ISPRS Journal of Photogrammetry and Remote Sensing 76 (2013) 17–29

Contents lists available at SciVerse ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Model driven reconstruction of roofs from sparse LIDAR point clouds

André Henn *, Gerhard Gröger, Viktor Stroh, Lutz Plümer

University of Bonn, Institute for Geodesy and Geoinformation, Meckenheimer Allee 172, D-53115 Bonn, Germany

ARTICLE INFO

Article history:
Received 1 June 2012
Received in revised form 15 November 2012
Accepted 16 November 2012

Keywords:
Reconstruction
Building
Data mining
Classification
LIDAR
Three-dimensional

ABSTRACT

This article presents a novel, fully automatic method for the reconstruction of three-dimensional building models with prototypical roofs (CityGML LoD2) from LIDAR data and building footprints. The proposed method derives accurate results from sparse point data sets and is suitable for large area reconstruction. Sparse LIDAR data are widely available nowadays. Robust estimation methods such as RANSAC/MSAC, are applied to derive best fitting roof models in a model-driven way. For the identification of the most probable roof model, supervised machine learning methods (Support Vector Machines) are used. In contrast to standard approaches (where the best model is selected via MDL or AIC), supervised classification is able to incorporate additional features enabling a significant improvement in model selection accuracy.

1. Introduction

Three-dimensional building models with prototypical roofs (CityGML Level-of-Detail 2, LoD2 (see Gröger et al., 2012)) are an essential prerequisite for many applications of 3D city models, such as urban and telecommunications planning (Knapp and Coors, 2008), environmental planning, particularly solar radiation calculations or noise emission simulation (Czerwinski et al., 2007), and visualization (Döllner et al., 2006). For those applications, 3D city models covering large areas, such as state territories, are required. However, those models are not available at present. The most detailed level LoD3 currently can be only reconstructed manually or semi-automatically and is available only for small areas such as city centers and prominent landmarks. For LoD2, the need for a high degree of automation for the derivation of LoD2 building models is much higher. Input data for LoD2 derivation are 2D building footprints and LIDAR data. Building footprints are available at a national level (at least in European countries or other developed countries), either from cadastral or tax-related data sets or from other sources such as OpenStreetMap. However, the feasibility, quality and reliability of the results of LoD2 reconstruction methods highly depend on the LIDAR point density. Currently, the point density does not allow for a covering derivation of LoD2. In the United Kingdom, for example, covering data sets exist only for England and Wales with a resolution of 2 m, i.e., 0.5 points/m^2. In Germany, the point density for the whole country is 1 point/m^2 or less. As a notable exception, in the Netherlands the AHN-2 dataset covering the area of the entire country with a point density of at least 8 points/m^2 is available (Vosselman, 2012). Hence, a method for deriving a country-wide LoD2 building model has to cope with low LIDAR point density. This article presents a robust method for the derivation of LoD2 building models from building footprints and sparse LIDAR data.

1.1. Related work

In the past decades, numerous approaches for detection and reconstruction of buildings from (high resolution) aerial images (Cord and Declercq, 1999; Scholze et al., 2001), LIDAR-data (Vosselman and Dijkman, 2001) used as active sensors and a combination of heterogeneous data sources (Rottensteiner and Briese, 2003; Suveg and Vosselman, 2004) have been presented. This abundance of approaches shows, that building reconstruction is challenging (Tarsha-Kurdi et al., 2007; Haala and Kada, 2010). An overview of several methods is given in Brenner (2005), Hu et al. (2003), Tarsha-Kurdi et al. (2007), Haala and Kada (2010) or Vosselman and Maas (2010).

According to Maas and Vosselman (1999) there are basically two approaches for roof reconstruction: *data-driven* (non-parametric or bottom-up strategy) and *model-driven* (parametric or top-down strategy). Tarsha-Kurdi et al. (2007) give a comparison between these approaches. In the data-driven case, buildings are assumed to be an aggregation of several estimated/segmented roof planes as unstructured objects (Tarsha-Kurdi et al., 2007; Rottensteiner and Briese, 2003). The complete point cloud is analysed

* Corresponding author. Tel.: +49 228 73 3164; fax: +49 228 73 1753.
E-mail addresses: henn@igg.uni-bonn.de (A. Henn), groeger@igg.uni-bonn.de (G. Gröger), stroh@igg.uni-bonn.de (V. Stroh), pluemer@igg.uni-bonn.de (L. Plümer).

0924-2716/$ - see front matter © 2012 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.isprsjprs.2012.11.004

and the roof planes are determined. In this vein, several approaches are proposed, applying different methods such as 3D-Hough-Transform (Vosselman and Dijkman, 2001; Overby et al., 2004), Random Sample Consensus (RANSAC) (Tarsha-Kurdi et al., 2008; Ameri and Fritsch, 2000; Brenner, 2000) or Region Growing (Rottensteiner et al., 2005; Rottensteiner, 2006; Verma et al., 2006). Quite a different approach is used by Oude Elberink (2009) who uses matching of graphs for the reconstruction of buildings. Here, a high point density is required since the approach is based on segmentation and a subsequent detection of intersection lines and step edges. Another technique proposed by Tarsha-Kurdi et al. (2008) uses RANSAC for plane detection in laser point clouds. The pure mathematical principle is enhanced by adjacency matrices and takes the minimum size of roof planes into consideration. This is done by a modification of the original RANSAC (Fischler and Bolles, 1981) by a restriction on the inliers (threshold for minimum number of points belonging to a roof plane) and improving the plane quality by removing noisy points. After the segmentation or determination of the planar roof planes, the neighborhood relationships of the roof planes are determined.

Zhou and Neumann (2009) propose an approach for the automatic reconstruction of large areas where the LIDAR data is streamed into memory. The LIDAR data is classified based on local point features to find roof points. Afterwards these roof points are segmented and their boundaries are extracted to generate the building outline.

In contrast to data driven methods, model-driven approaches consider parametrized roof models which have to be evaluated for a given point cloud. To this end, simple and complex roofs have to be distinguished (Tarsha-Kurdi et al., 2007). For simple roofs, e.g the standard roof types like gabled roofs, flat roofs or hipped roofs, the most appropriate model fitting the point cloud is determined. For complex roof models, additional data (for example footprints) have been included in order to split the complex roof into primitives. Suveg and Vosselman (2004) noticed that the shape of the footprint has a high influence on the roof type, especially on the structure of the roof. According to Brenner (2000), most complex buildings can be disaggregated into simple primitives. Usually, building footprints are decomposed into primitives which are the base of the reconstruction. Since standard roof types usually have rectangular shape, the footprints are decomposed as a rule into rectangles at first, like in Vosselman and Dijkman (2001) or Brenner and Haala (1998). Kada (2007) and Kada and McKinley (2009) propose a method for decomposition of footprint with an additional generalization of the footprint. This method is based on the generalization of 3D building models in boundary representation. The building model is subdivided into non-intersecting cells along its facade polygons using vertical planes. Parallelism and orthogonality can be enforced using angle thresholds for the plane normal vectors. The resulting cells are simplified and merged using buffers and distance thresholds. Further, Vallet et al. (2011) provide an approach where the footprint decomposition is triggered by a digital surface model derived from the laser points.

Model-driven approaches either use a single parametrized model or a model catalog resulting in parameter estimation and/or model selection problems. For model estimation and selection several approaches have been proposed. In particular Maas and Vosselman (1999) suggest a closed form solution for gabled roof estimation from laser points using invariant moments. Brenner and Haala (1998) decompose the footprint in rectangular primitives and instantiate each of their parametric primitives (standard roof types). After parameter estimation the best model is selected based on area and slope thresholds and the final fit error. In the context of building reconstruction from aerial images, Suveg and Vosselman (2004) propose a scheme for generation and verification of hypotheses. A building can be reconstructed by multiple primitives in multiple ways. For every footprint (rectangle/partition) multiple hypotheses are created and a ranking of the hypotheses is made using the minimum description length criterion (MDL). Furthermore, a score function based on mutual information is introduced for the evaluation of the building models within the image. Haala et al. (1998) analyze the surface normals of a digital surface model (DSM) and segment it into regions with same direction. The best fitting model from the roof catalog is the one which fits best to this segmentation. The quality of the resulting model is to be computed using root mean square error. The approach of Kada and McKinley (2009) uses normal vectors for determination of the models which fits best to the point cloud. The model whose corresponding point normals fit to most of the directions of the primitive is selected. Recently, Huang et al. (2011) presented a statistical model-driven approach for roof reconstruction using generative models. They allow overlapping primitives and estimate the parameters using reversible jump Markov Chain Monte Carlo (rjMCMC). To overcome the problem of processing too many small models which are too complex, they present an information criterion adapted from the Akaike information criterion (Akaike, 1974). Here, the influence of the model size and the number of involved laser points is included. Neighboring primitives can be combined using context-sensitive rules. A similar approach for a DSM has been proposed by Lafarge et al. (2010). For 2D primitives, which are automatically/interactively determined as quadrilaterals/triangles, a library of 3D parametric blocks including variants like connectors or hipped roofs is defined. The placement of these building blocks on 2D primitives is achieved using stochastic methods. The optimal configuration of the primitives and their parameters is achieved using rjMCMC. Prior knowledge is modeled in a Bayesian formulation taking the interactions between neighboring objects into account. Poullis and You (2009) present a method for roof type identification with a flexible parametrized geometric primitive. It consists of only two parameters, α and β, with which several building primitives can be created. The parameters are determined using a non-linear, bound-constraint minimization by means of Gaussian Mixture Models (GMM) and the Expectation Maximization (EM) algorithm. Here, by the help of GMM, it is possible to detect outliers and exclude them from fitting the roof planes. This approach determines the best model concurrently with the estimation of the model's parameters.

In order to reconstruct complex and irregular roofs, Lafarge and Mallet (2012) propose a method for the combination of 3D primitives (planes, cones, cylinders) with meshes. The points of the point cloud are classified in different classes, such as "building" and "tree". Afterwards, the classified 3D-points are mapped onto a 2D grid. The labels of the grid-cells are propagated by the help of Markov random fields. Afterwards a triangulation based on the label map (with cells marked as roof) is generated, followed by a mesh simplification.

Especially for the reconstruction from point clouds with low densities, model-based reconstruction has the advantage that the roof geometry is always topologically correct, as Dorninger and Pfeifer (2008) indicate. In data-driven approaches, however, it is generally not guaranteed that "water-tight" models are constructed (Vosselman and Maas, 2010). The imposition of geometrical constraints (e.g. rectangularity, parallelism) on the roof geometry while estimating its geometric parameters ensures topological correctness, as Brenner (2005) reports. After the determination of roof faces by region growing on a digital surface model, Rottensteiner et al. (2005) use statistical tests to derive jump and step edges. In a following global adjustment of the parameters, constraints are included in order to get regularized models using uncertain projective geometry. Verma et al. (2006) reconstruct the roof topology graph and fit the building geometry with constraints, in particular constraints for roof planes (e.g. orientation).

The parameters of the final building model are estimated using Gauß–Newton optimization. Beside the detection of locally fitted plane primitives, Zhou and Neumann (2012) detect relations between roof and wall polygons, such as orientation, placement, parallelism as well as orthogonality and equality of roof planes. Brenner (2005) stresses that constraints generally should be used wherever possible.

As a rule, in most approaches the model selection problem is solved with information criteria, such as AIC, BIC and MDL. A promising alternative which is proposed in this paper is classification. A specifically robust classification method is Support Vector Machines (SVM), a state-of-the-art, supervised classifier (Vapnik, 1998; Schölkopf and Smola, 2002), which reduces the risk of overfitting and has a high generalization ability. Originally SVM was developed by Vapnik based on structural risk minimization (Vapnik, 1998). SVM use a special form of linear hyperplanes, maximum margin hyperplanes, for class separation. By maximizing the margin, the distance from the nearest point to the separating hyperplane, the generalization ability is enhanced. Non-linear class boundaries are modeled using Kernels, e.g. the radial based function kernel (Schölkopf and Smola, 2002). SVM has successfully been used for classification of spatial datasets (Lodha et al., 2006; Haitao et al., 2007; Huang et al., 2010; Römer and Plümer, 2010; Henn et al., 2012). In the context of roofs, SVM has been applied to hyperspectral and aerial images for classification of roof materials Braun et al. (2011a); Braun et al. (2011b).

1.2. Our contribution

In this paper we present a fully automatic, model-driven approach for building reconstruction for a large area based on sparse LIDAR point clouds and building footprints. The focus lies on the robust reconstruction of roofs, in particular simple roofs and roof parts. The reconstruction can even be achieved for sparse point clouds with a very low number of points per roof parts (e.g. $\leqslant 20$). The construction and implementation of roof models is shown for the standard roof types as defined in CityGML (Gröger et al., 2012). Our method assumes that an appropriate decomposition of the building footprint into rectangles is available. Basically any appropriate decomposition method can be used. Currently, we employ the method of Kada and McKinley (2009) because it provides a generalization of the footprint in addition. Note that in our approach the 2D footprint is decomposed into cells along the footprint segments with the help of 2D decomposition lines (Kada and McKinley, 2009). For a rectangle $\mathcal{R}$ and a model $\mathcal{M}(\boldsymbol{\theta})$, the optimal parameter values $\tilde{\boldsymbol{\theta}}$ are robustly estimated. To this end, we apply RANSAC (Fischler and Bolles, 1981). In contrast to existing approaches estimating geometrical primitives, such as planes, the proposed method models semantic objects such as roofs directly. From this point of view, roofs are aggregations of geometric primitives which satisfy a number of constraints, namely symmetry, parallelism and orthogonality with respect to the surrounding rectangle. Consequently, the estimated models are not only topologically correct, as in Dorninger and Pfeifer (2008), but they also satisfy symmetry and orthogonality constraints a-priori that are common to all roofs. The presented approach uses the constraints as prior knowledge for roof model definition and initialization. In contrast to Rottensteiner et al. (2005) or Zhou and Neumann (2012), the constraints are actively used to reduce the number of estimable parameters. In the sense of Rottensteiner (2006) these assumptions are "soft-constraints" since we include observations in model initialization only for parameters which cannot be covered by constraints. Using constraints in an active, controlled way has a number of significant advantages. First, the number of model parameters to be estimated is reduced considerably. This improves efficiency and robustness and facilitates the handling of outliers. With regard to RANSAC, the minimal number of observations needed to uniquely define a model is decreased. The search for the best model converges more rapidly. This allows the reliable reconstruction of roofs even in cases where data are rather sparse and noisy. As a final advantage, semantical constraints are satisfied from the beginning and have not been imposed as an additional feature. Even for point clouds with very few points an appropriate reconstruction can be achieved. Fig. 1 provides an example. RANSAC allows for the deriving of near optimal parameters for a given model. Even for point clouds with many outliers (e.g. points on dormers, chimneys) the model can be robustly estimated.

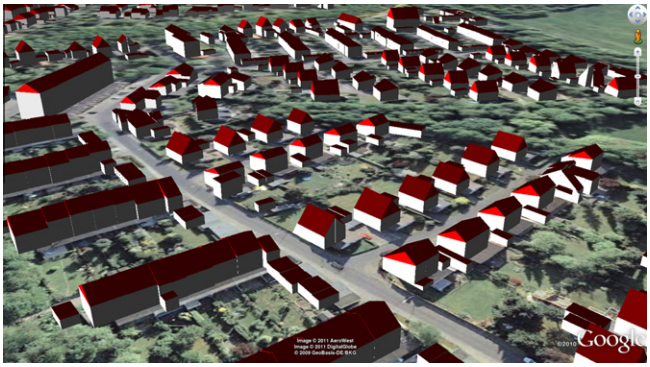


Fig. 1. Reconstruction of a suburb of Cologne, Germany, from LIDAR points, with an average density of $\approx$1.2 points/m^2 visualized in GoogleEarth. The 3D-model is derived from data provided by GEObasis.nrw.

Adequacy of this result depends on the suitability of the given model. Identification of the optimum model using as much information as available is crucial for this approach. In contrast to existing procedures, we approach model selection from a supervised classification point of view. In contrast to classical methods like AIC, BIC or MDL this allows us to incorporate a larger number of available features which may help to identify the best model. The incorporation of further features reduces the error rate, especially for shed roofs the error rate decreased by about 42% (see Section 3.3).

In contrast to Zhou and Neumann (2009) we achieve efficient storage of the data with the help of spatial databases, in particular PostGIS.

To the best of our knowledge, supervised learning, i.e. classification, has not been applied to model selection in the context of building or roof reconstruction up to the present.

The remainder of this paper is as follows: Our method for robust roof model classification will be presented in Section 2. In Section 3 the results of our method applied to a dataset of Cologne, Germany will be presented. The article ends in Section 4 with concluding remarks, a discussion of open questions and suggestions for further work.

2. A method for robust roof model classification

Airborne LIDAR datasets covering large areas, if they exist, are coarse or even sparse in most cases. On the other hand, building footprints of the authoritative real estate cadastre with high geometrical accuracy are available for nearly all regions. If cadastral data are not available they can be derived from freely-available datasets, such as OpenStreetMap or Google Maps. Their accuracy however, is certainly not as good as that of from cadastral datasets.

The histograms in Fig. 2 confirm the observation of sparseness, using a dataset from the federal state of North Rhine-Westphalia, Germany, as an example. It contains approx. 1/3 of North Rhine-Westphalia including the large cities of Cologne, Duisburg, Essen

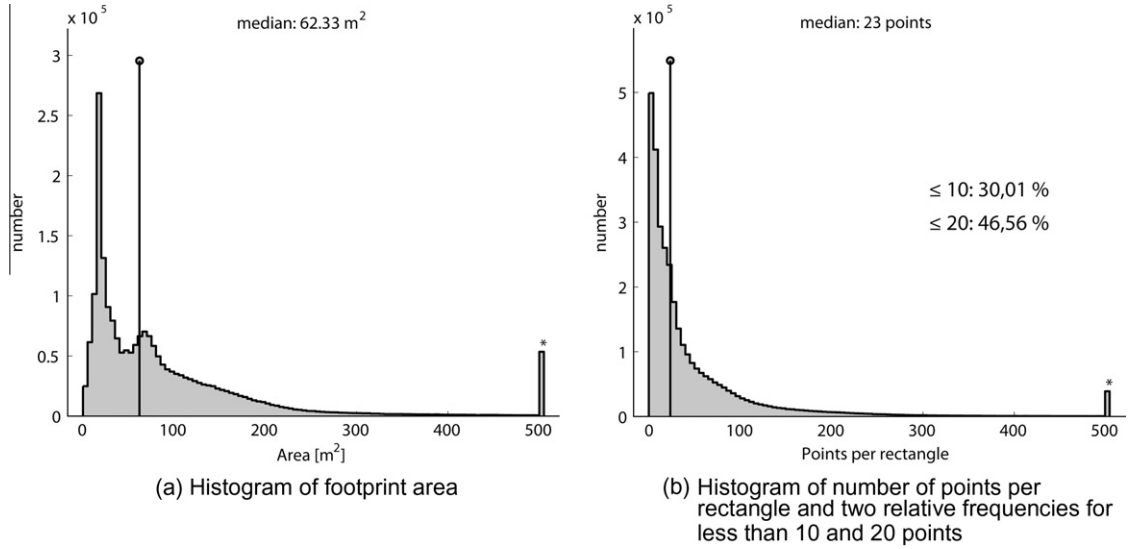


Fig. 2. Histograms of the used dataset of approx. one third of the federal state of North Rhine-Westphalia, Germany indicating the need for robust roof reconstruction method for sparse point clouds; $*$ indicates, that remaining values bigger than the maximum of the abscissa have been collected to one single bar.

as well as suburban regions. Fig. 2a shows, that most of the ≈2.2 mio. buildings are comparatively small. The median of the area is 62.33 m^2, leading at an average point density of ≈1 point/m^2 to approximately 62 points per footprint. However, the classification is made based on rectangles as a result of an appropriate footprint decomposition. The median number of points per rectangle (23 points, see Fig. 2b) is smaller than expected. In nearly 30% of the rectangles the number of points is less or equal than 10 and in 18.16% the number of points is less or equal than 5. Furthermore, outliers (e.g. points on walls, dormers, chimneys) reduce the number of points suitable for roof estimation. Hence, we are faced with very sparse point clouds in a significant number of cases.

In the following we present a method which robustly reconstructs the roof under the given circumstances.

2.1. Robust estimation of structured roof models

Roof model estimation is robustly performed by an enhanced version of RANSAC (Fischler and Bolles, 1981), implementing iteratively a hypothesize-verify-principle. RANSAC is a generic, highly robust algorithm for model estimation with the only requirement that the free model parameters can be instantiated by few observables. Its advantage of robustness (even if applied to datasets with more than 50% outliers) is evident if the number of estimable parameters is significantly lower than the number of observations (Schnabel et al., 2008). Model parameters can robustly be estimated even for small buildings with sparse LIDAR data. For the used dataset (see Fig. 2), the number of points is low in many cases. Algorithm 1 shows the pseudo-code of RANSAC and some adaptions.

Algorithm 1. Pseudo-code of RANSAC-based estimation of structural roof models.

Input: $\mathcal{R}$ ... Rectangle
Input: X ... LIDAR point cloud, clipped by $\mathcal{R}$; x ... 3D point
Input: $\mathcal{M}(\boldsymbol{\theta}, \mathcal{R})$ ... Roof model
Input: $\rho(x, \mathcal{M}(\boldsymbol{\theta}, \mathcal{R}))$ ... Cost function
Output: Robust estimation of model parameters $\tilde{\boldsymbol{\theta}}$
Output: Consensus set of the best model $CS(\tilde{\boldsymbol{\theta}})$

```
1  begin
2    k = 0;
3    Estimation of number N of iterations;
4    C_best = Inf;
5    while k < N do
6      k = k + 1;
7      Hypothesize – Draw randomly minimal sample S_K
       from X_C of size m, dependent of M
8      Hypothesize – Initialize model with drawn sample S_K,
       Result is M_0(θ_0, R)
9      Verify – Which points are consistent with model
       M_0(θ_0, R)?
       C = Σ_{x∈X_C} ρ(x, M_0(θ_0, R));
10     if C < C_best then
11       C_best = C;
12       θ̃ = θ_0;
13       CS(θ̃) = {x ∈ X_C : d(x; M(θ̃, R)) ⩽ ε};
14   end
```

In line 9 of Algorithm 1, the cost function of MSAC (M-Estimator Sample Consensus), as defined in Torr and Zisserman (2000) is used. In contrast to the original version of RANSAC scoring the number of inliers, MSAC minimizes the following cost function:

$$C = \sum_{x \in \boldsymbol{X}} \rho(x, \mathcal{M}(\boldsymbol{\theta}, \mathcal{R})) \tag{1}$$

with

$$\rho(x, \mathcal{M}(\boldsymbol{\theta}, \mathcal{R})) = \begin{cases} d(x; \mathcal{M}(\boldsymbol{\theta}, \mathcal{R}))^2 & |d(x; \mathcal{M}(\boldsymbol{\theta}, \mathcal{R}))| \leqslant \epsilon \\ \epsilon^2 & \text{otherwise} \end{cases} \tag{2}$$

The cost function of MSAC, based on M-estimators (Huber, 1981), is used to rank the consensus sets, the sets of points marked as inliers. Inliers are observations for which $|d(x; \mathcal{M}(\boldsymbol{\theta}, \mathcal{R}))| \leqslant \epsilon$ holds. Here, $d(x; \mathcal{M}(\boldsymbol{\theta}, \mathcal{R}))$ is the distance of x to the model $\mathcal{M}$, in particular the orthogonal distance to the corresponding roof plane. The points within the expected tolerance ϵ are consistent with the model and scored by their squared distance to the roof model, their fitness. In contrast, the outliers, the rest of the laser points, are penalized by a constant value ϵ^2.

In order to assess the number of iterations N (Line 3 of Algorithm 1), Fischler and Bolles (1981) provide a formula based on the probability of a point being an outlier and the probability "that at least one sample of m points is free from outliers" (Hartley and Zisserman, 2004). However, this measure has been proven to be too optimistic in our context, since the number of iterations is too small. Fischler and Bolles (1981) assume that the noise is distributed randomly. In our case however, the noise is highly structured (e.g. dormers or competitive models). This influence, though, is not estimable a-priori and cannot be learned on the fly. For this purpose, we empirically determined a generous upper bound of 1000 iterations.

The most important aspect of our usage of RANSAC is the specification of roof models. Roof models are *structured, aggregated objects* in contrast to existing RANSAC-based approaches, which use geometric primitives such as planes, lines and cylinders. For this purpose, notable examples are Tarsha-Kurdi et al. (2008) who presents an extended RANSAC algorithm for roof plane detection, Schnabel et al. (2008) using RANSAC for shape recognition and Beder and Förstner (2006) who use RANSAC for the estimation of cylinders in 3D point clouds. To the best of our knowledge, RANSAC has not been used so far in order to estimate structured roof models in such a way.

2.1.1. Definition of structured roof models

Roof models can be seen as an aggregation of planar meshes with a well-defined structure. This representation is by construction topologically consistent and well-defined since the roof face adjacencies are given a-priori. We observed that the space of different roof models has a underlying hierarchical structure, which is beneficial for the reconstruction and model selection process, as indicated in Brenner (2005). For example, a flat roof is a subtype of a roof with exactly one roof plane. The resulting hierarchy used here is shown in Fig. 3. Note that this hierarchy is intimately linked with the rectangular structure of single roof models.

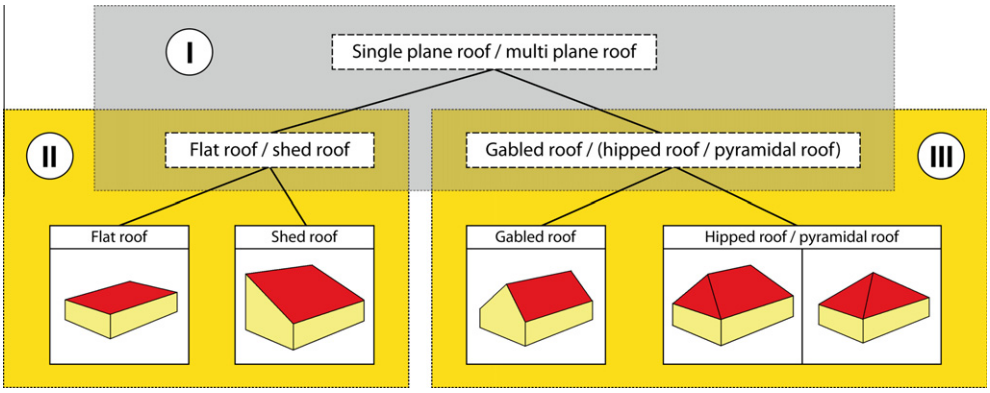


Fig. 3. Hierarchy for roof models.

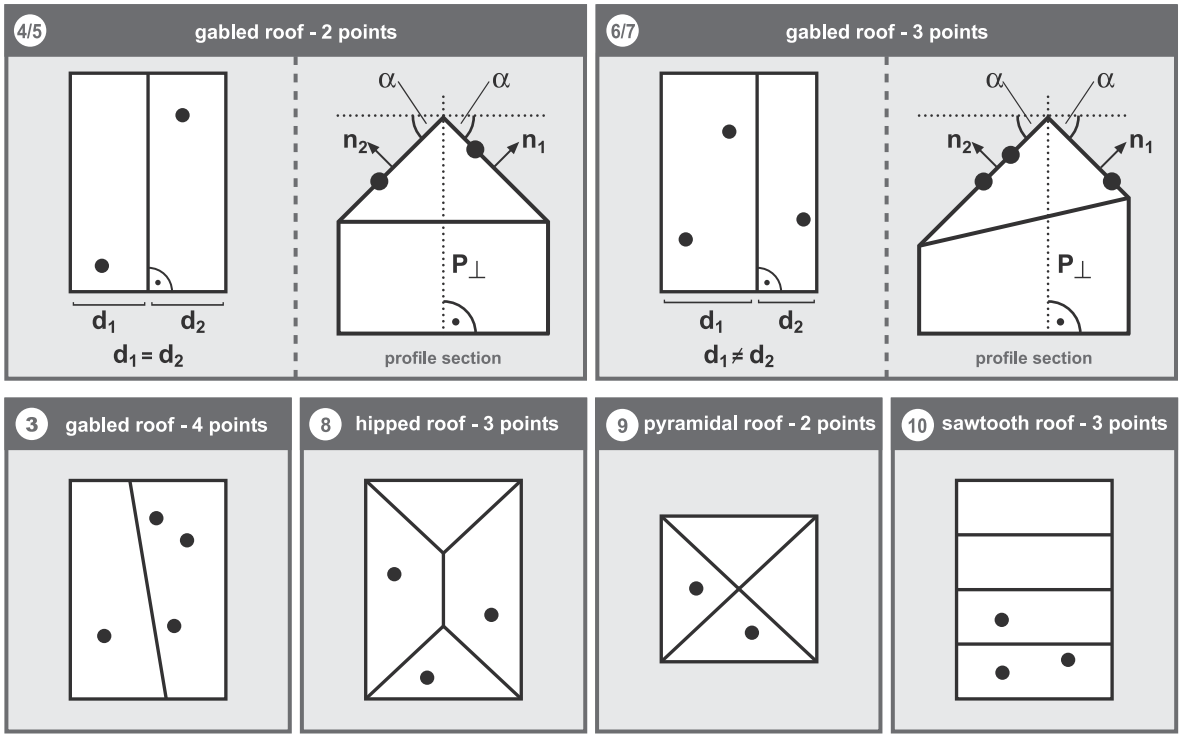


Fig. 4. Overview of the roof model catalog (without flat and shed roof) as defined in the international CityGML standard (Gröger et al., 2012) and points needed for initialization. In the first line the situation for two gabled roofs is shown (top-view/profile section). In both cases, the ridge is always orthogonal to a segment. For model 4/5 the ridge is also in the middle of the segment. Note that in all cases the marked points need not to be on the same roof faces.

Symmetry, perpendicularity and declination are highly relevant for roof reconstruction in order to achieve consistency of the model (Rottensteiner, 2006). In order to reduce the number of parameters, this knowledge is modeled a-priori, especially by representing the perpendicularity of the ridge to the longest/shortest edge of the corresponding rectangle and by representing the location of the ridge relative to the rectangle. Moreover, ridges are assumed to be horizontal and walls are assumed to be vertical and to share one edge with the rectangle. In the case of a gabled roof, the declination α of both roof planes is identical (as this is the case for nearly all existing gabled roofs). For hipped roofs in addition, the declinations of the hip planes (triangles in Fig. 4, model no. 8) have to be equal. These constraints are considered before initialization of the model and not in an optimization step after reconstruction of the building. By the use of these self-evident natural and common model assumptions, the roof model constructed by Algorithm 1 requires only a minimal number of points (see Fig. 4, where for example two points are needed to construct a gabled roof in the most prototypical case). The initialization of the model and the number of sample points for RANSAC-based estimation of structured roof models will be shown for some versions of gabled roofs and the sawtooth roof. Transfer to flat roof and shed roof is immediate.

The easiest case (standard gabled roof with perpendicular ridge in the middle of the shortest edge of a rectangle) is given in Fig. 4 (see model no. 4/5). The observables are the LIDAR points within the rectangle. The model has several interior and exterior parameters. The exterior parameters, translation and orientation, are already implied by the underlying rectangle, as well as the longest and the shortest edge (respectively the length and width of the roof). The remaining parameters of the most prototypical gabled roof model (model no. 4/5, see Fig. 4) considered here, the one with the most restrictions, are the declination of the two roof planes, the height and the location of the ridge. In this case, the location of the ridge is fixed by the constraint that the ridge is in the middle and orthogonal to one segment of the rectangle. Another constraint enforces that the declination of both roof planes is equal. The remaining interior parameters to be estimated from the point cloud are the declination α and the height of the ridge.

Under these conditions a model is already determined by two points. These two points must have different heights but need not to be on opposing planes. Using the plane in the axis of symmetry crossing the ridge $P_\perp$ as shown in Fig. 4 (profile section of model 4/5), the opposing point is mirrored. Consequently there are two points on one half of the roof. Herewith, a point on the ridge and hence the ridge height can be determined by crossing the line through the two points and the plane $P_\perp$. The declination α is the angle between the normal of the plane constructed by the two laser points and the calculated ridge point, and the Z-axis. The entire roof is constructed by rotating the constructed plane by 180° and subsequent intersection with the wall planes (planes orthogonal to the rectangle and sharing one edge with the rectangle). Now, all parameters and plane equations are known. By cutting the estimated roof planes with the wall planes the roof planes were bounded to meshes. Using these meshes the resulting boundary representation of the building can be created.

In contrast, model no. 6/7 in the middle of Fig. 4 does not assume that the ridge lies in the middle of the opposing edges of the rectangle. The distances d_1 and d_2 from the ridge to the left/right eave are not necessarily equal. An additional parameter has to be estimated resulting in the need of a further point. At first, by intersection of the line passing through two points and a wall plane (construction as described above) parallel to the fixed ridge, a point on the eave and the eave height can be determined. The first roof plane is constructed by these three points and the second one is simply the plane through the third point with the normal rotated by 180°.

The construction of the sawtooth roof model (see model no. 10 in Fig. 4) is only slightly more complicated than the construction of primitive roof models. At a first glance, it seems to be impossible to construct a model with an arbitrary number n of meshes as opposed to a model with a fixed number of meshes. However, RANSAC is also able to estimate models with a fixed but finite number of meshes, as long as they can be parametrized by a given number of parameters. Beside the exterior parameters (which are derived from the given rectangle) the interior parameters are the declination of a simple sawtooth, the number of sawteeth and the height of the ridge. The extent is given by the rectangle. The height of the ridge and the declination of a simple sawtooth can be calculated as explained above for the gabled roof. The estimation of the number of sawteeth remains to be done. The solution of this problem follows the work of Schmittwilken and Plümer (2010), and Schmittwilken (2012) in the context of stairs estimation from terrestrial laserscans. There, the sampling strategy for RANSAC has been adapted in order to search for samples in a terrestrial laser point cloud that most likely represent the stairs. In particular, stochastic universal sampling (Baker, 1987) has been used. In robotics this kind of sampling is known as *importance* sampling (Thrun et al., 2005). For every point, a fitness value, a "probability of being a good point", has been computed using non-parametric density estimation techniques. We use probability density functions (PDF) in order to estimate the rough distance to roof planes that are parallel to the initially constructed plane by two points (two points in lowest plane in Fig. 4, model no. 10). A third point lying on the plane adjacent and parallel to the initial plane is to be sampled. By use of the extrema of the PDF of the distances to the initial plane, an estimate of the number of steps can be obtained. Note that this is a mixed real-integer problem which can be solved by Constraint Logic Programming (Schmittwilken, 2012; Schmittwilken et al., 2007). By use of symmetry and parallelism, the sawtooth roof can be completely constructed.

The ability to construct quite complex models such as a sawtooth roof model shows the generality of our approach.

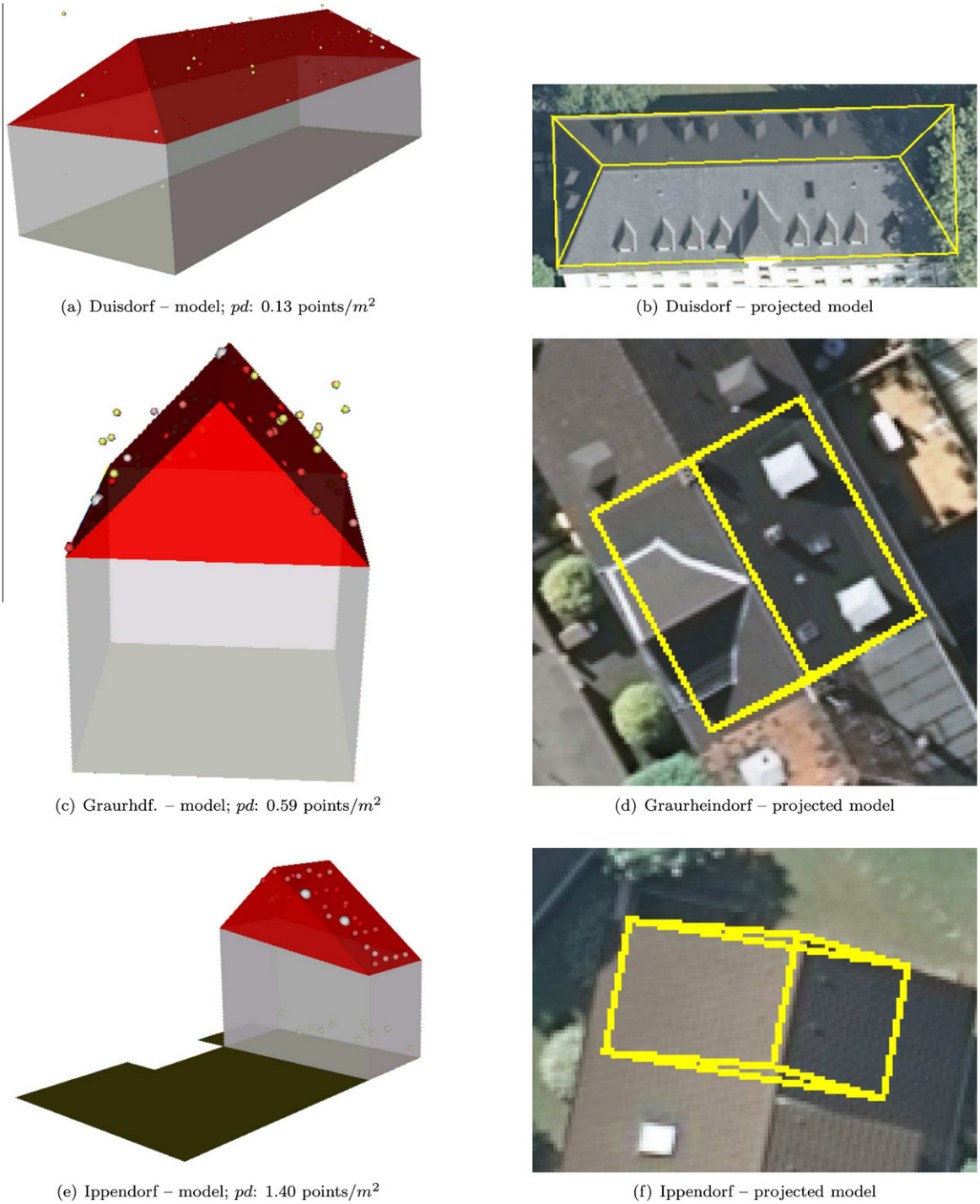
Some observations allow the early rejection of point samples. Points used to estimate parameters of a plane must not have the same height and must not be collinear (in case of a sample with more than two points). If the angle α of the roof plane(s) is too steep or the ridge height or the eave height is lower than the footprint height, the sample is rejected. For gabled roofs, the ridge has to intersect opposing segments of the rectangle. Pyramidal roofs can only be constructed if the rectangle is nearly quadratic.

As we have strong models with logical and sensible assumptions, the problem of low point density is alleviated. Table 1 gives an overview of the implemented roof model catalog.

Fig. 5 shows the estimated models of three exemplary selected buildings with different point densities. Even for buildings having a

Table 1
Roof model catalog with model numbers and number of points needed for model construction.

Model no.	Roof model	Number of points
1	Flat roof	1
2	Shed roof	3
3	Gabled roof (ridge arbitrary, horizontal ridge)	4
4	Gabled roof (ridge in middle+orthogonal to longest segment)	2
5	Gabled roof (ridge in middle+orthogonal to shortest segment)	2
6	Gabled roof (ridge orthogonal to longest segment)	3
7	Gabled roof (ridge orthogonal to shortest segment)	3
8	Hipped roof (symmetrical)	3
9	Pyramidal roof	2
10	Sawtooth roof	3



(a) Duisdorf – model; *pd*: 0.13 points/m^2

(b) Duisdorf – projected model

(c) Graurhdf. – model; *pd*: 0.59 points/m^2

(d) Graurheindorf – projected model

(e) Ippendorf – model; *pd*: 1.40 points/m^2

(f) Ippendorf – projected model

Fig. 5. Robust reconstruction of roofs for different situations and projection of the roof model onto an aerial image: (a) + (b) a very small number of points, resulting in low point density *pd*; (c) + (d) presence of outliers caused by huge dormers on the roof; (e) + (f) complex footprint has been decomposed into several rectangles resulting in a small number of points.

very low point density of 0.13 Points/m^2 (see example in Bonn-Duisdorf, Germany in Fig. 5a), a good estimation of the hipped roof has been achieved. In this case the asymmetrical gabled roof was confidently estimated.

After the robust estimation of every roof in the roof model catalog, the entire geometry of each candidate hypothesis is known. Several hypotheses are estimated for a single primitive or rectangle. Now we are faced with the problem – which of these hypotheses is the best one, given the point cloud, the model complexity and the shape of the footprint?

2.2. Classification of structured roof models

For the selection of the best model we applied a supervised machine learning approach. Basically, any supervised classifier could be used. The method Support Vector Machines (SVM) is particularly suitable because of their robustness and low risk of overfitting (Vapnik, 1998; Schölkopf and Smola, 2002). Originally, SVM is a linear classifier. In contrast to other linear classifiers (e.g. linear discriminant analysis), there are no assumptions on the underlying distribution of the data. SVM uses a special form of linear hyper-

planes, maximum margin hyperplanes, for class separation. By maximizing the margin, the distance of the nearest point to the separating hyperplane, the generalization ability is enhanced. For the solution of the optimization problem resulting from margin maximization, efficient solvers exist (such as sequential minimal optimization (Platt, 1999a)).

It is now of prime importance that this linear classifier can be transferred into a non-linear classifier using the kernel trick (Schölkopf and Smola, 2002). The non-linear classifier fits to the structure of the data automatically and non-linear class boundaries can be modeled. In principle the kernel trick consists in the definition of a new distance measure. Here, the radial based function kernel is a well established one.

Supervised machine learning methods, as well as the SVM, allow for the incorporation of a larger number of features, in contrast to classical model selection criteria. AIC, BIC and MDL can be regarded as special cases of supervised classification. In particular, AIC can be regarded as a linear classifier in two features with constant weight. In contrast, SVM cannot only deal with additional features, they can also model non-linear discriminants using Kernel functions.

2.2.1. AIC_C as a special case of a linear classifier

In order to understand why AIC_C, the small-sample adaption of the AIC (Hurvich and Tsai, 1989), can be regarded as a linear classifier with two features, consider distinct models $\mathcal{M}_+$ and $\mathcal{M}_-$ from which the "best" model is derived: AIC_C derives the model with minimal AIC_C value as the best one. AIC_C however only uses the information of the likelihood $\mathcal{L}(\hat{\theta}_i)$ and the number K_i of estimable model parameters for determination of the best model. The number n of observations is constant for all models. In contrast to supervised learning, attention is only restricted to the models of a current set, e.g. a set of roof models in a rectangle, are used for discrimination. We now rephrase AIC_C as a binary classification problem based on the standard definition:

$$AIC_C(\mathcal{M}_+) = -2\log(\mathcal{L}(\hat{\theta}_+)) + 2K_+ + \frac{2K_+(K_+ + 1)}{n - K_+ - 1}$$

$$AIC_C(\mathcal{M}_-) = -2\log(\mathcal{L}(\hat{\theta}_-)) + 2K_- + \frac{2K_-(K_- + 1)}{n - K_- - 1}$$

$$y = \begin{cases} +, & \text{if } AIC_C(\mathcal{M}_+) < AIC_C(\mathcal{M}_-) \\ -, & \text{otherwise.} \end{cases}$$

Table 3
F-Score-rating for the features in Table 2 divided by step in the hierarchy in Fig. 3.

Feature	F-Score – Rating
(a) Step I: single plane roof/multiple plane roof	
Inlierrate – onePlane	2.569559
Std – onePlane	1.638780
Declination – onePlane	0.581575
Median point cloud height	0.291285
Complexity – onePlane	0.052536
Slimness	0.040871
Area	0.020908
Complexity of footprint	0.004422
Degree of perpendicularity	0.000090
Azimuth	0.000034
Number of neighboring buildings	0.000009
(b) Step II: flat roof/shed roof	
Declination – onePlane	0.499380
Median point cloud height	0.030153
Std – onePlane	0.024276
Inlierrate – onePlane	0.011076
Complexity – onePlane	0.008662
Degree of perpendicularity	0.002411
Complexity of footprint	0.001207
Number of neighboring buildings	0.000811
Area	0.000120
Slimness	0.000013
(c) Step III: gabled roof/hipped roof	
Std – hipped	0.358398
Inlierrate – hipped	0.229233
Declination – hipped	0.057631
Inlierrate – gabled	0.044881
Declination – gabled	0.043834
Std – gabled	0.037393
Median point cloud height	0.024373
Number of neighboring buildings	0.023773
Complexity of footprint	0.004971
Slimness	0.004148
Complexity – hipped	0.004125
Degree of perpendicularity	0.003750
Complexity – gabled	0.003051
Area	0.002947
Azimuth	0.000669

Table 2
Features for roof type classification.

Feature	Algorithm 1	Step in Fig. 3	Description
Complexity of footprint			Number of nodes in footprint
Degree of perpendicularity			Divergence of polygon angles to 90°
Number of neighboring buildings			Number of topologically touching buildings
Azimuth angle	$\boldsymbol{x}_{FP}$	I, II, III	Azimuth angle of longest edge in rectangle
Slimness			ratio length/width of rectangle
Area			rectangle area
Median point cloud height			median height of laser points in rectangle
Complexity – onePlane	$\boldsymbol{x}_{C1,2}$	I, II	Complexity flat roof/shed roof
Complexity – gabled	$\boldsymbol{x}_{C3-7}$	III	Complexity gabled roof
Complexity – hipped	$\boldsymbol{x}_{C8,9}$	III	Complexity hipped roof
Inlier rate – onePlane		I, II	Inlier-rate (flat roof/shed roof)
Std – onePlane	$\boldsymbol{x}_{m_{1,2}}$	I, II	Standard deviation (flat roof/shed roof)
Declination – onePlane		I, II	Declination α of (flat roof/shed roof)
Inlier rate – gabled		III	Inlier-rate of gabled roof model
Std – gabled	$\boldsymbol{x}_{m_{3-7}}$	III	Standard deviation of gabled roof model
Declination – gabled		III	Declination α of gabled roof model
Inlier rate – hipped		III	Inlier-rate of hipped roof model
Std – hipped	$\boldsymbol{x}_{m_{8,9}}$	III	Standard deviation of hipped roof model
Declination – hipped		III	Declination α of hipped roof model

which can be transformed to:

$$y = \text{sign}(AIC_C(\mathcal{M}_+) - AIC_C(\mathcal{M}_-)) \quad (3)$$

Eq. (3) mimics the form at a linear classifier as defined in Bishop (2007):

$$y = \text{sign}(\boldsymbol{w}^T\boldsymbol{x} + b)$$

With Eq. (3) (offset $b = 0$) following representation can be obtained:

$$\boldsymbol{w}_{AICC}^T = [1 \; -1]$$

$$\boldsymbol{x} = \begin{bmatrix} AIC_C(\mathcal{M}_+) \\ AIC_C(\mathcal{M}_-) \end{bmatrix} = \begin{bmatrix} \overbrace{-2\log(\mathcal{L}(\hat{\theta}_+))}^{a_1^+} + \overbrace{2K_+ + \frac{2K_+(K_+ + 1)}{n - K_+ - 1}}^{a_2^+} \\ \underbrace{-2\log(\mathcal{L}(\hat{\theta}_-))}_{a_1^-} + \underbrace{2K_- + \frac{2K_-(K_- + 1)}{n - K_- - 1}}_{a_2^-} \end{bmatrix}$$

$$= \begin{bmatrix} a_1^+ + a_2^+ \\ a_1^- + a_2^- \end{bmatrix}$$

$$\Rightarrow y = \text{sign}\left(\boldsymbol{w}_{AICC}^T \begin{bmatrix} a_1^+ \\ a_1^- \end{bmatrix} + \boldsymbol{w}_{AICC}^T \begin{bmatrix} a_2^+ \\ a_2^- \end{bmatrix} \right)$$

The terms based on the likelihood and the model complexity in AIC_C formula have been split into a_1^i and a_2^i in order to emphasize that the weights $\boldsymbol{w}_{AICC}^T$ are constant terms. Consequently, AIC_C is a linear classifier in two features which does not fully exploit the information in the available data. In Section 3 we will demonstrate, that the proposed classification approach is superior to AIC_C in the given context.

2.2.2. Multi-model classification with SVM

The starting point of our approach is a set of labeled instances called training data. The label indicates the class, i.e. the specific roof type. Given attributes form a feature vector $\boldsymbol{x}$.

We observed that some characteristics of the footprint or rectangle suggest that the building has a specific roof type. A single rectangle with small area or low median point height is very likely to be a garage with a flat roof. The available features were split into those independent from the roof model $\boldsymbol{x}_{\text{FP}}$ (e.g. area of footprint, median height of points) and those depending on the roof model $\boldsymbol{x}_{\text{mi}}$ (e.g. inlier rate, complexity, standard deviation). The complexity term a_2^i was used as a feature value $\boldsymbol{x}_{\boldsymbol{Ci}}$. Model specific features derived by reconstruction attempts (see Section 2.1) were additionally incorporated. The standard deviation and the inlier rate of RANSAC were explicitly taken into account. An overview of the features used for roof type classification is given in Table 2.

The presence of more than two roof types results in a multi-classification problem. Many approaches reducing multi-classification to binary classification have been described in the literature, for example one-vs.-one or one-vs.-all (Bishop, 2007). In our case a sequential approach as in Rumpf et al. (2012), exploiting the hierarchical structure of the model space, proved to be most appropriate. Firstly, using this type of multi-class approach the hierarchical structure of model space (see Fig. 3) can be used. Secondly, it is possible to use different sets of features in each step. This makes the hierarchical classification superior to classical multi-class approaches in this context.

The first step is the discrimination between the roof models based on the number of roof planes. This discrimination is the root of the hierarchy (see root step I (gray) in Fig. 3) and is the discrimination of single plane roofs (*SPR*; flat roof and shed roof) vs. multiple plane roofs (*MPR*; gabled roof, hipped roof, etc.). For the instances classified as *SPR* (step II in Fig. 3) a further binary classification for discrimination between flat (*FR*) and shed roofs (*SR*) was learned. The instances in *MPR* contain, corresponding to our model catalog, gabled roof (*GR*), hipped roof (*HR*) and pyramidal roof (*PR*). Since a pyramidal roof is a special case of hipped roof (hipped roof with a ridge length of zero), they can be merged in the binary classification (*HRPR*, see step III in Fig. 3). The class of hipped roofs can be split into "real" hipped roofs and pyramidal roofs.

An overview of our proposed method for roof reconstruction based on a single rectangle and its corresponding point cloud is shown in Algorithm 2.

After estimation and reconstruction, the roof primitives within a building footprint may be aggregated. However, this is only performed if the parameters (e.g. height of the ridge/eave or declination) of neighboring roofs with the same type do not vary significantly.

Our approach, as it stands, is limited to roof models that are linear, i.e. consist of a small number of planar faces. Curved roofs which are typical for old castles or mosques, have not been included so far. This complies with the definition in the international standard CityGML. More complex, non-linear geometries (such as cones) can be approximated by a composition of planar faces in CityGML. This topic is outside the scope of this paper. However, as far as non-linear shapes can be specified by a limited number of parameters, extensions are straightforward. Another limitation lies in the assumption that roofs (and footprints) are orthogonally dominant. Non-orthogonal roofs/footprints are not covered by the rectangle decomposition we currently employ. Buildings for which the footprint cannot be decomposed are reconstructed as a flat roof with a uniform height (e.g. median height of clipped point cloud).

Algorithm 2. Roof model classification for rectangle and point cloud.

Input: $\mathcal{R}$ ... Rectangle
Input: X ... LIDAR point cloud, clipped by $\mathcal{R}$; x ... 3D point
Input: $\mathscr{M} = \mathcal{M}_1(\boldsymbol{\theta_1}, \mathcal{R}), \ldots, \mathcal{M}_m(\boldsymbol{\theta_m}, \mathcal{R})$ ... Roof model catalog
Input: $\mathcal{C}$ ... classifier
Output: most likely model $\widetilde{\mathcal{M}}$, corresponding probability $P(\widetilde{\mathcal{M}}|\boldsymbol{X}, \mathcal{R}, \mathcal{C})$ and corresponding standard deviation $\sigma_{\widetilde{\mathcal{M}}}$

begin
  initialize feature vector $\boldsymbol{x_b} = (\boldsymbol{x_{FP}}, \boldsymbol{x}_{\mathscr{M}_1}, \ldots, \boldsymbol{x}_{\mathscr{M}_l})^T$, see Table 2;
  determine model independent features x_{FP} (first block in Table 2);
  for $\mathcal{M}_i \in \mathscr{M}$ **do**
    robust estimation of $\mathcal{M}_i$; see Algorithm 1;
    determine model specific features $\boldsymbol{x}_{\mathscr{M}_i}$;
    determine model complexity $\boldsymbol{x}_{\boldsymbol{C_i}}$;
  group features corresponding to hierarchy; see Fig. 3 and Section 2.2;
  classify $\boldsymbol{x}_b$ with $\mathcal{C}$ to derive the best model $\widetilde{\mathcal{M}}$, $P(\widetilde{\mathcal{M}}|\boldsymbol{X}, \mathcal{R}, \mathcal{C}) = P(\widetilde{\mathcal{M}}|\boldsymbol{x_b})$ and $\sigma_{\widetilde{\mathcal{M}}}$
end

3. Experimental results

The dataset used in this study contained footprints and LIDAR point clouds of Cologne, Germany (1.13 points/m^2) and Oberhausen, Germany (1.08 points/m^2). As the footprints are cadastral datasets from the authoritative real estate cadastre in the federal state of North Rhine-Westphalia, Germany, they are assumed to have a high geometrical accuracy. The LIDAR points were derived from aerial surveys in around 2005 to 2007.

Model estimation was performed as described in Section 2.1. A standard deviation σ of 0.16 m was used for MSAC. The threshold ϵ

Table 4
Classification results; accuracies derived by 10 fold cross validation.

	True MPR	True SPR	Precision
(a) Step I: RBF-Kernel: C = 315.173, γ = 0.5359; w_{SPR} = 1.0943; w_{MPR} = 1; acc.: 95.29%; κ = 0.906			
pred. MPR	3056	170	94.73%
pred. SPR	146	3334	95.80%
Recall (%)	95.44	95.15	

(b) Step II: RBF-Kernel C = 3821.703, γ = 0.066; w_{FR} = 1; w_{SR} = 4.9508; acc.: 93.21%; κ = 0.755	True FR	True SR	Precision
pred. FR	2,803	125	95.73%
pred. SR	113	464	80.42%
Recall (%)	96.12	78.78	

(c) Step III: RBF-Kernel C = 955.426, γ = 1.310; w_{GR} = 1; w_{HRPR} = 10.0038; acc.: 96.47%; κ = 0.759	True GR	True HRPR	Precision
pred. GR	2,580	80	96.99%
pred. HRPR	21	180	89.55%
Recall (%)	99.19	69.23	

in MSAC was set to 3σ. All computations have been performed using a consumer-level PC (Intel Core i7 2.67 GHz CPU with 8 Gb RAM). All-in-all we needed ≈6 s per building for classification and reconstruction, including import and export from/to the database.

The training dataset of 6696 roofs has been selected from a large dataset of ≈2.2 million buildings. Each example in the training dataset has been labeled manually. For that purpose all reconstructed models have been projected onto an aerial image and afterwards they have been assessed and labeled, taking the corresponding point cloud into account. The selection of the training examples considered roofs with different heights as well as the footprint area. Hence the spectrum of roofs and buildings is covered as much as possible and a robust set of labeled buildings is available.

This training dataset was used for F-Score – Rating in Section 3.1 and for the training of the SVM models in Section 3.2.

3.1. Feature rating

In order to obtain the relevance of the features for roof type classification, they were ranked using F-Score – Rating (Chen and Lin, 2006). F-score – Rating is based on the F-Score, also known as the F_1 measure. Consequently, the discrimination of the positive and negative set in our binary classification problems on each step was measured. Precision and recall are taken into account since they are more appropriate than accuracy in case of unbalanced class distributions. This is the case in our context since certain roof types appear more often then others, for example gabled roofs compared to hipped roofs. After determination of the F-Scores per feature, the features below a certain threshold (i.e. less than 5% influence compared to feature with maximal F-Score) were dropped. With the remaining ones, the accuracy was determined using 10-fold cross validation (including the optimization process of the SVM parameters C and γ). The set of feature with the lowest classification error was used for further classifications.

In all steps, standard deviations and inlier rates resulting from the MSAC estimation have proved to be crucial for the classification of the roof type (see Table 3). Furthermore, the median height of the point cloud as further additional information has been relevant in every step. This confirms the assumption that the height of a building helps to discriminate between the roof types. The complexity term of AIC_C has been found to be less informative than the term including the residuals. However, building specific features have not been as significant as expected.

In the first step, classification of single plane roof (*SPR*) against multiple plane roof (*MPR*), the inlier rate and the standard deviation of the flat (shed) roof model dominated. Obviously, in the case of a gabled roof one makes a huge mistake when estimating this to be a flat (shed) roof. The choice of the wrong model leads to large residuals, a high standard deviation and a low percentage of inliers.

The declination of the roof plane turned out to be the most relevant feature in step II, flat roof (*FR*) vs. shed roof (*SR*). In reality, the declination is the indicator discriminating between flat roofs and shed roofs.

Owing to the residuals made at the additional roof planes of hipped roofs when a gabled roof has been estimated, the features of the hipped roof model have been the most relevant for discrimination of gabled roof vs. hipped (and pyramidal) roofs.

Table 3(a)–(c) shows the features and their ranking separated for each step in the hierarchy. The ranked features were used after all to determine subsets of features, leading to a higher classification accuracy (Chen and Lin, 2006). For the subsets (*italic* in Table 3(a)–(c)) the final SVM models were trained and validated. The results are presented in the next section.

3.2. Classification results

Before the SVM is used for learning purposes one has to choose the kernel function. In our case we selected the RBF-Kernel. This kernel needs a parameter γ in order to specify the influence of the support vector on the other examples. Using Soft-Margin SVM (Schölkopf and Smola, 2002), the penalty parameter C has to be chosen, too. However, this two parameters of the SVM are not known a-priori. For that purpose they were determined using a grid parameter optimization, as proposed in Hsu et al. (2010). The parameter set with the lowest error rate (10-fold cross validation) has been selected for building the SVM-Model (training). As a toolbox, we used the libSVM implementation of Chang and Lin (2011). Note that this optimization and training procedure was performed for steps I–III.

Prior knowledge of the roof type frequency has been included in the training process by weights w_i (see e.g. (Chang and Lin, 2011; Vapnik, 1998)) for every roof type i. An estimate for the weights has been determined from the training data. Table 4 presents the results of the roof type classification. It has to be mentioned here, that Table 4a–c[1] shows the results of stand-alone learning and validation phases.

[1] We used SVM with RBF-Kernel. Basically the classifier is freely selectable. The results (accuracy) of the applications of selected state-of-the-art methods (decision tree (DT; e.g. in Quinlan (1993)), Naive Bayes (NB; e.g. in Bishop (2007)), SVM with linear kernel (linSVM)) were as follows: step I: DT: 93.89%, NB: 94.20%, linSVM: 94.08%; step II: DT: 90.76%, NB: 87.93%, linSVM: 89.07%; step III: DT: 95.49%, NB: 94.09%, linSVM: 97.52%

The confusion matrix in Table 4a indicates that the first classification step of the hierarchy can be calculated highly successfully. The achieved accuracy was 95.29% and the κ-value, a measure of agreement of label and prediction under consideration of class frequencies (Cohen, 1960), emphasizes the good classification quality. A high level of accuracy is vital here, because for classification of unseen instances the following classification is chosen based on this decision.

For the classification of flat roofs as compared with shed roofs, a lower degree of accuracy was manifested. Shed roofs were not identified well. For that reason, the κ-coefficient is comparatively low. The low classification performance originates from the small number of points available for model estimation. Sometimes, each roof plane of a gabled roof is regarded as an independent roof and so they are estimated as shed roofs.

Table 4c presents the results of the classification of gabled roof vs. the set of hipped and pyramidal roofs. Here, also a high level of accuracy was achieved and the κ-coefficient in addition confirms this. However, hipped/pyramidal roofs have not been predicted as well as gabled roofs (only 69.23%). Certainly, the precision of hipped/pyramidal roofs was high. So, the classifier preferred the class gabled roof resulting from integrating semi-hipped roofs in the training data.

The classification accuracy measures the appropriateness of the derived model, i.e. the roof type. The estimation of the geometrical accuracy, i.e. the precision of the derived model parameters, affords ground truth data, which were not available. For that purpose, the root mean square error (RMSE) of the best classified models was computed. This quantity was used as a measure of geometrical correctness or internal accuracy. Fig. 6 shows the histogram of the RMSE of the classified models for the dataset of Cologne, Germany. The mean RMSE $\mu_{RMSE} = 0.1468$ meters was slightly lower than the standard deviation assumed in MSAC.

For the classification of new buildings we used the predicted class labels at each step and a combination of Platt's posterior probabilities (Platt, 1999b). According an empirical study in Duan and Keerthi (2005), Platt's posterior probabilities deliver a good measurement for the probability of a sample belonging to a certain class. The resulting posterior probability of a roof model $P(\widetilde{\mathcal{M}}|\boldsymbol{X}, \mathcal{R}, \mathcal{C})$ was calculated by the probability of the first step multiplied by the probability of the corresponding left/right step in Fig. 3. Fig. 7 shows an exemplary reconstruction for Cologne, Germany.

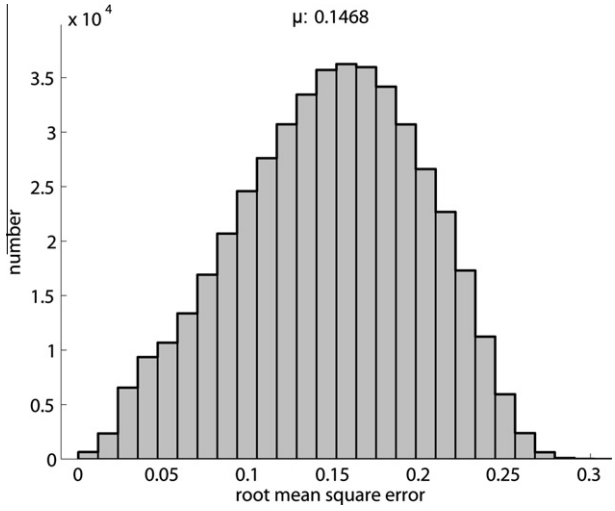


Fig. 6. Histogram of the root mean squared error for the best classified models for dataset Cologne, Germany.

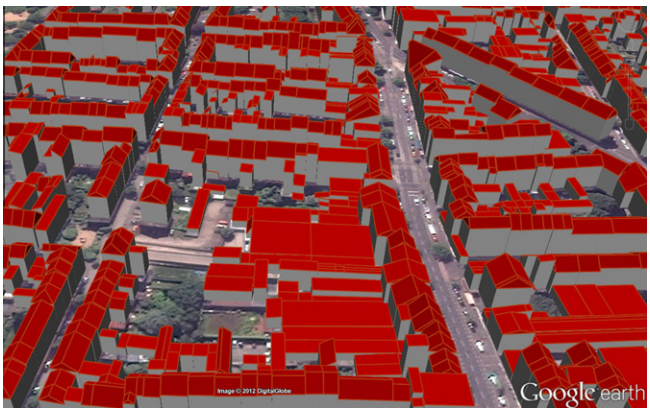


Fig. 7. Reconstruction of an urban scene in Cologne, Germany, visualized in GoogleEarth. The 3D-model is derived from data provided by GEObasis.nrw.

Table 5
Comparison of AIC_C vs. SVM – achieved accuracy and recalls for the validation dataset of 1050 rectangles in Oberhausen, Germany.

	SVM	AICc
Recall flat roof	94.50	91.01
Recall shed roof	82.61	40.43
Recall gabled roof	96.98	96.22
Recall hipped roof/pyramidal roof	70.00	70.00
Accuracy	95.77	92.10

In order to improve the final estimation of the roof model, a Gauß-Helmert-model (Niemeier, 2008) should be used in order to improve the estimation of the roof parameters under consideration of the point's uncertainty. For that purpose, the results of the RANSAC estimation could be used as initial values of the roof parameters.

There is a strong dependency between the footprint decomposition and the roof model estimation. If the decomposition fails or leads to rectangles which do not correspond to the roof structure, the roof model estimation will not be accurate. In this case, an improved version of the footprint decomposition method would lead to better results.

3.3. Comparison of model selection with AIC/AIC_C against SVM

In order to compare the accuracy of our approach to the results of the stand-alone AIC_C - rating, a validation dataset of 1050 rectangles in Oberhausen, Germany has been created. On the one hand, the buildings were classified using the SVM and on the other hand using the (grouped) AIC_C providing its best model. Table 5 shows the recalls and the overall accuracy obtained by applying the SVM models of Section 3.2 and AIC_C classification to the validation dataset. Please note, that in contrast to Section 3.2, the class recalls do not refer to the step as such, but to the confusion matrices obtained by application of the complete method.

The SVM method reduced the error rate for almost all roof types. Especially for shed roofs, the error rate has been significantly reduced. Only for hipped roofs did the error rate remain unchanged. All-in-all our assumption that incorporating further features leads to a better classification of roof types was vindicated.

4. Conclusions

In this article, a new method for the robust and fully automatic estimation, classification and reconstruction of LoD2 building models from sparse LIDAR data is presented. Based on a segmenta-

tion of the footprints into rectangles, roof models from a catalog of standard roof types provided by the international standard CityGML are derived. The model-based approach can cope with sparse LIDAR data (up to a point density of 1 point/m^2) which are widely available, enabling the automatic derivation of LoD2 models covering the area of large regions, such as the territory of a hole country. Our roof models ensure constraints which are observable for real-world roofs, such as symmetry and horizontal ridges. The RANSAC procedure, a stochastical method being robust to outliers, was used to estimate the roof models. With the help of the constraints, the number of the models' parameters have been significantly reduced, increasing the robustness of RANSAC. The classification problem for roofs was solved by applying supervised machine learning methods, which are superior to state of the art model selection methods such as AIC_C or MDL in practice as well as conceptually. In particular, we showed that those methods are a special case of our method, which in addition incorporates further information such as the roof inclination or the median point cloud height, and learned the coefficients (significance) of features from manually labelled data. In AIC_C, those coefficients are fixed. A hierarchical classification approach allows for successively refining the results by considering more features in each step and again contributes to the robustness of the method. To the best of our knowledge, the combination of model-based approaches, robust stochastical estimation methods such as RANSAC and supervised machine learning methods have not been applied in the field of roof reconstruction so far. The theoretical advantages of our method are also corroborated by the empirical results: For a data set that is independent of the training data set, the accuracy (percentage where the predicted and the actual roof type coincides) is about 95%.

The next step will be the improvement of the segmentation of the footprints by improving the incorporation of the distribution of the laser points. The combination of a model-based approach, stochastical methods, such as clustering, and prior information about roof characteristics (e.g. slope, length, width) represented as mixture models will allow for the segmentation of complex roofs into simple roof models. This will be the topic of a subsequent paper.

The method is designed to cope with sparse data, but in some cases the laser point density is not sufficient to derive reliable results. An extension of the algorithm described here exploits the roofs in the local neighborhood in order to predict the most probable roof type and roof parameters. We suggest a non-parametric maximum a posteriori (MAP) classification procedure using the neighborhood as prior information. As features, simple geometrical properties of the neighboring buildings (length/width of the footprint; rough height of the building) are used. The likelihoods can be determined by class conditional multivariate kernel density estimation using the geometric features.

Further improvements of this method are achieved by the extension of the roof model catalog by e.g., semi-hipped roofs or connectors. These models also will be included in our hierarchical approach. Furthermore interfaces between neighboring roof models resulting from the intersection of neighboring 3D-meshes or 3D-lines deserve to be considered more in detail. In addition, the roof parameters of neighboring buildings described in this article (e.g. height, ridge position, declination) can be harmonized in order to obtain a more homogeneous and realistic visualization.

Acknowledgements

The authors appreciate the helpful and valuable comments given by the anonymous reviewers. This work was partially funded by the German Federal Ministry of Economics and Technology by a Central Innovation Program (ZIM; No. KF2212501SS9) due to a resolution of the German Bundestag. This work has benefited from support of the German Research Foundation (DFG) in several projects in the past and present, namely "Semantic Modeling for the Acquisition of Topographic Information from Images and Maps", "Ontological scales for automated detection, efficient processing and fast visualization of landscape models" (FO 180-10) and "Mapping on Demand" (FOR 1505, PL 188/10-1). We thank GEObasis.nrw (surveying and mapping authority of the federal state of North Rhine-Westphalia) for providing LIDAR data and building footprints. We thank Youness Dehbi for valuable cooperation and Michael Kneuper for his assistance in preparing the illustrations.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19 (6), 716–723.

Ameri, B., Fritsch, D., 2000. Automatic 3D building reconstruction using plane roof structures. In: Proc. of ASPRS Annual Conference, Washington, DC, 21–26 May.

Baker, J.E., 1987. Reducing bias and inefficiency in the selection algorithm. In: Proc. 2nd International Conference on Genetic Algorithms (ICGA), Cambridge, MA, USA, July, pp. 14–21.

Beder, C., Förstner, W., 2006. Direct solutions for computing cylinders from minimal sets of 3D points. In: Leonardis, A., Bischof, H., Pinz, A. (Eds.), Proc. of European Conference on Computer Vision (ECCV), Graz, Austria, 7–13 May, pp. 135–146.

Bishop, C.M., 2007. Pattern Recognition and Machine Learning. Information science and statistics. Springer, New York.

Braun, A.C., Weidner, U., Hinz, S., 2011a. Classifying roof materials using data fusion through kernel composition – comparing ν-SVM and one-class SVM. In: Proc. of JURSE – Joint Urban Remote Sensing Event, Munich, Germany, 11–13 April, pp. 377–380.

Braun, A.C., Weidner, U., Jutzi, B., Hinz, S., 2011a. Integrating model knowledge into SVM classification – Fusing hyperspectral and laserscanning data by kernel composition. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 4/W19), 57–62.

Brenner, C., 2000. Towards fully automatic generation of city models. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 33 (Part B3/1), 85–92.

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. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 32 (Part 5), 323–330.

Chang, C.C., Lin, C.J., 2011. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2 (3), 27:1–27:27, <http://www.csie.ntu.edu.tw/ cjlin/libsvm> (accessed 15.11.12)..

Chen, Y.W., Lin, C.J., 2006. Combining SVMs with various feature selection strategies. In: Guyon, I., Nikravesh, M., Gunn, S., Zadeh, L. (Eds.), Feature Extraction. Springer Berlin/Heidelberg, Germany, pp. 315–324.

Cohen, J., 1960. A coefficent of agreement for nominal scales. Educational and Psychological Measurement 20 (1), 37–46.

Cord, M., Declercq, D., 1999. Bayesian model identification: application to building reconstruction in aerial imagery. In: Proc. of International Conference on Image Processing (ICIP '99), Kobe, Japan, 24–28 October, pp. 140–144.

Czerwinski, A., Sandmann, S., Stöcker-Meier, E., Plümer, L., 2007. Sustainable SDI for EU noise mapping in NRW – best practice for INSPIRE. International Journal for Spatial Data Infrastructure Research 2 (1), 90–111.

Döllner, J., Baumann, K., Buchholz, H., 2006. Virtual 3D city models as foundation of complex urban information spaces. In: Schrenk, M. (Ed.), Proc. 11. International conference on Urban Planning and Spatial Development in the Information Society (REAL CORP). CORP Competence Center of Urban and Regional Planning, Vienna, Austria, pp. 107–112.

Dorninger, P., Pfeifer, N., 2008. A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. Sensors 8 (11), 7323–7343.

Duan, K.B., Keerthi, S.S., 2005. Which is the best multiclass SVM method? An empirical study. In: Proc of 6th International Workshop on Multiple Classifier Systems (MCS), Seaside, CA, USA, June 13–15, pp. 732–760.

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.

Gröger, G., Kolbe, T.H., Nagel, C., Häfele, K.H., 2012. OpenGIS City Geography Markup Language (CityGML) Encoding Standard. Version 2.0.0, Open Geospatial Consortium, OGC Doc. No. 12-019.

Haala, N., Brenner, C., Anders, K.H., 1998. 3D urban GIS from laser altimeter and 2D map data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 32 (Part 3/1), 339–346.

Haala, N., Kada, M., 2010. An update on automatic 3D building reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing 65 (6), 570–580.

Haitao, L., Haiyan, G., Yanshun, H., Jinghui, Y., 2007. Fusion of high-resolution aerial imagery and LIDAR data for object-oriented urban land-cover classification based on SVM. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 4/W54), 179–184.

Hartley, R.I., Zisserman, A., 2004. Multiple View Geometry in Computer Vision, second ed. Cambridge University Press, New York, USA.
Henn, A., Römer, C., Gröger, G., Plümer, L., 2012. Automatic classification of building types in 3D city models – using SVMs for semantic enrichment of low resolution building data. GeoInformatica 16 (2), 281–306.
Hsu, C.W., Chang, C.C., Lin., C.J., 2010. A practical guide to support vector classification. Technical Report. Department of Computer Science, National Taiwan University.
Hu, J., You, S., Neumann, U., 2003. Approaches to large-scale urban modeling. IEEE Computer Graphics and Applications 23 (6), 62–69.
Huang, B., Xie, C., Tay, R., 2010. Support Vector Machines for urban growth modelling. GeoInformatica 14 (1), 83–99.
Huang, H., Brenner, C., Sester, M., 2011. 3D building roof reconstruction from point clouds via generative models. In: Proc. of 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Chicago, Illinois, USA, 1–4 November, pp. 16–24.
Huber, P.J., 1981. Robust Statistics. John Wiley & Sons, New York, USA.
Hurvich, C.M., Tsai, C.L., 1989. Regression and time series model selection in small samples. Biometrika 76 (2), 297–307.
Kada, M., 2007. Scale-dependent simplification of 3D building models based on cell decomposition and primitive instancing. In: Proc. of the International Conference on Spatial Information Theory: COSIT 07, Melbourne, Australia, 19–23 September, pp. 222–237.
Kada, M., McKinley, L., 2009. 3D building reconstruction from LIDAR based on a cell decomposition approach. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3/W4), 47–52.
Knapp, S., Coors, V., 2008. The use of eparticipation in public participation: the VEPs example. In: Coors, V., Rumor, M., Fendel, E., Zlatanova, S. (Eds.), Proc. of the Urban Data Management Society Symposium 2007 (UDMS Annual 2007), Stuttgart, Germany, 10–12 October, pp. 93–104.
Lafarge, F., Descombes, X., Zerubia, J., Pierrot-Deseilligny, M., 2010. Structural approach for building reconstruction from a single DSM. IEEE Transactions on Pattern Analysis and Machine Intelligence 32 (1), 135–147.
Lafarge, F., Mallet, C., 2012. Creating large-scale city models from 3D-point clouds: a robust approach with hybrid representation. International Journal of Computer Vision 99 (1), 69–85.
Lodha, K.S., Kreps, J.E., Helmbold, P.D., Fitzpatrick, M.D., 2006. Aerial LiDAR data classification using support vector machines (SVM). In: Proc. of 3rd International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT). Chapel Hill, NC, USA, pp. 567–574, 14–16 June.
Maas, H.G., Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), 153–163.
Niemeier, W., 2008. Ausgleichungsrechnung: Statistische Auswertemethoden. de Gruyter, Berlin, Germany (in German).
Oude Elberink, S., 2009. Target graph matching for building reconstruction. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3/W8), 49–54.
Overby, J., Bodum, L., Kjems, E., Ilsø, P.M., 2004. Automatic 3D building reconstruction from airborne laser scanning and cadastral data using hough transform. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 35 (Part B3), 296–301.
Platt, J.C., 1999a. Fast training of support vector machines using sequential minimal optimization. In: Schölkopf, B., Burges, C., Smola, A.J. (Eds.), A dvances in Kernel Methods – Support Vector Learning. MIT Press, Cambridge, Cambridge, MA, USA, pp. 185–208.
Platt, J.C., 1999b. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In: Smola, A.J., Bartlett, P., Schölkopf, B., Schuurmans, D. (Eds.), Advances In Large-Margin Classifiers. MIT Press, Cambridge, MA, USA, pp. 61–74.
Poullis, C., You, S., 2009. Photorealistic large-scale urban city model reconstruction. IEEE Transactions on Visualization and Computer Graphics 15 (4), 654–669.
Quinlan, J.R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann, San Mateo, CA, USA.
Römer, C., Plümer, L., 2010. Identifying Architectural Style in 3D City Models with Support Vector Machines. Photogrammetrie, Fernerkundung, Geoinformation (PFG) 5, 369–384.
Rottensteiner, F., 2006. Consistent estimation of building parameters considering geometric regularities by soft constraints. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3), 93–106.
Rottensteiner, F., Briese, C., 2003. Automatic generation of building models from LIDAR data and the integration of aerial images. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part 3-W13), 174–180.
Rottensteiner, F., Trinder, J., Clode, S., Kubik, K., 2005. Automated Delineation of Roof Planes from LIDAR Data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W19), 221–226.
Rumpf, T., Römer, C., Weis, M., Sökefeld, M., Gerhards, R., Plümer, L., 2012. Sequential support vector machine classification for small-grain weed species discrimination with special regard to cirsium arvense and galium aparine. Computers and Electronics in Agriculture 80, 89–96.
Schmittwilken, J., 2012. Attributierte Grammatiken zur Rekonstruktion und Interpretation von Fassaden. Ph.D. thesis. Institute for Geodesy and Geoinformation, University of Bonn, Bonn, Germany. <http://hss.ulb.uni-bonn.de/2012/2780/2780.htm> (in German).
Schmittwilken, J., Plümer, L., 2010. Model-based reconstruction and classification of facade parts in 3D point clouds. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3A), 269–274.
Schmittwilken, J., Saatkamp, J., Förstner, W., Kolbe, T.H., Plümer, L., 2007. A Semantic Model of Stairs in Building Collars. Photogrammetrie, Fernerkundung, Geoinformation (PFG) 6, 415–427.
Schnabel, R., Wessel, R., Wahl, R., Klein, R., 2008. Shape recognition in 3D point-clouds. In: Skala, V. (Ed.), Proc. of the 16-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Plzen-Bory, Czech Republic, 4–7 February, pp. 65–72.
Schölkopf, B., Smola, A.J., 2002. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.. MIT Press, Cambridge, USA.
Scholze, S., Moons, T., van Gool, L., 2001. A probabilistic approach to roof patch extraction and reconstruction. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part 3B), 195–204.
Suveg, I., Vosselman, G., 2004. Reconstruction of 3D building models from aerial images and maps. ISPRS Journal of Photogrammetry and Remote Sensing 58 (3-4), 202–224.
Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P., 2008. Extended RANSAC algorithm for automatic detection of building roof planes from LIDAR data. The Photogrammetric Journal of Finland 21 (1), 97–109.
Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P., Koehl, M., 2007. Model-driven and data-driven approaches using LIDAR data : analysis and comparison. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W49A), 87–92.
Thrun, S., Burgard, W., Fox, D., 2005. Probabilistic Robotics. The MIT Press, Cambridge, USA.
Torr, P.H.S., Zisserman, A., 2000. MLESAC: a new robust estimator with application to estimating image geometry. Journal of Computer Vision and Image Understanding 78 (1), 138–156.
Vallet, B., Pierrot-Deseilligny, M., Boldo, D., Brédif, M., 2011. Building footprint database improvement for 3D reconstruction: a split and merge approach and its evaluation. ISPRS Journal of Photogrammetry and Remote Sensing 66 (5), 732–742.
Vapnik, V.N., 1998. Statistical Learning Theory. A Wiley-Interscience publication, Wiley, New York, USA.
Verma, V., Kumar, R., Hsu, S.C., 2006. 3D building detection and modeling from aerial LIDAR data. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), New York, NY, USA, 17–22 June, pp. 2213–2220.
Vosselman, G., 2012. Automated planimetric quality control in high accuracy airborne laser scanning surveys. ISPRS Journal of Photogrammetry and Remote Sensing 74, 90–100.
Vosselman, G., Dijkman, E., 2001. 3D building model reconstruction from point clouds and ground plans. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part 3/W4), 37–43.
Vosselman, G., Maas, H.G., 2010. Airborne and Terrestial Laser Scanning. Whittles Publishing, CRC Press LLC, Boca Raton, FL, USA.
Zhou, Q.Y., Neumann, U., 2009. A streaming framework for seamless building reconstruction from large-scale aerial LiDAR data. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, 20–25 June, pp. 2759–2766.
Zhou, Q.Y., Neumann, U., 2012. 2.5D building modeling by discovering global regularities. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, USA, 26–21 June, pp. 326–333.