Model Selection for Composite Objects with Attribute Grammars

Jörg Schmittwilken and Lutz Plümer
Institute of Geodesy and Geoinformation, University of Bonn
{schmittwilken, pluemer}@igg.uni-bonn.de

ABSTRACT

In this paper we present a concept for the reconstruction of composite objects from LIDAR 3D point clouds with focus on symmetric and partly recursive structures. The concept is top-down and bases on a generative model given by an attribute grammar. It uses the well known random sample consensus paradigm for model selection.

The attribute grammars formalism is used for the modeling of 3D objects. It models composition by production rules and it can also incorporate constraints on parameters and geometric dependencies between parts by attributes.

The constraints which are given by the grammar advise to preprocess the data in order to improve the RANSAC algorithm. In the preprocessing step the normal vector of each plane defined by a point and his n neighbors is estimated. The samples are drawn from a subset of these 'needles'.

1. INTRODUCTION

Top-down approaches for the reconstruction of 3D urban scenes, particularly 3D buildings, base on abstract models. Therefore the objects are described by detailed parameterized models. The main task of their reconstruction from any observation is to find a set of parameters that instantiates an optimal model, i.e. the model that optimally fits the observation. Top-down approaches can be used for the interpretation of images and also for LIDAR based reconstructions. In this paper we present a concept for the reconstruction of symmetric objects from 3D point clouds measured by terrestrial laser scanners.

For the sake of clarity we will first explain the meaning of the words *object*, *model* and *instance* as they are used in this paper. (1) an *object* is a (generalised) representation of a real world item e.g. talking about the object 'building' we do not think about each brick or cornice but of a more generalized representation, (2) a *model* is an abstract and parameterised representation of a real world item e.g. a cuboid with unknown scale is a low detailed model of a building, and (3) an *instance* is a model with fixed parameters.

The random sample consensus (RANSAC) presented by Fischler and Bolles (1981) is a robust and well known top-down algorithm for the classification of 3D point clouds. It divides the given data set into inliers and outliers. The algorithm needs a parametrical model as input and then estimates sets of parameters that define an instance of the model.

In the presented concept the parameterized model which is used by the random sample consensus is given by an attribute grammar. The original usage of formal grammars is the definition of formal languages. Therefore grammars can either be used to generate words of the particular language or to parse given words in order to validate them. But attribute grammars can also be used to model geometric objects. Due to recursive production rules the variety of words that can be derived from grammars is quite high – actually infinite, but it can be limited by constraints defined in the semantic rules. The geometrical constraints of the model are also defined in the semantic rules. They are used to pre-process the observation data and to reduce the size of the set from which the samples are drawn. The geometric constraints also imply an iterative model selection which estimates one parameter at a time. In each loop one more parameter is estimated. Finally the model selection chooses the model with highest support i.e. the largest consensus set.

In this paper we present a concept for the reconstruction of recursive and symmetrical 3D objects from 3D point clouds. The rest of the paper is structured as follows. We give an overview of the related work in the following section (2). The model selection and first results are presented in section 3. In the last section (4) we summarize and discuss our work and give an outlook for future works.

2. RELATED WORK

A lot of work for automatic grouping or clustering of features has been presented since decades. The task is an autonomous segmentation of feature sets. The methods are not specialized to geometric features but reconstruction of 3D building (parts) is so. Well known and powerful methods are k-means (Lloyd 1982), the expectation maximization algorithm (Dempster et al. 1977) and the principal component analysis (Pearson 1901). Since these algorithms base on simple implicit models, it is hard to implement more elaborate models. However this is essential for the reconstruction of complex and composite objects in a top-down approach where a given model is used to limit the search space.

More elaborated approaches base on model selection methods. In principle they compare an instance of a given model with the observations and calculate the goodness-of-fit. Some methods like Markov Chain Monte Carlo (Metropolis and Ulam 1949) base on sampling from given distribution. Graphical models like Bayesian networks base on probabilistic modelling. The random sample consensus (Fischler and Bolles 1981) uses a geometrical model to calculate the distance of each point to a particular instance of the model. The model selecting function usually returns that set of parameters with the highest support, i.e. a high number of inliers. However, more elaborated functions can be used for the model selection.

Information criterions like the Akaike Information Criterion (Akaike et al. 1974) can be used to avoid over fitting or under fitting by statistically founded similarity measures.
Geman et al. (2002) give a mathematical formulation of compositionality which is, to their mind, fundamental to the human ability of cognition.

Formal grammars have been introduced by Chomsky (1956; 1959) to describe the linguistic structure of natural and formal languages. In the sense of Chomsky a formal grammar is defined as the quadruple

$$G = \{T, N, S, P\}$$

of a finite set of terminal symbols T, a finite set of non-terminal symbols N with $T \cap N = \varnothing$, a start symbol S with $S \in N$ and a finite set of production rules P of the form

$$X \rightarrow x \text{ with } X \in N,\ x \in (N \cup T)^*$$

Knuth (1968; 1971) described how to assign semantics to context-free languages by attribute grammars. Therefore a finite set of attributes $A = \{a_1, a_2, \ldots, a_n\}$ is added to each symbol Y with $Y \in (N \cup T)$. The processing of the attributes in the derivation tree is defined by a finite set of semantic rules $R(P)$ attached to each production rule P in the following way:

$$P\colon Y_0 \rightarrow Y_1 \ldots Y_m$$

$$R(P)\colon Y_i.a_p := f(Y_j.a_q, \ldots, Y_k.a_r) \text{ with } \{i, j, \ldots, k\} \in [0,m]$$

We use attributes for the information flow between separated nodes of the derivation tree. In contrast to the usage of attribute grammars in compiler design we do not restrict semantic rules on symbols but heavily use numbers and functions for the coding of geometric relations or probabilistic information.

Müller et al. (2006) extend split grammars (Wonka et al. 2003) for procedural modelling of buildings. They present a 3D grammar enabled to do affine transformations such as rotation. Symbols are represented by volumetric objects which interact like in constructive solid geometry (Mäntylä 1988).

A concept for the modelling of symmetry and topology of geometric objects using attribute grammars has been presented by Schmittwilken et al. (2009). Krückhans & Schmittwilken (2009) present a grammar based approach for the synthetic generation of facades.

Ripperda and Brenner (2006) use formal grammars in combination with reversible jump Markov Chain Monte Carlo (rjMCMC) methods for the reconstruction of facades from terrestrial image data. They generate model assumptions which optimally fit the observations by guiding the derivation process by rjMCMC. Müller et al. (2007) use terrestrial images for a grammar based reconstruction of facades. They present a four step top-down method which ends up in a 3D facade model. The work of Ripperda and Brenner and Müller et al. mainly use grammars for a two dimensional description of the facades.

3. MODEL SELECTION WITH ATTRIBUTED GRAMMARS

In this section the concept of grammar based model selection for symmetric objects is introduced and exemplarily explained on the reconstruction of stairs from terrestrial LIDAR data.

We first illustrate the usage of attribute grammars as a modelling language for 3D objects (3.1). After that, we give a closer look to our improvements of the random sample consensus which are used for the model selection (3.2).

3.1 Attribute grammars for 3D objects

Attribute grammars can be is used to define models of 3D objects. Generally the object and its parts are represented by the grammar symbols (terminal and non-terminal). The parameters of each object part, particularly the form parameter and location parameter, are represented by the attributes of the according symbol. Attributes are propagated by the semantic rules which are attached to the production rules.

Table 1 shows an exemplary grammar of stairs. Upper case letter are used for non-terminal symbols, lower case letters denote terminal symbols. The superscript indices are used to differentiate between multiple occurrences of the same symbol. The following abbreviations are used for the symbols and will also be used in the following of this section.

S… staircase
F… flight (unbroken series of steps)
L… landing
r… riser (vertical part of a step)
t… tread (horizontal part of a step)
w… winder (curved tread)

ID	Production rule	Description
P_1	$S \rightarrow F$	A flight is a (simple) stairway.
P_2	$S \rightarrow F^1\ L\ F^2$	Two flights connected by a landing are also a stairway.
P_3	$F^1 \rightarrow r\ t\ F^2$	Riser and tread are recursively aggregated to a flight.
P_4	$F^1 \rightarrow r\ w\ F^2$	Riser and winder are recursively aggregated to a curved flight (special case: spiral stair)
P_5	$F \rightarrow r\ t$	Termination of the flight (production)
P_6	$L^1 \rightarrow L^2\ F$	Multiple flights are connected by a landing.
P_7	$L \rightarrow l$	Termination of the landing (production)

Table 1: Production rules of the stair grammar.

Table 2 shows a subset of the semantic rules of the stair grammar. The rules $R_1 – R_9$ propagate the reference point from the left side symbol of the production rule to the right side symbols like it is illustrated in Fig. 1.

ID	Semantic rule
$R_1(P_3)$	$r.x = F^1.x$
$R_2(P_3)$	$r.y = F^1.y$
$R_3(P_3)$	$r.z = F^1.z$
$R_4(P_3)$	$t.x = F^1.x$
$R_5(P_3)$	$t.y = F^1.y$
$R_6(P_3)$	$t.z = F^1.z + F^1.rise$
$R_7(P_3)$	$F^2.x = F^1.x$
$R_8(P_3)$	$F^2.y = F^1.y + F^1.treadDepth$
$R_9(P_3)$	$F^2.z = F^1.z + F^1.rise$
$R_{10}(P_3)$	$r.rise = F^1.rise$
$R_{11}(P_3)$	$t.depth = F^1.treadDepth$
$R_{12}(P_3)$	$F^1.numberOfSteps = F^2.numberOfSteps + 1$

Table 2: Excerpt from the semantic rules of the stair grammar.

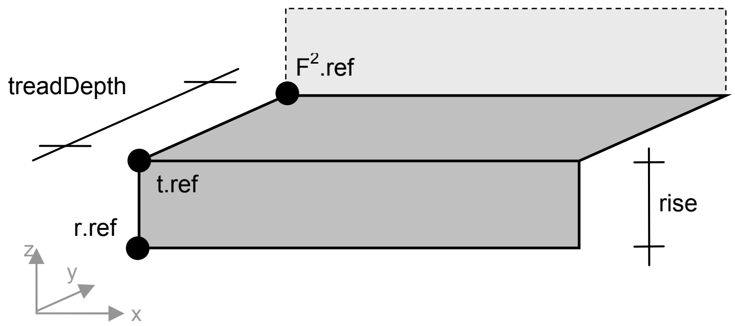

Figure 1: Location parameter and form parameter of a single step.

ID	Semantic rule	Substituted rules
$R_{1a}(P_3)$	*$r.ref = identity(F^1.ref)$*	$R_1(P_3)$, $R_2(P_3)$, $R_3(P_3)$
$R_{2a}(P_3)$	*$t.ref = transform(F^1.ref, 0, 0, F^1.rise)$*	$R_4(P_3)$, $R_5(P_3)$, $R_6(P_3)$
$R_{3a}(P_3)$	*$F^2.ref = transform(F^1.ref,0,F^1.treadDepth,F^1.rise)$*	$R_7(P_3)$, $R_8(P_3)$, $R_9(P_3)$

Table 3: Simplified semantic rules.

For the sake of readability and simplification three rules at a time are combined using predefined functions (cf. Table 3).

Since the geometry of the objects is exclusively represented by location parameters and form parameters, which are stored in attributes, the order of symbols in the derived word becomes less important. Adjacencies of symbols in the word need not result in adjacencies of their geometric interpretation.

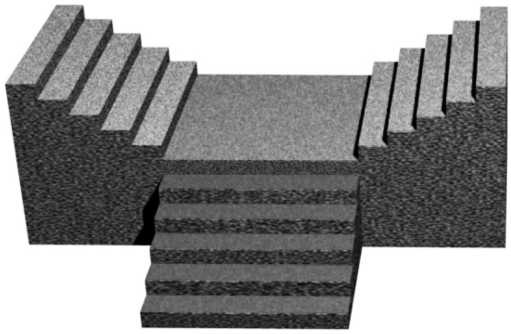

Figure 2: A triple-run staircase.

For example, the derivation of a stair with three flights converging at one landing (cf. Fig. 2) first involves the production rule P_2 to generate a landing with one oncoming flight and one flight going off. The production rule P_6 adds a third flight to the landing. We derive the word $F^1\ L\ F^2\ F^3$ (the superscript indices are used to differentiate between multiple occurrences of the same symbol) where the flights F^2 and F^3 are both topological correctly connected (adjacent) to the landing L, even if F^3 is not a neighbour of L in the word. Applying P_3 and P_5 to F^2 and F^3 generates a series of $r\ t$: $(r\ t)^n\ (r\ t)^m$ which is a notation for n or m repetitions of the sequence $r\ t$. After the termination of F^2 by P_5 the adjacency of $(r\ t)^m$ and L is even less obvious.

In addition to the rules given in Table 3 a lot of semantic rules remain. They are beyond the scope of this paper and therefore only the aims of the remaining rules are sketched as follows:

- Inheritance of all form parameters and location parameters
- Topological correct instances
- Constraining parameters and the derivation:
  - Maximum number of steps
  - Thresholds for rise, tread, and winder angle
  - Curved parts of flights normally sum to 90° at a stretch
- Correct connection of multiple flights to a landing

Applying the grammar, the semantic rules are applied consecutively. Each semantic rule propagates attribute values between father and sons. The father-son-metaphor is suitable concerning the derivation tree. As the stair example illustrates, the semantic rules realize an inheritance of attribute values rather than a simple propagation. Sons often inherit properties like their orientation and form parameter (R_{10} and R_{11}) from their fathers. This is essential because the form parameter of steps do not change within a flight.

3.2 Reconstruction

A subset of the semantic rules given above specifies suffices to specify the general concept of a staircase. The most general property of a staircase is the horizontal alignment of treads, the vertical alignment of risers, and the invariability of rise and tread depth within a staircase. The latter is given by $R_{10}(P_3)$ and $R_{11}(P_3)$ inter alia. This general model of a staircase ensures that treads and risers both lie in parallel, equidistant planes - treads in horizontal ones and risers in vertical ones.

Rise (distance of horizontal planes) and *treadDepth* (distance of vertical planes) are the parameters of this general model. The estimation of *rise* and *treadDepth* is done in the model selection step using random sample consensus, namely by finding the most probable instance of the general model. This is presented in the following sub sections. The presented results refer to the 'entrance stair data' set shown in figure 3.

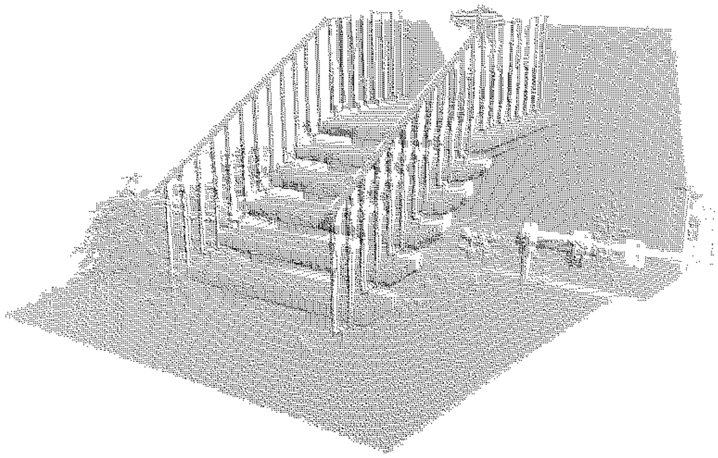

Figure 3: 3D point cloud of the 'entrance stair data' set.

Pre-processing

The two sets of parallel planes are an important property of the general model. The parallelism of the planes can be used for an improvement of the RANSAC algorithm. The efficiency of the algorithm mainly depends on the complexity of the model, i.e. the number of samples. The more complex the model is the more probable is the failure of a sample.

Although the model of parallel planes is quite simple, the sampling is improved by a reduction of the set of points from which the sample is drawn. Therefore the 'normal vector of each point' $\boldsymbol{n}_p$, is estimated i.e. the normal vector of the plane which is given by $\boldsymbol{n}_p$ and its k neighbours. The points and the corresponding normal vectors will be named *needles* in the following. Although this is an expensive process without spatial indexing, it has to be done only once and the number of failed sample sets will be reduced.

Our Java based implementation uses an octree spatial indexing of the 3D point cloud and we calculate the needles by a principal component analysis (PCA) (Pearson 1901). The octree of the 3D point cloud shown in figure 3 (entrance stair data set - ca 52.000 points) is build ca 1.100 times faster than the calculation of all needles by a PCA. The first lasts ca 30ms the latter 34s on a 1.6GH dual core CPU.

Two subsets of the needles P are additionally stored in H and V containing horizontal and vertical needles.

$$H = \left\{ p \in P \mid \left| \vec{n}_p \bullet \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \left| \vec{n}_p \right| \bullet \left| \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right| \right| \le \varepsilon_h \right\} \tag{1}$$

$$V = \left\{ p \in P \,\middle|\, \left| \vec{n}_p \bullet \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right| \leq \varepsilon_v \right\}$$

Model selection

In the following we will give an overview about the sampling strategy and model instantiation. We first focus on the sampling and model instantiation of horizontal planes (parameter *rise*) which is more general. The more elaborated computation of vertical planes (parameter *treadDepth*) is presented after that.

Horizontal planes are invariant against rotations around the z-axis i.e. the position of the laser scanner relative to the staircase i.e. the rotation of the point cloud within its relative coordinate system. Furthermore they are invariant against landings and winders (curves) within the stair. Thus the sampling and model instantiation of horizontal planes is done in the first stage which can be computed without consideration of any transformations of the 3D point cloud. Thus the presented model selection even works for winded stairs or spiral stairs.

Therefore two points are randomly drawn from the set of vertical needles V. The parameter rise is calculated by solving equation 2.

$$\Delta h = n * rise \tag{2}$$

Δh is the height difference of the two points, n is the number of steps and hence it is integral. The parameter rise is real-valued. Thus it states a mixed real integer problem. However, rise is constrained by thresholds. We have measured 120 stairs and determined the two thresholds $rise \in [0.14, 0.20]$ and $treadDepth \in [0.18, 0.38]$. Inserting the minimum rise and maximum rise into equation 2 leads to a minimum and maximum number of steps.

$$\begin{aligned} n_{\min} &= \lceil \Delta h / rise_{\max} \rceil \\ n_{\max} &= \lfloor \Delta h / rise_{\min} \rfloor \end{aligned} \tag{3}$$

For each possible number n of steps the parameter *rise* is calculated by solving equation 2.
For example, let a random sample of two points with a height difference of *1.53m* be given. The problem is solved by four tuples of the type *{n,rise}*: *{{8,0.19}, {9,0.17,}, {10,0.15}}*. This states the number of models which have to be validated independently.

Additionally we limit the sampling algorithm to sample locally. Since stairs are usually local features of the 3D point cloud and a large percentage of points are not onto the steps we force the two points of the sample to be within a maximum distance. The sampling algorithm is given below.

```
function getSampleSet(V) {
  εd = tolerance
  maxDist = maximum 3d distance
  Randomly select p1 from V
  Randomly select p2 from V
     {p ∈ V | |p.z - p1.z| < maxDist + 2 * εd}
  return {p1, p2}
}
```

Finally the model with the highest support i.e. the largest consensus set is selected. Therefore each observed point is tested whether or not it fits the model. Certainly the consensus set is computed from all points. They are not limited to horizontal or vertical needles! The score function is derived from equation 2 as follows: The height difference of the point that has to be tested and the first sample point has to fulfill equation 2 i.e. the number of steps *n* has to be integral. Since modulo operation is also defined for real numbers, equation 4 is derived. The tolerance measure ε is used to control accuracy.

$$(\Delta h \bmod rise) * rise \leq \varepsilon \tag{4}$$

In our implementation we use the mathematical definition of modulo operation instead. We have noticed that it is much faster in Java:

$$((\Delta h / rise) - \lfloor \Delta h / rise \rfloor) * rise \leq \varepsilon \tag{5}$$

Figure 4 shows the results of the classification. It is obvious that parallel horizontal planes have been estimated. It is also obvious that per plane the set of inliers still contains points that do not lie on treads, e.g. points of the handrails or of the backmost wall (c.f. figure 5). But however, the parameter *rise* is estimated accurately.

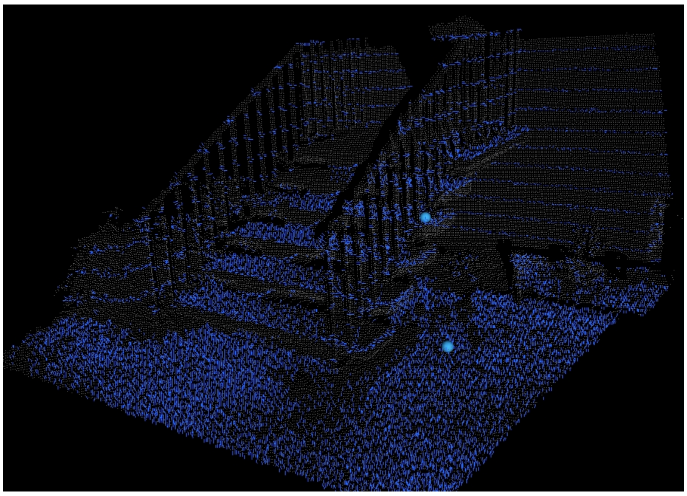

Figure 4: Estimation of parallel horizontal planes covering the treads of the entrance stair data set (light blue spheres = random sample, blue needles = inliers, grey points = outliers).

The estimation of *treadDepth* and parallel vertical planes is analogous. The sample set is drawn from horizontal needles. However, one more constraint is necessary to ensure parallelism of the planes: two randomly sampled needles have to be parallel and so each inlier has to do.

The model selection of curved steps also is analogous to the estimation of straight stairways. Like *rise* and *treadDepth* the angle of the winder is constrained by thresholds. According to straight stairs the vertical planes of a winded staircase have to draw this angle. The estimation of *rise* (horizontal planes) is independent from the type of staircase. Figure 6 shows the estimation of parallel horizontal planes from an artificial data set without noise, except the column in the centre of the stairway.

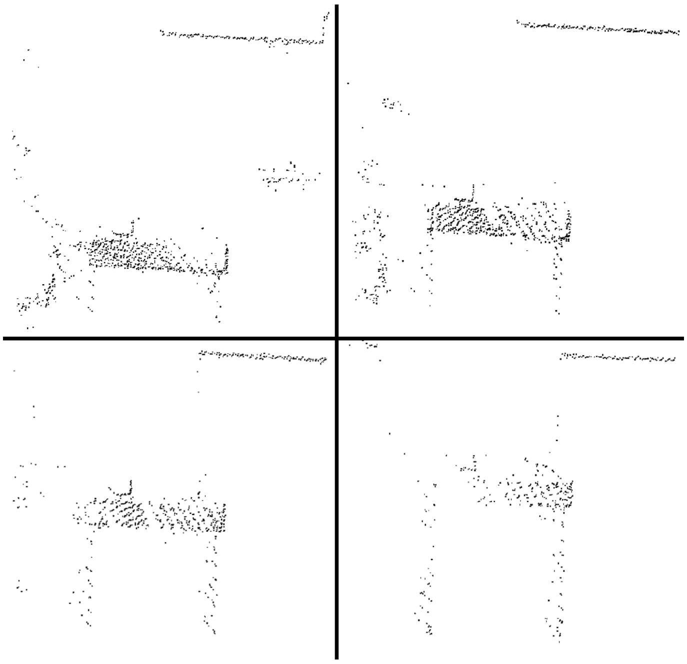

Figure 5: Four nadir views onto single planes of the entrance stair data set.

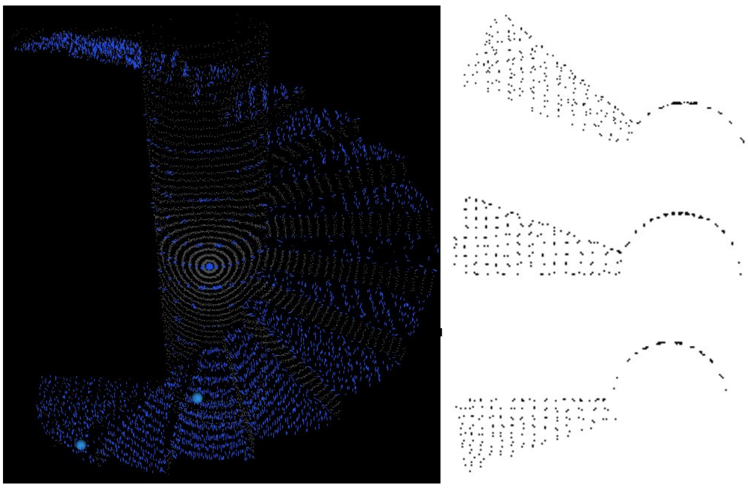

Figure 6: Left: Estimation of parallel horizontal planes covering the treads of a spiral staircase (light blue spheres = random sample, blue needles = inliers, grey points = outliers). Right: Three nadir views onto single planes.

4. CONCLUSION AND FUTURE WORK

In this paper we have presented a grammar based top-down approach for the reconstruction of symmetric objects by model selection. Exemplarily we have focused on stairs.

We have shown how attribute grammars can be used for generative modelling of geometric objects. Therefore form parameter and location parameter are represented by attributes and their propagation within the model is organized by semantic rules. The composition of object parts can neatly be defined by a single production rule which also covers recursive and symmetric compositions. Constraints on the geometric properties of the composition itself are also defined by semantic rules.

An improved sampling strategy for the random samples and consensus paradigm has been presented. The observations are pre-processed before starting the model selection with the RANSAC algorithm. Furthermore the samples are drawn from only a subset of the original data which is defined by the constraints specified in the grammar. Finally the model with largest consensus set is chosen.

The derived model parameter rise and *treadDepth* are estimated independently in two autonomous runs of the RANSAC algorithm although they are highly correlated by their geometry. Particularly the direction of the vertical planes sometimes does not fit the risers. This could be illustrated as follows: (1) Samples are drawn from horizontal needles, so the direction of the stair is unknown respectively it is free. (2) Particularly for open stairs without faced risers there are only a small number of inliers for vertical plane models.

Therefore a combined model selection which considers the geometric constraints of risers and treads should be preferred.

Our actual work deals with the reduction of the horizontal planes in order to receive treads. For each plane the subset of points which lie on treads is computed. These subsets are constraint by the parallelism of their principal components. Given the subsets and their principal components, the sampling of points for the model selection of vertical planes (*treadDepth*) will be limited to horizontal needles parallel to the second principal component of the treads. The set of ‘tread points’ also gives a stronger threshold for the depth of the treads. Furthermore it gives an estimation of the width of the stair.

The presented concept should also be transferred to other symmetrical objects, e.g. arcades or rows of windows, and afterwards be generalised to any composite object. Therefore the linkage between grammar based modelling and model selection has to be deepened.

Acknowledgements

The authors thank Dr. Claus Brenner from the Institute of Cartography and Geoinformatics, University of Hannover for the provision of data sets (including entrance stair data set).

This work is funded by the German Research Foundation (DFG) as part of the sino-german joint cooperation on ‘Interoperation of 3D Urban Geoinformation’.

BIBLIOGRAPHY

Akaike H, Petrov B, Csaki F (1974) A new look at the statistical model identification. IEEE Transactions on Automatic Control, 6:716-723

Chomsky N (1956) Three models for the description of language. Information Theory, IEEE Transactions, 3:113-124

Chomsky N (1959) On Certain Formal Properties of Grammars. Information and Control, 137-167

Dempster AP, Laird NM, Rubin DB (1977) Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society. Series B (Methodological), 1:1-38

Fischler MA, Bolles RC (1981) Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Communications of the ACM, 6:381-395

Geman S, Potter D, Chi Z (2002) Composition Systems. Quarterly of Applied Mathematics, 707-736

Knuth DE (1968) Semantics of context-free languages. Theory of Computing Systems, 2:127-145

Knuth DE (1971) Top-down Syntax Analysis. Acta Informatica, 2:79-110

Krückhans M, Schmittwilken J (2009) Synthetische Verfeinerung von 3D-Stadtmodellen. In: Proceedings of DGPF-Jahrestagung, March 2009, Jena, Germany

Lloyd S (1982) Least squares quantization in PCM. IEEE Transactions on Information Theory, 2:129-137

Mäntylä M (1988) An Introduction to Solid Modeling. Principles of Computer Science. Computer Science Press

Metropolis N, Ulam S (1949) The Monte Carlo Method. Journal of the American Statistical Association, 247:335-341

Müller P, Wonka P, Haegler S, Ulmer A, van Gool L (2006) Procedural Modeling of Buildings. In: (eds): Proceedings of ACM SIGGRAPH 2006 / ACM Transactions on Graphics (TOG) ACM Press, pp 614-623

Müller P, Zeng G, Wonka P, van Gool L (2007) Image-based Procedural Modeling of Facades. In: Proceedings of ACM SIGGRAPH 2007 / ACM Transactions on Graphics. ACM Press, New York, NY, USA, p 85

Pearson K (1901) On lines and planes of closest fit to systems of points in space. Philosophical Magazine, 6:559-572

Ripperda N, Brenner C (2006) Reconstruction of Facade Structures Using a Formal Grammar and RjMCMC. In: Franke K, Müller K-, Nickolay B, Schäfer R (eds): Proceedings of 8th DAGM Symposium, Berlin, Germany, September 12-14 2006. Springer, pp 750-759

Schmittwilken J, Dörschlag D, Plümer L (2009) Attribute Grammar for 3D City Models. In: Proceedings of Urban Data Management Symposium (UDMS), June 2009 Ljubljana, Slovenia. Taylor & Francis

Wonka P, Wimmer M, William FS (2003) Instant Architecture. ACM Transactions on Graphics, 4:669-677