A GENERIC 3D MODEL FOR AUTOMATED BUILDING ROOF RECONSTRUCTION

S. Scholze\textsuperscript{a,}\textsuperscript{c}, T. Moons\textsuperscript{b,}\textsuperscript{c}, L. Van Gool\textsuperscript{b,}\textsuperscript{c}

\textsuperscript{a}ETH Zurich, Computer Vision Laboratory (BIWI), Zurich, Switzerland
\textsuperscript{b}Katholieke Universiteit Leuven, Center for Processing Speech and Images (ESAT / PSI), Leuven, Belgium
\textsuperscript{c}Katholieke Universiteit Brussel, Group of Exact Sciences, Brussel, Belgium

Working Group V/2

KEY WORDS: Modelling, Reconstruction, Building, Three-dimensional, Aerial, Vision

ABSTRACT

This paper investigates into model-based reconstruction of complex polyhedral building roofs. A roof is modelled as a structured ensemble of planar polygonal faces. The modelling is done in two different regimes. One focuses on geometry, whereas the other is ruled by semantics. Inside the geometric regime, which is the primary topic of this paper, 3D line segments are grouped into planes and further into faces using a Bayesian analysis. In the second regime, the preliminary geometric models are subject to a semantic interpretation. It is shown, how the knowledge gained in this step can be used to infer missing parts of the roof model (by invoking the geometric regime once more) and to adjust the overall roof topology. Several successfully reconstructed complex roof structures corroborate the potential of the approach.

1 INTRODUCTION

Three-dimensional, multiview modelling is a powerful approach for automated scene interpretation and reconstruction. Although good results have been obtained from no other information than image content, applications like 3D roof reconstruction for buildings from aerial images still pose a serious challenge. There is often little texture for correspondence search to work well, and essential features like straight lines often go undetected. On the other hand, the class of typical roof shapes is sufficiently constrained as to supply strong prior knowledge to the system. Dealing with complex man-made objects such as building roofs in urban areas requires a generic and expressive model. Finding the right balance between flexibility and predictive power of the model is of key importance to the entire reconstruction. Although considerable research has been conducted on model-based building reconstruction, less effort has been invested in the development of the models themselves. This paper proposes a parametric model which is applicable to a broad range of roof types, together with probabilistic and semantic reasoning procedures to instantiate and finetune the model.

1.1 Previous Work

Despite the long lasting demand for cheap, reliable and up-to-date 3D city models for technical and environmental planning tasks, automated building reconstruction is still an only partially solved task. An excellent overview of important developments and state-of-the-art in the field can be found in the Proceedings of the Ascona Workshops on Automatic Extraction of Man-Made Objects from Aerial and Space Images (Grün et al., 1995, Grün et al., 1997, Baltzavias et al., 2001). Some previous approaches, able to reconstruct complex buildings with minimal user interaction include (Henricsson, 1998, Moons et al., 1998, Baillard and Zisserman, 2000, Fischer et al., 1998).

The roof models used in the literature range from very generic ones (polyhedral structures) up to very specific ones (parametrized models of different building types). For an overview see (Mayer, 1999). For handling uncertainty and imprecision of the input data, some building reconstruction systems already have a probabilistic underpinning (Cord et al., 2001, Kulschewski, 1997, Heuel and Förstner, 2001).

1.2 A Geometric Model in a Probabilistic Setup

In this paper, a probabilistic formulation for geometric building reconstruction is proposed. Its key component is a generic and expressive roof model. As a novelty, various constraints can be explicitly imposed onto the model. The conjunction of robust geometric reasoning in 3D space together with a semantic interpretation allows to reconstruct complex building roofs completely and with correct topology. Figure 1 gives an overview of the presented methodology.

The organisation of this paper is as follows. A detailed description of the proposed roof model is presented in Sect. 2. Section 3 mainly states the processing steps from the image data up to reconstructed 3D line segments. Additionally, the notion of semantic labels and the used test dataset are introduced here. In Section 4, the steps leading to a preliminary geometric roof model are described. A summary on how a semantic interpretation of this geometric model is obtained and used to complete and refine the roof model is given in Section 5. Finally, results of the approach are presented in Sect. 6.

2 GEOMETRIC ROOF MODEL

Guided by the observation that the broad majority of building roofs (at least in the western world) consist of planar surfaces, and that both, internal roof boundaries and the outline of the entire roof are delineated by straight lines, a roof model is proposed, which describes a roof as ensemble of planar polygonal faces (patches). It suffices to distinguish between triangular and quadrangular patches, since more complex patch shapes (e.g. L-shapes) are obtained by patch composition. Theoretically, it would suf-
A roof patch is a triangular or quadrangular face in 3D space. It can be accessed through two types of representations, which are kept in parallel. On the one hand, a parametric representation is provided, allowing direct inference of quantities such as angles or lengths. Three coordinate systems are involved in the parametrization: a world coordinate system $O$, a patch centered coordinate system $O'$ and a 2D coordinate system in the patch plane $O''$. The relation between $O$ and $O'$ is illustrated in Figure 2. Figures 3(a,b) show the parametrization of the triangular and quadrangular patch models in the patch plane coordinate system $O''$. The advantages of the chosen parametric representation are that the quantities involved have an obvious meaning. Probability distributions have been obtained from a test dataset. Additionally, this representation allows to incorporate symmetries between different patches of a roof model.

In parallel to the parametric representation, a representation based on the 3D world coordinates of the corner points $P_0, \ldots, P_3$ of a patch is kept as well. Although this dual representation is redundant, its importance transpires when modelling the relations between the patches of a roof. Possible relations are given by coincidence or equality constraints and can hold either between coordinates or parameters of different patches. Consequently, the dual representation allows to integrate diverse constraints such as geometric symmetry or topological connectivity simultaneously and at the same level of complexity.

![Image of Figure 1: Structure of geometric and semantic modelling system](image1)

Figure 1: Structure of geometric and semantic modelling system. From the input data (a), consisting of 3D line segments, a set of seed segments (b) is chosen by virtue of some geometric properties. In the neighbourhood of these, preliminary geometric models consisting of planar faces in 3D space are reconstructed (c). Firstly, the semantic labelling performed in (e) allows to detect missing parts of the model and thus to determine additional seed segments (d) which in turn are used for geometric model instantiation. Secondly, the semantic labels are used to improve the overall model topology (f).

![Image of Figure 2: Patch model; global and patch centered coordinate systems](image2)

Figure 2: Patch model; global and patch centered coordinate systems. $T$ is the transformation from world coordinates to the local patch coordinate system including a translation $t$ and a rotation around the $z$-axis by $\phi_2$. $\phi_1$ is the angle between the $x', y'$-plane and the $x, y$-plane, $z$ and $z'$ are parallel; the $x, y$-plane is horizontal. The slant of the patch plane is given by $\phi_2$, which corresponds to the rotation performed to get the 2D patch coordinate system $O''$. $\{P_1\}$ denote the corner points of the patch ordered in mathematical positive sense.

![Image of Figure 3: Patch models; patch plane coordinate system](image3)

Figure 3: Patch models; patch plane coordinate system. $O''$ is the origin of the 2D coordinate system in the patch plane. $\{p\}$ denote the projections of the corner points into the patch plane. $l_{ij}$ denotes the line segments given by $p_i$ and $p_j$, $l_{12}$ and $l_{30}$ are parallel to $x''$. The inclination of the bordering segments is denoted by $\alpha_1$ and $\alpha_2$ respectively. The full parametrization of a patch is given by specifying the height $h$, and the width $w$ in case of a quadrangular patch.

2.2 Constraints

A roof is described by its constituting patches and the relations between them. These relations are modelled as constraints between parameters or coordinates of different patches in a roof similar to what has been done to describe dependencies within CAD-models (Seybold et al., 1997, Ault, 1999). Once a relation between patches has been es-
established, changes to one patch are propagated to all other affected patches in the roof. In keeping with the dual patch representation, two kinds of constraints exist. Firstly, parametric constraints can be used to couple a specific parameter of two patches. As example consider a symmetric roof consisting of two patches (Figure 4(a)). By setting the parametric constraint $\phi_2^{(1)} = \phi_2^{(2)}$, this symmetry of the roof is always preserved; if the inclination of one of the patches is changed, the inclination of the other is updated automatically.

Secondly, coordinate constraints allow to model coincidence relations between patch corner points. As an example for both types of constraints, consider the L-shaped patch in Figure 4(b), constructed out of two quadrangular patches which have collinear upper borders. In order to compose the two patches to represent one L-shaped patch, two constraints are imposed. By making use of the parametric representation, a unique slant angle can be achieved by requiring $\phi_2^{(1)} = \phi_2^{(2)}$. Additionally, the coordinate based representation allows to glue corner points together by setting $P_0^{(1)} = P_3^{(2)}$. Constraints are also important during topological model refinement (Section 5.2).

3.1 3D Line Segment Reconstruction

The raw input data for our application consist of high resolution aerial colour images of densely built up urban areas. (Figure 5(a,b) shows an example stereo pair.) The images are taken with mutual overlap in order to allow a 3D reconstruction of the scene. The precise camera calibration is known. After edge detection and straight line fitting, corresponding 2D line segments are sought in the stereo pairs. Besides the geometric constraints given by epipolar geometry and the specific image acquisition setup, chromatic constraints are also exploited to disambiguate the matching and thus reduce the number of wrong matches (Scholze et al., 2000). The chromatic constraints are obtained by comparing the image neighbourhood of the line segments under consideration in order to reject chromatically incompatible matches. Some of the remaining matches are purged by considering the trifocal constraint, and the chromatic constraints already mentioned between multiple views. If a match passes all statistical and geometrical tests, its 3D position is recalculated using bundle adjustment to enhance the geometric precision of the reconstructed 3D line segment. The resulting 3D line segments (see Figure 5(c)) form the input for the roof reconstruction process described in this paper.

3.2 Semantic Labels by Geometric Measurements

The semantic interpretation of 3D line segments is a key feature of the presented methodology. Five different semantic labels for patch segments are distinguished. The
five possible labels form the set $\Omega$:

$$\Omega = \{ \omega^{(1)} = \text{ridge}, \omega^{(2)} = \text{gutter}, \omega^{(3)} = \text{gable}, \omega^{(4)} = \text{convex}, \omega^{(5)} = \text{concave} \} \quad (1)$$

The names of the labels are chosen to be coherent, although they should not be taken literally. For example, a gutter segment just corresponds to the lower boundary of a patch, no matter if there is a gutter in the scene or not. Figure 6 gives an overview. For convenience each semantic label (e.g. gutter) is represented by a variable $\omega^{(i)}$, where $i$ encodes the actual label. In order to assign the semantic labels to the segments in a patch, geometric measurements are used. Measurements characterizing individual segments $u_i$ (unary attributes) and measurements between adjacent segments $b_{ij}$ (binary attributes) are distinguished. The actual measurements of

$$u_i^{(1)} : \text{length of segment } i$$
$$u_i^{(2)} : \text{slant angle of segment } i \text{ relative to a horizontal plane}$$
$$b_{ij}^{(1)} : \text{angle enclosed by adjacent (coplanar) segments } i \text{ and } j$$
$$b_{ij}^{(2)} : \text{(signed) mid point height difference of adjacent segments } i \text{ and } j \quad (2)$$

form the components of the attribute vectors $u_i = (u_i^{(1)}, u_i^{(2)})$ and $b_{ij} = (b_{ij}^{(1)}, b_{ij}^{(2)})$.

Figure 6: Functional parts of a roof and their semantic labels. The label ridge is generally used for the upper boundary of a patch, gutter for the lower one. The boundaries of a patch are either labelled gable, convex connection or concave connection depending on the neighbourhood.

### 3.3 Test Dataset

A semi-automatically reconstructed test dataset has been used to learn the statistics of the geometric measurements. The dataset shows urban and sub-urban areas with four-fold overlap at an image scale of approximately 1:5000 (IGP, 2001). Nearly 150 roofs, consisting of about 80 triangular and 270 quadrangular patches have been labelled manually using the semantic labels from Equation (1). Additionally to the semantic labels, the values of the geometric measurements (unary and binary attributes) have been recorded.

### 3.4 Initial Semantic Interpretation

As will be described in the next Section, the patch reconstruction algorithm makes use of seed line segments for its search of patch hypotheses. Given the set of reconstructed 3D line segments, a set of seed segments has to be derived which is used for the first pass of the reconstruction algorithm (cf. Figure 1(b)). In order to invoke the appropriate instantiation procedure, the label of the seed segment in a roof should be known in advance. Hence, the segments have to be attributed semantic labels from Equation (1). Of course, at this stage it is impossible to unambiguously assign a label to an arbitrary 3D line segment. Nevertheless, class membership probabilities can be derived. Using the distribution of only the unary attributes (line segment length and slant) of the test dataset, the probability distributions for the respective labels are determined. For each class, the algorithm creates a list of seed line segments sorted by probability.

### 4 PATCH RECONSTRUCTION

Geometric patch reconstruction poses a twofold task. First, planar 3D line segment configuration corresponding to planar roof patches are detected. Second, the outlines of patches, corresponding to the actual roof faces lying in these planes, have to be determined.

#### 4.1 Probabilistic Plane Detection

For plane detection, the general approach is to rotate a half plane around a seed line segment, e.g. a tentative ridge line (cf. Figure 7). For each inclination of the half plane, the line segments in its neighbourhood which approximately lie in this plane are collected. To determine the number of planes and their positions, a Bayesian model selection procedure is applied. The plane model selection procedure maximizes the posterior probability

$$p(m|x) \rightarrow \max$$

with $x$ being the set of neighbouring 3D line segments an $m$ the actual plane model. The plane model describes the number, positions and orientations of planes currently considered. Model selection is done taking into account model complexity. That is, a more complex model (containing more planes) may well describe the data better; but might also tend to over-fit the data. Consequently, a simpler model should be preferred if the more complex one does not describe the observations significantly better. This qualitative statement is quantified in form of an Ockham factor. A detailed explanation of the procedure is available in (Scholze, 2002). The result is a small number (usually 1–4) of sets of plane hypotheses in the neighbourhood of a seed line together with their probabilities. The different sets of plane hypotheses correspond to different interpretations of the scene.

#### 4.2 Maximum Expected Utility Patch Reconstruction

During the Bayesian plane selection procedure, 3D line segments are associated with different plane hypotheses.
The next step now is to determine the outlines of possible patches, containing the segments in the individual plane hypotheses. The outlines are found by computing the extremal points of the convex hull of the segments in a tentative plane. The extremal points correspond to the end points of the associated line segments, projected into the patch plane, with minimal/maximal coordinates. Then, the patch parameters are adjusted to contain the extremal points.

At this stage of the reconstruction, different patch and plane configurations exist in parallel. To select the optimal configuration Utility Theory is applied. The Utility Function used consists of two parts. One part quantifies the reliability of the patch hypothesis (Scholze et al., 2001), whereas the second part takes into account the compatibility between different patches. Finally, the patch and plane configuration with Maximum Expected Utility is entered into a preliminary roof model. (A detailed description of plane and patch instantiation will be published elsewhere.)

5 ROOF RECONSTRUCTION

The patch hypothesis generation procedure described in Section 4 is sufficient to reconstruct roofs whose patches lie on planes which contain the seed line. To complete the reconstruction of more complicated roofs, a semantic interpretation of already reconstructed patches is performed.

5.1 Semantic Model Completion

Geometry based reconstruction results in a preliminary roof model where patches might be missing. To locate the missing patches the outline of the reconstruction is scrutinized. The key idea is now to identify patch segments on the outline which actually should correspond to an internal boundary of the roof – that is, a concave or convex joint of roof patches. If such segments could be identified (if present at all), these in turn form a set of seed segments which are fed into the reconstruction algorithm again (cf. Figure 1(d)).

More precisely, in order to assign the semantic labels from Equation (1) to the patch segments, the unary and binary attribute vectors are computed for each patch individually in a first step. This involves a classification of the measurements according to the attribute distributions of the test dataset. For the unary attributes, the classes correspond to the semantic labels given in Equation (1). Since the binary attributes describe relations between pairs of line segments, their classes are given by compatible label combinations (e.g. \((\text{ridge};\text{gable})\)). A nonparametric classification technique, namely Linear Discriminant Analysis (LDA) is used to classify the unknown observations (Fisher, 1936). During LDA, the input data (unary and binary attributes respectively) are transformed as to obtain maximal separation between the classes. Optimal classification results would be achieved, when the decision boundaries perfectly separate the classes. However, the clusters corresponding to the different semantic labels in this space are overlapping for both unary and binary attributes. This makes an unambiguous label assignment impossible. This is overcome by an iterative procedure which determines the labels for all segments in a patch in such a way, that the entire label assignment (per patch) has maximal probability, exploiting unary and binary attributes simultaneously (Christmas et al., 1995).

5.2 Semantic Model Refinement

It is desirable to have a topologically correct, that is, a point-wise connected reconstruction of the roof. Therefore, corner points of patches with compatible semantic labels are forced to coincide in a final gluing step. Thus, in this step, topological correctness is preferred over geometric precision. Due to coincidence constraints introduced into the model, the gluing step does not open gaps at other locations in the model.

6 RESULTS AND CONCLUSION

The presented results are obtained using a state-of-the-art dataset, produced by Eurosense Belfotop n.v. The image characteristics are: 1:4000 image scale and geometrically accurate film scanning with \(20 \times 20 \ \mu m^2\) pixel size corresponding to \(8 \times 8 \ cm^2\) on the ground. The 3D line segments are obtained using three overlapping views. The precise sensor orientation is known. To emphasise the quality of the reconstructed roof geometry no texture mapping is applied.

Figure 8(a) shows a reconstruction result for the upper left building in Figure 5(c), which was completely reconstructed from its seed (here: ridge) line by one pass of the reconstruction algorithm. The reconstruction results are detailed and topologically correct. For instance the small difference in the slope of the two patches on the right side of the roof has been correctly detected. Figure 8(b) shows the reconstruction of the roof of the corner building in Figure 5(c). The triangular patches on the front side do not lie in a plane given by the seed line. However, driven by the semantic labels attributed to the partially reconstructed roof, the missing patches could be successfully found.

Concluding, in this paper we have presented a novel approach for the probabilistic modelling of building roofs. The proposed theoretical foundation and technical implementation leads to stable and reliable results. Future work will focus on exploiting the knowledge available in the form of test datasets to a broader extent, with the goal to
initialize roof models with even less evidence in 3D space. Another research direction, which is pursued at the moment, is oriented toward the optimization of the geometric accuracy of the models using back-projection into the input images.

Acknowledgement: The authors gracefully acknowledge support of the ETH project AMOBE II and the European IST project CogViSys.

REFERENCES


Figure 8: (a,b): Three-dimensional view of reconstructed building roofs. The non-trivial roof structures are captured to their full extent. Thin lines show the input 3D segments.