Photogrammetric Week '11   Dieter Fritsch (Ed.)   Wichmann/VDE Verlag, Belin & Offenbach, 2011

Towards Complete LOD3 Models – Automatic Interpretation of Building Structures

SUSANNE BECKER, Stuttgart

ABSTRACT

The paper presents our work on the automatic reconstruction of complete LOD3 building models. Existing LOD2 models are refined by adding 3D façade geometry with explicit semantic information to the planar facades. Aiming at robust and realistic façade structures, independent of the quality and availabilty of sensor data, object knowledge in the form of a grammar is automatically inferred and applied.

The grammar-based modeling concept is integrated in a learning system environment. Within a data-driven algorithm, basic façade elements such as windows and doors are automatically extracted and modeled from terrestrial LiDAR point clouds. After this interpretation step, superior structures and hierarchical relationships are derived from the obtained 3D façade model and transferred to a grammar-based description. This is implemented by means of a specially developed formal grammar, called *façade grammar*, which is appropriate to express the architectural style of a building through individually adapted rules. By applying such a façade grammar realistic façade structures can be generated even for building regions for which no sensor data or only measurements of poor quality are available.

However, opportunities for grammar-based interpretation and modeling of building structures do not end at this point. The concept of describing geometrical and semantic relations in a formal way can also be applied for the representation of indoor environments as well as large-scale relationships between whole buildings or districts of a city. The final part of the paper proposes how the façade grammar could be integrated in a more general graph-oriented structure which allows for grammar-based modeling of buildings in various levels of details and different contextual environments.

1. INTRODUCTION

Developments in the area of acquisition, visualization and interpretation of 3D geodata have been advancing rapidly for several years. New systems and algorithms simplify the handling of huge and complex data. Once being offered, the steady availability of masses of information induces at the same time further demands of the users. No doubt, researchers are facing a wide range of topics to deal with. Even in daily life, 3D geodata are of growing relevance: Digital globes such as Google Earth or Microsoft Bing Maps 3D attract the public's attention. When urban environments are to be explored, the existence of 3D city models is of great importance.

Mostly generated from airborne LiDAR or image data, 3D building models are usually available as LOD2 models. These building representations consist of detailed roof shapes but planar, unstructured façades often enriched by image textures (Kolbe et al., 2005). While 2D façade textures provide additional information about the visual appearance of the buildings, images are, though, just a set of gray values requiring the observer's intelligence to be understood. However, façade interpretation – especially knowledge about the position and size of functional façade elements such as windows and doors – is of increasing interest. Driving forces are considerable progress in the areas of computer graphics, virtual reality, detailed urban planning or 3D navigation. They all require at least LOD3 models in which façade structures are represented as distinct 3D geometries which are added as indentations or protrusions to the façade plane. Nevertheless, the moment necessity of applications goes beyond pure visualizations, semantic information occurs to be essential. Knowledge about functional aspects of façade elements make the difference between façade structures that are explicitly modeled and identified as functional entities, and façade structures which, for example, are added as uninterpreted gray values or 2.5D relief to the façade plane. As soon as facades offer detailed geometric and semantic properties, a huge field of potential applications is being opened. Complex search queries embedded in a 3D context, energetic calculations such as the estimation of a building's heat loss, or navigation tasks adapted to the needs of certain groups of people – for instance leading a blind person to the entry of a building –, represent just a few aspects of what could be possible. Even for civil engineering, explicitly modeled façade geometries may support the simulation of a building's reaction to tremors caused by earthquakes etc. Interpreted façade models provide not least a valuable input to intelligent environments in which spatial objects are multiply cross-linked based on geometric, topological or semantic properties. Such virtual environments are the framework for the modeling and simulation of complex processes and work flows which need to consider a number of dependencies between the objects involved and possible decisions and interactions by the users. The knowledge of façade structures could particularly contribute to Building Information Models (BIM). There, semantic information about buildings is consolidated in order to gain greater insight into specific building aspects, and to ease the development of efficient solutions for the management of the buildings.

In order to reconstruct detailed façade elements such as windows and doors, frequently terrestrial LiDAR point clouds and images are used. By combining several laser scanners and/or image sensors to an integrated measurement device mounted on top of a vehicle, mobile mapping systems allow for efficient data collection in street scenes (Kutterer, 2010). While today's systems are able to already generate dense point clouds of sufficient quality, ongoing improvements of the sensors and calibration algorithms feed expectations for even better results. Beside great advances in sensor technology, recent progress in image matching algorithms further expands the range of possibilities in generating dense 3D point clouds of high quality. Rothermel and Haala (2011), for example, prove the great potential of the Semi-Global Matching method for the automatic derivation of Digital Elevation Models from aerial image data.

However, in spite of the continuous improvement of data quality concerning the accuracy and density of 3D point clouds, there still remains an intractable problem in terrestrial data acquisition, i.e. the problem of partial occlusions which are due to obstacles in the line of sight leading to significant holes in the point clouds. Especially when buildings are observed from vehicle-based systems which are forced to follow the course of the streets, obstacles such as parking cars, benches or trees cannot be avoided. As a consequence, 3D points measured on building facades will show an inhomogeneous distribution hindering the extraction of building structures in consistently good quality. For instance, as a result of a test study dealing with mobile mapping systems, Rutzinger et al. (2009) report that only 56% of the street-facing facades could be identified as wall faces. The authors give partial occlusions caused by vegetation and fences as a reason.

Dependence on data quality and completeness is a general drawback with data-driven reconstruction methods. Since geometrical features are directly extracted and modeled from the measurements, such approaches are relatively sensitive to erroneous or incomplete data. In this respect, model-based reconstructions are much more robust. Algorithms cope with data uncertainty by the integration of knowledge about the appearance and arrangement of object structures. Object knowledge can be represented in different ways. One possibility is implicit shape models, which, for example, Reznik and Mayer (2008) employ for the appearance based detection and delineation of windows in façade imagery. An appearance based approach is also applied by Wenzel and Förstner (2008) for façade interpretation. Another possibility to integrate knowledge into building reconstruction is generative modeling, where the parameters of a predefined 3D template model are fitted to observation data during a stochastic process. Dick et al. (2004) used this technique for the modeling and interpretation of architecture from images.

Beyond that, knowledge, which is inherently available in models, can also be stored in form of a grammar. Formal grammars – a generic concept that has its origin in informatics where it is needed to define programming languages and compilers – also have been applied successfully for building reconstruction for several years. Grammars offer the advantage that they are more flexible towards the great variety of real world objects than other knowledge representations. By means of a formal grammar, manifold relationships, hierarchies and symmetries between the different parts of a building can be described efficiently and introduced as a priori knowledge in the reconstruction.

Generally, knowledge-based reconstruction techniques provide a means for procedurally creating architecture in a predefined style. However, this style first has to be constituted and translated into some kind of model or grammar description manually, which takes a lot of effort and requires expert knowledge of a specially skilled person. Therefore, we developed a generic template grammar being most flexible to describe arbitrary configurations of façade elements such as windows and doors. By integrating this generic grammar in a learning system environment, individual instances of façade grammars can be derived automatically from building structures that have been extracted from sensor data beforehand. Since such an individual façade grammar holds all the information which is necessary to reconstruct façade geometries in the style of the observed building, realistic façade structures can be generated even for building regions for which no sensor data or only measurements of poor quality are available. Combining data-driven and knowledge-based techniques in this way, we obtain a fully automatic and robust method for façade reconstruction.

The paper is organized as follows. Section 2 gives an overview of the usage of formal grammars in architectural modeling and introduces our idea of a generic façade grammar. Section 3 describes the main steps of the developed reconstruction algorithm. Results are presented for exemplary buildings showing different architectural characteristics. Finally, section 4 demonstrates opportunities for integrating the façade grammar in a more general graph-oriented structure which allows for grammar-based modeling of buildings in various levels of details and different contextual environments.

2. GRAMMAR IN ARCHITECTURE

Architecture is subject to a number of conditions and principles which arise from the intention to create habitable space while considering aesthetic aspects. Functional restrictions, for example minimum heights for doors or floors, have to be reconciled with statical requirements. Beyond that, there are also topological relations defining typical floor and room arrangements. Geometrical and topological constraints have an impact on the design decisions of an architect and, thus, are finally reflected in the layout of the constructed building. Mitchell (1990) understands architecture as a process of transforming an amorphous world into an ordered state. Following this view, the focus is on the partitioning of space in room entities and, thus, on the structural concept of a building. The visual impression of a façade is affected by the form elements, which occur on the façade plane as indentations or protrusions, as well as the way how these geometric shapes are arranged. Frequently, façade designs either feature horizontally or vertically accentuated window configurations (see Figure 1). Such predominant structures emerge when single elements of a façade are perceived as parts of a superior geometric entity. Such grouping phenomena have been discussed thoroughly in Gestalt psychology; so-called Gestalt laws were formulated describing how a human being anticipates structural coherence in a configuration of distinct shapes (Wertheimer, 1923; Arnheim, 1974). Groupings based on Gestalt laws reveal high-order patterns of structuring which have a considerable impact on the appearance of the façade.

Efforts to detect logic relations and regularities and to express them mathematically go back to the 1960s (Alexander, 1964). While originally the focus was set to formalize patterns of thought and reasoning, meanwhile also geometric objects and construction processes are taken into account. This is due to the fact that geometrical and topological constraints as well as principles of composing a high-order structure from single elements are highly appropriate to be translated into rule-based descriptions. In this context, formal grammars are a powerful tool. After giving a short overview of recent developments in grammar-based façade modeling in section 2.1, our concept of a generic façade grammar is introduced in section 2.2.

Figure 1: Horizontally and vertically accentuated façade designs: (left) Zeppelin-Bau, (right) Staatstheater Großes Haus.

2.1. Formal grammars and their application to geometric modeling

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.

A famous example for formal grammars is given by Lindenmayer systems (L-systems) (Prusinkiewicz and Lindenmayer, 1990), which can be used to model the growth processes of plants. Even though L-systems allow for the procedural modeling of complex objects, they are not appropriate for the modeling of buildings. In contrast to plants, buildings do not grow in free space, and modeling can be described rather by a space partitioning than a growth-like process. A more suitable grammar for the modeling of architecture is given by Stiny and Gips (1972). They introduce shape grammars, which define rules for the specification and transformation of 2D and 3D shapes. A variation on shape grammar, called split grammar, is employed by Wonka et al. (2003) in order to automatically generate architectural structures from a database of rules and attributes. Following this idea, Müller et al. (2006) develop the CGA shape grammar and present a procedural modeling approach for the generation of detailed building shells without considering any sensor data. They solely use context-sensitive shape rules to implement splits along the main axes of the facades.

However, while approaches like these are able to create 3D building models in a high level of detail, the formal rules applied have to be set up manually which is a very difficult and time consuming process. Another problem is that the variety of façade structures to be generated is restricted to the knowledge base inherent in the grammar rules or model libraries. The appearance of façade elements is limited to prespecified types, even when leaving some freedom in the values of their parameters. Also, the arrangement of façade structures is at least roughly given in advance due to the predefined partitioning procedures established in the grammar rules. In order to be more flexible in this respect, several systems aim at deriving some kind of knowledge from observed or given data. For example, Ripperda (2008) derives prior façade information from a set of façade images in order to support the stochastic modeling process. Methods which try to infer procedural rules from given images as proposed by Müller et al. (2007) and 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. 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 façade interpretation, has to be done manually.

2.2. Façade grammar

Our development of a formal grammar, called *façade grammar*, operates on existing coarse LOD2 building models in order to enrich them by procedurally adding 3D geometries to the planar facades. Similar to the CGA shape grammar, the façade grammar applies basic principles of the split grammar and the L-systems. However, the CGA shape grammar requires the manual definition of rules and elementary façade structures which is an extremely complex process especially when the facades to be reconstructed show a high level of detail. By contrast, our façade grammar is designed in a way which allows for the automatic derivation of both grammar rules and the geometric façade elements. Thus, manual interaction is becoming obsolete. Based on observed 3D data, basic façade elements and configurations are automatically detected and stored as terminals and production rules. For the purpose of being most flexible towards the great variety of possible façade geometries, the grammar components are of generic nature so that they can react to real world situations and express not only regular arrangements but also irregular ones.

The façade grammar can be described by its non-terminals, terminals, production rules and the axiom. The non-terminals and terminals symbolize finite regions on the façade, in the following also referred to as *tiles*. Two main types of tile occur: *wall tiles* and *geometry tiles*. While a wall tile describes a homogeneous part on the façade, a geometry tile contains a 3D geometry which can be either an indentation, for instance a window or door, or a protrusion such as a balcony. Figure 2 illustrates a façade part consisting of two wall tiles on the left and the right, and a geometry tile with a window in the middle. 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 abbreviated by the symbols G and g_i in case of non-terminals and terminals, respectively. The axiom of our grammar represents the planar façade polygon of a LOD2 building which is to be enriched by additional 3D façade geometries.

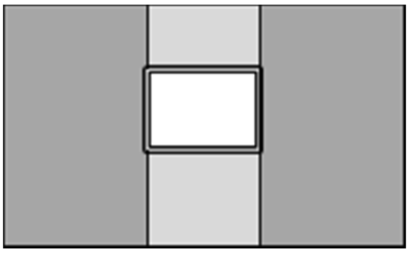


Figure 2: Façade part with two wall tiles (*dark gray*) and one geometry tile (*light gray*).

Concerning the rules, we can distinguish between the so-called *split rules* and *instantiation rules*. The application of split rules effect the recursive partitioning of the façade area into tiles. However, unlike the split rules of the CGA shape grammar and the split grammar, here these rules are kept general enough to also model inhomogeneous distributions of façade elements. Instantiation rules implement the transition from non-terminals to terminals and, thus, create geometric instances for the geometric entities which are inherent in the geometry tiles. The rules are parametric, stochastic and partially context-sensitive. In their general form they can be written as follows:

$$\textit{left context} < \textit{predecessor} > \textit{right context} : \textit{condition} \rightarrow \textit{successor} : \textit{probability}$$

The grammar description given in this section defines a kind of template grammar which can be fitted to real facades leading to explicit instantiations of individual façade grammars. Each façade grammar reflects the architectural style of the respective building. The process of grammar inference, which is discussed in section 3.2, runs fully automatically.

3. FACADE RECONSTRUCTION

Our approach for façade reconstruction aims at refining an existing coarse LOD2 building model by adding 3D geometries to the planar facades. The focus is set on functional façade elements such as windows, doors or balconies. For the purpose of being flexible to the great variety of possibly occurring structures on the one hand, but showing robustness against noisy and incomplete sensor data on the other hand, we combine data-driven and model-based techniques. Section 3.1 gives a short overview of the data-driven component. 3D LiDAR points measured at façade areas are interpreted geometrically and semantically. The resulting façade model contains structural knowledge about the design and configuration of the reconstructed façade geometries. Thus, it can be used as knowledge base for further modeling. During an inference process described in section 3.2, dominant and repetitive structures are learned from the data-driven façade model and mapped to the basic components of our façade grammar, i.e., an individual instance of the façade grammar is created. Applications and reconstruction results for different façade grammars are presented in sections 3.3 and 3.4. Further examples as well as more detailed descriptions of the proposed approach can be found in Becker (2009) and Becker (2011).

3.1. Data-driven extraction of basic façade structures

The structural information which is needed to refine the planar facades of LOD2 models is extracted from terrestrial LiDAR point clouds. In this paper, we will refer to exemplary data sets collected by the Leica scanner HDS3000 (Schuhmacher and Böhm, 2005) and the mobile mapping system StreetMapper (Haala et al. 2008). The respective point clouds were measured at a point spacing of approximately 4 to 10 cm. Figure 3(a) depicts a part of the point cloud acquired by the HDS3000 at the historic Schillerplatz in the pedestrian area of Stuttgart. The observed points are overlaid to the corresponding LOD2 building models provided by the City Surveying Office of Stuttgart.

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. 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 which represent either a homogeneous part of the façade or a window area. 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 laser scanning usually no points are measured in the façade plane at window areas since LiDAR pulses are either reflected or the glass is penetrated. If only the points are considered that lie on or in front of the façade, the windows will describe areas with no point measurements. Thus, window edges can be detected by searching for holes in the point cloud.

In a next step, horizontal and vertical lines are estimated from non-isolated edge points. Figure 3(b) 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 façade. 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 façade, 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. Each 3D cell covers either a blank wall area of the façade or is part of a window region. Therefore, they have to be differentiated based on the availability of LiDAR points lying in the façade plane.

After this classification step, the result of which is illustrated in Figure 3(c), the window cells are cut out from the existing coarse building model. Thus, windows and doors appear as indentations in the building façade which is depicted in Figure 3(d). Moreover, the reconstruction approach is not limited to indentations. Details can also be added as protrusions to the façade (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. This is exemplarily shown for the reconstruction of window crossbars in Becker and Haala (2007).

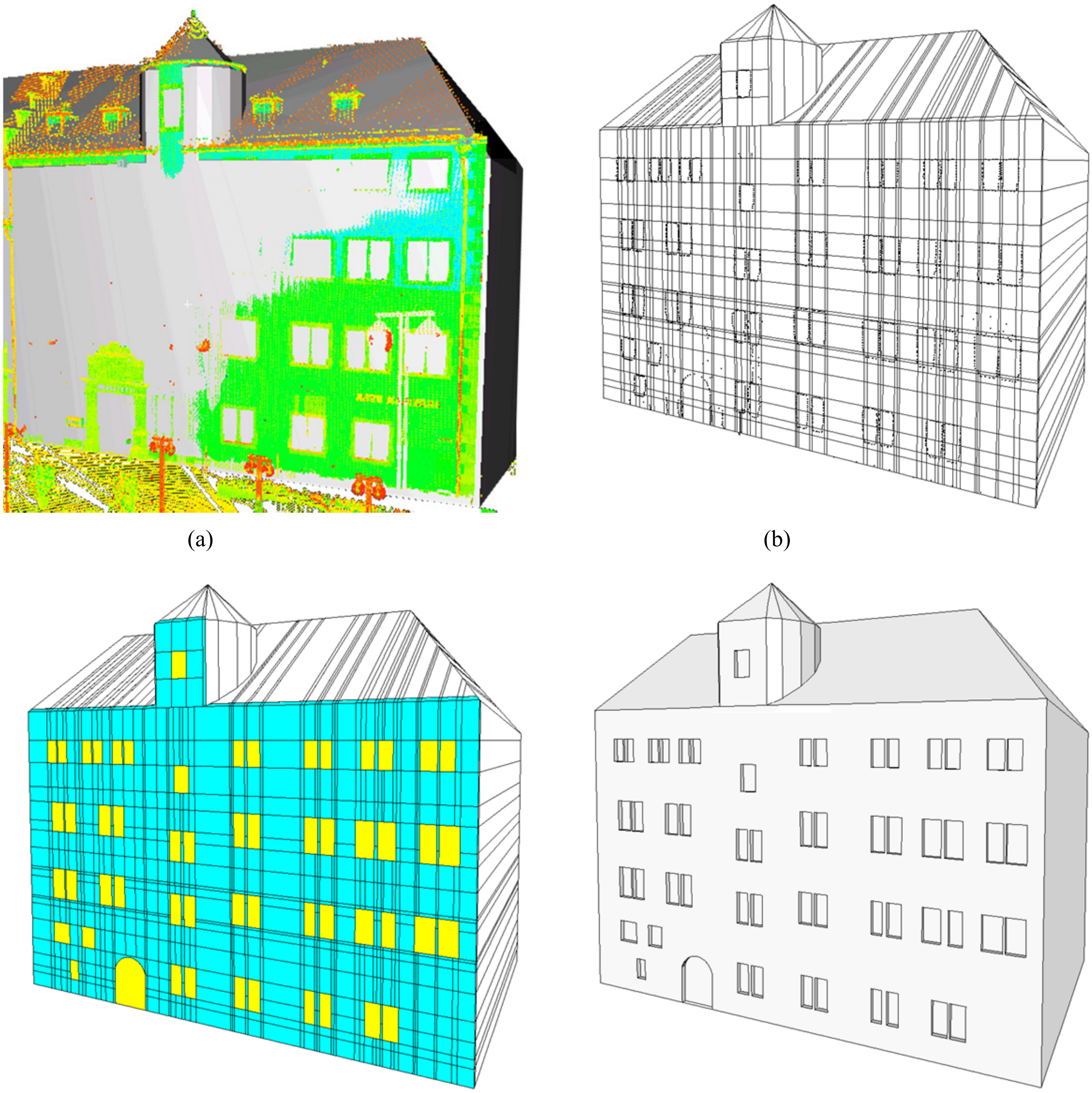


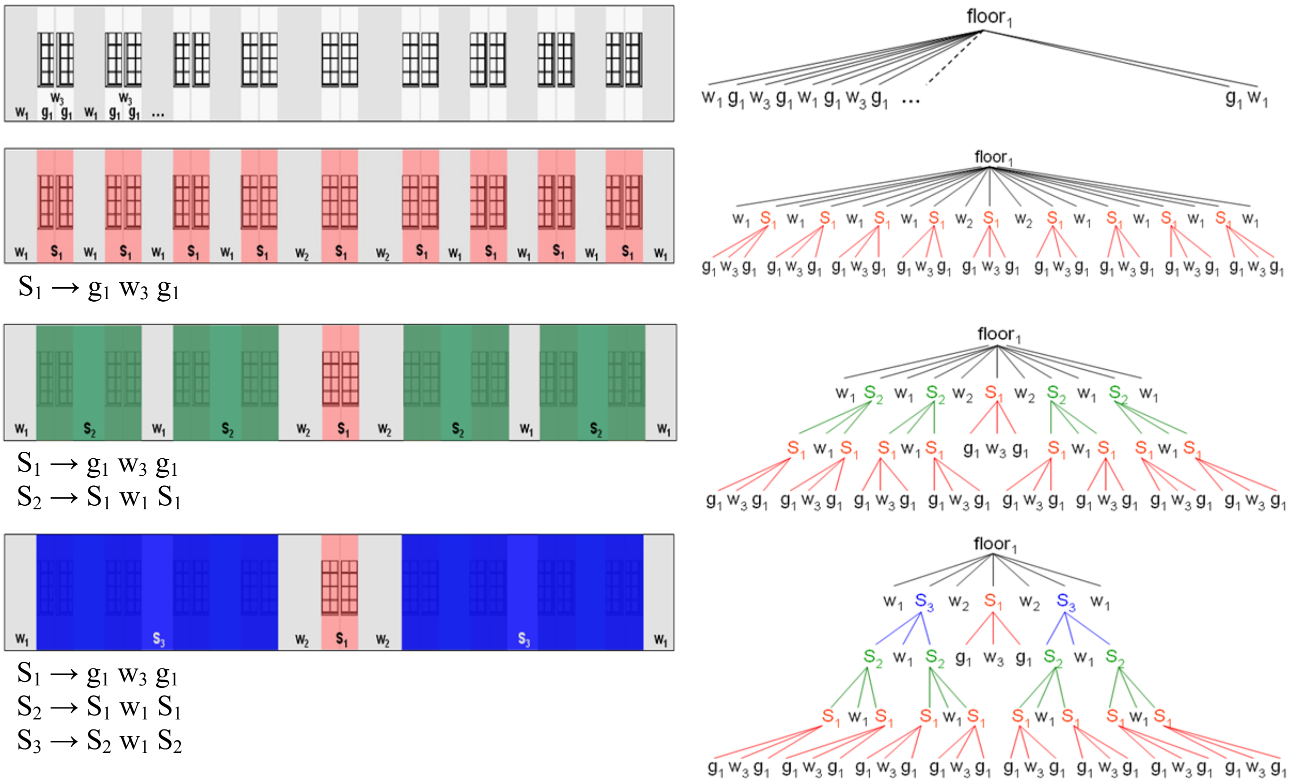
Figure 3: Data based reconstruction process for the example „Alte Kanzlei, Stuttgart“: (a) LiDAR point cloud and coarse building model, (b) detected edge points and window lines, (c) classified 3D cells, (d) refined 3D façade model.

3.2. Automatic instantiation of individual façade grammars

The result of the data-driven façade reconstruction provides information on the building’s architectural style which is affected by both the basic façade elements (e.g. windows and doors), and their interrelationship. Thus, a 3D façade model as obtained in the previous section can be used to infer an individual façade grammar representing the characteristic architecture of the respective building. While formal grammars are frequently applied during object reconstruction to ensure the plausibility and the topological correctness of the reconstructed elements (Müller et al., 2006) we will use our façade grammar for the generation of realistic façade structures where only partially or no sensor data is available.

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 windows and doors generated during the data-driven reconstruction process (section 3.1). The resulting sets of wall tiles and geometry tiles, which are assigned character symbols as described in section 2.2, serve as terminals within the individual façade grammar.

Having distinguished elementary parts of the façade we now aim at giving further structure to the perceived basic tiles by grouping them into superior 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. Fig. 4 shows how hierarchical structures are extracted stepwise by performing this process iteratively to an exemplary floor of a historic building called “Prinzenbau”. While the left of Figure 4 contains 3D representations of the tile sequences together with the currently detected structures, the right of Figure 4 illustrates the hierarchical relations by means of parse trees.



$S_1 \rightarrow g_1\ w_3\ g_1$

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

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

Figure 4: Inference of hierarchical relationships for the example „Prinzenbau, Stuttgart“.

Based on the extracted terminals and structures, the production rules for our façade grammar are automatically inferred. In the following, the five most important types of production rule are listed:

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

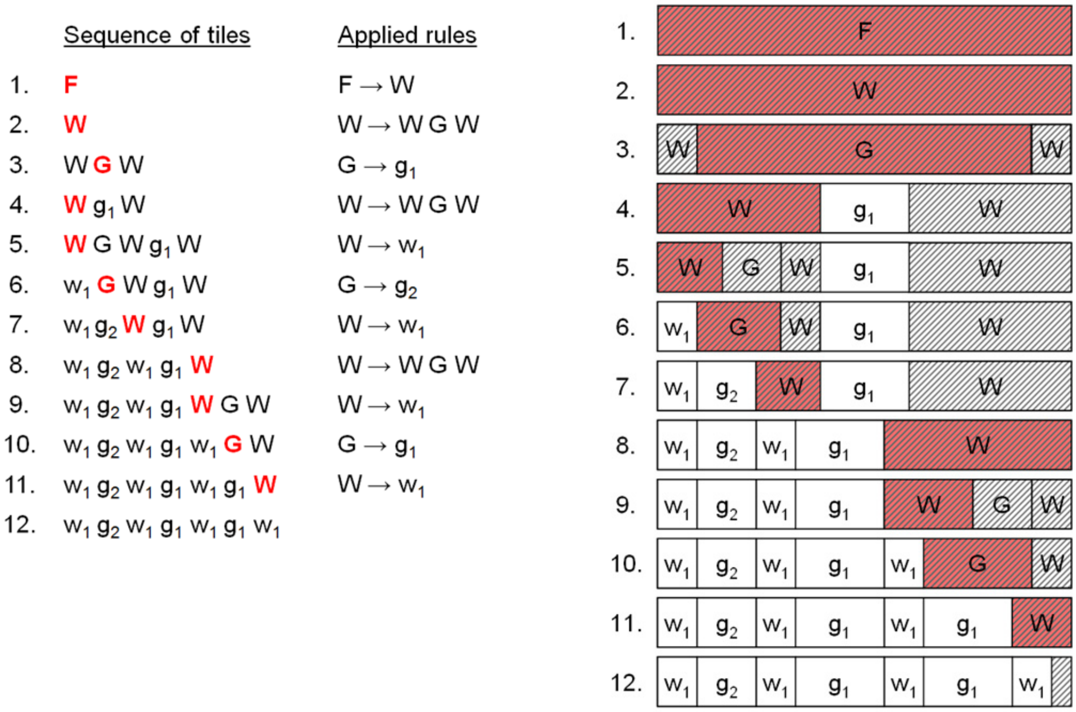
The production rules p_1 and p_2 are split rules stemming from the spatial partitioning of the façade. p_1 corresponds to the horizontal segmentation of the façade into a set of floors. The vertical partitioning into tiles is reflected in rule p_2. A wall tile, which in the first instance can stand for a whole floor, is replaced by the sequence wall tile, geometry tile, wall tile. All remaining rules are of type "instantiation rule". 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 $P(\mathbf{x}|p_i)$ assigned to them can be found in Becker (2011).

3.3. Application of façade grammar

Our façade grammar derived in the previous section implies information on the architectural configuration of the observed façade concerning its basic façade elements and their interrelationships. Based on this knowledge façade hypotheses can be generated within a so-called production process. This 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). The geometrical result of the production process depends on the order in which the non-terminals are selected. Usually, best results are obtained when façade structures which are likely to appear in the middle of the façade are placed first, and the remaining spaces to the left and the right side are filled afterwards.

Figure 5(a) illustrates this principle by means of a successively increasing tile sequence. The geometric interpretation of the tile string is given in Figure 5(b). Exemplarily, a simple façade grammar shall be applied consisting of the terminals, w_1, g_1, g_2 and the production rules $F \rightarrow W^*$, $W \rightarrow WGW$, $W \rightarrow w_1$, $G \rightarrow g_1$, $G \rightarrow g_2$. For clearness, we assume a façade with only one floor. In each step, the non-terminal selected for the next substitution is marked in bold red.

As long as the façade 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 façade 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 line 8 in Figure 5), the left part of the façade 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 façade string contains only terminals, the production is completed and the string can be transferred into a 3D representation.



	Sequence of tiles	Applied rules
1.	F	F → W
2.	W	W → W G W
3.	W G W	G → g_1
4.	W g_1 W	W → W G W
5.	W G W g_1 W	W → w_1
6.	w_1 G W g_1 W	G → g_2
7.	w_1 g_2 W g_1 W	W → w_1
8.	w_1 g_2 w_1 g_1 W	W → W G W
9.	w_1 g_2 w_1 g_1 W G W	W → w_1
10.	w_1 g_2 w_1 g_1 w_1 G W	G → g_1
11.	w_1 g_2 w_1 g_1 w_1 g_1 W	W → w_1
12.	w_1 g_2 w_1 g_1 w_1 g_1 w_1	

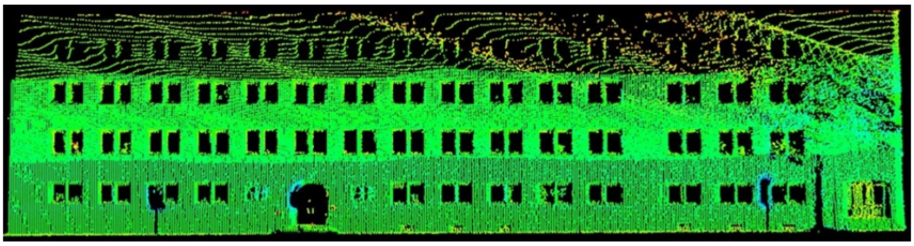
(a) (b)

Figure 5: Non-terminal selection.

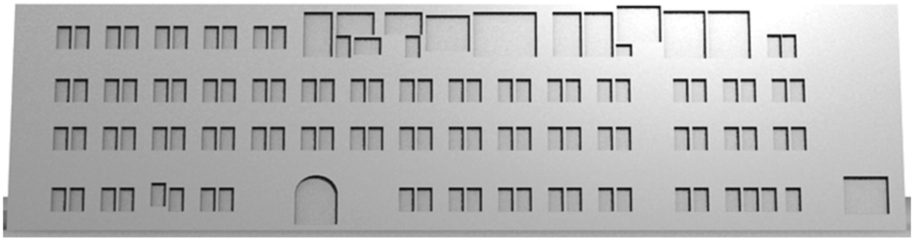
Within the production process, the grammar is applied to generate hypotheses about possible positions of each geometry tile and thereby synthesize façade geometry for given coarse building models. This process can for example be used to generate façade structure at areas, where sensor data is only available at limited quality. Furthermore, the production process can also be used to synthesize totally unobserved building objects.

3.4. Results

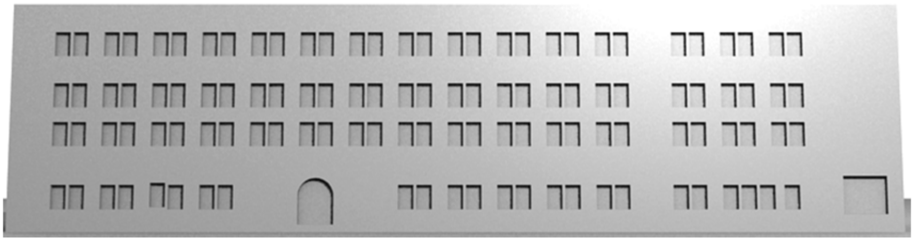
Facades and building parts which are covered by noisy or incomplete sensor data are usually subject to inaccurate and false reconstructions which are due to problems of the data-driven reconstruction process. For such regions possible façade geometry can be proposed in order to improve and complete façade structures. Figure 6(a) depicts a scenario where a StreetMapper point cloud has been acquired for an exemplary façade during one epoch (Haala et al., 2008). The point sampling distance varies strongly due to occlusions and oblique scanning views to the upper part of the building. For this reason, the facade contains areas where no or only little sensor data is available. In those regions, an accurate extraction of windows and doors cannot be guaranteed anymore (see the false reconstructions in the data-driven result in Figure 6(b)). Nevertheless, a grammar-based façade completion allows for meaningful reconstructions even in those areas. For this purpose, the façade grammar is derived 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 evaluating the sampling distances of the points lying on the façade surface. In the given example, the extracted convex dense area comprises the lower three floors. Since the inference process is restricted to this dense area, a façade grammar of good quality can be provided, which is then used to synthesize the remaining façade regions during the production step (Figure 6(c)).



(a) LiDAR point cloud



(b) 3D façade model after data-driven reconstruction



(c) 3D façade model after grammar-based completion

Figure 6: Improvement potential through grammar-based modeling for an exemplary office building in Stuttgart.

In the example of Figure 6 the grammar is applied in order to obtain realistic structures in areas where sensor data is only available at limited quality. But aiming at complete LOD3 building models even in situations where only one single façade has been observed – which is often the case for mobile mapping scenarios in street scenes –, façade structures are also required for building walls for which no measurement is available at all. However, due to the concept of automatically inferring individual façade grammars, it can be ensured that the grammar-based structures proposed for the unobserved building facades will match the architectural style of the respective building. Figure 7 illustrates the reconstruction results for the facades of two historic buildings of Stuttgart called Linden-Museum and Alte Kanzlei. In both data sets only the front façade, marked by a rectangle in Figure 7, was covered by dense LiDAR data. Basic façade geometries like windows and doors have been extracted from these areas and used as knowledge base to automatically derive an individual façade grammar for each building. Afterwards, based on the respective grammar, structure hypotheses have been generated for all side and back facades of the two buildings. Affected by high regularity in case of the Linden-Museum but structural variety in case of the Alte Kanzlei, the buildings feature different architectural styles. Thus, these two data sets are appropriate to estimate the achievable quality of procedurally created façade models in consideration of the architectural complexity of the building. The reconstructed facades of the Linden-Museum reveal the high quality which is feasible for homogeneous buildings. The shape, number and configurations of the hypothesized windows match the reality up to nearly 100 %. For the example Alte Kanzlei, differences between the modeled and the real facades are more obvious. This is due to the fact that the observed front façade only covers a subset of all the structures that occur in the whole building. However, even when the modeled windows do not fit the real distribution exactly, the overall style of the building is still maintained. Architectural inconsistencies can be avoided which is a special benefit of the concept of individual façade grammars. More detailed evaluations, also confirmed by applying quantitative quality measures, can be found in Becker (2011).

(a) Linden-Museum

(b) Alte Kanzlei

Figure 7: 3D reconstruction results for the facades of the buildings Linden-Museum and Alte Kanzlei.

4. EXTENSION AND ABSTRACTION OF GRAMMAR-BASED MODELING

A city is more than the sum of its buildings. Typical cityscapes principally arise from the interaction of adjacent buildings affecting each other's visual appearance. Beyond that, vegetation and street furniture considerably contribute to the architectural design of the city. For instance, trees, benches or bus shelters not only serve aesthetic purposes but also act as spatial and functional elements and, thus, are important for a number of applications. So far, research within the area of city modeling has brought forth numerous individual grammar-based concepts each focusing on a certain building aspect to be reconstructed such as roof types, façade structures or stairs (Milde et al. 2008; Brenner and Ripperda, 2006; Schmittwilken et al., 2006). However, there is still a lack of a superordinate rule-based description for urban areas which reflect both detailed views on specific building parts, and high-order relationships between buildings and their environment. Such an integrated representation of urban objects, though, is essential for modeling and deriving superior semantic knowledge. For this reason, a first modeling concept for urban environments is proposed in section 4.1. Its potential usage is discussed in section 4.2.

4.1. Hierarchical graph-based modeling structure for urban environments

The proposed façade grammar – especially the concepts for the automatic mapping of geometric primitives and structures into terminals and formal rules – can serve as basis for further developments. The façade grammar is designed to describe and generate configurations of windows and doors and, thus, deals with the very specific application of modeling façade structures. However, striving for an integrated grammar-based description for arbitrary objects of urban environments requires the generalization of the façade grammar's main concepts in order to be flexible towards the objects and relationships to be represented. In this regard, especially the topological dimension of the objects has to be considered. Concerning the interpretation of façade structures, the strategy of exploring the floors separately makes it possible to understand window configurations as linear sequences of façade elements, and to express them as a word of the grammar. However, such a one-dimensional topological dimension does not hold for most of the urban objects such as ground plans, whole buildings, vegetation or street furniture. For instance, parcels of land are usually arranged like pieces in a mosaic instead of following a linear structure. As a consequence, neighborhood relationships become much more complex not only for the parcels of land but also for the buildings which are built on those parcels. A linear description by means of character strings is not feasible any more. A solution to this problem might be embedding the grammar components into a hierarchically designed, graph-oriented structure. Using graphs,

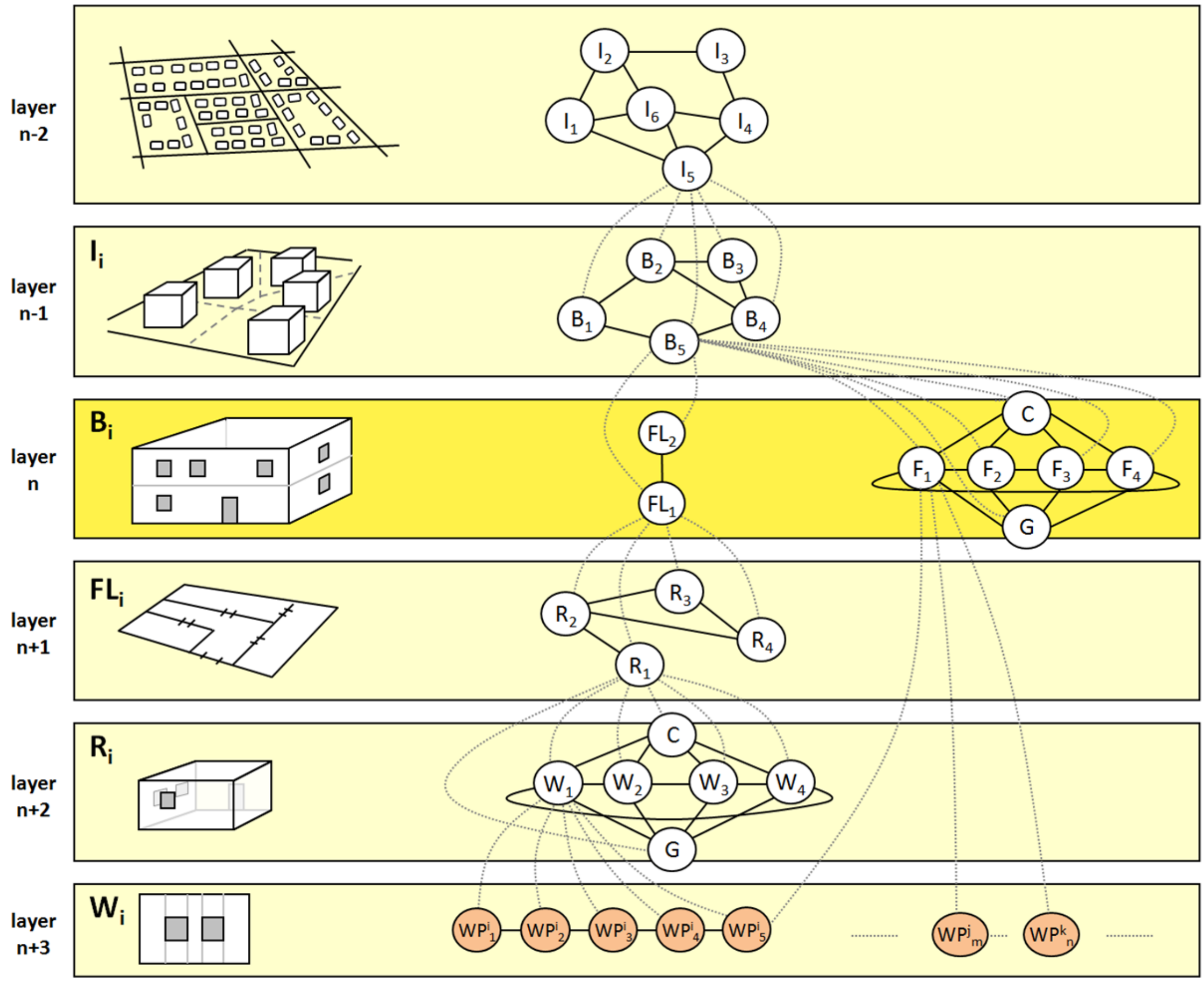


Fig. 8: Graph-oriented hierarchical modeling structure for urban environments.

arbitrary topological relationships can be modeled. The hierarchical structure supports the representation of a 3D city model in different levels of detail. Each node of the graph may contain a further graph describing knowledge about the parts the object is composed of. Thus, it should be possible to assign the components of a 3D city model to stepwise arranged layers in which the geometrical detail is increasing from layer to layer. In formal terms, such a modeling concept can be implemented by a so-called graph grammar where the basic elements are nodes and edges instead of characters or shapes, and where the production rules define the replacement with subgraphs.

Fig. 8 shows a first idea of how the components of a graph grammar could be incorporated in a layered model structure which enables an integrated grammar-based modeling of buildings in various levels of details and different contextual environments. From the perspective of the façade grammar a building is described by its outer shell as a set of planar surfaces. An exemplary building B_i and its corresponding graph, which contains nodes for the roof, façade and ground surfaces, are illustrated in the right part of layer *n*. However, a building can also be interpreted as a solid which defines a finite space in 3D. Based on the coarse segmentation along the floor planes, the building appears as a configuration of vertically arranged 3D blocks, each representing the scope of one single floor. The corresponding graph consisting of the floor-nodes FL_1 and FL_2 are illustrated in the middle of layer *n*. The interpretation of the building as a composition of volumes allows for the transition to a detailed description of the interior. Each floor-node of layer *n* can be expanded to a subgraph which represents the configuration of rooms belonging to this floor (see layer *n+1*). A further layer can be established when each room-node R_i is not only regarded as volumetric entity but also as surface describing the boundary of the room. Such a surface usually consists of planes for the ceiling, the ground and the vertical walls (see layer *n+2*). A next level of detail can be reached by segmenting each vertical wall in a linear sequence of wall parts which describe either a homogeneous wall area or contain an opening such as a window or door (see layer *n+3*). The resulting wall parts correspond to the definition of tiles, i.e. the basic elements of our façade grammar. Thus, it is possible to apply the façade grammar not only to whole building facades but also to the boundary walls of a room regardless whether being an exterior wall, which is part of the façade, or an interior wall. In this way, the loop to layer *n* is closed showing how the components of the façade grammar – in Fig. 8 highlighted as filled nodes – can be integrated in a more general view of building structures.

Generally, the substitution of a single node by a subgraph denotes a change in perspective switching from a rather coarse object description to a more detailed one which provides an insight into structural properties of the object. In this way, object knowledge is added inducing a new layer with a higher level of detail. However, the hierarchical modeling structure can also be traversed in the opposite direction in order to model an object in its superior context. This can be implemented by expressing a whole subgraph of layer *i* as a single node in layer *i-1*. In this new layer, the node is in turn part of a subgraph which reveals the superior context of the object. For example, while layer *n* focuses on a single building unaware of its environment, layer *n-1* considers the neighborhood relationships between those buildings that are part of the same building block. Since a block of buildings is usually bounded by streets forming a kind of an "island", such a group of buildings is also known as island site. A further degree of abstraction can be reached by modeling the topological relationships of island sites that result from the segmentation by the street network. Even if this is the highest layer displayed in Fig. 8, the opportunities for grammar-based modeling do not end here. Street networks can be efficiently described and generated by specially designed L-systems (Müller, 2001). Furthermore, since the model hierarchy can be extended by introducing more layers in both directions, the modeling of new abstraction or detail levels is nearly unlimited.

4.2. Potential usage

When a 3D city model is organized through a hierarchical graph-oriented structure as described above, urban objects are no longer modeled as single, isolated geometries. Instead, they are embedded in a network of geometrical and topological relationships which is the basis for deriving and modeling superordinate dependencies between them. The benefits resulting thereof are manifold. Relations modeled for objects and object components facilitate the analysis and the preservation of geometrical and topological consistency; constraints can be formulated in terms of rules in order to ensure consistent 3D city models. As an example, windows are not only structure generating elements of a façade; they are also essential parts of a room configuration. Thus, they serve as connectors between two different points of view, i.e. façade modeling and indoor modeling. Within each approach, a window has to meet certain constraints which, however, must not be regarded separately from each other when consistency problems are to be avoided. For instance, while the façade's context is necessary to enforce geometrical dependencies between those windows that belong to the same floor, knowledge about the room configuration of the building's interior is needed to prevent windows from intersecting inner walls. The modeling concept proposed provides an integrated view of the conditions into which the windows are involved and, consequently, allows for efficient solutions for the generation of consistent 3D models.

Beyond that, the enhanced object knowledge inherent in the hierarchical graph-based structure can be used to make the extraction and reconstruction of new objects more robust. For example, relationships between a building's size and façade design as well as correlations between the façade style and the proximity of buildings could be derived from a representative part of the city and translated into formal rules. By including these rules in the grammar-based façade modeling, the generation and completion of realistic façade structures could be further improved.

In each layer of the graph-based modeling concept dominant and repetitive patterns can be found which represent characteristic configurations of building structures. However, such relational patterns are not confined to buildings but, instead, can comprise relationships between any kind of urban objects. Consequently, one could imagine to also model typical arrangements of landscape elements such as street lights, fences or plants. For example, similar to the façade modeling approach, distinctive configurations of urban trees could be extracted from observation data and transferred into formal rules. These rules then can be used to support the extraction and reconstruction of trees from noisy or incomplete sensor data, or to procedurally generate realistic representations of urban vegetation for areas, where no sensor data is available.

To show the feasibility, the graph-oriented modeling concept has been partially implemented for a historic part of Stuttgart called "Bohnenviertel" for which façade structures had already been produced using the grammar-based approach described in section 3. As a first example, three characteristic tree configurations have been identified and translated into graph rewriting rules. These rules either generate an unstructured cluster of trees, a regular grid-like pattern, or a linear arrangement of trees running parallel to a dominant building façade. Depending on the dedication of an urban district which, e.g., can be dominated by historic, residential or industrial buildings, trees in general and tree configurations in particular occur with different probabilities in different parts of the city. In order to obtain realistic 3D representations for the selected Bohnenviertel, appropriate probability values have been manually derived from an aerial photo of this region. Based on the neighborhood relationships modeled for the buildings of an island site, contiguous, non built-up areas can be determined. Subsequently, the tree rules are applied to each of these areas according to their probability. Subject to the configuration that a currently selected tree rule describes, such a region is decomposed into discrete cells each of them big enough to offer space for a tree of average size. The result can be seen in Fig. 9. The positions of the trees have been automatically determined applying the tree rules as delineated above. Afterwards, instances of an exemplary 3D tree model have been inserted at the calculated positions.

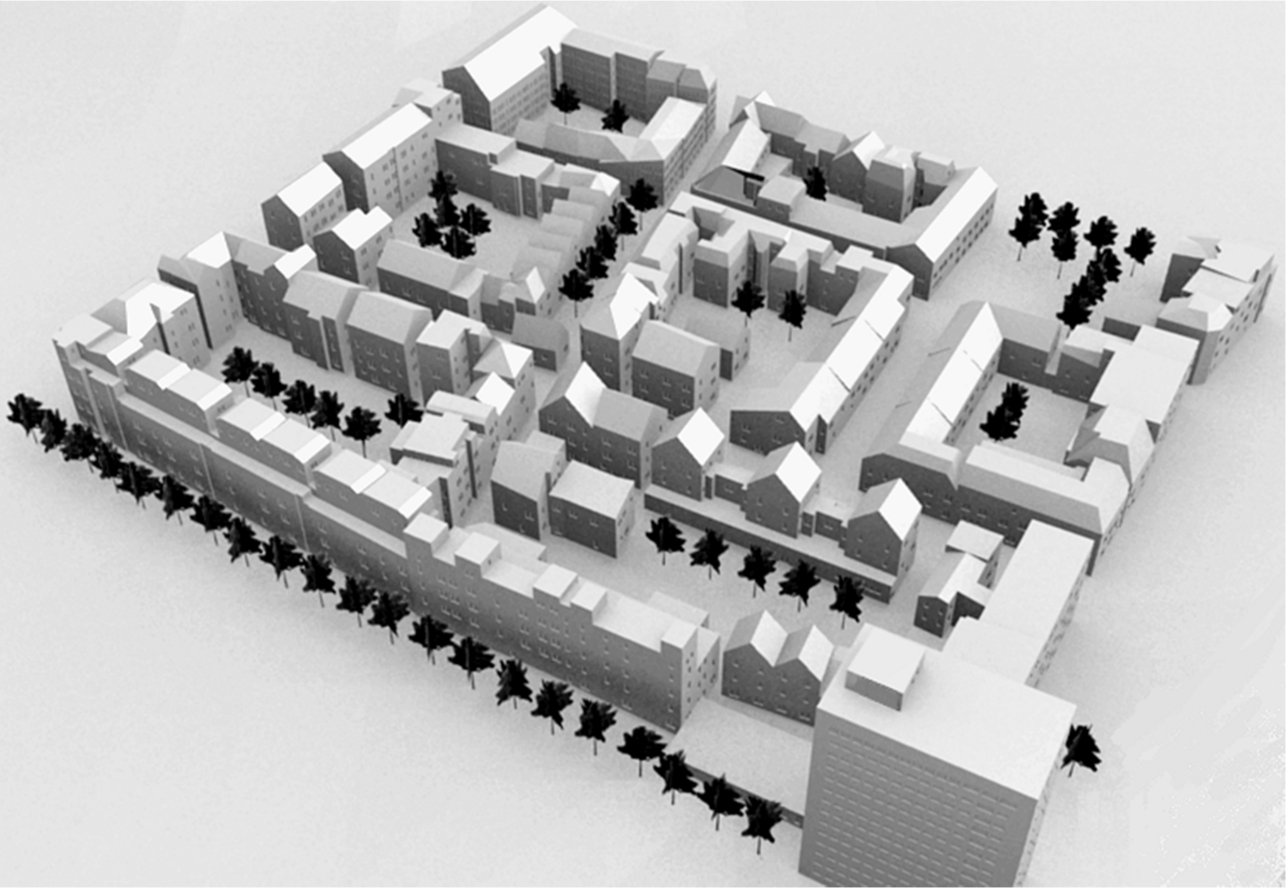


Fig. 9: Application of tree rules and façade grammars to a historic district of Stuttgart called “Bohnenviertel”.

5. CONCLUSIONS

We proposed an automatic approach for the reconstruction of complete 3D façade models. Grammar rules are extracted from observed 3D façade geometries and applied for the generation of synthetic façade structures for unobserved building parts. The inference of individual façade grammars ensures realistic reconstruction results: If a new façade to be modeled is covered by inaccurate, noisy or incomplete sensor data, grammar rules can be used for the improvement and completion of façade structures. In the case of facades that have not been observed at all, the grammar allows for the prediction of structural information in the style of the respective building.

Moreover, the proposed concept for grammar-based modeling of building structures is not limited to facades but can serve as basis for numerous further fields of application. This contains the extension and abstraction of the façade scenario to a more general and hierarchically designed graph-oriented modeling structure for urban environments. Based on the huge number of relationships that can be represented through such a modeling concept, new semantic information can be derived taking us another step towards the vision of a fully interpreted 3D city model.

6. REFERENCES

Alexander, C.W. (1964): Notes on the Synthesis of Form. Harvard University Press.

Aliaga, D., Rosen, P., Bekins, D. (2007): Style grammars for interactive visualization of architecture. IEEE Transactions on Visualization and Computer Graphics, 13 (4), 786-797.

Arnheim, R. (1974): Art and Visual Perception: A Psychology of the Creative Eye. University of California Press.

Becker, S. (2009): Generation and application of rules for quality dependent façade reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing, 64, 640-653.

Becker, S. (2011): Automatische Ableitung und Anwendung von Regeln für die Rekonstruktion von Fassaden aus heterogenen Sensordaten. Dissertation. Deutsche Geodätische Kommission, Reihe C, Nr. 658, München, ISBN 978-3-7696-5070-9, ISSN 0065-5325.

Becker, S., Haala, N. (2007): Refinement of building façades by integrated processing of LIDAR and image data. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (Part 3/W49A), 7-12.

Brenner, C., Ripperda, N. (2006): Extraction of facades using rjMCMC and constraint equations. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (3), 155-160.

Dick, A., Torr, P., Cipolla, R. (2004): Modelling and interpretation of architecture from several images. International Journal of Computer Vision, 60 (2), 111-134.

Haala, N., Peter, M., Kremer, J., Hunter, G. (2008): Mobile LiDAR mapping for 3D point cloud collection in urban areas – a performance test. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (B5).

Kolbe, T. H., Gröger, G., Plümer, L. (2005): CityGML – interoperable access to 3D city models. In: Oosterom, Zlatanova, Fendel, Editors, Proceedings of the International Symposium on Geoinformation for Disaster Management. Springer-Verlag.

Kutterer, H. (2010): Airborne and Terrestrial Laser Scanning, Chapter: Mobile Mapping, 293-311. Whittles Publishing.

Milde, J., Zhang, Y., Brenner, C., Plümer, L., Sester, M. (2008): Building Reconstruction using a structural description based on a formal grammar. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37 (B3b).

Mitchell, W.J. (1990): The Logic of Architecture: Design, Computation, and Cognition. M.I.T. Press, Cambridge, Massachusetts.

Müller, P. (2001): Design und Implementation einer Preprocessing Pipeline zur Visualisierung prozedural erzeugter Stadtmodelle. Master thesis, ETH Zürich.

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

Müller, P., Zeng, G., Wonka, P., Van Gool, L. (2007): Image-based procedural modeling of façades. ACM Transactions on Graphics (TOG), 26 (3), 1-9. Article 85.

Prusinkiewicz, P., Lindenmayer, A. (1990): The Algorithmic Beauty of Plants. Springer, New York, NY.

Reznik, S., Mayer, H. (2008): Implicit shape models, self-diagnosis, and model selection for 3D façade interpretation. Photogrammetrie, Fernerkundung, Geoinformation, 3/2008, 187-196.

Ripperda, N. (2008): Determination of façade attributes for façade reconstruction. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37 (B3a), 285-290.

Rothermel, M., Haala, N. (2011): Potential of Dense Matching for the Generation of High Quality Digital Elevation models. In: Proceedings of ISPRS Hannover Workshop High-Resoultion Earth Imaging for Geospatial Information, 6 pages.

Rutzinger, M., Oude Elberink, S., Pu, S., Vosselman, G. (2009): Automatic extraction of vertical walls from mobile and airborne laser scanning data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 38 (3/W8), 7-11.

Schmittwilken, J., Kolbe, T. H., Plümer, L. (2006): Der Gebäudekragen – Eine detaillierte Betrachtung des Übergangs von Gebäude und Gelände. In: Geoinformatik und Erdbeobachtung. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation, 127-135.

Schuhmacher, S., Böhm, J. (2005): Georeferencing of terrestrial laserscanning data for applications in architectural modeling. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (5/W17).

Stiny, G., Gips, J. (1972): Shape grammars and the generative specification of painting and sculpture. In: Freiman, C.V. (Ed.), Information Processing 71: Proceedings of IFIP Congress. North-Holland, Amsterdam, 1460-1465.

Van Gool, L., Zeng, G., Van den Borre, F., Müller, P. (2007): Towards mass-produced building models. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W49A), 209-220.

Wenzel, S., Förstner, W. (2008): Semi-supervised incremental learning of hierarchical appearance models. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 37 (Part B3a), 309-404.

Wertheimer, M. (1923): Untersuchungen zur Lehre der Gestalt II. Psychologische Forschung, 4, 301-350.

Wonka, P., Wimmer, M., Sillion, F., Ribarsky, W. (2003): Instant architecture. ACM Transaction on Graphics, 22 (3), 669-677.