Grammar Based Facade Reconstruction using rjMCMC

NORA RIPPERDA, Hannover

Keywords: facade modelling, grammar, Markov Chain Monte Carlo

Summary: These days 3d models are used in a huge variety of applications and the demands in quality and quantity are steadily growing. At the same time, the extraction of man-made objects from measurement data is quite traditional. Often, the processes are still point based, with the exception of a few systems which allow to automatically fit simple primitives to measurement data. The need to be able to automatically transform object representations, for example, in order to generalize their geometry, enforces a structurally rich object description. Likewise, the trend towards more and more detailed representations requires to exploit structurally repetitive and symmetric patterns present in man-made objects, in order to make extraction cost-effective. In this paper, we address the extraction of building facades in terms of a structural description. Our reconstruction is based on a formal grammar to derive a structural facade description in the form of a derivation tree and uses a stochastic process based on reversible jump Markov Chain Monte Carlo (rjMCMC) to guide the application of derivation steps during the construction of the tree.

Zusammenfassung: *Grammatik-basierte Fassadenrekonstruktion mittels rjMCMC.* 3d-Modelle werden heutzutage in vielen Anwendungen gebraucht und die Anforderungen an sie steigen ständig an. Gleichzeitig werden aber großteils noch die klassischen Extraktionsverfahren verwendet, die meist punktbasiert arbeiten. Für viele Anwendungen wird ein Modell in unterschiedlich detailreicher Darstellungen benötigt. Die hierzu hilfreiche automatische Transformation des Modells in verschiedene Darstellungen kann durch eine strukturelle Beschreibung des Objekts ermöglicht werden. Zusätzlich kann die Beschreibung von sich strukturell wiederholenden oder symmetrischen Mustern eine effektive Modellierung begünstigen. In diesem Artikel wird eine Methode zur automatischen Rekonstruktion von Fassaden aus Bild- und Entfernungsdaten vorgestellt. Das strukturelle Modell ist durch eine Fassaden-Grammatik gegeben und der Modellierungsprozess wird durch ein rjMCMC-Verfahren gesteuert.

1 Introduction

The extraction of man-made objects from sensor data has a long history in research (BALTSAVIAS 2004). Especially for the modelling of 3D buildings, numerous approaches have been reported, based on monoscopic, stereoscopic, multi-image, and laser scan techniques. While most of the effort has gone into sensor-specific extraction procedures, very little work has been done on the structural description of objects. Modelling structure though is very important for downstream usability of the data, especially for the automatic derivation of coarser levels of detail from detailed models.

Representing structure is not only important for the later usability of the derived data, but also as a means to support the extraction process itself. A fixed set of structural patterns allows to span a certain subspace of all possible object patterns, thus forms the model required to interpret the scene.

Grammars have been extensively used to model structures. For modelling plants, Lindenmayer systems were developed by PRUSINKIEWICZ & LINDENMAYER (1990). They have also been used for modelling streets and buildings (PARISH & MÜLLER 2001, MARVIE et al. 2005). But Lindenmayer systems are not necessarily appropriate for modelling buildings. Buildings differ in structure from plants and streets, in that they don't grow in free space and modelling is more a partition of space than a growth-like process.

For this reason, other types of grammars have been proposed for architectural objects. STINY & GIPS (1972) introduced shape grammars which operate on shapes directly. The rules replace patterns at a point marked by a special symbol. MITCHELL (1990) describes how grammars are used in architecture. The derivation is usually done manually which is why the grammars are not readily applicable for automatic modelling tools.

ALEGRE & DALLAERT (2004) use a stochastic context free attribute grammar to reconstruct facades from image data by applying horizontal and vertical cuts. WONKA et al. (2003) developed a method for automatic modelling which allows reconstructing different kinds of buildings using one rule set. The approach is composed of a split grammar, a large set of rules which divide the building into parts, and a control grammar which guides the propagation and distribution of attributes. During construction, a stochastic process selects among all applicable rules. VAN GOOL et al. (2007) discuss different facade reconstruction algorithms and show the use of repetitions in the structure for the reconstruction with shape grammars.

Our aim is to extract facade elements from image and range data automatically. The facade model is defined by a grammar which comprise the structure of facades. Each grammar rule subdivides a part of the facade in smaller parts according to the layout of the facade. The derivation process is guided by a reversible jump Markov Chain Monte Carlo (rjMCMC) process.

DICK et al. (2004) introduce a method which generates building models from measured data, i.e. several images. This approach is also based on the rjMCMC method. In a stochastic process, 3D models with semantic information are built. MAYER & REZNIK (2006) also use a MCMC method for the facade reconstruction from images.

The rjMCMC algorithm is used for other applications e.g. detection of road marks (TOURNAIRE et al. 2007) as well. In general rjMCMC is a top-down-approach, but TU (2005) integrated generative and discriminative methods and used a data driven MCMC (DDMCMC) for image parsing.

We also present a way to use information about the facade structure from the data. We derive distributions of facade attributes like the position of windows. These distributions are used for the rule proposal additionally to the general prior knowledge, which was used in our previous work on facade reconstruction (RIPPERDA & BRENNER 2006). The extra information from the data causes to evade the large number of wrong proposals which occur using only general prior knowledge on facades.

For the facade reconstruction we need a structural model that describes the facade. In the presented approach the model is given by a facade grammar. A derivation tree of a word of the grammar represents the model of a given facade.

A stochastic process, the rjMCMC process, guides the reconstruction process. Section 2 introduces the facade grammar and section 3 gives an idea of the rjMCMC process and shows how to adapt it to the grammar.

2 The Facade Grammar

A formal grammar G consists of an alphabet of terminal T and nonterminal N symbols, a start symbol S and a set of production rules P. We use a context-free grammar, this means that P contains rules of the form $N \rightarrow (T \cup N)^+$. All words that can be derived from S with rules from P build the language $L(G)$ of the grammar G.

For facade reconstruction we define a grammar G_F which language $L(G_F)$ contains possible facades (for details see RIPPERDA & BRENNER 2006). In the derivation process the model of the facade should be developed further in each step. Therefore each rule splits the part of the facade corresponding to the left side symbol in a variable number of facade parts corresponding to the right side symbols. So the derivation process is a partitioning process of the facade. The start symbol S is an empty facade. This is subdivided in further derivation steps.

A split can be caused by different reasons. The first is a difference in the facade structure. If a facade contains different structural parts it is split into part facades according to the structure and the parts are modelled individually. This change in structure often occurs in ground floor and upper floors.

The other reason for a split is similarity or repetition. If a facade is symmetric or contains repetitions the repeated pattern needs to be stored only once. Additional information like number of repetitions completes the model.

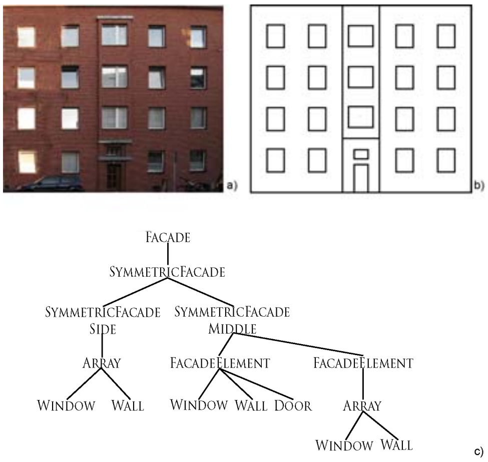

Fig. 1: Example facade (a) with a partition according to the facade grammar (b) and the corresponding derivation tree (c).

Fig. 1 illustrates an example of a facade reconstruction. Part a) shows the image of the facade and part b) a partitioning according to the facade grammar. The corresponding derivation tree c) and additional attributes build the reconstruction of the facade. The example contains splits of both kinds, based on similarities and based on differences. Similarities are arising in the symmetric part and in the arrays of windows. So for example the rules SymmetricFacadeSide→Array and FacadeElement→Array are based on the repetitions of the facade elements. The rule Facade→SymmetricFacadeSide SymmetricFacadeMiddle contains a bit of both. The SymmetricFacadeSide is the similarity part but the additional SymmetricFacadeMiddle is due to differences in the middle of the facade. Another rule based on differences is SymmetricFacadeMiddle→FacadeElement FacadeElement.

The structure of the grammar is shown is Fig. 2. There are three levels in the grammar. The first one contains the symbols which have no information about the structure of the facade. For example the start symbol Facade. The only information at this stage is the outline of the building. In the second level structural information is added. The symbols can express symmetries, repetitions and so on. The terminal symbols, which are the real facade elements like WINDOW or DOOR, belong to the third level.

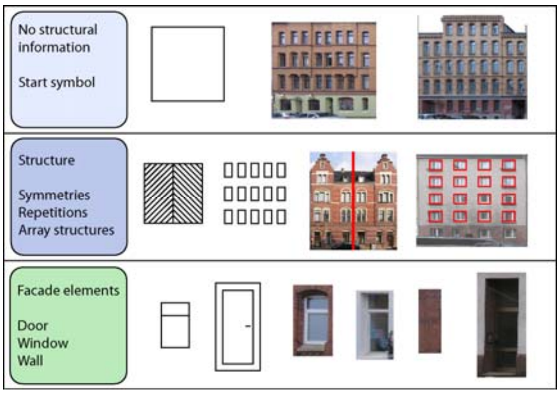

Fig. 2: Structure of the facade grammar.

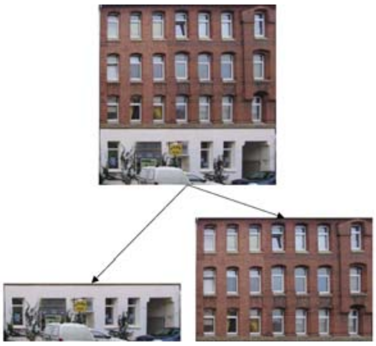

Fig. 3: Subdivision of a facade in an upper and a lower part.

The model is described by a parameter vector θ which contains the derivation tree and the attributes of the symbols. E.g. the parameter vector of the configuration in Fig. 3 is represented by the hierarchic structure θ = FACADE(0,0,*w*,*h*,(PARTFACADE(0,0,*w*,h_s),PARTFACADE(0, h_s,*w*,*h*-h_s))), where *w* and *h* are the width and height of the facade and h_s is the height of the split.

3 Facade Reconstruction using RjMCMC

We obtain the model of the facade using a stochastic process. We are searching for the model given by parameter vector θ with the highest probability $p(\theta|D_S D_I)$ under given scan (D_S) and image data (D_I) where the parameter vector θ encodes the current state of the derivation tree, including attributes.

So we search for an unknown probability distribution $p(\theta|D_S D_I)$. To sample from such a distribution MCMC methods are often used. A Markov Chain that simulates a random walk in the space of θ is constructed. The transition kernel assigns a probability to each change from one state to another. After a proposed change an acceptance probability decides whether the change is accepted or not. The acceptance probability is defined in a way that the system converges to the target distribution $p(\theta|D_S D_I)$. In our case the dimension of θ changes during the process. This is not possible in the basic MCMC method. Therefore we use rjMCMC which contains jumps (dimensions changes) of θ. The probability of a dimension change is added to the transition kernel.

For the rjMCMC process with target distribution $p(\theta|D_S D_I)$ we have to define a transition kernel $J(\theta_t|\theta_{t-1})$ and the acceptance probability α.

The transition kernel $J(\theta_t|\theta_{t-1})$ assigns a probability to each rule and is made up from the commonness of the result in a dataset of facade images and some functions of the processed facade, which are described below. With the transition kernel in each iteration a rule is proposed. This is accepted with the acceptance probability

$$\alpha = \min\left\{1, \frac{p(\theta_t \mid D_S D_I) \cdot J(\theta_{t-1} \mid \theta_t)}{p(\theta_{t-1} \mid D_S D_I) \cdot J(\theta_t \mid \theta_{t-1})}\right\} \qquad (1)$$

This depends on the unknown distribution $p(\theta_t|D_S D_I)$. Using Bayes' law, this is proportional to $p(D_S D_I|\theta_t) \cdot p(\theta_t)$, a product of likelihood and prior of the facade. In the following sections the jumping distribution and the acceptance probability are described in detail.

3.1 Jumping Distribution

The jumping distribution assigns a probability to each possible change in the facade structure. According to this probability a change is proposed. The method contains changes of different kinds. The first one is the application of a grammar rule. This splits the facade in different parts based on differences or repetitions in the facade. For this kind of change additional parameters must be proposed as well. These are for example the cut position or the number of parts in the facade. The distribution of these parameters is important for the acceptance of the change. To ensure reversibility, each rule can be applied from left to right and vice versa. This is a difference to the way split grammars are used, but is a requirement for the rjMCMC approach.

The second kind of change is a rearrangement in the structure. The symbols stay the same but the parameters are modified. The position of a parting line can change or the size or number of windows alters.

To build the transition kernel two kinds of distributions have to be defined. The first one is the probability to choose a rule and the second one defines the parameter like the position of a split line or the number of windows. Presently the probability for rules is assigned manually depending on an assumed likelihood of the result. For example, a change FACADE→IDENTICALFACADEARRAY is more likely than FACADE→FACADEARRAY because facades build regular structures of similar elements. Some hints for the assumptions are taken from a database of facade images from Hannover.

We need information about the distribution of colour or depth on the facade to control the split operation and to determine the distribution of the windows. Both depend on regularities and differences. For window grids we use autocorrelation and for splits a function based on a norm.

Fig. 4: Smoothed image maintains only large changes in facade structure (left). Clustered facade calculated by colour value and depth (right).

For splitting the facade into parts a change in colour or depth on a large part of the facade or irregularities in structure are needed. The changes of colour and depth occur in different scales. We search for changes which influence a large part of the facade, for example a horizontal colour change is often associated with a change in the window structure, or alternatively changes caused by windows. Smaller artefacts in the facade may disturb the result. So we have different ways to score splits but in each we have to mask the small changes which falsify the result. One way to suppress such unwanted changes is to use a smoothed image (see Fig. 4, left). Another possibility is to cluster the facade depending on the colour value and in another step depending on the depth value. The colour and depth image clustered with k-means with manually chosen k are shown in Fig. 4 (right). From these images we can derive distributions for the additional parameters.

To get the distribution of a split line we move the proposed split line from bottom to top of the facade (see Fig. 5) and look at the regions above R_u and below R_l the line. Differences between the regions score for the split. To evaluate the split line we compute the norm of the difference of both regions $\|R_u - R_l\|_2 = \sqrt{\sum_{x,y} (R_u(x,y) - R_l(x,y))^2}$, where $R_u(x,y)$ is the rgb value at position (x,y).

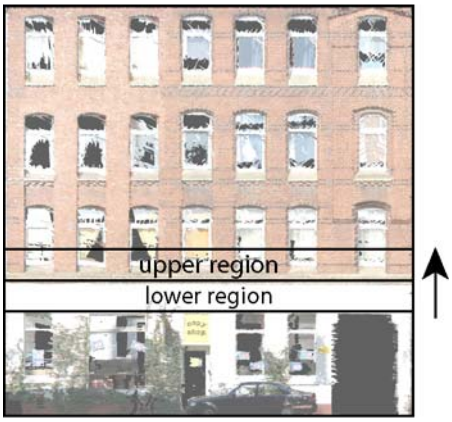

Fig. 5: Two regions above and below the tested split line were moved over the facade.

The results are shown in Fig. 6. For a better visual understanding the original facade image is overlaid to the resulting graph. With the cluster image (blue line) we achieve better results than with the scaled image (red line) because on the scale image lines at top edges of windows are scored better than colour changes throughout the entire facade. This is because the colour differences between black window area and grey or red wall area is greater than the difference between grey and red wall area. This happens for many facades with different colours in ground floor and upper floors. Therefore we use the norm of cluster images to get the split line distribution.

To reduce the number of false proposals we integrate a general assumption to the distribution. The split occurs between the ground floor and the upper floors and most of the facades in the test area have four or five floors. So we introduce the assumption that the position of the split line is normally distributed with a mean at one quarter of the height of the facade. This masks the high scores in the upper part of the facade out (see Fig. 6 green line).

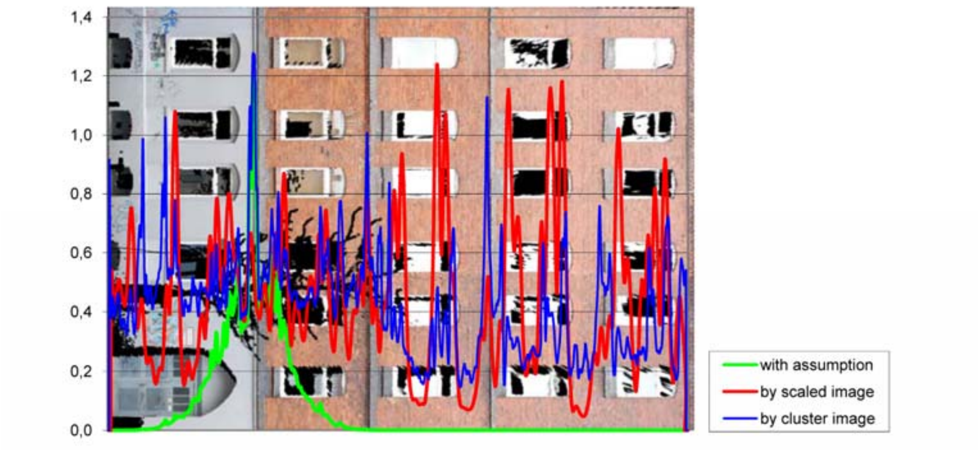

Fig. 6: Facade image overlaid with the probability of splits evaluated by a scaled image and cluster image. Additionally the probability derived from the clustered image is combined with a general assumption to reduce high scores at false positions.

To predict the distribution of windows we use autocorrelation. We correlate the overlapping parts of the facade image and a copy of it which we shift horizontally resp. vertically. Fig. 7 shows the resulting graphs. In the case of a regular window grid the correlation values show peaks in a regular distance. The number of peaks is the number of window rows resp. columns including one peak for the identical image. If the margins of the image are alike one additional peak for the case when the overlap tends towards zero arises. In the example the horizontal correlation shows nine peaks because of the eight window columns plus one for identical and border case. This pattern is not so clear for the vertical correlation because of the different ground floor.

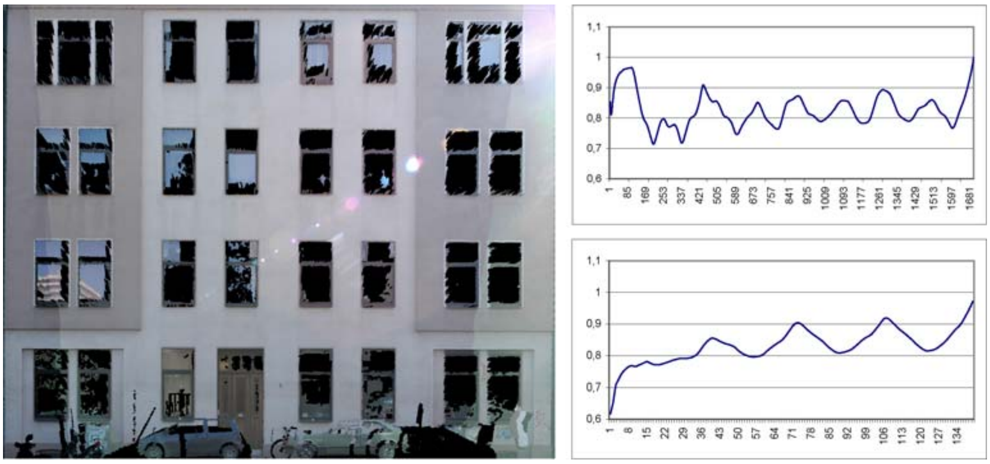

Fig. 7: Autocorrelation coefficient of a facade in horizontal and vertical direction.

3.2 Scoring Functions

The evaluation if a change is accepted is based on the scan and image data as well as the general knowledge of facades. The scoring functions affect the acceptance probability (1) in the term $p(D_S D_I|\theta_t) \cdot p(\theta_t)$ respectively $p(D_S D_I|\theta_{t-1}) \cdot p(\theta_{t-1})$.

The general plausibility of the model of the facade is given by the second term $p(\theta_t)$, the prior. It depends on the alignment, the extent and the position of the facade elements. Here we use the same scoring functions as given in (DICK et al. 2004) which where described in (RIPPERDA &BRENNER 2006) as well.

The second group evaluates how good the model fits the data by comparing it to range and image data. This corresponds to the likelihood term $p(D_S D_I|\theta_t)$. In any case, the evaluation functions return a score which builds an acceptance probability for the change.

To determine $p(D_S D_I|\theta_t)$ we have different possibilities which use scan and image data. We develop measures for depth and colour and use correlation, entropy and variance as well.

First we look at a method to score a single window. For colour images we use the fact that windows have a different colour from facades. Typically they appear darker than the facade but in some cases also brighter because of reflections. If we use depth images we have the information that the windows typically lie behind the facade. This leads us to a method working on the clustered images. Therefore we consider one region for the window and one for the boundary (see Fig. 9 left) and look at the clusters inside these regions. Let N_{max} be the number of pixels of the largest cluster inside the proposed window region, N_0 the number of unclassified pixels, A_{win} the area of the window, A_{bound} the area of the boundary and N_{bound} the number of pixels of the boundary which belong to the largest cluster inside the window. α_C gives a measure for the window.

$$\alpha_C = \frac{1 + \frac{N_{\max} + N_0}{A_{win}} - \frac{N_{bound}}{A_{bound}}}{2}$$

To test this method separately we cut out a single window from a facade. For this small data set we compute the score α_C for each possible position of the window (see Fig. 8 a)). Width and height are usually estimated in the process as well, but we show only the position here because of the 2d visualisation. The position is the lower left corner of the window and the plot of the score shows the lower left part of the test area where the possible positions are located. Then we run the MCMC process for a single window (see Fig. 8 b)) and compare the results with the distribution given by the score function. In both plots red colour means high values and blue colour low values. To give an idea of the changes between two states Fig. 8 c) shows a part of the random walk. Fig. 8 d) shows the most frequent window position marked in the colour cluster image where different colours indicate different clusters.

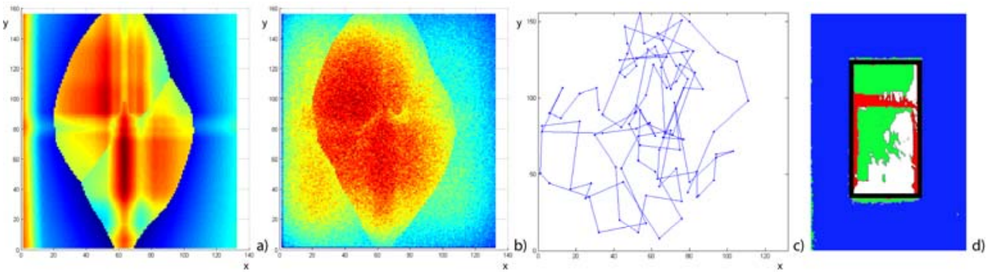

Fig. 8: Reconstruction of a single window from a colour cluster image. a: Score function for all possible positions, b: Frequency of positions sampled with MCMC, c: Extract of the random walk, d: Most frequent window position drawn in the colour cluster image.

To score the distribution of windows we use a homogeneity measure. Here we give the priority to the similarity within a region instead the difference of two regions. We define one region for all windows and one for the surroundings (see. Fig. 9, right). If both regions are homogeneous the score for the window distribution is high. As a measure for homogeneity we use entropy or variance. Here we discuss entropy in detail.

Entropy is

$$I = \sum_{i=1}^{n} \frac{|C_i|}{A} \log_2 \frac{A}{|C_i|},$$

where n is the number of clusters, A the total area and $|C_i|$ the number of points in the i-th cluster. We calculate the entropy for the proposed window area and the surrounding separately and use the sum for the score function. Fig. 10 a) shows the score function for different grid positions. We fix the number of grid points and the distance between them for a better visualisation.

Because entropy gives high values for disorder and low values for homogeneous regions we invert the function. Before that we normalize it by $log_2\, n$ which is the highest possible value. So the probability is given by $\alpha_I = 1\text{-}I/log_2$ (see Fig. 10 b)).

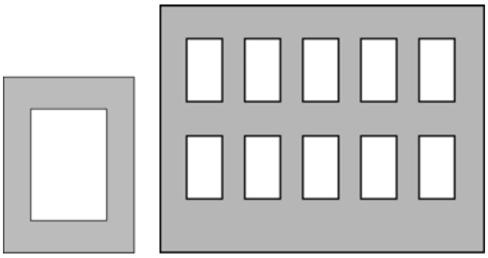

Fig. 9: Mask for a single window (left) and an array of windows (right). The window area is white and the boundary area grey.

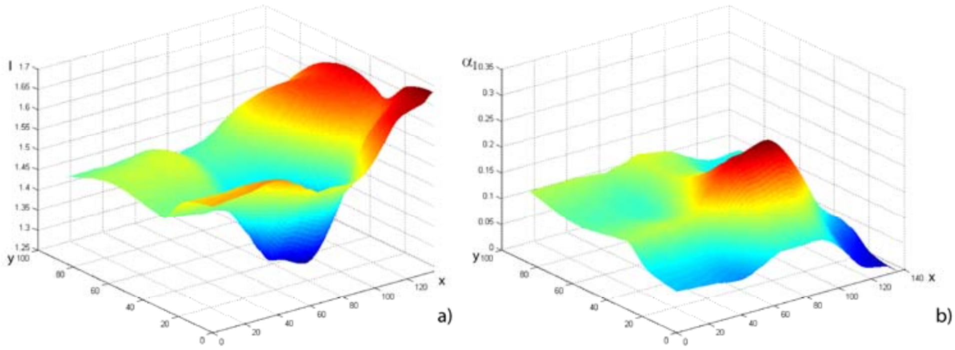

Fig. 10: Sum of entropy of window and boundary area for different grid positions (a) and the probability function derived from the entropy (b).

4 Results

We've tested the method on facades of dwelling houses. The input data are the point cloud and an orthophoto which is generated with the RiScanPro software. The other required data are computed in a first step.

Fig. 11 shows some results of the reconstruction. In the facade on the left the model consists of a regular grid of window pairs. The size of the windows is modelled properly but not all windows are modelled at the right position. This is because the windows are not exactly arranged in a regular grid. In the second facade the vertical split line (green line) between ground floor and upper floors is modelled at the correct position. For a similar reason as in the first facade not all windows are at the right position. But after a vertical split the windows in the regular region are modelled correctly (Fig. 11, right).

Fig. 11: Reconstruction of facades. A regular grid of double windows is modelled for the left facade. In the middle the horizontal split line is reconstructed correctly (green line) but not all

windows are modelled at the right position. After splitting the regular area of the upper part the window grid is modelled at the right position.

5 Conclusion and Outlook

In this paper, we have presented a method for automatic facade reconstruction from scan and image data. It combines the generation of artificial facade structures using grammars, and the reconstruction of facades using rjMCMC. Compared to existing grammar-based approaches, we gain the ability to reconstruct facades based on measurement data. Compared to existing rjMCMC approaches, by using a grammar, we obtain a hierarchical facade description and the ability to evaluate superstructures such as regularity and symmetry at an early stage, i.e., before terminal symbols such as WINDOW are instantiated.

For further work we want to enlarge our knowledge of facades to improve the proposal of facade elements. Therefore we analyse a set of facade images to get information about average window size, distance or style. Furthermore we plan to extend the facade grammar in order to be able to model a wider class of facade elements like balconies or ornaments.

Acknowledgements

This work was done within in the scope of the junior research group "Automatic methods for the fusion, reduction and consistent combination of complex, heterogeneous geoinformation", funded by the VolkswagenStiftung, Germany.

References

ALEGRE, F. & DALLAERT, F., 2004: A probabilistic approach to the semantic interpretation of building facades. In: International Workshop on Vision Techniques Applied to the Rehabilitation of City Centers.

BALTSAVIAS, E. P., 2004: Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. ISPRS Journal of Photogrammetry and Remote Sensing 58, pp. 129–151.

DICK, A., TORR, P., CIPOLLA, R. & RIBARSKY, W., 2004. Modelling and interpretation of architecture from several images. International Journal of Computer Vision 60(2), pp. 111–134.

MARVIE, J.-E., PERRET, J. & BOUATOUCH, K., 2005. The fl-system: a functional l-system for procedural geometric modeling. The Visual Computer 21(5), pp. 329 – 339.

MAYER, H. & REZNIK, S., 2006. Building facade interpretation from uncalibrated wide-baseline image sequences. ISPRS Journal of Photogrammetry & Remote Sensing 61, pp. 371–380.

MITCHELL, W. J., 1990. The Logic of Architecture: Design, Computation, and Cognition. Cambridge, Mass.: The MIT Press.

PARISH, Y. & MÜLLER, P., 2001. Procedural modeling of cities. In: E. Fiume (ed.), ACM SIGGRAPH, ACM Press.

PRUSINKIEWICZ, P. & LINDENMAYER, A., 1990. The algorithmic beauty of plants. New York, NY: Springer.

RIPPERDA, N. & BRENNER, C., 2006. Reconstruction of facade structures using a formal grammar and rjmcmc. In: K. Franke, K.-R. Müller, B. Nickolay, R. Schäfer (Eds.), Pattern Recognition, Proceedings of the 28th DAGM Symposium, pp. 750-759.

STINY, G. & GIPS, J., 1972. Shape Grammars and the Generative Specification of Painting and Sculpture. Auerbach, Philadelphia, pp. 125–135.

TOURNAIRE, O., PAPARODITIS, N. & LAFARGE, F., 2007. Rectangular Road Marking Detection with Marked Point Processes. In: Stilla U. et al. (Eds.) PIA07. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (3/W49A), pp. 149-154.

TU, Z., CHEN, X., YUILLE, A. & ZHU, S., 2005. Image Parsing: Unifying segmentation, detection and recognition. International Journal of Computer Vision 63(2), pp. 113-140.

VAN GOOL, L., ZENG, G., VAN DEN BORRE, F. & MÜLLER, P., 2007. Towards Mass-Produced Building Models. In: Stilla U. et al. (Eds.) PIA07. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (3/W49A), pp. 209-220.

WONKA, P., WIMMER, M., SILLION, F. & RIBARSKY, W., 2003. Instant architecture. ACM Transaction on Graphics 22(3), pp. 669–677.

Anschrift des Autors:

Dipl.-Math. Nora Ripperda

Leibniz Universität Hannover, Institut für Kartographie und Geoinformatik,
Appelstraße 9A, D-30167 Hannover
Tel.: +49-511-762-19436, Fax: +49-511-762-2780
e-mail: Nora.Ripperda@ikg.uni-hannover.de