven reconstruction process, which is the basis for knowledge inference, may contain false façade structures and therefore be partly incorrect. Some of these errors can be eliminated during an iterative grammar based verification. For this purpose, a rectified image of the façade is used. The grammar is applied to generate hypotheses about possible positions of each geometry tile and project them onto this orthophoto. An image correlation process decides whether a proposed position is accepted or rejected. In case of acceptance, the geometry can be inserted; existing geometries that intersect with the new one are deleted. Afterwards, the resulting improved façade model is used to update the set of terminals and production rules.

Fig. 3 depicts the orthophoto of the data set "Prinzenbau, Stuttgart" as well as parts of the reconstruction result before and after verification. The grilles of the arched windows in the ground floor cause reconstruction errors (Fig. 3b). Only the window to the right of the door could be reconstructed correctly. Its corresponding image mask and the hypothesised and accepted positions in the orthophoto are marked by a yellow rectangle and white crosses, respectively (Fig. 3a).

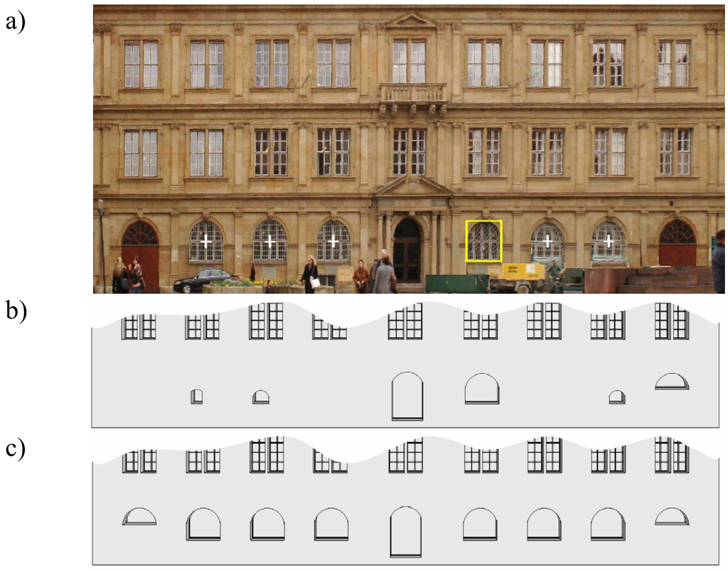


Fig. 3. Prinzenbau, Stuttgart: Orthophoto (a), 3D model before (b) and after verification (c).

3.2 Completion

Due to the scan configuration during data acquisition, façades may contain areas where no or only little sensor data is available. In such regions, an accurate and reliable extraction of windows and doors cannot be guaranteed anymore. Nevertheless, a grammar based façade completion allows for meaningful reconstructions even in those areas. The main idea is to derive the façade grammar solely from façade parts for which dense sensor data and thus accurate window and door reconstructions are available. The detection of such 'dense areas' is based on a heuristic approach which evaluates the sampling distances of the points lying on the façade surface. The restriction of the grammar inference to dense areas ensures a façade grammar of good quality. It can then be used to synthesise the remaining façade regions during a production process.

3.3 Production

The production process starts with an arbitrary façade, called the axiom, and proceeds as follows: (1) Select a non-terminal in the current string, (2) choose a production rule with this non-terminal as predecessor, (3) replace the non-terminal with the rule's successor, (4) terminate the production process if all non-terminals are substituted, otherwise continue with step (1).

During the production, non-terminals are successively rewritten by the application of appropriate production rules. When more than one production rule is possible for the replacement of the current non-terminal, the rule with the highest probability value is chosen. As soon as the façade string contains only terminals, the production is completed and the string can be transferred into a 3D representation.

4 Quality evaluation

The façade models derived in previous sections represent 3D geo data that feature geometric and semantic properties. They consist of 3D cells where each cell is assigned a special meaning, for example 'façade-cell' or 'window-cell'. Both the geometric and semantic aspects of the façade models are uncertain depending on the quality of the sensor data and the reconstruction approach applied. Estimating the quality of such façade models is important, especially when they are to be combined with other data or used within subsequent applications. In principle, two basic strategies for quality assessment can be distinguished: external and internal evaluation.

External evaluation is based on controlled tests using simulated or real data. Reconstructed objects are compared with reference data of superior accuracy yielding quality estimations that are independent of the reconstruction approach. This is necessary to prove the accuracy potential of the algorithm or the adequateness of the applied model. Existing approaches for quality evaluation of 3D building models use either manually generated 3D models or independent measurements with a low uncertainty as reference data [6], [7], [8].

Internal evaluation, also referred to as self-diagnosis, means the estimation of geometric and semantic quality aspects of objects during the object generation process. This requires additional information, for example redundancies in the underlying data or a priori knowledge about the generated object in terms of geometric, topologic or semantic criteria. However, knowledge inherent in the object model can only be applied for an unbiased evaluation if it has not already been used within the reconstruction [9]. Thus, self-diagnosis is strongly linked to the object generation process. It achieves autonomy within a chain of automatic procedures where generally no reference data is available [10]. Error propagation applied to the stochastic properties of input data results in precision measures which are appropriate indicators of the accuracy if the estimation model can be considered correct.

We aim at the quality assessment of 3D façade structures as derived in previous sections for both bottom-up and top-down reconstruction. Due to the lack of highly detailed reference models, external evaluation based on real 3D buildings is not possible. On the other hand, internal evaluation is only appropriate for reconstruction processes that can be described mathematically and thus are the basis for applying error propagation. Though this is the case for the bottom-up reconstruction (section 2.1), the top-down modelling (section 3) rather represents a black box instead of a clear functional model. Thus, we strive for different approaches for quality estimation depending on the reconstruction algorithms proposed. While this is still ongoing work, first considerations are presented in sections 4.1 and 4.2.

4.1 Quality evaluation for data driven reconstruction

The data driven façade reconstruction (section 2.1) is based on cell decomposition. The delimiters of the cells are planes which are determined through least squares adjustment from LiDAR points at window borders. Thus, the quality of the 3D façade structures is mainly influenced by the uncertainty of the estimated planes and the applicability of the model assumptions inherent in the algorithm.

The prediction of the uncertainty in the plane parameters requires information about the stochastic properties of the LiDAR points. Two main error sources affect the quality of the laser data: the uncertainty due to the georeferencing procedure, and the point noise inherent in the scanning process. The accuracy potential of the georeferencing of terrestrial laser data depends on the scan configuration, the availability of control and tie points as well as the accuracy potential of the sensors [11]. However, within the area of interest, which is a single façade, each of these error sources has the same influence on the LiDAR points. They act as systematic errors in the least squares adjustment without any stochastic effects on the quality of the plane equation [12]. When focussing on the relative accuracy of façade structures instead of the absolute one, these factors can be neglected.

The uncertainty of the plane parameters purely caused by point noise can be estimated through the least squares adjustment in terms of the cofactor matrix for the unknown plane parameters: $\boldsymbol{M}^{-1}=(\boldsymbol{A}'\boldsymbol{Q}_{ll}^{-1}\boldsymbol{A})^{-1}$ where $\boldsymbol{A}$ stores the functional and $\boldsymbol{Q}_{ll}$ the stochastic model of the plane estimation. Geometrically, the matrix $\boldsymbol{M}$ defines an elliptic bipartite hyperboloid which is the uncertainty region of the estimated plane. Its parameters can be used as quality measures for the planes. Error influences based on inappropriate model assumptions have to be considered separately, which will be part of our future work.

4.2 Quality evaluation for grammar supported reconstruction

The proposed method for top-down reconstruction aims at the modelling of façade parts for which no or only a few point measurements are available. It is based on grammar rules which are automatically derived from façade structures reconstructed during the bottom-up process. The application of the rules is guided by a stochastic process and thus cannot be expressed in functional terms. Consequently, quality evaluation requires external reference data. Since, in our case, highly detailed 3D reference data is not available, we use additional image data showing the reconstructed façades in reality as additional, independent observations. Images are converted into binary images in which wall regions are distinguished from façade structures such as windows and doors. Additionally, similar binary images are generated from each of the modelled 3D façades. Corresponding images are compared by means of image correlation. The resulting correlation values describe the difference between the 3D model and the information derived from real data. Hence, it can also be interpreted as a quality estimate for the reconstructed façade.

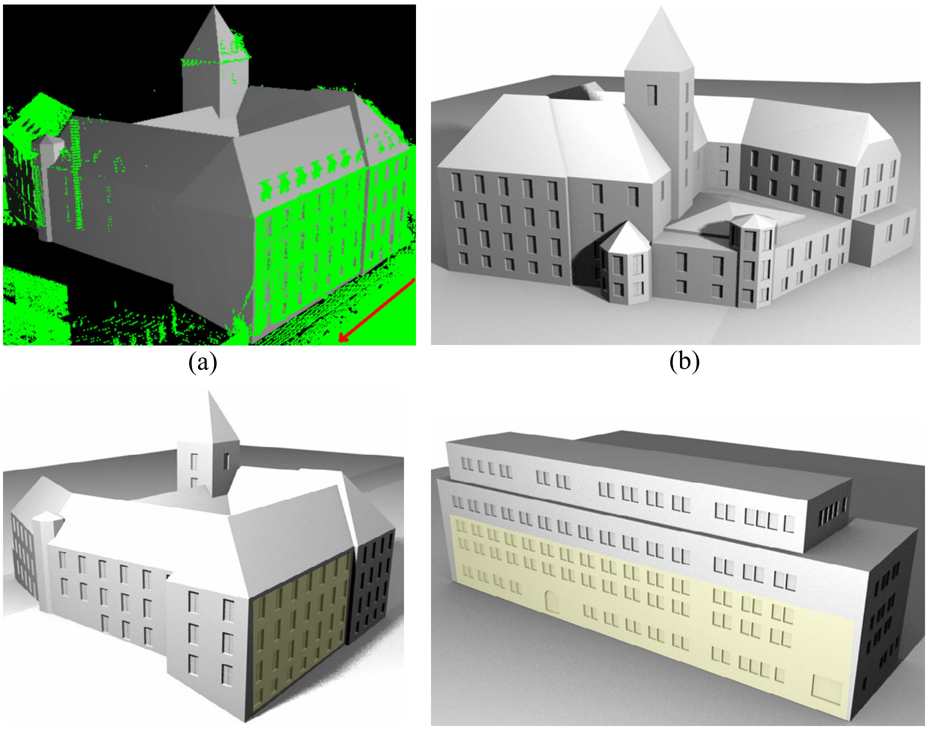
This quality measure is just one possibility to roughly assess the accuracy potential of the grammar supported reconstruction process. Other evaluation methods are currently analysed in order to find the most meaningful quality estimation. Furthermore, future work will go into more detail by searching for factors that influence the quality of the façade model. It will be examined whether certain properties of the grammar can be identified as affecting this quality significantly.

5 Results and conclusions

LiDAR data as it is used within our data driven reconstruction process is usually acquired by either static terrestrial laser scanning or mobile mapping from vehicle based systems. Despite the good geometric accuracy, which can be realised by mobile mapping systems [13], the unavailability of measurement data for building parts is particularly a problem with these systems since only street-facing façades can be observed. Fig. 4a depicts the configuration for the “Lindenmuseum, Stuttgart” where the LiDAR points are obtained by the mobile mapping system “StreetMapper” [13] for a single façade. The direction of driving is marked by a red arrow.

Restricted to the extent of those measurement areas, the output of the data driven bottom-up modelling is used to complement the rest of the building. Fig. 4b and Fig. 4c show the resulting 3D model from different viewpoints. Although the building is quite complex, geometrical structures derived from the street-facing façade (yellow shaded area in Fig. 4c) are sufficient for synthesising the whole building. While in this example at least one façade could be observed completely, typical problems arising with data acquisition by mobile laser scanning are demonstrated in Fig. 4d. Narrow streets with high buildings inevitably lead to oblique viewing angles and thus insufficient point densities in upper building parts. The office building in Fig. 4d represents such an example where only the lower floors could be reliably reconstructed based on LiDAR data. Nevertheless, using the grammar derived from the marked region, the façade and the superstructure on the roof could be completed.

Thus, due to the combination of bottom-up and top-down modelling, our approach is highly flexible towards data of different quality. In this regard, the completion of façade structures at areas of limited sensor quality has been demonstrated. Moreover, façade reconstruction is also possible for whole districts featuring uniform architectural styles if a small set of façade grammars is derived from just a few observed buildings. In addition to the good visual quality, which could be realised for all of our reconstructed 3D models, a number of other applications using the resulting building structures will require a decent quality assessment of the results. While different strategies already were proposed within the paper, this issue is still ongoing work.



(a) (b)

(c) (d)

Fig. 4. Coarse building model and available LiDAR points for the Lindenmuseum (a), results from grammar supported reconstruction for the Lindenmuseum (b,c) and an office building (d).

References

1. Becker, S., Haala, N.: Refinement of Building Façades by Integrated Processing of LIDAR and Image Data. IAPRS & SIS Vol 36 (3/W49A), pp. 7-12 (2007)
2. Müller, P., Wonka, P., Haegler, S.,Ulmer, A.,Van Gool, L.: Procedural Modeling of Buildings. ACM Transactions on Graphics (TOG) 25 (3), pp 614-623 (2006)
3. Bekins, D., Aliaga, D.: Build-by-number: Rearranging the real world to visualize novel architectural spaces. In IEEE Visualization, pp. 143-150 (2005)
4. Müller, P., Zeng, G., Wonka, P., Van Gool, L.: Image-based Procedural Modeling of Façades. ACM Trans. Graph. Vol. 26 (3), article 85, 9 pages (2007)
5. Van Gool, L., Zeng, G., Van den Borre, F., Müller, P.: Towards mass-produced building models. In IAPRS & SIS, Vol. 36 (3/W49A), pp. 209–220 (2007)
6. Henricsson, O. and Baltsavias, E.: 3-D building reconstruction with ARUBA: a qualitative and quantitative evaluation. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Ascona, pp. 65-76 (1997)
7. Akca D.,Freeman M.,Gruen A.,Sargent I.: Quality Assessment of 3D Building Data by 3D Surface Matching. ISPRS Congress Beijing 2008, Proceedings of Commission II, Vol. 37, Part B2, Com. II, ISSN 1682-1750 p.771 ff (2008)

8. Freeman M.,Sargent I.: Quantifying and visualising the uncertainty in 3D building model walls using terrestrial lidar data. Proc. of the Remote Sensing and Photogrammetry Society Conference 2008 'Measuring change in the Earth system'. Univ. of Exeter, 15-17.9. (2008)
9. Hinz, S., Wiedemann, C.: Increasing Efficiency of Road Extraction by Self-diagnosis. PE&RS, 70(12), p.1457 ff. (2004)
10. Förstner, W.: Diagnostics and performance evaluation in computer vision. NSF/ARPA Workshop: Performance vs. Methodology in Comp. Vision, Seattle, WA, pp. 11-25 (1994)
11. Schuhmacher, S., Böhm, J.: Georeferencing of Terrestrial Laser scanner Data for Applications in Architectural Modeling IAPRS Vol 36, Part 5/W17 (2005)
12. Oude Elberink, S., Vosselman, G.: Quality analysis of 3D road reconstruction. In: Laserscanning 2007. IAPRS & SIS, Espoo, Finland (2007)
13. Haala, N., Peter, M., Kremer, J., Hunter, G.: Mobile LiDAR Mapping for 3D Point Cloud Collection in Urban Areas - a Performance Test. IAPRS & SIS Vol. 37, Part B5, Com. 5. ISPRS Congress 2008, Beijing, China. P. 1119ff. (2008)