GRAMMAR SUPPORTED FACADE RECONSTRUCTION FROM MOBILE LIDAR MAPPING

Susanne Becker, Norbert Haala

Institute for Photogrammetry, University of Stuttgart
Geschwister-Scholl-Straße 24D, D-70174 Stuttgart
forename.lastname@ifp.uni-stuttgart.de

KEY WORDS: Architecture, Point Cloud, Urban, LiDAR, Facade Interpretation

ABSTRACT:

The paper describes an approach for the quality dependent reconstruction of building facades using 3D point clouds from mobile terrestrial laser scanning and coarse building models. Due to changing viewing conditions such measurements frequently suffer from different point densities at the respective building facades. In order to support the automatic generation of facade structure in regions where no or only limited LiDAR measurements are available, a quality dependent processing is implemented. For this purpose, facades are reconstructed at areas of sufficient LiDAR point densities in a first processing step. Based on this reconstruction, rules are derived automatically, which together with the respective facade elements constitute a so-called facade grammar. This grammar holds all the information that is necessary to reconstruct facades in the style of the given building. Thus, it can be used as knowledge base in order to improve and complete facade reconstructions at areas of limited sensor data. Even for parts where no LiDAR measurements are available at all synthetic facade structures can be hypothesized providing detailed building geometry.

1. INTRODUCTION

Due to the growing need for visualization and modelling of 3D urban landscapes numerous tools for the area covering production of virtual city models were made available, which are usually based on 3D measurements from airborne stereo imagery or LiDAR. This airborne data collection, which mainly provides the outline and roof shape of buildings, is frequently complemented by terrestrial laser scanning (TLS). However, the applicability of standard TLS is usually limited to the 3D data capturing of smaller scenes from a limited number of static viewpoints. In contrast, the application of dynamic TLS from moving platforms allows the complete coverage of spatially complex urban environments from multiple viewpoints. One example of such a mobile mapping system, which combines terrestrial laser scanners with suitable sensors for direct georeferencing, is the StreetMapper system (Kremer and Hunter, 2007). This system enables the rapid and area covering measurement of dense 3D point clouds by integrating four 2D laser scanners with a high performance GNSS/inertial navigation system. By these means accuracy levels better than 30mm have been demonstrated for point measurement in urban areas (Haala et al., 2008).

In general, such systems allow for an efficient measurement of larger street sections including the facades of the neighbouring buildings. However, depending on the look angle during the scanning process, strong variations of the available point densities at the building facades can occur. Such viewpoint limitations and occlusions will subject the collected point cloud to significant changes of accuracy, coverage and amount of detail. For this reason, the following interpretation of the measured point clouds will be hampered by considerable changes in data quality. Thus, algorithms for automatic facade reconstruction have to be robust against potentially incomplete data sets of heterogeneous quality. For this purpose, dense point cloud measurements for facades with good visibility are used in our approach to extract rules on dominant or repetitive features as well as regularities. These rules then are used as knowledge

base to generate facade structure for parts or buildings where no sensor data is available. By these means bottom-up and top-down propagation of knowledge can be combined in order to profit from both reconstruction techniques. The production rules, which are automatically inferred from well observed and modelled facades, are represented by a formal grammar.

Such formal grammars are frequently used within knowledge based object reconstruction to ensure the plausibility and the topological correctness of the reconstructed object elements. Lindenmayer-systems (L-systems), which can be applied to model the growth processes of plants, are well known examples of formal grammars (Prusinkiewicz and Lindenmayer, 1990). So-called split grammars are introduced by Wonka et al. (2003) to automatically generate architectural structures from a database of rules and attributes. Similarly, Müller et al. (2006) present a procedural modelling approach for the generation of detailed building architecture in a predefined style. However, the variety of facade structures which can be generated is restricted to the knowledge base inherent in the grammar rules or model libraries. Thus, the appearance of facade elements is limited to prespecified types, even when leaving some freedom in the values of their parameters. Another problem while applying such approaches for object reconstruction is that manual interaction is required to constitute suitable building styles and translate them into some kind of model or grammar description. For this reason, several approaches aim at deriving such kind of knowledge from observed or given data. For example, Ripperda (2008) derives prior facade information from a set of facade images in order to support the stochastic modelling process. However, existing methods which try to derive procedural rules from given images as proposed by Müller et al. (2007) or Van Gool et al. (2007) still resort to semi-automatic methods. The same holds true for the work of Aliaga et al. (2007). They present an interactive system for both the creation of new buildings in the style of others and the modification of existing buildings. At first, the user manually subdivides a building into its basic external features. This segmentation is then employed to automatically infer a grammar

which captures the repetitive patterns and particularities of the building. Finally, new buildings can be generated in the architectural style defined by the derived grammar. Even though this approach provides individually representative grammars instead of predefined ones, the crucial part of the inference process, the facade interpretation, has to be done manually. In contrast we pursue an approach which runs fully automatically during all processing steps.

The automatic generation of a facade grammar, which is derived from 3D point cloud measurements of a mobile mapping system, are discussed in section 2. As demonstrated in section 3 top-down predictions can be activated and used for the improvement and completion of the reconstruction result that has already been derived from the observed measurements during the bottom-up modelling. Moreover, the facade grammar can be applied to synthesize facades for which no sensor data is available. The discussion of 3D reconstruction results demonstrated in section 4 will conclude the paper.

2. GENERATION OF FACADE GRAMMAR

The automatic generation of a facade grammar based on terrestrial LiDAR data is the core of our facade modelling approach. The first step is a data driven reconstruction process aiming at the detection of geometric facade structures in the observed point clouds. In this regard, a facade defines a planar polygon with holes. Such holes indicate either windows, which will be modelled as indentations, or salient structures such as balconies, oriels or windowsills, which will be attached in the form of protrusions. The result of the data driven facade reconstruction serves as knowledge base for the generation of facade geometries where no sensor data is available. This knowledge, which includes information on dominant or repetitive structures as well as their interrelationships, can be inferred fully automatically and stored as a facade grammar. While data collection will be described as a pre-processing step in section 2.1, the basic concepts of the data driven reconstruction and the subsequent grammar inference will be addressed in section 2.2 and 2.3, respectively.

2.1 Data Collection

The StreetMapper mobile laser scanning system which was used for our experiments collects 3D point clouds at a full 360° field of view by operating four 2D-laser scanners simultaneously. The required direct georeferencing during 3D point cloud collection is realized by the integration of observations from GPS and Inertial Measurement Units (IMU). Figure 1 shows a 3D visualisation of the measured trajectory overlaid to the 3D city model which was also used for the following tests. This 3D city model is maintained by the City Surveying Office of Stuttgart. The roof geometry of the respective buildings was modelled based on photogrammetric stereo measurement while the walls trace back to given building footprints. The trajectory was captured during our tests within an area in the city centre of Stuttgart at a size of 1.5 km x 2km. The respective point clouds were measured at a point spacing of approximately 4cm. Figure 2 depicts a part of the StreetMapper point cloud at the historic Schillerplatz in the pedestrian area of Stuttgart. The observed points are overlaid to the corresponding 3D building models in order to show the quality and amount of detail of the available data. Another measured point cloud overlaid to an existing coarse building model is shown in

Figure 3. This example is used in the following to illustrate our bottom-up process for facade reconstruction. Within this process, the geometric information inherent in the available point cloud is exemplarily extracted for the facade marked by the white polygon.
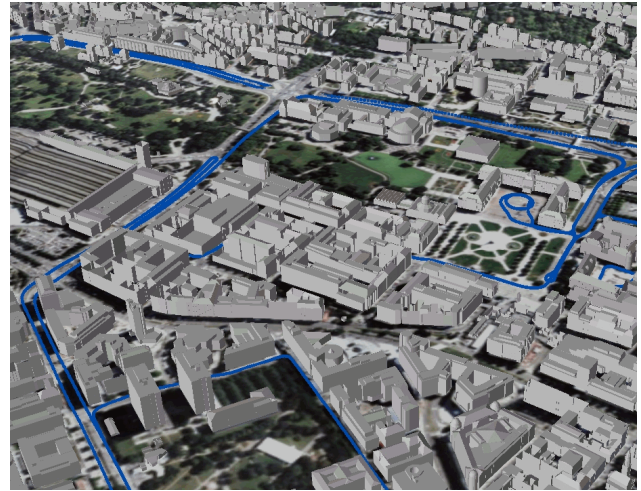

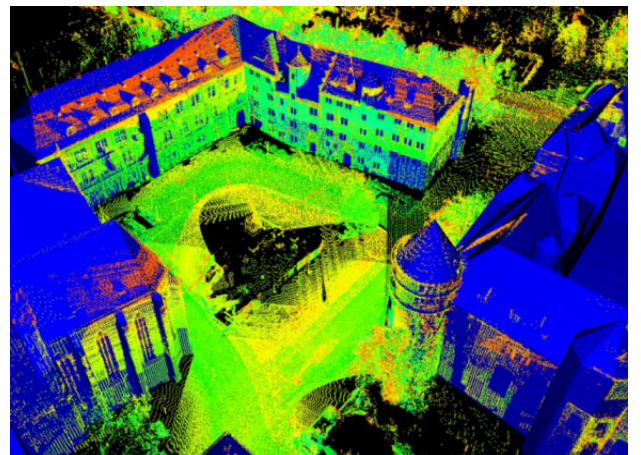Figure 1. 3D city model with overlaid trajectory from mobile TLS

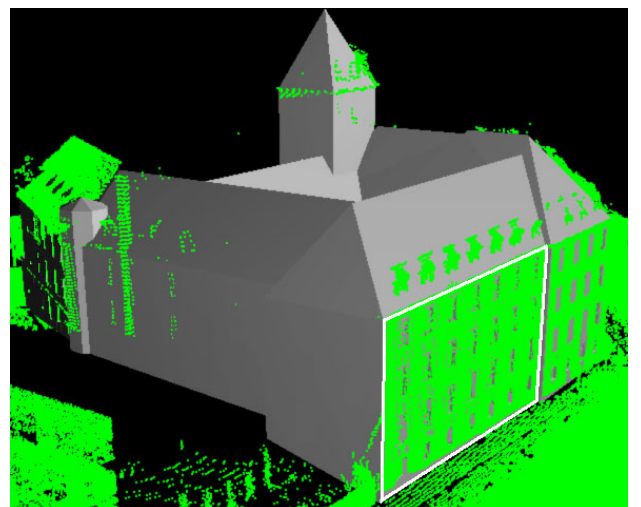Figure 2. Point cloud from mobile TLS aligned with virtual city model

Figure 3. Lindenmuseum, Stuttgart: point cloud from TLS aligned with existing coarse building model

2.2 Data Driven Reconstruction

Frequently, the representation of buildings is based on constructive solid geometry (CSG) or boundary representation (B-Rep). In contrast, we apply a representation of the buildings by cell decomposition (Haala et al., 2006). By these means, problems which can occur during the generation of topologically correct boundary representations can be avoided. Additionally, the implementation of geometric constraints such as meeting surfaces, parallelism and rectangularity is simplified. Due to the applied representation scheme, the idea of our reconstruction algorithm 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 facade or a window area. Therefore, they have to be differentiated depending on the availability of measured LiDAR points. After this classification step, window cells are eliminated while the remaining facade cells are glued together to generate the refined 3D building model. These steps are depicted exemplarily within Figure 4 and Figure 5, and will be explained in the following sections. The processing is based on the facade and point cloud marked by the white polygon in Figure 3.
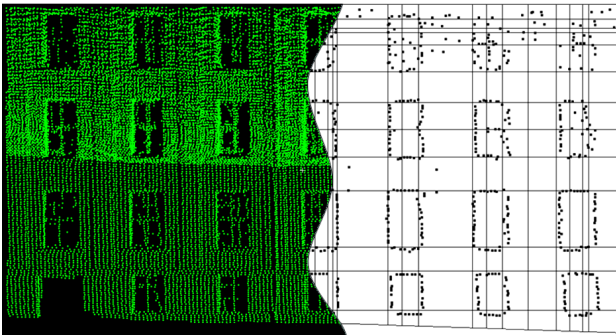

Figure 4. Lindenmuseum, Stuttgart: LiDAR point cloud (left), and detected edge points and window lines (right)
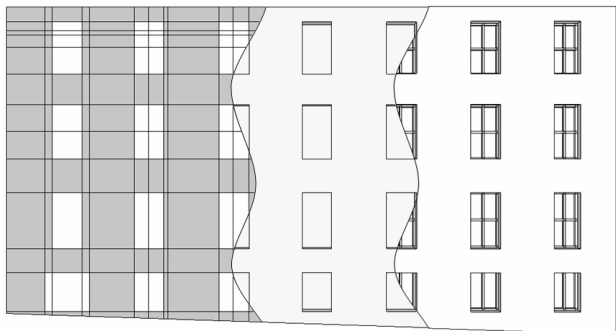

Figure 5. Lindenmuseum, Stuttgart: classified 3D cells (left), 3D facade model (middle), and refined 3D facade model (right)

2.2.1 Point Cloud Segmentation

At glass LiDAR pulses are either reflected or the glass is penetrated. Thus, as it can be seen in Figure 4(left), by laser scanning usually no points are measured in the facade plane at window areas. If only the points are considered that lie on or in front of the facade, the windows will describe areas with no point measurements. These no-data areas can be used for the point cloud segmentation which aims at the detection of window edges. For example, the edge points of a left window border are detected if no neighbour measurements to their right side can be found in a predefined search radius. In a next step, horizontal and vertical lines are estimated from non-isolated edge points. Figure 4(right) shows the extracted edge points at the window borders as well as the derived horizontal and vertical lines. Based on these window lines, planar delimiters can be generated for a subsequent spatial partitioning. Each boundary line defines a partition plane which is perpendicular to the facade. For the determination of the window depth, an additional partition plane can be estimated from the LiDAR points measured at the window crossbars. These points are detected by searching a plane parallel to the facade, which is shifted in its normal direction. The set of partition planes provides the structural information for the cell decomposition process. It is used to intersect the existing building model producing a set of small non-overlapping 3D cells.

2.2.2 Classification and Reconstruction

In order to classify the 3D cells into facade and window cells, a point-availability-map is generated. It is a binary image with low resolution where each pixel defines a grid element on the facade. The optimal grid size is a value a little higher than the point sampling distance on the facade. Grid elements on the facade where LiDAR points are available produce black pixels (facade pixels), while white pixels (non-facade pixels) refer to no-data areas. The classification is implemented by computing the ratio of facade to non-facade pixels for each 3D cell. Cells including more than 70% facade pixels are defined as facade solids, whereas 3D cells with less than 10% facade pixels are assumed to be window solids. While most of the 3D cells can be classified reliably, the result is uncertain especially at window borders or in areas with little point coverage. However, the integration of neighbourhood relationships and constraints concerning the simplicity of the resulting window objects allows for a final classification of such uncertain cells. Figure 5(left) shows the classified 3D cells: facade cells (grey) and window cells (white).

Within a subsequent modelling process, the window cells are cut out from the existing coarse building model. Thus, windows and doors appear as indentations in the building facade which is depicted in Figure 5(middle). Moreover, the reconstruction approach is not limited to indentations. Details can also be added as protrusions to the facade (Becker and Haala, 2007). However, the achievable level of detail for 3D objects that are derived from terrestrial laser scanning depends on the point sampling distance. Small structures are either difficult to detect or even not represented in the data. Nevertheless, by integrating image data with a high resolution in the reconstruction process the amount of detail can be increased (Becker and Haala, 2007). This is exemplarily shown for the reconstruction of window crossbars in Figure 5(right).

2.3 Automatic Inference of Facade Grammar

As it is already visible in Figure 3, the given scan configuration resulted in considerable variations of the available point coverage for the respective building. Thus, the bottom-up facade reconstruction presented in the previous section was realized for a facade, which is relatively well observed. This overall result is now used to infer the facade grammar. Frequently, such formal grammars are applied during object reconstruction to ensure the plausibility and the topological correctness of the reconstructed elements (Müller et al., 2006). In our application, a formal grammar will be used for the generation of facade 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 gets 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 facade grammar $\Gamma^{facade}(N,T,P,\omega)$ which allows us to synthesize new facades of various extents and shapes. The axiom ω refers to the new facade to be modelled and, thus, holds information on the facade polygon. The sets of terminals and non-terminals, *T* and *N*, as well as the production rules *P* are automatically inferred from the reconstructed facade as obtained by the data driven reconstruction process (section 2.2).

2.3.1 Searching for Terminals

In order to yield a meaningful set of terminals for the facade grammar, the building facade 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 facade 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 facade that has been generated during the data driven reconstruction process (section 2.2). 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 facade 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.

2.3.2 Interrelationship between Terminals

Having distinguished elementary parts of the facade 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. For more information on this process please refer to Becker et al. (2008). As an example, Figure 6a shows a modelled floor. While Figure 6b depicts the corresponding tile string in its original version, the compressed string and the extracted structures are given in Figure 6c. The hierarchical relations between the facade elements can be stored in a parse tree illustrated in Figure 6d.
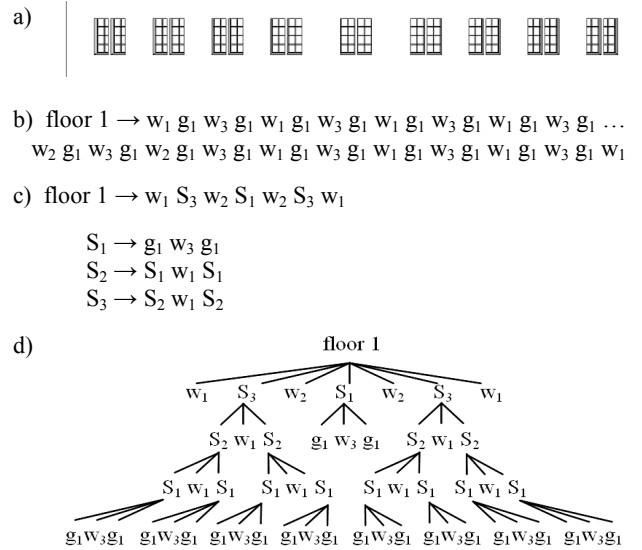
a)

b) floor $1 \rightarrow w_1 g_1 w_3 g_1 w_1 g_1 w_3 g_1 w_1 g_1 w_3 g_1 w_1 g_1 w_3 g_1 \cdots$
   $w_2 g_1 w_3 g_1 w_2 g_1 w_3 g_1 w_1 g_1 w_3 g_1 w_1 g_1 w_3 g_1 w_1 g_1 w_3 g_1 w_1$

c) floor $1 \rightarrow w_1 S_3 w_2 S_1 w_2 S_3 w_1$

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

d)

Figure 6. Modelled floor (a), corresponding tile string (b), compressed tile string and extracted structures (c), parse tree (d)

2.3.3 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 facade grammar can be inferred. Following types of production rules are obtained during the inference process:

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

The production rules p_1 and p_2 stem from the spatial partitioning of the facade. p_1 corresponds to the horizontal segmentation of the facade 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. A more detailed description of all rule types p_i and the probabilities $P(\boldsymbol{x}|p_i)$ assigned to them can be found in Becker et al. (2008).

3. APPLICATION OF FACADE GRAMMAR

Our facade grammar derived in the previous section implies information on the architectural configuration of the observed facade concerning its basic facade elements and their interrelationships. Based on this knowledge facade hypotheses can be generated as described in section 3.1. Section 3.2 presents different application scenarios. Facades and building parts which are covered by noisy or incomplete sensor data are usually subject to inaccurate and false reconstructions which are due to problems of the data driven reconstruction process. For such regions possible facade geometry can be proposed in order to improve and complete facade structures. Furthermore, the production process can also be used to synthesize totally unobserved building objects.

3.1 Production of Facade Hypothesis

The production process starts with an arbitrary facade, 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). The geometrical result of the production process depends on the order in which the non-terminals are selected. Usually, best results are obtained when facade structures which are likely to appear in the middle of the facade are placed first, and the remaining spaces to the left and the right side are filled afterwards. As it is illustrated in Figure 7, the non-terminal selection refers to this principle. For clearness, we here assume a facade with only one floor. In each step, the non-terminal selected for the next substitution is marked in red.

Figure 7. Non-terminal selection

As long as the facade string consists of only one symbol, the non-terminal selection is trivial. In the third line, substitution starts with the non-terminal G in the middle of the string. According to this replacement, the chosen geometry tile g_i will be placed about in the middle of the facade floor. The following replacements are taken from the left to the right of the string. When there is only one non-terminal left on the right end of the string (see the last line in Figure 7), the left part of the facade floor is completely filled with a sequence of wall and geometry tiles. At this stage, symmetry can be enforced by substituting the remaining non-terminal W by a mirrored version of the left terminal string. If no symmetry is required, the replacement can be continued as described before. 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 facade string contains only terminals, the production is completed and the string can be transferred into a 3D representation.

3.2 Application Scenarios

Within the production process, the grammar is applied to generate hypotheses about possible positions of each geometry tile and thereby synthesize facade geometry for given coarse building models. This process can for example be used to generate facade structure at areas, where sensor data is only available at limited quality. Such a scenario is depicted exemplarily in Figure 8, which shows a StreetMapper point cloud for an exemplary facade acquired during two epochs. The colours encode the different scanners mounted on the StreetMapper. Points that stem from the upward facing laser scanner are marked in yellow; points that are measured by the

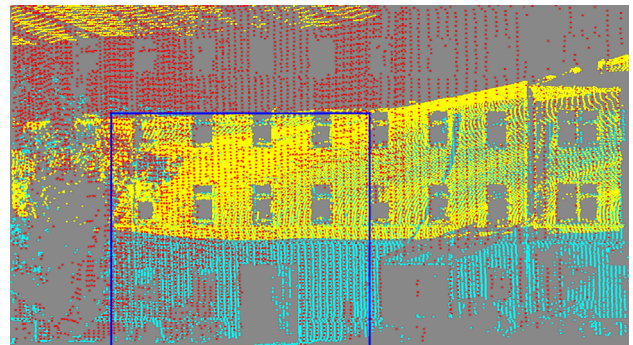side facing scanners are blue (right scanner) and red (left scanner).

Figure 8. Measured facade points and determined convex 'dense area' (blue rectangle)

As it is visible in Figure 8, the point sampling distance varies strongly due to occlusions and oblique scanning views to the upper part of the building. For this reason, facades may contain areas where no or only little sensor data is available. In such regions, an accurate extraction of windows and doors cannot be guaranteed anymore. Nevertheless, a grammar based facade completion allows for meaningful reconstructions even in those areas. The main idea is to derive the facade grammar solely from facade 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 evaluating the sampling distances of the points lying on the facade surface. In Figure 8 the extracted convex dense area is marked by a blue rectangle. Since the inference process is restricted to this dense area, a facade grammar of good quality can be provided, which is then used to synthesize the remaining facade regions during the production step.
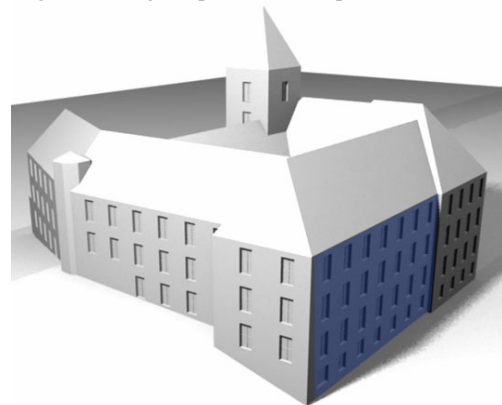

Figure 9. Facade reconstruction for the "Lindenmuseum"

Figure 10. Facade geometry synthesized from grammar library

As an example, this process is depicted in Figure 9 for Stuttgart's Lindenmuseum which has already been illustrated in Figure 3. There, the original coarse model is shown in combination with the overlaid 3D point cloud whereas Figure 9 demonstrates the reconstructed facade geometry. The blue shaded region corresponds to the white polygon in Figure 3 and indicates the facade geometry that has been generated during the data driven reconstruction process. All remaining building parts are modelled based on the grammar inferred from the marked region.

While in that example the grammar is applied for the completion of facade structure, it can also be used as a "library" to generate building facades for objects, where no measurement is available at all. This step is demonstrated in Figure 10 where facade geometry was synthesized for a number of residential houses. Since these building were not covered by any sensor data at all, a range of grammars was derived in advance from a few buildings in the near environment. Similarly, the applicability of our facade grammars to a larger scene using grammars which represent compatible architectural styles is shown in Figure 11.

4. DISCUSSION

Within the paper an automatic approach for the geometric modelling of 3D building facades was proposed. Based on observed 3D point clouds from a mobile mapping system, grammar rules are extracted, which can then be used to generate synthetic facade structures for unobserved building parts. Despite the good geometric accuracy which is feasible for terrestrial point clouds such data frequently suffer from the unavailability of measurements for hidden building parts. This problem is solved by extracting the grammar from observed street-facing facades and then applying it for the improvement and completion of remaining facade structure in the style of the respective building. Moreover, knowledge propagation is not restricted to facades of one single building. Based on a small set of facade grammars derived from just a few observed buildings, facade reconstruction is also possible for whole districts featuring uniform architectural styles. Due to these reasons the proposed algorithm is very flexible towards different data quality and incomplete sensor data.

5. REFERENCES

Aliaga, D., Rosen, P., Bekins, D., 2007. Style Grammars for Interactive Visualization of Architecture. IEEE TVCG 13 (4).

Becker, S., Haala, N., 2007. Refinement of Building Facades by Integrated Processing of LIDAR and Image Data. IAPRS & SIS Vol. 36 (3/W49A), pp. 7-12.

Becker, S., Haala, N., Fritsch, D., 2008. Combined Knowledge Propagation for Facade Reconstruction. IAPRS & SIS Vol. 37 (B5), pp. 1682-1750.

Haala, N., Becker, S., Kada, M., 2006. Cell Decomposition for the Generation of Building Models at Multiple Scales. IAPRS Vol. 36 (3), pp. 19-24.

Haala, N., Peter, M., Kremer, J., Hunter, G., 2008. Mobile LiDAR Mapping for 3D Point Cloud Collection in Urban Areas - a Performance Test. IAPRS, Vol. 37, (B5), pp. 1119f.

Kremer, J., Hunter, G., 2007. Performance of the StreetMapper Mobile LIDAR Mapping System in "Real World" Projects. Photogrammetric Week '07, pp. 215-225.

Müller, P., Wonka, P., Haegler, S., Ulmer, A., Van Gool, L., 2006. Procedural Modeling of Buildings. ACM Transactions on Graphics (TOG) 25 (3), pp 614-623.

Müller, P., Zeng, G., Wonka, P., Van Gool, L., 2007. Image-based Procedural Modeling of Facades. ACM Transactions on Graphics (TOG) 26 (3), article 85, 9 pages.

Prusinkiewicz, P., Lindenmayer, A., 1990. The algorithmic beauty of plants. New York, NY: Springer.

Ripperda, N., 2008. Determination of Facade Attributes for Facade Reconstruction. IAPRS & SIS Vol. 37 (B3a), 6 pages.

Van Gool, L., Zeng, G., Van den Borre, F., Müller, P., 2007. Towards mass-produced building models. IAPRS & SIS, Vol. 36 (3/W49A), pp. 209–220.

Wonka, P., Wimmer, M., Sillion, F., Ribarsky, W., 2003. Instant architecture. ACM TOG 22 (3), pp. 669–677.

Figure 11. Facade geometry for larger area