ISPRS Journal of Photogrammetry and Remote Sensing 65 (2010) 570–580

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

An update on automatic 3D building reconstruction

Norbert Haala *, Martin Kada

Institute for Photogrammetry, University of Stuttgart, 70174 Stuttgart, Germany

ARTICLE INFO

Article history:
Received 29 January 2010
Received in revised form
6 September 2010
Accepted 14 September 2010
Available online 20 October 2010

Keywords:
Reconstruction
Three-dimensional
Building
Urban
Automation

ABSTRACT

The development of tools for the generation of 3D city models started almost two decades ago. From the beginning, fully automatic reconstruction systems were envisioned to fulfil the need for efficient data collection. However, research on automatic city modelling is still a very active area. The paper will review a number of current approaches in order to comprehensively elaborate the state of the art of reconstruction methods and their respective principles. Originally, automatic city modelling only aimed at polyhedral building objects, which mainly reflects the respective roof shapes and building footprints. For this purpose, airborne images or laser scans are used. In addition to these developments, the paper will also review current approaches for the generation of more detailed facade geometries from terrestrial data collection.

1. Introduction

The automatic reconstruction of urban 3D models became an important part of photogrammetric research almost two decades ago (e.g. Grün et al., 1995, 1997). Since these early beginnings numerous research papers on different reconstruction methods were published with quite a number of approaches emerging to commercial services and software (Brenner, 2005). As e.g. documented by the EuroSDR Building Extraction project (Kaartinen and Hyyppä, 2006), which aimed at a comprehensive test of commercial products and services, areas covering sets of 3D building models are commonly collected from photogrammetric 3D measurement using airborne stereo imagery or LiDAR. Some systems additionally support the extraction of building outlines as 2D map data. Available methods usually record the roof shapes and building footprints at the required detail and accuracy and then use this information to generate a geometric representation of the building in a subsequent step.

However, as Habib et al. (2010) point out: "digital building model generation of complex structures still remains to be a challenging issue". Since fully automatic image understanding is very hard to solve, semi-automatic components are usually required to at least support the recognition of very complex buildings by a human operator. The difficulties of aerial image interpretation also motivated the increasing use of 3D point clouds from laser altimetry as an alternative data source. By these means, the interpretation task can be restricted to explicit geometric information, which helps to facilitate the development of automatic tools for 3D building reconstruction. In the past, the success of approaches based on elevation data was also supported by the continuously increasing density and accuracy of point clouds as a result of the fast evolution in LiDAR technology. Meanwhile, suitable image matching software can alternatively generate 3D point clouds and 2.5D raster representations at an accuracy, reliability and amount of detail, which was only feasible by LiDAR measurements. This is especially true if high quality imagery from digital airborne cameras is used, which usually provides good radiometric quality and high redundancy due to large image overlap (Haala, 2009; Hirschmüller and Bucher, 2010). The 3D city model of Las Vegas in Figs. 1 and 2 is e.g. automatically generated from high-resolution images. As can be seen from Fig. 2, the result is a detailed 3D surface mesh, which can be textured with the images accordingly (Fig. 1).

If image based surface reconstruction is applied, both geometric and radiometric information is available from one sensor. The integration of these two complementary data types can then be used for the extraction of three dimensional features or for large scale classification as pre-processing for 3D urban modelling (Vosselman, 2002; Zebedin et al., 2006). The joint availability and combination of geometric and radiometric information are also required for visualization applications. There the building geometry as provided from dense elevation data is enriched by surface texture from aerial imagery.

The interactive visualizations of 3D city models were opened to a general public mainly by applications such as Google Earth and Bing Maps (Leberl et al., 2009). Such visualizations at large and medium scale are feasible by relatively coarse building models, which are usually limited to roof structures and planar facades.

* Corresponding author.
E-mail addresses: norbert.haala@ifp.uni-stuttgart.de (N. Haala), martin.kada@ifp.uni-stuttgart.de (M. Kada).

0924-2716/$ – see front matter
doi:10.1016/j.isprsjprs.2010.09.006

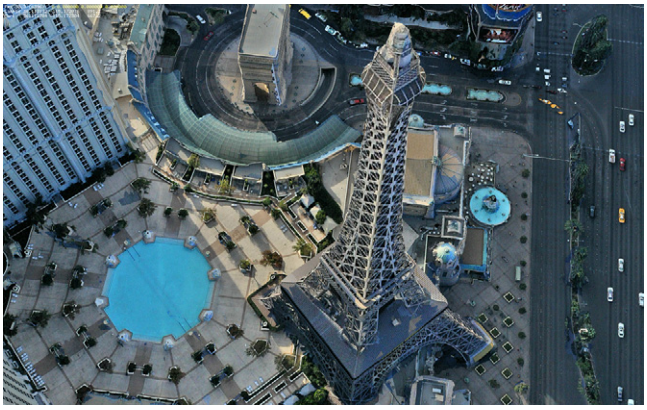

Fig. 1. 3D city model of Las Vegas, USA (courtesy of C3 technologies).

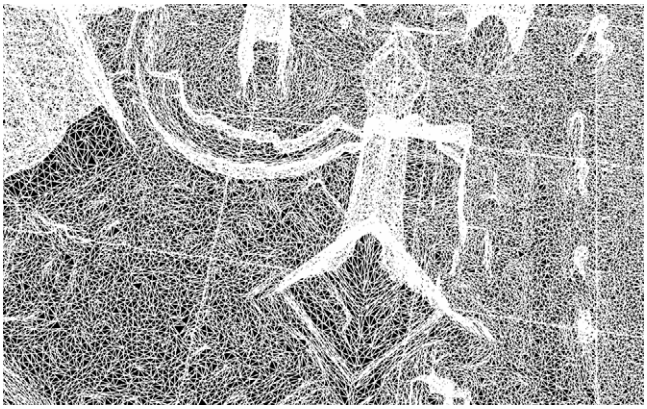

Fig. 2. Wireframe version of the 3D city model of Las Vegas, USA (courtesy of C3 technologies).

Thus, polyhedral models are sufficient, which represent the building shapes by rather simple planar surfaces. Since a number of operational tools have been developed for the automatic reconstruction of such polyhedral models from airborne data, large areas can be captured by fully automatic city modelling systems while interactive components can usually be limited to relatively complex landmark buildings or highly detailed reconstructions. Within Section 2, these developments on the automatic reconstruction of building shapes from elevation data – either provided from airborne LiDAR or automatic matching of highly overlapping imagery – are grouped according to the underlying principles of the respective approaches and discussed in detail.

While the outcome of these approaches can for example be used very well for visualizations of large areas from elevated viewpoints, an increasing demand for ground based presentations is currently evolving. This presumes very detailed geometric reconstructions including the building facades, which are frequently made available from terrestrial data collection. In this context, terrestrial images are e.g. used by suitable texturing methods to improve the visual appearance of the respective building facades. However, as discussed in Section 3, a number of approaches also aim at an explicit geometric modelling of features like doors or windows in order to enrich the respective facades. In addition to pleasing visualizations, such representations also allow for "location-aware" applications of city models. Such more complex search and navigation applications within urban environments require fully interpreted urban scenes with knowledge of doors and windows, but also roads, sidewalks, trees or parking spaces. Thus, technologies and algorithms are required to automatically describe urban areas in much higher detail, maybe even the building's interior. This will be discussed in the final part of the paper.

2. Roof shapes from elevation data

In this article, we want to focus primarily on the developments of the last couple of years with the purpose to close the gap between today and the thorough overviews given by Brenner (2005) and Baltsavias (2004) and the EuroSDR project on building extraction (Kaartinen and Hyyppä, 2006). A great number of approaches have since then been presented, which will briefly be described and put in context to one another in order to show both the past and present trends in building roof reconstruction from elevation data. As mentioned in the introduction, this type of input data can originate from various sources like LiDAR or image matching. Also, it is assumed that footprints are available or can be automatically extracted beforehand. Footprints have recently been derived from digital elevation models (DEM) e.g. based on marked point processes (Ortner et al., 2007), by combining them with aerial images (Li and Wu, 2008) or high-resolution satellite images (Sohn and Dowman, 2007). They are then delineated by a graph-based point reduction of the segmented building points (Neidhart and Sester, 2008), hierarchical least squares with perpendicularity constraints (Sampath and Shan, 2007). An evaluation of different methods on the detection of building footprints for the update of 2D databases is given by Champion (2009). Especially in Western European countries, footprints are available nationwide as cadastral data and the governmental authorities more and more request from the data providers to deliver 3D building models that are consistent with them.

One has to keep in mind that the point density greatly increased during that time period, but there is still a huge difference between what data is available for large area production purposes and what current sensor technology is able to deliver. While no one expects anything better to result from low density data other than rough and generalized roof shapes, the expectations are increasing along with the quality of the input data. Also the architectural style varies among rural, suburban, and inner city areas and geographical regions. Therefore, many different procedural methods are still being proposed, motivated by previous work, their planned application area, quality of input data or just by the urge to follow new ideas. But it seems that at least for high density data, the reconstruction of building roofs converges towards a uniform process that is based on a segmentation process of the elevation data (see Section 2.2).

The remainder follows the developments from simple parametric buildings, their combination to more complex ones (Section 2.1) to the construction of general roof structures that are based on point cloud segmentation (Section 2.2). Also an alternative reconstruction approach is discussed (Section 2.3), which does not assemble or construct building models in the traditional sense, but rather simplifies the meshed raw data until it suffices certain geometric and semantic criteria. As we will see from the reported results, detailed roof shapes that are close to reality are already within grasp.

2.1. Reconstruction with parametric shapes

A great number of buildings in rural and suburban areas are rather simple. They can be approximated by rectangular footprints and parameterized standard roof shapes. Most common are the saddleback roof, sometimes with hip ends on one or either sides, pent, flat, tent, and mansard roof. A description of common roof shapes are e.g. given in Milde and Brenner (2009) and Kada (2009). Our experience is that if roof details like dormers and chimneys are not required, these buildings can be automatically and reliably recognized and their parameters exactly determined even from low density data since the early works on building reconstruction; e.g. with the approach described in Brenner and Haala (1998).

From completed large-area reconstruction projects of Western European cities, we conclude that approximately 40%–50% of the buildings in this particular geographic area fall into that category. It must also be noted, that trueness to given footprints can easily be achieved for these buildings by expanding the parametric models to the objects' bounding rectangles and intersecting them with the extruded footprints.

However, these basic parametric shapes are not expressive enough to generate precise building models with arbitrarily shaped footprints. To overcome this limitation, Taillandier (2005) reconstructs buildings by extruding given footprints to a uniform building eaves height. It is then hypothesized that sloped roof faces pass through every extruded line segment, which an exhaustive search over all possible plane intersections verifies. Incorrect hypotheses are discarded and gables assumed instead. From the DEM, the precise eaves height and all slopes are determined by a least squares estimation and a polyhedral building model is constructed from the intersection of the verified planes. The approach is capable of reconstructing building roof structures with preceding wall dormers and diagonal roofs if they are apparent in the footprint. A visual inspection of approximately 600 buildings reconstructed during a test study of the city centre of Amien in France revealed that around 90% of the roof shapes could be correctly reconstructed with this approach.

Because the slopes of all hypothesized roof faces are initially fixed at a 45°angle, the intersection lines are sometimes rather far off from their true location. This obviously results in errors verifying the slope hypotheses. Durupt and Taillandier (2006) therefore extract the plane parameters before the exhaustive search using a robust random sampling consensus (RANSAC) (Fischler and Bolles, 1981) estimation. This modification is reported to improve the acceptance rate by another 4%.

Regardless of footprint complexity, both approaches perform well on symmetric roof shapes with global eaves height and central ridge, but obviously fail for other shapes, especially in the presence of height discontinuities. So some approaches try to circumvent this problem by hypothesizing these height discontinuities from the footprint and decomposing it accordingly. Thus, existing algorithms can be used to determine the respective roof shapes for the resulting decomposition elements, usually called cells if they are non-overlapping. The reconstruction of building roofs with complex footprints are therewith reduced to simple subtasks for which solutions already exist. After the simple 3D primitives have been constructed for all elements, they are then glued together to form a complex building model.

The reconstruction approach presented in Kada and McKinley (2009) and Kada (2009) takes this idea and decomposes the given building footprints into wings and fits basic roof shapes into these. Special connecting elements are used where the wings meet. Due to the complexity of building footprints and the limited geometric accuracy of cadastral maps, exact decompositions that fit well with the roof shapes are very hard to generate. The spatial partition is therefore not generated by decomposing a footprint itself, but is rather recreated from scratch by decomposing an enlarged block along the approximated linear features of the building. The resulting 2D cells are then compared with the original footprint and the ones with a low overlap are discarded. The remaining cells form together a generalized shape of the building's footprint. Points that are inside a cell are compared to a library of template roof shapes. A cell is given the shape for which most points fit with regard to their normal direction. Fig. 3 shows the 3D city model of Frankfurt am Main as reconstructed by this approach.

The approach has been used in many large-scale projects, e.g. Berlin, Cologne, Frankfurt am Main, Chemnitz (all in Germany), and Gent (Belgium) with approximately 450,000, 280,000, 200,000, 150,000, and 40,000 buildings. The mean point density for all projects was 4 points/m^2, with the exception of Gent, where a denser data set (16 points/m^2) was available. An overall automation rate in the order of 80% has been quantified, which gets slightly worse in inner city areas with complex roof structures and better for residential areas. The rest of the reconstructed buildings fail the automatic or visual inspection and have to be manually corrected. With the provided tools, the manual correction simply comes down to small modifications of the footprint's decomposition or choosing the appropriate roof type.

Fig. 3. 3D city model of Frankfurt am Main, Germany, reconstructed by footprint decomposition and parametric shapes (courtesy of VirtualCitySystems).

If the decomposition of the ground plan generates too many cells, then the roof parts are sometimes divided over two or more cells. This makes it difficult or even impossible to fit the parametric roof shapes. Vallet et al. (2009) therefore do not only split the footprints along their principal directions, but also merge them again if necessary. This process is controlled by an energy function that incorporates the horizontal and vertical gradient of the DEM. In contrast to the former approach, a generalization is not performed. For the resulting cell decomposition, the roof shape is reconstructed by the approach of Durupt and Taillandier (2006).

Lafarge et al. (2008) extract building footprints from DEMs that are generated from high resolution satellite images. The footprints are initially a collection of overlapping rectangles that are obtained via a reversible jump Markov chain Monte Carlo (RJMCMC) sampler (Green, 1995) embedded in a simulated annealing scheme (Kirkpatrick et al., 1983; Cerny, 1985). The primitives are then regularized by fusing neighbouring rectangles and splitting them at detected roof height discontinuities. The result is a spatial partitioning consisting of convex polygons that are perfectly connected and non-overlapping. From their skeletal structure, the ridge line and height are determined and the primitives constructed.

2.2. Reconstruction based on segmentation

If building height discontinuities are not apparent in the footprint, then they cannot get reconstructed by the so far described approaches. Also, the strict parametric roof shapes often do not exactly fit the roof's break lines, e.g. the ridge of a saddleback roof. The detection of such features must therefore be accomplished by other means. The common approach is to segment the point cloud or DEM (Section 2.2.1). Then most approaches perform a neighbourhood analysis of the segments in order to detect roof features like step edges, break lines or even shapes of roof sub-structures (Section 2.2.2). From this information, the building models are constructed (Section 2.2.3).

2.2.1. Segmentation

Vosselman et al. (2004) and Sampath and Shan (2010) e.g. give comprehensive overviews on the topic. Given an initial point cloud, the result of the segmentation process is a partition of the points, where all points in one segment belong to the same shape. Planar, cylindrical, and spherical structures can be recognized by known algorithms. However, reconstruction approaches only make use of segmented planar regions so far. If the points' local planes are needed for the segmentation, they can be approximated beforehand e.g. from their nearest neighbours.

In this article, the discussion of the segmentation is not comprehensive, but rather a compilation of examples showing how the reconstruction approaches that follow in this section handle this particular task. We think that even though the topic is still ongoing research, at least the segmentation of planar roof faces has long matured to a point where the results are of high quality. This is also attributed to the fact that the point densities continue to increase. So our emphasis is not the segmentation itself, but the subsequent steps that construct a 3D building model from it.

Sohn et al. (2008) and Novacheva (2008) use a 3D Hough transform (Vosselman, 1999), the latter based on the perpendicular distance to the origin and the polar coordinates of the plane's normal vector. Park et al. (2006), Dorninger and Pfeifer (2008), and Milde et al. (2008) perform a region growing process (Ballard and Brown, 1982) on surfaces. The seed points can be the ones that best fit their locally estimated plane. Dorninger and Nothegger (2007) determine seed clusters for region growing by a hierarchical clustering of the points' local regression planes in a 4D feature space. Then all points are added to the cluster that are adjacent to the cluster points and are statistically consistent. This process is repeated until no more points can be added to one cluster and no more clusters with minimal size can be generated. Sohn et al. (2008) additionally cluster points with upward pointing normal vectors using a histogram, followed by a region growing to find the connected components. Zhou and Neumann (2009) only work with horizontal planes, which they gain from a distance-based region growing process. Following Ma (2004) and Gurram et al. (2007) apply the mean shift procedure (Fukunaga and Hostetler, 1975) for segmentation, using the normal vectors and point locations as the defining features. An in-depth description on how this non-parametric segmentation can be applied to 3D point clouds is e.g. given in Melzer (2007). Lodha and Arun Kumar (2005) cluster points into appropriate planes using twice the k-means algorithm (MacQueen, 1967) with a distance metric first in the normal space and then in the coordinate space of the points. Outliers are removed with the k-nearest neighbour distance algorithm. Khoshelham (2005) on the other hand does a segmentation of images that are then integrated with height data to determine the 3D information of the segments and also to refine incomplete regions with a split and merge strategy. Drauschke et al. (2009) improve image segmentation by integrating depth information from multi-view analysis.

Once the segmentation is available, rough boundary polygons are sometimes generated from the segments' points. They allow for an efficient neighbourhood analysis as only the boundary points need to be considered. If generalized and adjusted to form a closed set of disjoint polygons, polyhedral building models can also be directly constructed. Because the roof segments are by definition non-vertical, the height component of the segment points can be ignored for the boundary generation. This reduces the problem to the 2D case, for which the algorithms become easier and perform faster. However, the resulting boundary still consists of 3D points, which can be pulled to the segment's plane to have a planar boundary, which might be more convenient for further processing.

Verma et al. (2006) identify segment boundaries with the ball-pivoting algorithm as described in Medeiros et al. (2004). A seed triangle is grown by collecting points that are inside a ball of fixed radius that pivots around the front edges of the current active boundary. The result is a triangulation of the given points. It is closely related to alpha shapes, which is what Park et al. (2006) and Dorninger and Pfeifer (2008) compute using e.g. the algorithm of Edelsbrunner et al. (1983). Zhou and Neumann (2008) generate their boundaries by distributing the points to a uniform grid and then connect the points that are closest to the grid sides that separate empty and non-empty grid cells. If the segment boundaries are generated independently from each other, there will be gaps between neighbouring planes e.g. in areas where the segmentation could not recognize planar segments. In order to avoid this problem, Rottensteiner et al. (2005) construct their boundaries based on Voronoi diagrams. For a refinement of the initial building boundaries, edge extraction methods from aerial images have also been successfully incorporated, e.g. by Habib et al. (2010), Novacheva (2008), and Chen et al. (2005).

The resulting boundaries consist of short line segments that are heavily jagged. This is sufficient for neighbourhood analyses, but cannot be used for modelling purposes. Therefore, some approaches generalize the segment boundaries at this point. Zhou and Neumann (2008) e.g. find the principal directions of the buildings by a statistical analysis of the boundary points' tangent vectors. For this purpose, a histogram is computed in which the peaks represent the principal directions. These directions are slightly adjusted pair-wise in order to possibly form angles that are multiples of 45°. The segments' boundary lines are then aligned with the principal directions and with neighbouring segments. Dorninger and Pfeifer (2008) first simplify the boundaries using an angular criterion and then regularize them to enforce orthogonal or parallelism.

2.2.2. Feature recognition

In image understanding, the topological relations between segmented regions are commonly described in a region adjacency graph. Regions are represented by nodes that are pair-wise connected by an edge if they share a common border. This formulization of the topological relations of the roof parts is very helpful in finding intersection lines, step edges or even the sub-shapes of the roof.

Verma et al. (2006) consider two segments adjacent if at least one pair of line segments from their rough 3D boundaries is close to one another. Milde et al. (2008) have a stricter notion and compute the perpendicular distance between their oriented bounding boxes that are aligned with the intersection line. The segments are considered adjacent if the distance is smaller than a given tolerance value and if sufficient points are found in the vicinity of the intersection line. For Oude Elberink (2009), adjacency relations are dependent on the length of the segments' intersection lines, which are determined from the points of both segments that are within a specific distance. To accommodate to differences in the data set's point density, both the length and the distance are not fixed, but computed from the local median point spacing. In addition to 3D adjacencies, Park et al. (2006) and Oude Elberink (2009) also store 2D adjacencies in the graph if the horizontal distance between two segments is smaller than a given threshold.

The properties of the topological relations are sometimes stored with the edges. Verma et al. (2006) label the edges orthogonal, parallel or without constraint according to the normal directions of the polygons, where the orthogonal label is further differentiated whether the vectors are pointing away from or towards each other. If relations between horizontal and non-horizontal segments are added and the strict orthogonality label is relaxed to a more flexible one, then a broader variety of roof shapes and their interconnections can be described (Oude Elberink, 2009). Common examples are gambrel, mansard, and general L-shaped saddleback

roofs. Milde et al. (2008) also keep track of symmetrical relations between segments, e.g. if they have the same slopes and if their intersection lines are horizontal.

From the segments, their boundaries and topological relations, the intersection lines and step edges of a roof can be determined. While intersection lines occur between neighbouring segments depending on their 3D adjacency, step edges occur at height discontinuities and are therefore defined in 2D. Rottensteiner et al. (2005) detect step edge candidates at segment boundaries by analyzing the profile of the DEM at each vertex position that is orthogonal to the segment. A step edge is confirmed at the building boundary if a point below the roof plane is found or at the building interior if two points, one below the plane and one above the neighbouring plane, are found. The position of the step edge is then at the point of maximum height difference. Sohn et al. (2008) modified Kirsch's compass kernels to work on irregularly distributed point spaces in order to detect step edges at fixed angle intervals. Once a step edge is confirmed, its real orientation is then determined with a Hough transformation.

With the purpose to identify the constituent parts of a complex roof shape, Verma et al. (2006) define simple saddleback and hipped roofs and their combination to L-and U-shaped structures as graphs. The complex roof structure can then be decomposed into these simple shapes by recognizing their topology in the adjacency graph by subgraph matching. Because graphs that represent roof topologies are rather small, a brute-force search is feasible even though the problem of subgraph matching is NP-complete. In order to avoid any ambiguities caused by two or more roof shapes that share a common subgraph, the matching is performed in decreasing order of roof complexity. The topology graph of a hipped roof e.g. has a saddleback roof as a subgraph, which concludes that a search must look for hipped roofs before saddleback roofs. Verma et al. (2006) applied their approach on a data set with a mean point density of 9 points/m^2. Most of the sub-shapes of the 248 buildings are reported as being correctly recognized.

In Oude Elberink (2009), the ambiguities are solved by a statistical analysis that determines the roof shapes that are most likely. A subgraph matching result is either complete if all nodes and edges of the roof graph are associated with subgraphs or incomplete otherwise. Oude Elberink and Vosselman (2009) analyze the reasons for incomplete matches and discuss how they can be solved. In 39% of the cases, the authors found the reason to be a missing segment, which would complete the shape. It is suggested that these segments could be automatically determined depending on a best match strategy. The authors show their reconstruction results for residential areas of Dutch cities, where the architectural style ranges from villas to apartment houses. Although 20% of the 728 buildings have incomplete matches, only 8% of all buildings fail to fit to initial laser points.

2.2.3. Modelling

With the segments, their boundaries, intersection lines, step edges, and roof type information being available at this point in the processing chain, the 3D building models can now be constructed. Some approaches perform yet again a modelling with parametric shapes. But even if these shapes are combined to complex structures, their capability to build exact roof models is limited; at least if only the Boolean union operation is used. Therefore, many approaches directly construct polyhedrons. This can, however, be very complicated, especially if the segments do not exactly meet at intersection lines or corners. Faces can either end up not being planar or small and unaesthetic substructures arise. Also it has to be ensured that the segments cover the entire footprint, as otherwise there will be holes in the building. Generating a ground plan decomposition that respects the segments' boundaries makes the modelling step somewhat easier, but cannot avoid all the discussed problems.

Verma et al. (2006) model complex buildings using parametric primitives. As a proof-of-concept, they model saddleback roofs with complex footprints by rectangular, L-and U-shaped primitives that have a uniform ridge and gutter height. Their parameters are estimated before the modelling step with the RANSAC algorithm, and adjusted to account for symmetries of faces that share a ridge segment. The remaining parts, for which the roof shapes could not be recognized, are modelled by rectilinear buildings that are aligned to the dominant orientations of the building. By fitting planes to the point cloud, the roof slopes and heights are determined for these building parts. For the construction of complex buildings with parametric shapes, the generated solids are united using the proper Boolean operation.

In quite a few approaches, the building models are directly generated as boundary representations from the roof segments. Intersection lines and step edges are extended, so that they connect at corners and form a closed polygon. Then vertical faces are inserted at step edges and the building's outer boundary (Park et al., 2006; Zhou and Neumann, 2008; Dorninger and Pfeifer, 2008). There are, however, ambiguities how the lines are to be connected. A thorough discussion about how the linear features are to be extended and connected is given by Oude Elberink (2009); and the use of existing 2D map data is suggested. Rottensteiner et al. (2005) describe the combination of the generalized roof polygon sections to a polyhedral building model, checking their consistency and inserting step edges if necessary.

A general problem is the alignment of the planar segments in order to avoid small inaccuracies in the resulting building models. Without regularization of the roof and wall planes, the intersection of more than three segments will in most cases not meet at exactly the same point. Dorninger and Pfeifer (2008) simply merge them together, which violates the assumption of planar faces. Oude Elberink (2009) takes the intersection point of the three largest segments and adjusts the smaller ones accordingly. Rottensteiner (2006) adds geometric regularities as soft constraints in the estimation process of the buildings' parameters; which is an idea first proposed by Brenner (2005).

As for the results, Zhou and Neumann (2008, 2009) tested their algorithm on the inner city areas of three major US cities (Oakland, Denver, and Atlanta). The data sets cover areas of 1, 12, and 35 km^2 with point densities of 17, 6, and 17 points/m^2. As it its common to such areas, most of the buildings have flat roofs, often with more than one height level. Under these conditions, the limitation of the algorithm to flat roofs does not become evident. It is able to automatically reconstruct 3D buildings that apparently fit very well the initial point cloud. Dorninger and Pfeifer (2008) show the validity of their reconstruction approach on small historical European cities. A suburban area and a rural settlement were generated from data sets with a mean point density of 6 points/m^2. Another data set of a small city that consists of approximately 2000 buildings was available with a density of 20 points/m^2. The buildings' complexities range from simple to medium with mostly sloped roof structures. According to the authors' exemplary strict quality evaluation, which includes a height comparison of the initial points and the resulting polyhedral model, 75% of the automatically reconstructed buildings are modelled correctly.

Sohn et al. (2008) elegantly avoid many problems in the modelling stage by constructing a 2D binary space partitioning. This can be regarded as a decomposition of the footprint, even though the footprint is also generated in this step. A rectangular area around the building is recursively divided along its linear features – the previously detected intersection lines and step edges – into two subspaces. As the outcome of the partitioning depends on the order of applied lines, the subdivision process is controlled

by a hypothesize-and-test scheme. With the help of a scoring function, a sequence is found that optimizes plane homogeneity, edge correspondence and geometric regularity of the partitioning. To counter the over-fragmentation of space, which always results for concave features in line-based (or plane-based) space partitioning, adjacent subspaces that feature the same roof plane are merged again afterwards. The result is a decomposition of the ground plan into cells, which are extruded, then capped by their uniform flat or sloped roof segment and united to form the final model. The modelling step is straightforward and always results in valid models with planar faces, even if the segments are not adjusted to form nice intersection lines or corners. Even though the approach of Sohn et al. (2008) reconstructs roofs with sloped faces, their 1 km^2 test area of downtown Toronto features only a few such buildings. Most of the 53 buildings have flat roofs, which makes it difficult to evaluate the algorithm's performance on complex non-flat roof shapes. However, the presented results look feasible and the authors themselves rate them as satisfactory, concluding that the technique is useful to provide fine rooftop models with very high shape complexity. They also explain that the quality is subject to the extracted linear features and that erroneous lines could be avoided by including a model-driven post-processing step.

2.3. Reconstruction by DSM simplification

In contrast to the reconstruction approaches described above, which construct building models from scratch to best fit the given elevation data, the idea of the following approach is, that the buildings are contained in highly detailed, meshed digital surface models (DSM) and that they only need to be simplified to the right abstraction level and extracted if necessary.

The technique used is surface simplification, which reduces a mesh's number of faces to speed up its loading, decoding and rendering in the context of real-time visualization applications. Wahl et al. (2008) put another meaning to it and regard the simplification of digital surface models in urban areas as a means to rapidly generate 3D cities models. Commonly, geometric error criteria are used to control the simplification process. However, they are insufficient to preserve the shape properties of man-made objects. Therefore, the simplification is additionally constrained by a semantic component, which takes the results from a shape detection algorithm that is based on connected components analysis. Common shapes like planes, spheres, cylinders, cones and tori are iteratively detected in the DSM by a probabilistic RANSAC-based algorithm as described in Schnabel et al. (2007).

As the algorithm works on uniform point clouds, additional points have first to be inserted at height discontinuities, e.g. at the facades. The DSM vertices are then associated with all shapes that they are in close proximity of, defining them edge or corner points if they are close to two or respectively three or more different shapes.

In addition to respecting a geometric error criterion, a DSM edge may only be collapsed during simplification if all feature edges and corners are preserved by the operation. This way, DSMs at various levels of detail can be automatically generated ranging from highly detailed versions for pixel correct real-time visualizations to low detailed ones for internet and mobile applications. If discrete building objects are needed, the relevant parts of the mesh can be extracted and converted to a closed boundary representation.

Due to the data-driven nature of the underlying surface simplification, the method is able to handle roof shapes of arbitrary complexity and always leads to valid representations that are geometrically close to the raw data. The transformation of the mesh via edge collapse operations thereby elegantly avoids the difficulty of constructing polyhedrons in boundary representations from the segmented shapes. By adding low-level model information to the process in the form of topological constraints, the quality of the simplification is improved by this semantic component. However, the geometric criteria are a reliable fallback when shapes cannot be detected or if their alignment is inaccurate. In such cases, the mesh might at worst become locally unaesthetic. In order to better deal with incomplete and inconsistent structural information, Möser et al. (2009) extend the approach, which results in models with fewer artifacts.

Fig. 4. Textured 3D city model of Berlin, Germany (courtesy of VirtualCitySystems).

3. Building facades

As discussed in the previous section, a considerable number of approaches have been developed to collect 3D building models at acceptable effort for complete city areas. Usually such building representations feature flat facades and distinctive roof structures according to level of detail 2 within the OGC standard CityGML (Kolbe et al., 2009). Due to their accessibility by platforms like Google Earth or Bing Maps, such data sets have become commonplace to a huge group of potential users. Furthermore, municipal administrations frequently maintain and manage this type of representation for planning purposes. However, in addition to the available footprint and roof shape, facades are the most important features that reflect the building's style and dimension. Nevertheless, the geometric representation of building facades is usually limited to planar structures if data collection is realized from aerial data.

3.1. Facade texture from airborne imagery

The visual appearance of such planar building facades is frequently enhanced using suitable image texture. As an example, Kada (2009) describes the texture mapping for the 3D model of Berlin realized from approximately 100,000 oblique aerial images (see Fig. 4). To guarantee a perfect fit to the underlying geometry, all images had to be manually adjusted. Then, for each surface polygon, a texture was automatically synthesized using the approach described by Lorenz and Döllner (2006). Depending on the distance, viewing angle and potential occluders, the best source image is determined for colouring a texture pixel. Heuristics additionally minimized the inclusion of vegetation.

If this facade texture is extracted from aerial images, oblique views are necessary in order to guarantee the required visibility of the respective building facades. This visibility is e.g. inherent in the wide view angle of standard vertical aerial images, especially if highly overlapping image blocks have been captured. This is

frequently realized to support image matching for elevation data generation in densely built-up areas. Even though facade texture can in principle be extracted from standard aerial images, texture mapping is frequently realized using oblique aerial photography, which of course provides a good visibility of building facades.

Traditionally, oblique images have been used in aerial mapping to overcome the optical limitation of wide-angle views during mapping, reconnaissance and surveillance (Petrie, 2009). Since they provide a quick 3D view of urban scenes without having to reconstruct the respective buildings, oblique images have meanwhile become an important photogrammetric product within urban applications. Since they can be used as a visual substitute for 3D building models, oblique images are e.g. integrated within applications like Bing Maps and Google Earth as an alternative or supplement to the presentation of 3D building models. Within these applications, large amounts of oblique frame images are currently made available for the consumer market. As an example, georeferenced oblique images are currently captured with the Pictometry camera system for the so-called bird's eye imagery within Bing Maps. For Western Europe, the imagery covers every major town and city with a population greater than 50,000 inhabitants, resulting in approximately 1000 cities and a total of almost 100,000 km^2 with 80% of the population (Karbøand Schroth, 2009). The data base will be updated biannually; the images have a typical resolution between 10 and 12 cm at a positional accuracy of approximately 50 cm. This type of oblique imagery was used for texture mapping during generation of Fig. 4. In principle, the matching between the faces of the building models and the respective image patches of the oblique scenes is implemented fully automatically. However, small errors in image orientation or available building geometry as well as occlusion from objects like trees, scaffolding etc. can require interactive control and correction.

3.2. Terrestrial data collection

Oblique imagery is widely used for texture mapping in order to enhance the visual quality of 3D city models as reconstructed from airborne images or LiDAR. However, the quality of airborne data collection as input for the generation of visually pleasing urban scenes is usually limited. It is adequate to generate virtual images from elevated viewpoints in order to facilitate overviews of larger areas. Nevertheless, visualizations at very high degree of realism are frequently required for pedestrian viewpoints especially in the context of urban planning or 3D navigation. Such road-side walkthroughs presume very detailed information especially for the building facades. Due to viewpoint restrictions, this is usually not available from airborne platforms. For this reason terrestrial data collection is frequently required.

The collection of images or laser scans from static viewpoints is in general limited to smaller areas while the efficient mapping of street scenes covering complete city areas presumes the use of mobile mapping systems. Starting from research platforms in the 1990s, a large number of commercial land-based mobile mapping systems are meanwhile available (Toth, 2009). Depending on the specific purpose, such systems combine mapping sensors like multiple CCD cameras and/or laser scanners with suitable navigation components. These navigation components include GNSS (Global Navigation Satellite Systems) and Inertial Measurement Units (IMU), which can be further aided by speed or wheel sensors.

3.3. Textured surface models

Applications like Google Street View or Microsoft Streetside use images collected from such mobile mapping systems to generate panoramic views of urban areas at street level. In this context, these views supplement the oblique images from airborne platforms. Georeferenced video streams as collected from mobile mapping systems can alternatively be used for more advanced real-time visualizations. As an example, large scale 3D reconstructions of street scenes are e.g. required for car navigation systems. Such an approach, which aims at the generation of realistically textured surface models at video frame rates, is e.g. described by Cornelis et al. (2008). Textured polygonal meshes allow for compact representations of 3D city models. This relatively simple representation of building geometry is especially useful for mobile applications since it allows for relatively simple distribution and visualization. Additionally, such memory efficient representations by textured meshes can be generated at video frame rate. Pollefeys et al. (2008) describe a 3D reconstruction system, where the georeferenced video stream is evaluated by a plane sweeping stereo algorithm with three orthogonal sweeping directions—one for the ground and two for the facades. This approach is especially tailored for street level visualization of dense urban scenes with tall buildings. After fusion of the generated depth maps, which helps to eliminate erroneous matches, triangular meshes with mapped image texture are generated. This simple representation of the building environment by textured meshes is sufficient, since it is mainly used for visualization. Thus, data processing is particularly driven by geometrical and graphical considerations. In contrast, object recognition is limited to the detection of cars to prevent disturbing effects from stereo matching. In this system, the localization of cars is used to replace these objects by virtual 3D car models in order to cover potential reconstruction artifacts and to augment the visual realism of the final city model.

3.4. Geometric facade models from terrestrial LiDAR

Mesh based representations, which are mainly generated for visualization purposes, provide only limited semantic information. In contrast, location-aware applications like advanced navigation and search tools presume building representations of much higher complexity. This additionally requires ontological structures for the respective building representations, which include thematic classes, attributes, and their interrelationships as well as building objects decomposed depending on their structures in the real world. Since such interpreted structures are also required for the building facades, the respective elements such as windows and doors have to be identified. As an example, Lee and Nevatia (2004) extract window candidates defined as homogeneous rectangles from ground view images and integrate them in a 3D building model. Image based approaches aiming at the interpretation of building facades are also described in Mayer and Reznik (2005) and Wenzel et al. (2008).

However, the difficulty of image understanding frequently aggravates purely image based approaches for automatic facade reconstruction if the application is not limited to visualization. For this reason, automatic approaches frequently extract the required features like windows and doors from 3D point clouds similar to airborne data collection. If larger areas have to be covered, ground-based mobile mapping systems with integrated terrestrial laser scanners are often used to provide dense 3D point coverage at facades and the neighbouring architecture. As an example, the collection of geo-referenced point clouds at point densities of 2–5 cm at absolute accuracies of 10–20 cm are reported by Haala et al. (2008) and Barber et al. (2008) for the commercial StreetMapper system.

The use of terrestrial point clouds to automatically recover building facades is e.g. described by Pu and Vosselman (2009). In order to roughly provide potential building features like walls, doors, roofs and protrusions, they first segment the point cloud into planar elements. Based on characteristics like size, position, orientation and topology, all extracted segments are then distinguished in different semantic feature types. Afterwards, geometric reconstruction is finalized by the fitting and merging of the extracted polygons to a complete 3D building model.

Fig. 5. LiDAR points measured at a building facade.

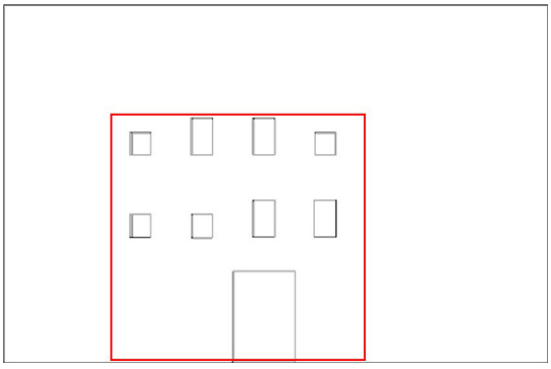

Fig. 6. Facade features extracted from LiDAR points.

3.5. Facade grammars

In addition to such data driven approaches, model driven reconstruction is especially suitable to benefit from the regular structure of facades. In this context, shape grammars originally introduced in architecture by Stiny and Gips (1972) have gained increasing importance as a tool for urban reconstruction. So-called split grammars as introduced by Wonka et al. (2003) allow for practical design by an automatic generation of architectural structures from a database of rules. Based on this work, Müller et al. (2006) therefore developed a system for the automatic generation of large scale city models. The use of grammars for facade reconstruction from terrestrial data is e.g. presented in Ripperda and Brenner (2007) and Hohmann et al. (2009).

An approach for a grammar supported facade reconstruction, which enables the generation and application of rules aiming at a quality dependent processing is presented by Becker (2009). The algorithm aims at the refinement of existing building models from airborne data collection by adding 3D geometries to the planar facade polygons. The required modifications are realized by using a cell decomposition to represent the respective building models.

A data driven reconstruction is realized in a first step (Becker and Haala, 2007). For this purpose, features of interest are extracted from the LiDAR points measured by the StreetMapper mobile mapping system at the respective building facade. As it is visible for the exemplary point cloud depicted in Fig. 5, window features can be extracted by searching for no-measurement areas. However, occlusions and oblique views result in strong variations of the sampling distance during scanning of the LiDAR points. An accurate extraction of windows and doors can therefore only be guaranteed at parts of the facade with a sufficient amount of point measurements.

For the example in Fig. 5, the corresponding area is marked by a red rectangle in Fig. 6. As it is visible, facade features are only extracted for this area with sufficient point density. From this part, dominant or repetitive features and regularities are detected in order to infer a facade grammar, which contains all the information necessary to produce facades in the style of the respective building. Since the inference process is restricted to this dense area, a facade grammar of good quality can be provided, which is then used to synthesize facade regions for which sensor data is only partially available or (even) non-existing.

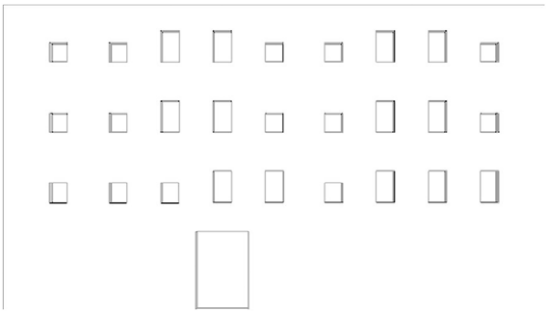

Fig. 7. Synthesized facade from inferred grammar.

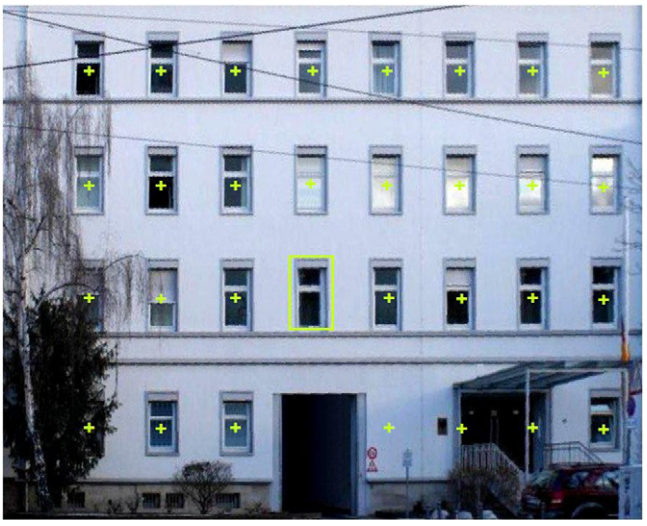

Fig. 8. Verification using rectified facade image.

This production step is demonstrated in Fig. 7, which shows the synthesized facade structure for the remaining parts. Such top-down predictions can also be verified if additional data like rectified facade imagery is available. Such an image is shown in Fig. 8, where the green rectangle represents an image region, which corresponds to a structure element as extracted from LiDAR point analysis. This area serves as a mask during a correlation based verification of possible positions of facade elements. In Fig. 8, these hypotheses are represented by green crosses. Hypotheses are accepted for sufficient correlation values and the corresponding structure is inserted at the respective position. This process allows an iterative refinement of the facade model and update of the respective production rules.

The final result after this verification step is shown in Fig. 9, where the relatively small region with dense point measurements used to generate the original grammar is marked in the predicted 3D building model.

The use of facade grammar is not limited to the completion at areas of insufficient data quality. As it is demonstrated in Fig. 10, it also provides a "library" to generate facade structures for building objects, where no measurement is available at all. In this example, a variety of grammars was derived in advance from a few buildings in the near environment and then used to enrich the 3D city model of larger scene with facades of compatible architectural styles.

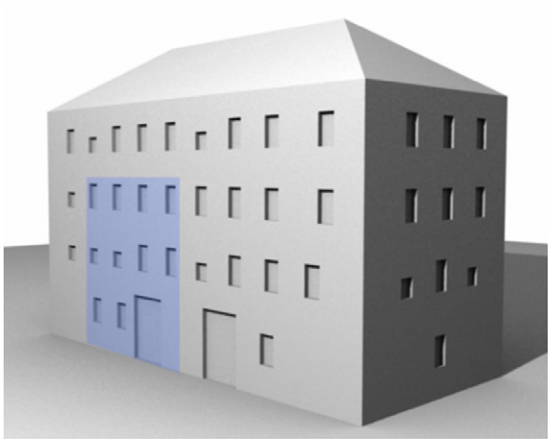

Fig. 9. 3D facade model from quality dependent knowledge propagation.

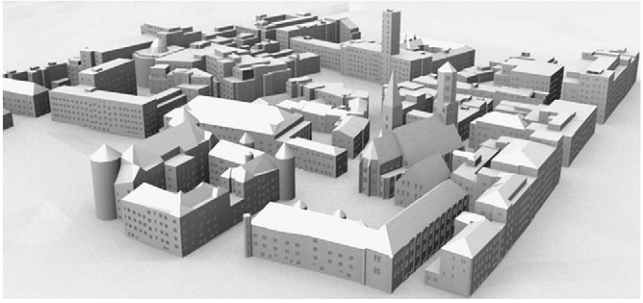

Fig. 10. City model with grammar induced facades.

4. Conclusions and outlook

Despite considerable effort, the difficulty of automatic interpretation enduringly limited operational 3D city modelling to systems with considerable manual operations and just some marginally tools for automation. Thus, the development of fully automatic algorithms is still a problem tackled by a large group of researchers. Nevertheless, the operability of available software tools meanwhile allows commercial vendors to collect large area 3D models of complete cities at a suitable amount of time, while the required interactive effort depends on the scene's complexity and the aspired quality and amount of detail for the respective city model.

Early applications of city models like 3D simulations for planning telecommunication networks or visualizations from bird's eye views in urban planning were limited to a few specialized users. For this purposes, relatively simple geometric representations, which mainly reflect the roof shape and footprint of the respective buildings were sufficient. As discussed in Section 3, algorithms to automatically generate this type of 3D models from suitable elevation data are available.

Since the mid-1990s the demand for 3D city models is continuously growing and has meanwhile become commonplace. Especially Internet applications like digital globes as provided by Google Earth or Bing Maps and advanced navigation systems carry 3D city models to a huge group of potential users. The spread and wide range of applications also result in more heterogeneous demands ranging from the provision of input data for photorealistic real-time visualization of terrestrial walkthroughs to interpreted scenes for advanced search and navigation.

In this context, the use of terrestrial data collection to provide detailed information on building facades becomes more and more important. As discussed in Section 3, modelling can be realized by shape grammars, which are suitable to represent the regular structure of these objects. Such approaches can also integrate measurements of heterogeneous quality. As an example, if 3D point clouds are used, differences in accuracy, coverage and density will occur due to variations in viewpoint and distance to the object of interest. These differences are inherent for terrestrial data collection even if measurement is realized by specialized surveying systems like mobile mapping systems. However, such problems will become even more evident if terrestrial data is collected by consumers and non-experts in web-based applications like the image based rendering of urban scenes as realized within PhotoSynth (Snavely, 2008). Thus, approaches will be required, which are much more flexible towards different data quality and incomplete sensor data.

Acknowledgement

Martin Kada is indebted to the Baden-Württemberg Stiftung for the financial support of this research project by the Eliteprogramme for Postdocs.

References

Ballard, D.H., Brown, C.M., 1982. Computer Vision. Prentice-Hall, Inc., Englewood Cliffs, NJ, USA.

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 (3–4), 129–151.

Barber, D., Mills, J., Smith-Voysey, S., 2008. Geometric validation of a ground-based mobile laser scanning system. ISPRS Journal of Photogrammetry and Remote Sensing 63 (1), 128–141.

Becker, S., 2009. Generation and application of rules for quality dependent facade reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing 64 (6), 640–653.

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

Brenner, C., 2005. Building reconstruction from images and laser scanning. International Journal of Applied Earth Observation and Geoinformation 6 (3–4), 187–198.

Brenner, C., Haala, N., 1998. Rapid acquisition of virtual reality city models from multiple data sources. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 32 (Part 5), 323–330. Chikatsu, H., Shimizu, E. (Eds.).

Cerny, V., 1985. Thermodynamical approach to the travelling salesman problem: an efficient simulation algorithm. Journal of Optimization Theory and Applications 45 (1), 41–51.

Champion, N., 2009. Detection of unregistered buildings for updating 2D databases. EuroSDR Official Publication. 56. pp. 7–54.

Chen, L., Teo, T., Rau, J., Liu, J., Hsu, W., 2005. Building reconstruction from LiDAR data and aerial imagery. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, vol. 4. pp. 2846–2849.

Cornelis, N., Leibe, B., Cornelis, K., Van Gool, L., 2008. 3D urban scene modeling integrating recognition and reconstruction. International Journal of Computer Vision 78 (2–3), 121–141.

Dorninger, P., Nothegger, C., 2007. 3D segmentation of unstructured point clouds for building modeling. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W49A), 191–196.

Dorninger, P., Pfeifer, N., 2008. A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. Sensors 2008 (8), 7323–7343.

Drauschke, M., Roscher, R., Läbe, T., Förstner, W., 2009. Improving image segmentation using multiple view analysis. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3/W4), 211–215.

Durupt, M., Taillandier, F., 2006. An automatic reconstruction from a digital elevation model and cadastral data: an operational approach. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W24) (on CD-ROM).

Edelsbrunner, H., Kirkpatrick, D., Seidel, R., 1983. On the shape of a set of points in the plane. IEEE Transactions on Information Theory 29 (4), 551–559.

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24 (6), 381–395.

Fukunaga, K., Hostetler, L.D., 1975. The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Transactions on Information Theory 21 (1), 32–40.

Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82 (4), 711–732.

Grün, A., Baltsavias, E., Henricsson, O. (Eds.), 1997. Automatic Extraction of Man-Made Objects from Aerial and Space Images (II). Birkhäuser, Basel.

Grün, A., Kübler, O., Agouris, P. (Eds.), 1995. Automatic Extraction of Man-Made Objects from Aerial and Space Images. Birkhäuser, Basel.

Gurram, P., Lach, S., Saber, E., Rhody, H., Kerekes, J., 2007. 3D scene reconstruction through a fusion of passive video and LiDAR imagery. In: Proceedings of the 36th Applied Imagery Pattern Recognition Workshop (AIPR). IEEE Computer Society, Washington, DC, pp. 133–138.

Haala, N., 2009. Comeback of digital image matching. In: Fritsch, D. (Ed.), Photogrammetric Week 2009. Wichmann Verlag, Heidelberg, pp. 289–301.

Haala, N., Peter, M., Kremer, J., Hunter, G., 2008. Mobile LiDAR mapping for 3D point cloud collection in urban areas—a performance test. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (Part B5) (on CD-ROM).

Habib, A.F., Zhai, R., Kim, C., 2010. Generation of complex polyhedral building models by integrating stereo-aerial imagery and LiDAR data. Photogrammetric Engineering & Remote Sensing 76 (5), 609–623.

Hirschmüller, H., Bucher, T., 2010. Evaluation of digital surface models by semi-global matching. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation 19, 571–580.

Hohmann, B., Krispel, U., Havemann, S., Fellner, D., 2009. CityFit—high-quality urban reconstructions by fitting shape grammars to images and derived textured point clouds. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 5/W1) (on CD-ROM).

Kaartinen, H., Hyyppä, J., 2006. EuroSDR—project commission 3 "Evaluation of building extraction". Final Report. EuroSDR—European Spatial Data Research. Official Publication. 50. pp. 9–77.

Kada, M., 2009. The 3D Berlin project. In: Fritsch, D. (Ed.), Photogrammetric Week 2009. Wichmann Verlag, Heidelberg, pp. 331–340.

Kada, M., McKinley, L., 2009. 3D building reconstruction from LiDAR based on a cell decomposition approach. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3/W4), 47–52.

Karbø, N., Schroth, R., 2009. Oblique aerial photography: a status review. In: Fritsch, D. (Ed.), Photogrammetric Week 2009. Wichmann Verlag, Heidelberg, pp. 119–125.

Khoshelham, K., 2005. Region refinement and parametric reconstruction of building roofs by integration of image and height data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W24), 3–8.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by Simulated Annealing. Science, New Series 220 (4598), 671–680.

Kolbe, T., Nagel, C., Stadler, A., 2009. CityGML—OGC standard for photogrammetry? In: Fritsch, D. (Ed.), Photogrammetric Week 2009. Wichmann Verlag, Heidelberg, pp. 265–277.

Lafarge, F., Descombes, X., Zerubia, J., Pierrot-Deseilligny, M., 2008. Automatic building extraction from DEMs using an object approach and application to the 3D-city modeling. ISPRS Journal of Photogrammetry and Remote Sensing 63 (3), 365–381.

Leberl, F., Kluckner, S., Bischof, H., 2009. Collection, processing and augmentation of VR. In: Fritsch, D. (Ed.), Photogrammetric Week 2009. Wichmann Verlag, Heidelberg, pp. 251–263.

Lee, S.C., Nevatia, R., 2004. Extraction and integration of window in a 3D building model from ground view images. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 2. pp. 113–120.

Li, Y., Wu, H., 2008. Adaptive building edge detection by combining LiDAR data and aerial images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (Part B1), 197–202.

Lodha, S.K., Arun Kumar, K.K., 2005. Semi-automatic roof reconstruction from aerial data using *k*-means with refined seeding. In: Proceedings of the ASPRS 2005 Annual Conference, Baltimore, Maryland, USA (on CD-ROM).

Lorenz, H., Döllner, J., 2006. Towards automating the generation of facade textures of virtual city models. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 2), (on CD-ROM).

Ma, R., 2004. Building model reconstruction from LiDAR data and aerial photographs. Ph.D. Thesis. The Ohio State University, USA.

MacQueen, J.B., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, vol. 1. University of California Press, pp. 281–297.

Mayer, H., Reznik, S., 2005. Building facade interpretation from image sequences. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W24), 55–60.

Medeiros, E., Velho, L., Lopes, H., 2004. Restricted BPA: applying ball-pivoting on the plane. In: Computer Graphics and Image Processing, XVII Brazilian Symposium on, SIBGRAPI'04. pp. 372–379.

Melzer, T., 2007. Non-parametric segmentation of ALS point clouds using mean shift. Journal of Applied Geodesy 1 (3), 159–170.

Milde, J., Brenner, C., 2009. Graph-based modeling of building roofs. In: Proceedings of the 12th AGILE Conference on GIScience, Hannover, Germany (on CD-ROM).

Milde, J., Zhang, Y., Brenner, C., Plümer, L., Sester, M., 2008. Building reconstruction using a structural description based on a formal grammar. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (Part B3b), 227–232.

Möser, S., Wahl, R., Klein, R., 2009. Out-of-core topologically constrained simplification for city modeling from digital surface models. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 5/W1) (on CD-ROM).

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

Neidhart, H., Sester, M., 2008. Extraction of building ground plans from LiDAR data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (Part B2), 405–410.

Novacheva, A., 2008. Building roof reconstruction from LiDAR data and aerial images through plane extraction and colour edge detection. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (Part B6b), 53–57.

Ortner, M., Descombes, X., Zerubia, J., 2007. Building outline extraction from digital elevation models using marked point processes. International Journal of Computer Vision 72 (2), 107–132.

Oude Elberink, S., 2009. Target graph matching for building reconstruction. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3/W8), 49–54.

Oude Elberink, S., Vosselman, G., 2009. Building reconstruction by target based graph matching on incomplete laser data: analysis and limitations. Sensors 9, 6101–6118.

Park, J., Lee, I., Choi, Y., Lee, Y.J., 2006. Automatic extraction of large complex buildings using LiDAR data and digital maps. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3), 148–154.

Petrie, G., 2009. Systematic oblique aerial photography using multiple digital frame cameras. Photogrammetric Engineering & Remote Sensing 75 (2), 102–107.

Pollefeys, M., Nister, D., Frahm, J.-M., Akbarzadeh, A., Mordohai, P., Clipp, B., Engels, C., Gallup, D., Kim, S.-J., Merrell, P., Salmi, C., Sinha, S., Talton, B., Wang, L., Yang, Q., Stewénius, H., Yang, R., Welch, G., Towles, H., 2008. Detailed real-time urban 3D reconstruction from video. International Journal of Computer Vision 78 (2), 143–167.

Pu, S., Vosselman, G., 2009. Knowledge based reconstruction of building models from terrestrial laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing 64 (6), 575–584.

Ripperda, N., Brenner, C., 2007. Data driven rule proposal for grammar based facade reconstruction. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W49A), 173–178.

Rottensteiner, F., 2006. Consistent estimation of building parameters considering geometric regularities by soft constraints. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3), 13–18.

Rottensteiner, F., Trinder, J., Clode, S., Kubik, K., 2005. Automated delineation of roof planes from LiDAR data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W4), 221–226.

Sampath, A., Shan, J., 2007. Building boundary tracing and regularization from airborne LiDAR point clouds. Photogrammetric Engineering & Remote Sensing 73 (7), 805–812.

Sampath, A., Shan, J., 2010. Segmentation and reconstruction of polyhedral building roofs from aerial LiDAR point clouds. IEEE Transactions on Geoscience and Remote Sensing 48 (3), 1554–1567.

Schnabel, R., Wahl, R., Klein, R., 2007. Efficient RANSAC for point-cloud shape detection. Computer Graphics Forum 26 (2), 214–226.

Snavely, N., 2008. Scene reconstruction and visualization from internet photo collections. Ph.D. Thesis. University of Washington, USA.

Sohn, G., Dowman, I., 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. ISPRS Journal of Photogrammetry and Remote Sensing 62 (1), 43–63.

Sohn, G., Huang, X., Tao, V., 2008. Using a binary space partitioning tree for reconstructing polyhedral building models from airborne LiDAR data. Photogrammetric Engineering & Remote Sensing 74 (11), 1425–1438.

Stiny, G., Gips, J., 1972. Shape grammars and the generative specification of painting and sculpture. In: Friedman, C.V. (Ed.), Information Processing 71. pp. 1460–1465.

Taillandier, F., 2005. Automatic building reconstruction from cadastral maps and aerial images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W24), 105–110.

Toth, C., 2009. R&D of mobile mapping and future trends. In: Proceedings of the ASPRS Annual Conference (on CD-ROM).

Vallet, B., Pierrot-Deseilligny, M., Boldo, D., 2009. Building footprint database improvement for 3D reconstruction: a direction aware split and merge approach. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3/W4), 139–144.

Verma, V., Kumar, R., Hsu, S., 2006. 3D building detection and modeling from aerial LiDAR data. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR'06. IEEE Computer Society, Washington, DC, pp. 2213–2220.

Vosselman, G., 1999. Building reconstruction using planar faces in very high density height data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 32 (Part 3/2W5), 87–92.

Vosselman, G., 2002. Fusion of laser scanning data, maps, and aerial photographs for building reconstruction. In: IEEE International Geoscience and Remote Sensing Symposium. IGARSS'02, vol. 1, 4 p. (on CD-ROM).

Vosselman, G., Gorte, B.G.H., Sithole, G., Rabbani, T., 2004. Recognising structure in laser scanner point clouds. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 8/W2), 33–38.

Wahl, R., Schnabel, R., Klein, R., 2008. From detailed digital surface models to city models using constraint simplification. Photogrammetrie, Fernerkundung, Geoinformation 12 (3), 207–215.

Wenzel, S., Drauschke, M., Förstner, W., 2008. Detection of repeated structures in facade images. Pattern Recognition and Image Analysis 18 (3), 406–411.

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

Zebedin, L., Klausa, A., Gruber-Geymayer, B., Karner, K., 2006. Towards 3D map generation from digital aerial images. ISPRS Journal of Photogrammetry and Remote Sensing 60 (6), 413–427.

Zhou, Q.Y., Neumann, U., 2008. Fast and extensible building modeling from airborne LiDAR data. In: Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM GIS 2008 (on CD-ROM).

Zhou, Q.Y., Neumann, U., 2009. A streaming framework for seamless building reconstruction from large-scale aerial LiDAR data. In: Proceedings of the 2009 IEEE Computer Vision and Pattern Recognition. CVPR'09. IEEE Computer Society, Washington, DC, pp. 2759–2766.