# 3D roof details by 3D aerial vision

Philipp Meixner, Franz Leberl, Mathieu Brédif

# 3D Roof Details by 3D Aerial Vision

Philipp Meixner
Institute for Computer Graphics and Vision
Graz University of Technology
meixner@icg.tugraz.at

Franz Leberl
Institute for Computer Graphics and Vision
Graz University of Technology
leberl@icg.tugraz.at

Mathieu Brédif
MATIS Laboratory
IGN, Paris, France
Mathieu.bredif@ign.fr

## Abstract

*Virtual cities in a 3D GIS will not be complete without the details of roofs with chimneys, dormer windows, skylights. We contribute a method of extracting such details from digital large format aerial photography. The dense point clouds produced by a photogrammetric process produce major and minor plane surfaces of a roof and enter into a classification of its superstructures. In real city data with a total of 1312 superstructures the proposed automated process finds 1024, thus achieving success at a rate of 78%.*

## 1. Introduction

We observe the emergence of virtual cities as the 2D geographic information system GIS morphs into a 3D GIS. This is being greatly helped along by innovations in sensors, computing, storage, software so that data can be created at low cost per building and per unit area. Innovations in the applications infrastructure with the Internet and smart telephony increase the applicability of 3D detail.

Advances in digital aerial cameras are the source of low-cost very dense point clouds that are necessary to generate 3D building models. "Very dense" refers to clouds of points at a density of 1 to 2 pixels in the source imagery. This has become possible through new algorithms in combination with multi-view geometry and high overlap imagery. In the process, 3D city modeling has become a lively research area within the photogrammetric community [5, 6, 1]. Whereas in the past city models often have been built for visualization purposes new applications have information needs beyond visual characteristics. This implies that the geometric information be augmented by semantics and topology to help with thematic queries, analysis tasks, automatic integration, validity checking, or spatial data mining.

Buildings are the most relevant elements of a 3D GIS and are in the focus of considerable research efforts. The detection of buildings and of their footprints, even mapping of the main roof faces, has become possible by automated means [5].

The required level of detail (LOD) depends on the application. This may range from the Lego-type parallelepipeds denoted as LOD-1 via models with general roof shapes as LOD-2 to building models with windows and roof details as LOD-3 and on to LOD-4 to include the building interior [8] (see figure 1). Figure 2 shows a French LOD-2 city model augmented with photo texture.

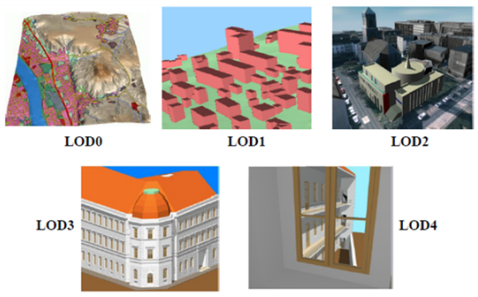


Figure 1. CityGML levels of details [3]

Figure 2. 3D model of a detail of the 12th district of Paris, GSD 15cm [Courtesy: Laboratoire MATIS, IGN France]

In its most sophisticated form, each building, tree, street detail, bridge and water body is modeled in three dimensions, details such as windows, doors, façade elements, sidewalks, manholes, parking meters, suspended wires,

street signs etc. exist as separate objects. For LOD- 2 to LOD-4 the roof and its extrusions and intrusions are of interest.

This paper contributes a method to map a roof's details as they extrude from, or sometimes intrude into, the major roof planes. The paper also addresses skylights as an additional element for the degree of development of a roof. The novel method uses high-resolution vertical aerial photography to segment and classify roof details, so that one obtains the type roof architecture, the number and location of chimneys, dormer windows and skylights. We segment the roofs into major roof planes, minor roof planes and planes associated with superstructures. The major roof planes provide information about the architectural roof style; another application would be the examination of the roof's suitability for the installation of solar panels. Of particular interest in our case are superstructures. We divide these into dormer windows, chimneys and other structures. Especially the existence of dormers and of skylights is of great interest for the level of use of a building's attic.

Experiments use aerial photography of an urban core with 186 buildings and at a GSD of 10cm. We show that our initial approach maps dormers, chimneys and skylights at a rate between 76% and 81%. We see significant avenues for improvement in the source data as well as algorithms to increase that rate of success in the future.

## 2. Approach

In order to analyze roofs, one will first want to separate the given input data per building. The aerial imagery with its derived 3D point clouds and classified land surface elements is being merged with the cadastral property data. The result consists of each individual property and the area occupied by a building. Its outline is from the image classification and the 3D point clouds. From the building one proceeds to the façades: building footprints become façade baselines and they in turn serve to refine the building's roof outlines. This has been developed by [10] and is used for this work.

Roofs and their superstructures now need to get mapped using the building's point cloud and image segments. The process consists of five steps (Figure 3). First is the smoothing of the noisy digital surface model (DSM) using total generalized variation TGV presented in [12]. Second is the extraction of all plane roof segments using a method introduced in [14] and resembling the RANSAC algorithm for plane detection, however with the consideration of neighborhoods. Third, each plane surface patch gets labeled as a major, minor or superstructure plane. This initially uses the DSM point cloud and then refines the segmentation by incorporating the RGB-image data. In a region growing approach, plane segments get merged into larger surfaces, or are set aside for further analysis.

A fourth process element is the search for skylights in the major and minor roof planes using a template matching method in the form of fast directional chamfer matching introduced by [9].

The fifth step is the interpretation of the data and assignment of labels to the isolated data elements and collected list of features. The roof gets an overall attribution to a specific architectural style, the "left-over" data are the basis for a classification into dormers, chimneys and sky lights.


Figure 3. Sequence of the five major processing steps

### 2.1. DSM Smoothing

Photogrammetrically measured elevation data (range data) are noisy. At a pixel to multi-pixel level, this will obstruct any search for plane surface patches. For this reason the 3D point clouds need smoothing. The filtering of elevation measurements is a traditional function of photogrammetric work in the form of an approximation and interpolation of terrain elevation grids from manually measured surface points. In the current application, we have very dense data and need a filter that preserves sharp surface discontinuities and small structures while producing smoothness within the roof. A recent innovation is "total generalized variation" TGV introduced by [12]. It satisfies the requirements of the roof surface analysis and is available in a GPU-implementation for very fast throughput.

Figure 4 is taken from our experiments. It illustrates that outliers and noise in the elevation data are minimized, yet the global structure of a roof is maintained.

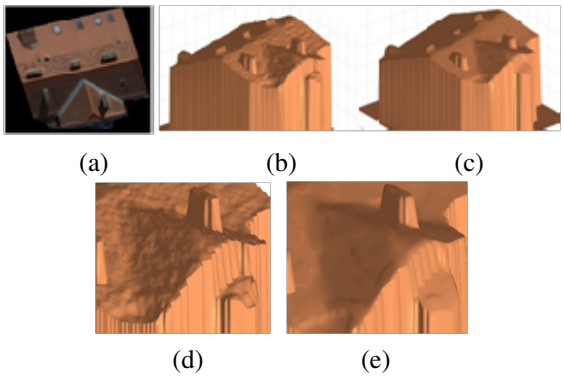


Figure 4. (a) RGB image of single building of the test dataset; (b) Raw range image prior to smoothing; (c) Range image using TGV; (d) Detail of raw range image; (e) Smoothed detail.

## 2.2. Plane Detection

Roof planes get detected in the smoothed point cloud. Of the many different methods available for the task, RANSAC plane detection has become a standard, in particular in the LiDAR literature [2, 13, 15]. Its principle is well explained in [11, 4]. A limitation of RANSAC results from its disregard for neighborhoods; it treats all points the same without considering that a point might be near another. For that reason [14] has introduced a variation on RANSAC that incorporates neighborhood relations. It starts by random sampling to generate model hypotheses. Minimal sets are constructed in a way that neighboring points are selected with higher probability. That means that if a point $x_i$ has already been selected, that $x_j$ has the following probability of being drawn:

$$P(x_i|x_j) = \begin{cases} \frac{1}{Z} exp - \frac{\|x_j - x_i\|^2}{\sigma^2} & \text{if } x_j \neq x_i \\ 0 & \text{if } x_j = x_i \end{cases} \quad (1)$$

where $Z$ is the normalization constant, $\sigma$ is a heuristically chosen constant, $x_i$ and $x_j$ are single observations. After the creation of all hypotheses, a preference set of preferred hypotheses is created for each point. Points belonging to the same structure have a similar preference set, thus they are close in the conceptual space. To find the models the method uses an agglomerative clustering procedure, where at each step the two clusters with the minimum pairwise distance are merged. This distance reaches from 0 (identical sets) to 1 and just elements are linked together whose preference sets overlap. Figure 5 shows the result of the plane detection for one building of our test dataset. The comparison of a plane with the point cloud produces a measure of accuracy. We find that the test data have elevation noise of $\pm$12cm; this is to be related to the GSD at 10 cm.

## 2.3. Classification

The roof plane detection is the basis for a classification into major planes, minor planes and superstructure planes. We further divide the category superstructures into chimneys, dormer windows and other structures. Our classification workflow, illustrated in figure 6, consists of an initial and a refined classification. The initial classification analyzes every plane segment to determine if it is linked to any other segment to build a larger "region". Then the size of the resulting region or assembly of plane segments gets considered.

Depending on the size of the regions with respect to the overall size of the roof we assign each region to an appropriate category.

A secondary analysis addresses the smaller planes and superstructures. Depending on their height values with respect to the neighboring pixels these regions are classified

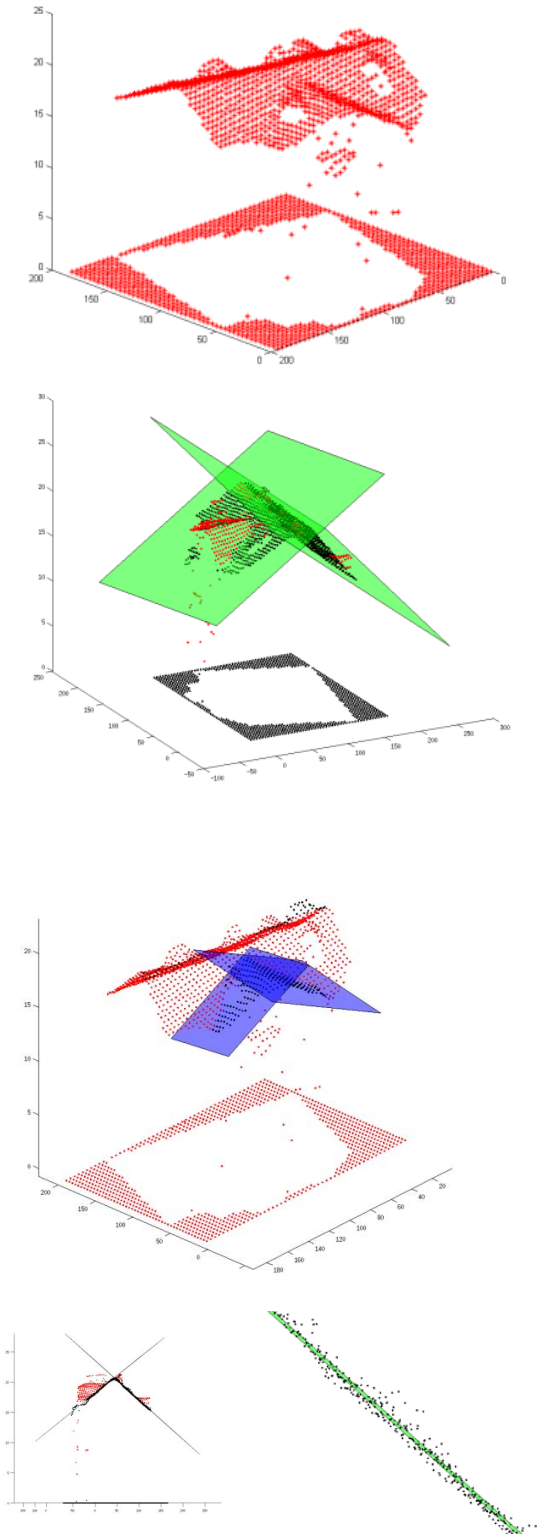

Figure 5. Axonometric view of (a) point cloud of building roof, points in red; (b) point cloud and detected major planes in green. Note the intersection line and points belonging to the planes in black (c) point cloud and detected minor planes in blue. Note intersection line and points belonging to the planes in black; (d) Side view of point cloud and major planes; at right is an enlarged detail showing standard deviation between points and plane at $\pm$12cm.

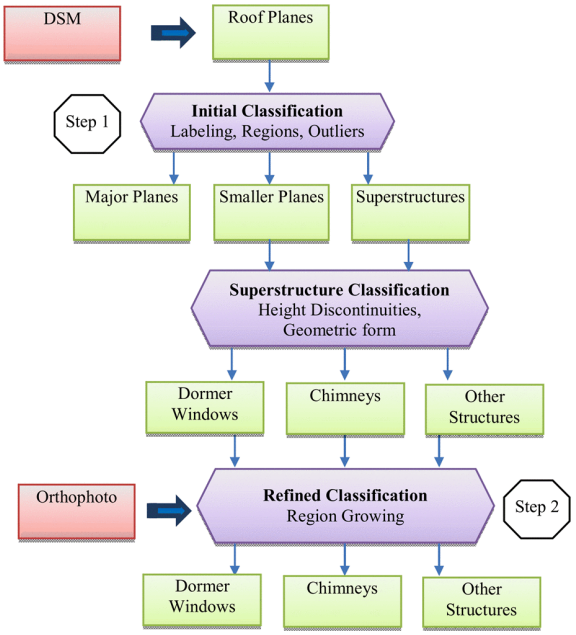


Figure 6. Proposed classification processing steps

as part of a smaller plane segment (regions without height continuities at the borders), or of a superstructure or are eliminated entirely. To achieve meaningful results we distinguish between height discontinuities at the edges of the roof and within the roof by using the information about a building from the building classification.

The refined classification for the superstructures uses the DSM and the RGB-values for a more accurate and robust segmentation of the roof parts as proposed by [7]. This is region growing with seed pixels for every region from the initial segmentation. The seed pixel locations are the centroid of each segment. The region is iteratively grown by comparing all unallocated neighboring pixels to the region. The difference between a pixel's intensity value and the region's mean is used as a measure of similarity.

The pixel with the smallest difference measured this way is allocated to the region. This process stops when the intensity difference between region mean and new pixel becomes larger than a certain threshold.

The classification of the single regions is performed using area, geometric form and height discontinuities at the region's borders.

For the example of chimneys one postulates that on all 4 edges of the region are height discontinuities and that the maximum height is not lower than the height of the roof's ridge. Dormer windows have height discontinuities on at least three edges. Moreover their area is larger than the chimney's and the geometric form is more "quadratic" than elongated. Chimneys have usually smaller and narrower forms with a maximum width of 50cm. In a last step we approximate the single chimneys by best fitting rectangles.

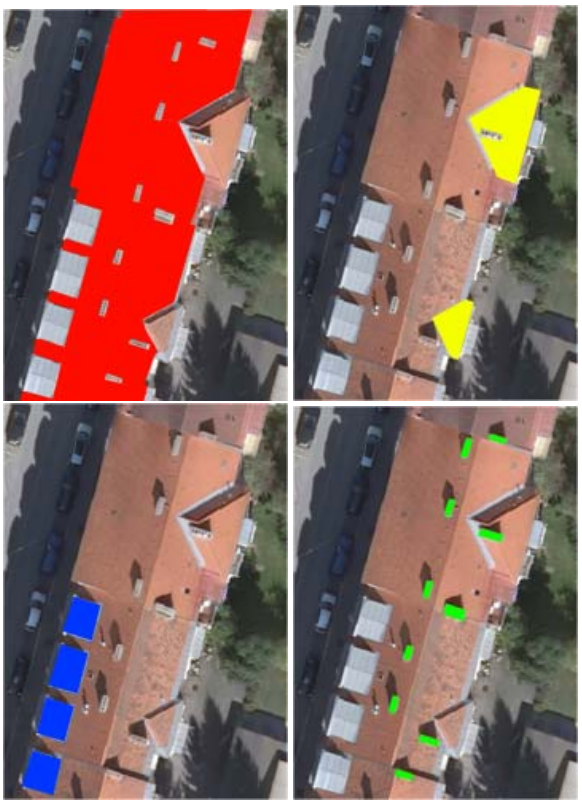

Figure 7. Classification of two test buildings; (a) major roof planes (in red); (b) smaller roof planes (in yellow); (c) dormer windows (in blue); (d) chimneys (in green);

Figure 7 shows the different generalized segmentation results for two example buildings.

## 2.4. Template Matching

A major question of roof analysis addresses the use of the attic. The existence of dormer windows strongly indicates that the attic has been extended. However, the absence of dormers does not indicate that the attic has not been extended. In lieu of dormers there might be skylights.

Skylights can be detected by matching a rectangular skylight template with an edge image computed from the aerial photography over the area of the major roof planes. The match is scale and rotation variant. Fast directional chamfer matching as proposed in [9] improves the traditional alignment between two edge images by incorporation of edge information into the matching costs. This improves the robustness of chamfer matching in the presence of background clutter. This method augments the chamfer distance with an additional cost for orientation mismatch which is given by the average difference in orientations between template edges and their nearest edge points in the query image. Instead of an explicit formulation of the orientation mismatch, the chamfer distance is generalized to points in $R^3$ for matching directional edge pixels. The directional cham-

fer distance score is given in [9] by

$$d_{DCM}(U,V) = \frac{1}{n} \sum_{u_i \in V} \min_{v_j \in V} |u_i - v_j| + \lambda |\phi(u_i) - \phi(v_j)|$$

$\lambda$ ... weighting factor between location and orientation terms;
$u_i$ ... set of template image edge map
$u_i$ ... set of query image edge map
$\phi()$ ... direction term

Figure 8 shows results for the detection of skylights on building roofs.

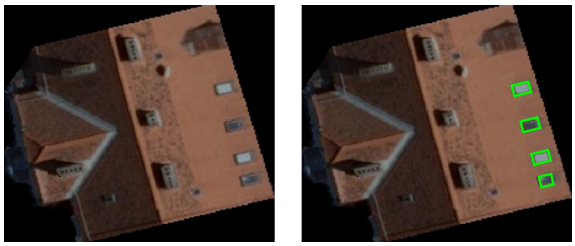

Figure 8. Left: original roof image; Right: detected skylights (marked in green) on a building roof.

### 2.5. Roof Meta-Interpretation

A meta-interpretation of the roof concerns its extended or not. This is based on the number and location of chimneys, the use of skylights and dormers. These may be of interest as escape routes in the event of a disaster.

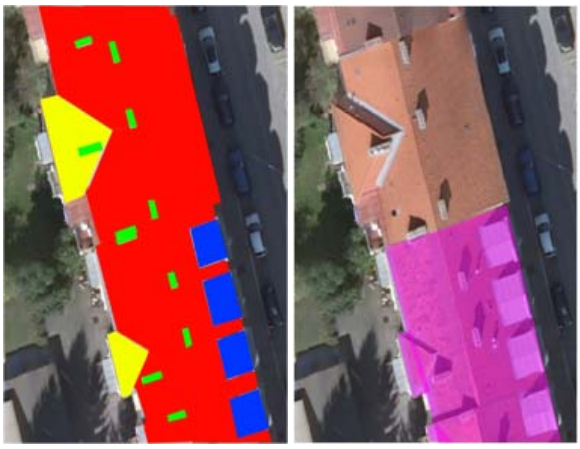

Figure 9. Left: original roof image; Right: detected skylights (marked in green) on a building roof.

Future work might address the existence of roof terraces, flat roofs with their elevator shafts and air conditioning compressors to interpret the uses of the building. This is interesting for urban planning, real estate development and transactions or it could be used for value assessments for property tax assessments. Figure 9 illustrates a segmented building roof with the conclusion whether the attic is living space or not.

## 3. Experimental Results

The test area covers 400m x 400m near a European urban core with 186 different buildings. The vertical aerial photography was taken with a GSD of 10 cm, 80% forward-, 60% side-overlap.

Ground truth was obtained by a manual interpretation of the 69 extended attics in the 186 building (see table 1). The extended attics are characterized in 20 instances by dormer windows, in 17 cases they have just skylights and 32 have both dormers as well as skylights. Additional ground truth was obtained manually for all major and minor planes, dormer windows, chimneys and skylights (see table 2).

| | Dormer windows | skylights | both | total |
|---|---|---|---|---|
| number | 20 | 17 | 32 | 69 |

Table 1. Ground truth for 69 extended attics in 186 buildings. Some attics have only dormer windows or skylights, and some have both.

| | Major planes | Smaller planes | Dormer windows | Chimneys | Skylights |
|---|---|---|---|---|---|
| number | 446 | 198 | 186 | 552 | 574 |

Table 2. Ground truth for the major and smaller planes, dormer windows, chimneys and skylights in the 186 buildings of the test area.

Part of the experimental work was to understand the effect of the smoothing operation. It not only eliminates outliers, but improves the throughput of plane detections by a factor two since the smoothness of the point cloud improves the size of the plane segments.

A differentiation between major and minor or smaller planes is heuristic and a function of architectural customs in an area. The threshold can be assigned as a function of every building's footprint and the size of the individual plane segments. But some smaller planes will get erroneously assigned to the major plane class and vice versa.

Table 3 summarizes the achieved results. Detected were 446 major and 198 smaller planes, 186 dormer windows, 552 chimneys and 574 skylights. Of the total of 644 planes, 579 were detected, thus achieving a detection rate of 90%. It was somewhat more successful to detect major planes than minor ones, with success rates at 92% and 86%.

Current limitations of the approach were found in cases where plane segments have very similar pitches and are thus merged erroneously into one plane when they should be kept separate. Another limitation exists with curved roof surfaces that exist in some regions extensively, such as in France, and they may also be found in modern buildings. In these cases the attempt at fitting roof planes produces many

| | Major planes | Smaller planes | Dormer windows | Chimneys | Skylights |
|---|---|---|---|---|---|
| Total number | 446 | 198 | 186 | 552 | 574 |
| Detected segments | 409 | 170 | 151 | 420 | 453 |
| Detection rate (%) | 92 | 86 | 81 | 76 | 79 |

Table 3. Experimental results for the building roof segmentation

small segments and a failure of the approach. A check will be needed that addresses curved roofs.

Of 738 superstructures, 571 were detected. This corresponds to a detection rate of 77%. The dormer window detection rate was at 81%, the chimneys' at 76%. Applying these superstructures to an interpretation of extended attics delivers success at a rate of 82%.

Misclassifications of superstructures occurred in the case of very complex roof shapes with non-allocatable roof structures, particularly when surfaces were curved. Since our approach had not been considering the existence of roof gardens, it failed in those cases.

It is thus obvious that further work is needed to address curved roofs and roof terraces.

## 4. Conclusion

We contribute a novel use of vertical aerial images to describe and characterize building roofs as a new content in a 3D GIS. Since the aerial photographs offer a very high overlap, they produce very dense 3D point clouds. These in turn support, jointly with the RGB-content of the images, the mapping of major roof planes and of roof superstructures. The rate of correctly detecting roofs, dormer windows, chimneys and skylights are at 76% to 81%. This in turns permit one to assign to a "normal" building a label of "extended attic" and assign specific architectural roof styles.

LiDAR today is the workhorse for the production of 3D point clouds, and research in building reconstruction is often considering LiDAR as the major data source. However, a search for LiDAR-literature addressing roof superstructures came up empty. We will therefore argue that vertical aerial photography is a valid data source for mapping building roofs, and that the same cannot be said at this time for LiDAR.

Our future work therefore will continue with vertical aerial imagery. Improvements of algorithms will have to address curved roofs and roof terraces, and will have to broaden the algorithmic basis to include more image- derived features to augment the point clouds.

A differentiation between major and minor or smaller planes is heuristic and a function of architectural customs in an area which makes the algorithm less robust. In order to solve this problem in future work we want to enter this information as prior knowledge of an area.

Very importantly, a widening of the range of architectural styles will be needed to go from coastal resort environments via historical small towns and mountainous urban areas to modern urban zones with skyscrapers and industrial facilities in various regions of the World.

## References


[1] Yushin Ahn. *Object space matching and reconstruction using multiple images*. PhD thesis, The Ohio State University, 2008.

[2] Claus Brenner, Norbert Haala, and Dieter Fritsch. Towards Fully Automated 3 D City Model Generation. 2001.

[3] Mathieu Brédif. *Modélisation 3D de bâtiments : reconstruction automatique de superstructures de toits et recalage cinétique de toits polyédriques prenant en compte la topologie*. thesis, Paris, Télécom ParisTech, January 2010.

[4] Martin A. Fischler and Robert C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM*, 24(6):381–395, June 1981.

[5] Norbert Haala and Martin Kada. An update on automatic 3d building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(6):570–580, November 2010.

[6] KyoHyouk Kim and Jie Shan. Building roof modeling from airborne laser scanning data based on level set approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4):484–497, July 2011.

[7] Visa Koivunen and Matti Pietikaeinen. Combined edge- and region-based method for range image segmentation. In *Intelligent Robots and Computer Vision IX: Algorithms and Techniques*, volume 1381, pages 501–513. International Society for Optics and Photonics, February 1991.

[8] Thomas H. Kolbe, Claus Nagel, and Alexandra Stadler. CityGML – OGC Standard for Photogrammetry ? 2009.

[9] M. Liu, O. Tuzel, A. Veeraraghavan, and R. Chellappa. Fast directional chamfer matching. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1696–1703, June 2010.

[10] Philipp Meixner and Franz Leberl. Characterizing Building Façades from Vertical Aerial Images. *The international archives of photogrammetry, remote sensing and spatial information sciences*, 38(3B):98–103, 2010.

[11] Roy Mullen. *Manual of Photogrammetry. Fifth Edition*. ASPRS, Bethesda, Md, 5 edition edition, 2004.

[12] Thomas Pock, Lukas Zebedin, and Horst Bischof. TGV-Fusion. In Cristian S. Calude, Grzegorz Rozenberg, and Arto Salomaa, editors, *Rainbow of Computer Science: Dedicated to Hermann Maurer on the Occasion of His 70th Birthday*, Lecture Notes in Computer Science, pages 245–258. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[13] Ruwen Schnabel, Roland Wahl, and Reinhard Klein. Efficient RANSAC for point-cloud shape detection. In *Computer graphics forum*, volume 26, pages 214–226. Wiley Online Library, 2007.

[14] Roberto Toldo and Andrea Fusiello. Robust Multiple Structures Estimation with J-Linkage. In David Forsyth, Philip Torr, and Andrew Zisserman, editors, *Computer Vision – ECCV 2008*, Lecture Notes in Computer Science, pages 537–547. Springer Berlin Heidelberg, 2008.

[15] Michael Ying Yang and Wolfgang Förstner. Plane Detection in Point Cloud Data. 2010.