

Quality analysis on 3D building models reconstructed from airborne laser scanning data

Sander Oude Elberink*, George Vosselman

Faculty of Geo-Information Science and Earth Observation, University of Twente, Hengelosestraat 99, P.O. Box 6, 7500 AA, Enschede, The Netherlands

ARTICLE INFO

Article history:

Available online 2 February 2011

Keywords:

Accuracy
Analysis
Three-dimensional
Building
Lidar

ABSTRACT

This paper presents a method to assess the geometric quality of 3D building models. The quality depends on properties of the input data and the processing steps. Insight in the quality of 3D models is important for users to judge whether the models can be used in their specific applications. Without a proper quality description it is likely that the building models are either treated as correct or considered as useless because the quality is unknown. In our research we analyse how the quality parameters of the input data affect the quality of the 3D models. The 3D models have been reconstructed from dense airborne laser scanner data of about 20 pts/m². A target based graph matching approach has been used to relate specific data features to general building knowledge. The paper presents a theoretical and an empirical approach to identify strong parts and shortcomings in 3D building models reconstructed from airborne laser scanning data without the use of reference measurements. Our method is tested on three different scenes to show that a proper quality description is essential to correctly judge the quality of the models.

© 2011 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

Reconstructing buildings in 3D has been a challenging research topic for at least ten years, and will be in future as long as acquisition systems are improving and model requirements are increasing. The tendency is that the reconstructed building models become more realistic and more detailed. 3D building models are requested as an input source for a variety of applications, such as urban planning, shadow modelling and analysing solar potential from roof directions. Recent developments in the production of 3D building models are the increasing capabilities of sensor data and the higher demands from the users' perspective. This leads to an increasing need to produce highly detailed 3D building models. Several researchers have described methods to speed up the production of 3D building models, such as roof models in Kada and McKinley (2009) and Dorninger and Pfeifer (2008), and building façades in Becker (2009), Ripperda and Brenner (2009) and Pu and Vosselman (2009). Most of the reconstruction methods work well if both the data is complete and the objects in the scene fit to assumptions made in the algorithm. However, the quality of 3D building models is actually determined by those situations where either data is lacking or the objects do not fit to general assumptions. Becker (2009) presents an approach for façade reconstruction that is flexible to variations in data quality

or even absence of data. The key element in their approach is the ability to use both a top-down and bottom-up approach. In case of erroneous or lacking data, the algorithm uses grammar rules to hypothesize the façade structure. The probability on the grammar rules is again dependent on the data. For reconstructing roof faces, Milde and Brenner (2009) describe the importance of grammar rules to hypothesize the roof structures. Algorithms as these are essential for building up a complete description of a scene, also when data is missing or erroneous. In this paper we explicitly describe the quality of features of 3D roof models. Our aim is to give insight in the construction of the error budget in these models.

If a clear explanation of the errors is lacking, eye-catching errors in city models will result in a decreasing confidence in the rest of the models. On the other hand, if the model visually appears to be attractive the risk is that the 3D models are considered to represent the truth. The necessity of a proper quality description is embedded in the geodetic and photogrammetric working fields for a long time. Most of the research on 3D modelling focuses on the algorithms that produce 3D information. Generally, the quality descriptions seem to be rather limited to indicating a Level of Detail (LoD) and a sample check on reference data, resulting in several absolute quality measures such as completeness and correctness or planimetric accuracies. Specifying the LoD does not mean that the geometric accuracy of the model has been determined. Mostly, general parameters such as minimum footprint size and positional accuracy values are mentioned for a certain LoD (Kolbe et al., 2005). However, this does not give an insight in the quality of the specific building models.

* Corresponding author. Tel.: +31 53 4874 350; fax: +31 53 4874 335.
E-mail address: oudeelberink@itc.nl (S. Oude Elberink).

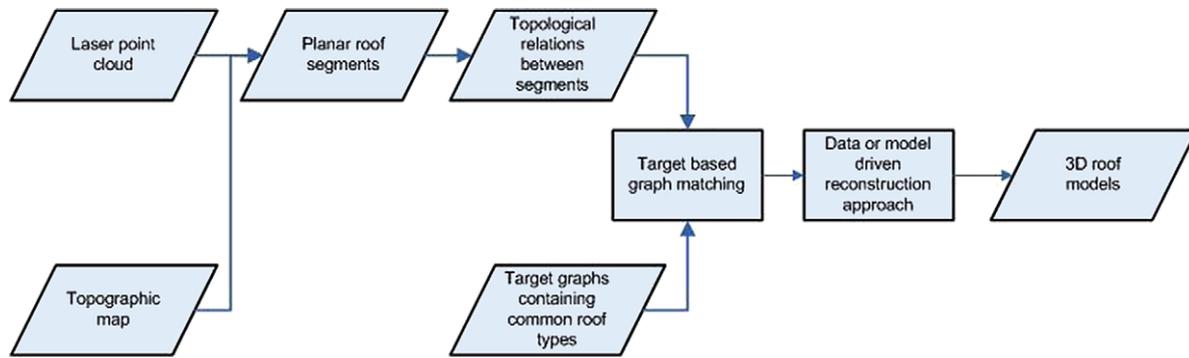


Fig. 1. Global workflow from point clouds to 3D models, using a target based graph matching approach.

For an absolute accuracy measure, an independent reference dataset would be needed in order to check for differences between the reconstructed models and the reference dataset. These differences contain information on the mean and local variation in 3D between reference data set and reconstructed model. Reference data can be acquired manually or semi-automatically, as in Rottensteiner (2006). If the reference data is considered to be the ground truth, its quality should be better than the reconstructed models. The problem is that such a detailed reference dataset might not be available (yet) at a large scale for detailed models. When using reference data to determine the quality of a certain set of reconstructed buildings, it is necessary to correctly analyse differences between the two datasets.

In this paper a building reconstruction evaluation method is presented that delivers quality parameters attached to building models. The quality parameters are all determined by performing internal measures. Reference data has not been used in our study. The structure of this paper is as follows: after a short description on the building reconstruction approach (Section 2), the quality of the input data is described in Section 3. Based on the properties of the reconstruction algorithm and the quality of the input the theoretical and empirical geometric quality of the extracted features are described in Section 4. Section 5 discusses the quality analysis of reconstructed 3D building models in three different scenes.

2. Building reconstruction algorithm

Before determining the quality of the features, a short description of our building reconstruction approach is given. The models are reconstructed using a target based graph matching algorithm as presented earlier in Oude Elberink and Vosselman (2009a), using airborne laser scanner data and topographic maps. The global workflow is that roof segments are extracted from the point cloud. The topographic maps have been used to select segments that (partly) fall within a building polygon. The topological relations between neighbouring segments are stored in a roof topology graph, which is matched with a limited number of target graphs of the most common roof types. This is called the target based graph matching algorithm. This step relates specific laser data features with generic database information. Using this relation it is possible to intersect segments that topologically correspond to a certain part of a common roof structure. Knowing that a certain combination of features corresponds to general building information increases the reliability of the individual data features. At the same time, segments that do not correspond to any common roof type are left out in the automatic building reconstruction, and can be processed in a semi-automatic reconstruction step. The map information has not been used to determine the boundaries of the

roof faces. In Fig. 1 a schematic overview is given of our workflow from input data to 3D roof models.

The advantage of having the relation between common roof types and data features is that the reconstruction itself can be either more data or model driven or a combination of both, as explained in Oude Elberink (2009).

3. Quality of input data

Starting point for analysing the quality of 3D building models is to determine the quality of the input data. In our study point clouds from airborne laser scanning data have been used as input. This section highlights two important elements of the quality of raw laser data: precision of laser point data and point density of the data.

3.1. Accuracy of laser point clouds

Generally, in a raw laser point cloud systematic and stochastic errors occur, depending on the configuration during the time of acquisition. Crombaghs et al. (2002) discriminate four main error types and corresponding scales as a result of measurement inaccuracies. The authors separate inaccuracies in the laser ranging measurements (1), GPS (2) and INS (3) measurements, and the quality of the adjusted strip offsets itself (4). The scale of each of the errors depends on the area that is influenced by each of the errors. In this configuration of four errors the area of influence are respectively: (1) a single laser echo, (2) one GPS epoch, (3) approximately one strip, and (4) the whole laser block. The study is based on detection and elimination of systematic errors by analysing features in strip overlapping areas, also used in Pfeifer et al. (2005) and Vosselman (2008). Karel et al. (2006) describe these errors' influences to the quality of a Digital Terrain Model. Reference data might be incorporated in this step, e.g. in a calibration or overall strip adjustment procedure, to achieve an absolute quality measure for the input data. Values for the planimetric accuracy of dense airborne laser scanner data range from 5 to 25 cm standard deviation (Rentsch and Krzystek, 2009; Vosselman and Maas, 2001), and 5–10 cm for the vertical accuracy (Crombaghs et al., 2002). Within the dataset variations can occur due to the earlier mentioned appearance of errors at various scales.

3.2. Laser point density

The point density of airborne laser data is an important property of the data in order to decide which objects can and which cannot be detected in the laser data. The goal of this section is to point out that the relation between the desired Level of Detail (LoD) should not only be related to the average point density of the dataset. For feature extraction purposes the variation in point

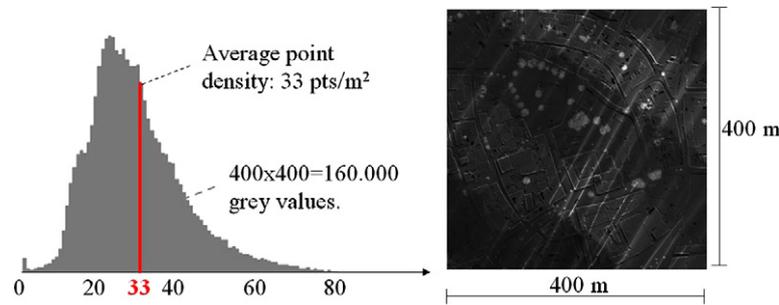


Fig. 2. Histogram showing the variation in point density in pts/m² (left) and corresponding density image (right). The range is from 0 pts/m² (black) to 80 pts/m² (white).

density is often more important than the mean point density. This means that in some areas the point density is higher or lower than the mean value, and that the processing parameters have to be chosen with care, or have to be made flexible, in order to correctly process all laser data (Oude Elberink and Vosselman, 2009b). Locally there can be large variations in point density. These variations are caused by irregular movements of the sensor platform, differences in number of overlapping strip coverage, or caused by different scatter properties of various surface types. Specifically, when using laser sensors that are capable of recording multiple returns, the point density in vegetated areas will be higher than at its surrounding. To visually show the variation in point density, the number of laser points per square meter (pts/m²) is represented in a histogram and in a density image, see Fig. 2. The histogram shows the relative amount of pixels (y-axis) from the density image per point density (x-axis). The density image (right) contains pixels of 1 m². Remarkable on the point density image is that the structure of the terrain and individual objects already becomes visible, without looking at the height information. This means that the point density partly depends on the object at the surface. For reconstructing objects it is therefore important to know if the mean point density value can be used for further processing. In this area of 400 m × 400 m, the sample dataset contains 5.4 million points, resulting in an average point density of 33 pts/m². However, this number is biased by the higher point density in vegetated areas, so for the building areas the average drops down to 25 pts/m², with local averages of less than 15 pts/m² in single strip coverage areas.

3.3. Data gaps

There are several reasons why laser scanning data can contain gaps. Data gaps that are of influence for building roof modelling, are caused by occlusion by neighbouring objects and the gaps due to absorption of the laser pulse by a layer of water on the (mostly flat) roof face. The absence of laser data is an important indication that the reconstructed models might not be complete. The quality measure that can be attached to a 3D building model is the size of the area per building that contains no laser points, if this area exceeds a certain threshold.

4. Geometric quality of data features

Now that the quality of the input data is described and the general processing steps are known, it is of interest to analyse the quality of the features extracted from that data. This can be done in two different ways. The first is an error modelling approach (Section 4.1) that describes how precise specific features can be determined from airborne laser data using a certain modelling strategy. This approach is important in the stage of setting up the user requirements in relation to the data properties and the processing steps. The second, more empirical, approach

(Section 4.2) of describing the quality of features is a relative and internal accuracy check on the extracted features. This means that the differences between the input laser data and the extracted features are checked and analysed.

4.1. Error modelling of features

The quality of the input dataset influences the quality of the extracted features in a systematic manner: if the whole laser data contains a certain offset, the extracted features inherit the same offset. As shown in 3.1 systematic errors occur at multiple scales in each dataset, for example for each GPS observation. In general, we assume that systematic errors are constant within the area of individual objects. For individual buildings this is a fair assumption, but the larger the object, the weaker this assumption becomes.

In this section we describe the theoretical quality of planes, intersection edges, height jumps and object points based on error propagation of input data and reconstruction assumptions.

4.1.1. Roof planes

Most of the man-made objects consist of a combination of planar faces. In laser scanning data these planar faces can be detected in a segmentation procedure. Finding planar segments for roof extraction is widely used in building reconstruction algorithms, (Brenner, 2000; Dorninger and Pfeifer, 2008; Hofmann, 2004; Jochem et al., 2009; Rottensteiner and Briese, 2003; Vosselman et al., 2005). The accuracy of the orientation of a plane fitted through the segment increases by the segment's size and its planarity. A problem is that some roof faces might not be captured at all in the laser data, let alone detected in the segments. Another problem arises if the assumption that each segment represents one roof face does not hold. This assumption does not hold in cases of under and over segmentation.

4.1.2. Boundaries of roof faces

When looking at boundaries of faces in 3D models reconstructed from airborne laser scanner data we can distinguish features that can be detected and determined with relatively high or low accuracy, as visualised in Fig. 3 and more deeply explained in Oude Elberink (2010). In Fig. 3 a schematic overview is presented of varying quality of edges and corner points, caused by the various ways how they are realised. A simple half hip roof, including a dormer and a flat shed attached to the building are shown. The figure is not supposed to be representative for all building reconstruction approaches, but it is an example to show the varying quality within one single building.

Corner points are points representing the nodes of an object's face boundary. We distinguish three main categories of extraction precision for corner points:

1. Corner points whose location is determined by intersection of three planes.

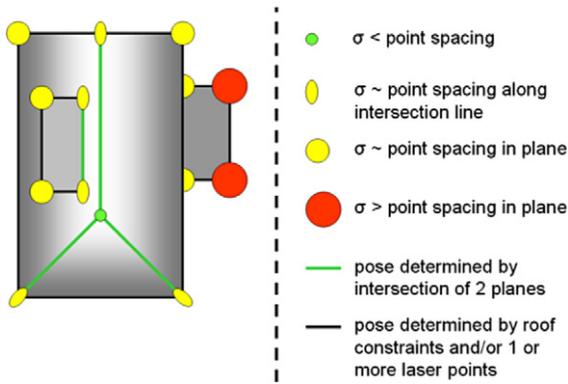


Fig. 3. Quality differences in roof edges and corner points.

2. Corner points on intersection line between two planes.
3. Corner points on the outline of a roof segment.

Category 1 represents the points that from a reconstruction point of view are most precise, as they are rather fixed by the intersection of three planes. The accuracy is expected to be better than the point spacing, although it depends on the size and planarity of the roof planes and the configuration of intersections between the three planes. In Fig. 3 one object point is determined by the intersection of three roof planes, and can be considered to be accurately determined. Points in category 2 are the end points of an intersection line between two roof faces. The accuracy can be visualised by ellipses as the location in the line direction is determined less accurately than perpendicular to the line direction. These are visualised by ellipses in Fig. 3. The exact determination of this location depends on the particular implementation of calculating the intersection line. However, it is likely that in every implementation the position of an end point along a 3D line somehow depends on the existence of nearby laser points in both segments. We assume that the geometric quality of the end points of intersection lines between two segments is in the order of the median point spacing along the intersection line. Category 3 corner points are located on the outline of roof segments. These points are considered to be less accurate than the first two categories' points, as these are not determined by an intersection of planes. Often, the location is based on analysing laser points that are located on the segment boundary, either by taking the most outer point or fitting a line through the points on the boundary. Constraints on preferred edge directions and symmetry of the roof also influence the location of these corner points. Two points within in this third category show a larger precision value, visualised by red circles, as they depend on the location of laser points on a flat surface. As the borders of horizontal segments may suffer from a lack of points due to water standing on that roof part, the position can be considered as the most uncertain of this example.

Next, several roof edges are discussed that differ in terms of determination. The first type is the edge represented by an intersection line between two roof faces, in Fig. 3 shown as green lines. The pose is determined by the intersection of two planes. The quality of the pose depends on the size of each of the two roof faces and the intersection angle between the two faces. This is the reason why intersection lines of gable and hip type roofs are generally better determined than the lower intersection lines at gambrel or mansard roofs. A second type of roof edges are the outer boundaries of roof faces. Examples are the roof gutter locations. Typical rules of thumb for the quality of these edge locations are in the order of point spacing (Rottensteiner, 2006; Kaartinen et al., 2005). Some approaches take the location of a map polygon for these roof edge locations, such as Vosselman et al. (2005) and Kada and McKinley (2009). If the map polygon represents the outer

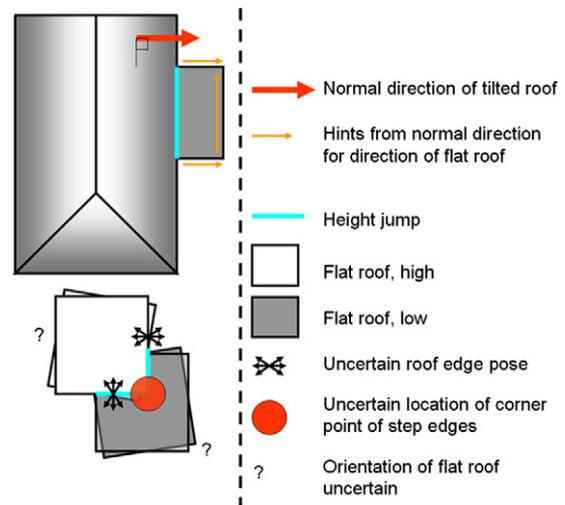


Fig. 4. Two types of height jumps: between a tilted roof and a flat roof (top) and between two flat roofs (bottom).

boundary of the roofs, i.e. not the walls, in an accurate manner this can be used to improve the accuracy and reliability of the outer edges.

A special kind of roof edge is the height jump between two roof faces. In Fig. 4 two types of height jumps are shown. The top figure shows a height jump between one tilted and one flat roof part. The bottom part of the figure shows a height jump between two flat roof faces. It is expected that the latter situation results in less reliable reconstructed roof edges. The reason is twofold:

1. The lower boundary of the tilted roof, e.g. the gutter, gives an approximate location for the step edge. The gutter location can be determined by observations, e.g. lowest laser point in that roof segment or the end of a tilted intersection line, and geometric constraints.
2. Next, it is shown that if there is a height jump between a tilted roof and a flat roof, the normal of the tilted roof can be helpful to limit the directions of the flat roof, although it is likely that the precision will still be larger than the point spacing because of the earlier mentioned lack of points on flat segments. When step edges occur between two flat roof faces, the normals of the planes do not contain information for the direction of the step edge and the outlines of both roof faces.

Depending on the reconstruction algorithm, other data sources, such as images or maps, can be incorporated to improve the location of edges or points in the model.

4.1.3. Abstraction precision

This section shortly explains the error budget of the abstraction precision in relation to the other inaccuracies in the reconstruction process. Many of the complexities of reality are ignored when reconstructing 3D models. Individual roof tiles, the exact shape of gutters and small chimneys are often ignored in 3D models. To give an example, in reality gutters contain an inclination of about 10 cm per 10 m to drain the water. This inclination is often not reconstructed in the model, and even not recognised in the data. Fig. 5 shows three pictures of roof elements. The pictures support the reasoning that there is a limitation to the quality of 3D models, just by looking at the abstraction level. It is especially important to keep these limitations in mind when analysing differences between a certain 3D model and reference data.

Not all the differences between the two data sets can be accountable to the reconstructed 3D model as both the modelled and the reference data sets contain a budget for the abstraction



Fig. 5. Left: Roof planes do have a height texture of several centimetres. Middle: Intersection of two planes is only an approximation of the real edge location. Right: Corner point is hard to identify.

Table 4.1

Indications for abstraction level of reconstructed roof features.

| | |
|---|----------|
| Orthogonal differences of roof planes to a mathematical plane | 3–5 cm |
| Edge approximation by intersection of roof planes | 5 cm |
| Corner point identification | 10–15 cm |

precision. In Table 4.1 indications are given to values for the abstraction level. The indicated values may differ from building to building as they strongly depend on the construction methods and material.

4.1.4. Total error budget

In the previous paragraphs several components have been described that influence the quality of extracted features. In this paragraph we summarize the components and analyse what are the crucial elements. As we have seen the quality of feature extraction and the abstraction precision depend on the individual feature type. The total error budget should therefore also be determined per individual feature. In Table 4.2 an example is shown that roughly estimates the total error budget of two types of corner points, of which one example is also calculated for less dense laser data. The values are based on experience. It is assumed that the components are independent from each other so that there is no correlation between the components. This assumption holds as long as the unknown systematic errors in the laser scanning data do not influence the quality of the feature extraction, which is reasonable when looking at the areas influenced by the errors. The unknown systematic errors in the laser data are caused by noise in GPS and INS measurements and are systematic for areas acquired in one GPS epoch; that is the strip width by about 50–100 m in flying direction. Quality of the feature extraction is considered to have a random behaviour for each extracted feature. The message is to show that there are multiple components in the determination of the total expected quality of 3D object points. When using a higher density data set mainly the quality of feature extraction is improved, whereas the other components are likely contributing the same amount of uncertainty. When analysing differences between 3D models and independent reference data, it is important to keep in mind which components are accountable in the differences. Remember from Fig. 2 that the point density is not a fixed value for a data set. The variation in point density influences the quality of the feature extraction.

4.2. Empirical quality of features

In the previous section we discussed the quality of the processed data features that can be expected when following a certain modelling approach. For specific groups of features

we can empirically get a quality indication, only by analysing the differences between the output features and the input data. Analysing orthogonal height differences can be found in Dorninger and Pfeifer (2008) for giving a global impression on how well the data fits to the final models. In this section the focus is on the quality of features at the level of individual buildings.

1. The first measure is the analysis of whether a segment actually represents a planar roof face. Systematic patterns in the residuals between segmented laser points and a plane fitted through these points indicate under-segmentation or the presence of non-planar roof faces. Discrepancies between data and model can be visualised and quantified easily. It is expected that in data driven approaches the majority of residuals is small and even the average residual is (nearly) zero as the model faces are constructed by fitting a plane through the same laser points. Obviously, large residuals are found on laser points or even complete segments that are left out from the reconstruction step. In model driven approaches, this quality check is helpful to detect non-symmetric roofs which are incorrectly modelled as symmetric ones.

The disadvantage of calculating the perpendicular distance between laser points and model faces is that it can be misleading in the sense that most of the laser points show a small residual. This is especially the case for laser points in data driven approaches that fit each individual roof face through a laser scanning segment. It does not show the quality of the location of the edges of the roofs.

2. The second measure calculates the distance between each corner point and its nearest laser point. Only laser points are considered that are assigned to the roof plane(s) of the corresponding corner points. Although this is not an independent check either, it is of added value to the previous height check, because it holds information on how assumptions and constraints on the edge locations fit to the data. The distance is important because it indicates whether the roof corners are within a certain distance to the laser data.
3. The third measure is the existence of segments that are not used in the automated reconstruction approach. Segments which neighbourhood relations did not match with any target graph were left out from the reconstruction approach. As explained in Oude Elberink and Vosselman (2009a) there are several reasons for incomplete matching results. One reason is that the segment actually is not representing a roof face, so it is correct to leave out that segment. However, in other cases the segment is representing a roof face but either the correct target graph is missing or another roof segment is missing that would give a complete match on a certain target graph. The problem is that it is hard to automatically detect whether leaving out the

Table 4.2

Summary of error components, for hip roof top points and gable gutter corner points. Quality measures are given in standard deviation values.

| | Hip roof, top corner point, density 16 pts/m ² | Corner point, gable gutter location, density 16 pts/m ² | Corner point, gable gutter location, density 8 pts/m ² | Remarks |
|---|--|---|--|--|
| Systematic errors in input laser scanning data | 20 cm x, y 10 cm z | 20 cm x, y 10 cm z | 20 cm x, y 10 cm z | Values indicate systematic but unknown errors in planimetry and height. |
| Quality of feature extraction | 5–10 cm x, y, z | 25 cm x, y, z | 35 cm x, y, z | Quality depends amongst others on the point density, which varies locally. |
| Abstraction precision (corner point identification, from Table 4.1) | 10–15 cm x, y, z | 10–15 cm x, y, z | 10–15 cm x, y, z | Depends on the object construction. |
| Total standard deviation | 25 cm x, y 15–20 cm z | 35 cm x, y 30 cm z | 40 cm x, y 40 cm z | |

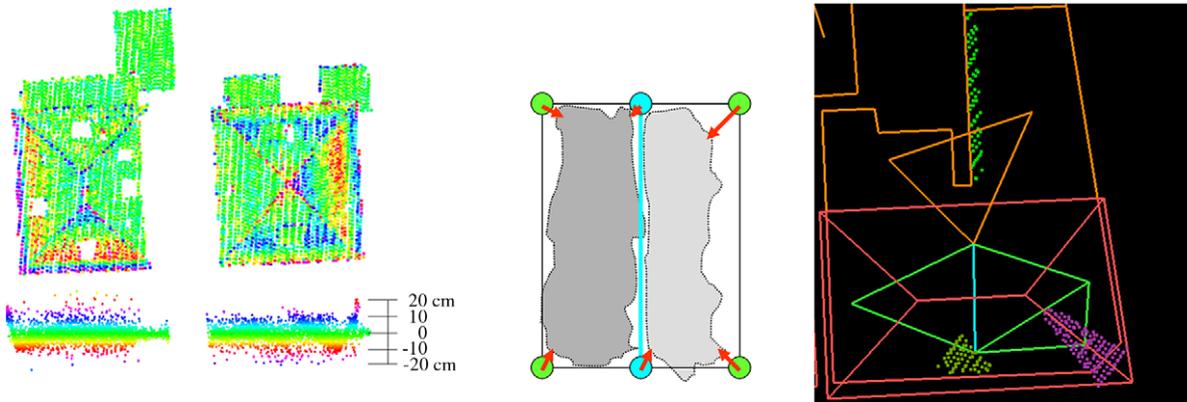


Fig. 6. Three empirical quality measures: orthogonal differences between roof plane and laser points (left) including the vertical profile of the differences (on the lower left), nearest distance between corner points and laser points (middle) and laser segments which neighbourhood relations do not fit to a certain target graph.

segment was correct or incorrect. On the right in Fig. 6 three segments are superimposed to the boundary representation of a building model including the neighbourhood relations between roof segments. These three segments and their relations did not match to any of the target graphs.

5. Quality analysis of 3D models

In this section it is shown how quality measures are integrated in our building reconstruction approach. The importance of this section is to visualise various quality indicators, attached to the reconstructed buildings. This is relevant information for the user of these models to be able to judge whether the models are of sufficient quality for their purpose or not.

5.1. Scene properties

Three areas have been selected and presented in this paper. The areas differ in terms of complexity or point density. The first two areas are located in Enschede, The Netherlands, where high point density laser data of, on average, more than 20 pts/m² is available. The difference between the two scenes is that the first scene is located in a suburban area whereas the second scene is in the city centre of Enschede. The third scene is a suburban area in Middelburg, The Netherlands, where laser data is available with a point density of, on average, about 8 pts/m². The level of detail aimed at is to reconstruct all roof faces larger than 2 m². The 3D building models have been reconstructed using a target graph matching algorithm as described in Oude Elberink (2009). The requirements of the models are comparable with the geometric component of CityGML's LoD 2 (Kolbe et al., 2005). The main difference is that we do not add textures to the roof faces.

5.2. Reconstructed 3D models

The three scenes contain respectively 60, 324 and 250 buildings. Detailed topographic maps from the cadastre have been used to select roof segments and to reconstruct the walls of the models. It is assumed that the cadastral map contains all buildings that are required in the 3D building model. The models are shown in Fig. 7.

5.3. Visual and quantitative analysis

The goal of this section is to present geometric quality measures to the models shown in Section 5.2. The quality measures are automatically generated during the reconstruction process itself. For each building the following quality measures have been calculated:

- the orthogonal distance between laser points to their corresponding roof face;
- the shortest distance from object points to laser points on corresponding roof faces;
- the segments that have and have not been used in the reconstruction process.

However, these measures do not explicitly contain information whether a model is correct or incorrect. From these measures several quality indicators can be derived, such as number of laser points that have a large orthogonal distance to their corresponding face. If the goal of a quality analysis would be to check whether a model is correct or incorrect, then the user requirements are to be introduced to judge whether the quality measures exceed a certain threshold and what the consequence would be. As this paper focuses on methods to analyse the quality, several ways of presenting and calculating the quality parameters are mentioned. Orthogonal distances and the minimal distance between object

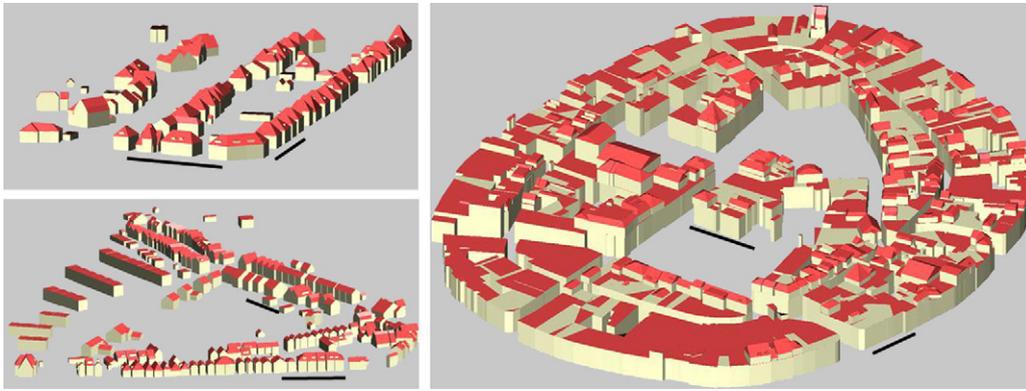


Fig. 7. 3D models reconstructed from airborne laser scanner data and topographic maps. Upper left Scene 1, right Scene 2 and lower left Scene 3. Black lines correspond to 50 m in object space.



Fig. 8. Differences between laser points and roof planes, coloured by residual: green (grey) < 20cm < yellow (light grey, hardly used) < 50cm < red (dark grey) for three scenes.

points and laser points are checked against a predefined threshold value. In the presented case the threshold values do not depend on the features' type or size. The threshold values for orthogonal distances are green < 20 cm < yellow < 50 cm < red. Fig. 8 gives a global overview on how well the laser data fits to the output models.

For a closer look, the building in the white rectangle on the left in Fig. 8 is enlarged in Fig. 9. The points that have a distance larger than 50 cm are located outside the roof face. The location of the gutter was affected by constraints of using the same gutter height of neighbouring roof faces. That is why some laser points fall outside the roof face polygon. Some laser points on the extension behind the building, upper left part of the left image, are located just beneath the roof face and therefore not visible when overlaying the model on the laser points. In fact there was a small height jump that was not reconstructed as the height jump was below 20 cm, so the laser points were grouped into the same segment. Distances between object point and nearest laser point (middle figure) are considered large (red) if they exceed the median point spacing of the roof segments. One large distance is highlighted by a larger red dot on the top ridge where locally there is a laser data gap due to a small chimney. At the right two segments have not been used during the reconstruction as they topologically did not match with target graphs in the data base. The reason was the lacking of laser points and therefore the lack of segments on two very small roof parts (indicated by red polygons) connecting the extrusions with the main roof shape. Because of the lack of points on those roof parts, the two segments on other roof parts, shown in the figure, could not be used.

The effect of different ways of interpretation of the same quality measure is shown in Fig. 10. The meaning of the figure is to highlight the locations that need extra attention. The message is to

show that something what seems to be a small issue (individual segments, left) may have a large impact (individual buildings, right). Both, on the left and the right, the same quality measure is shown: the segments that did not fit to a certain target graph. The difference between left and right is that on the left only the segments are highlighted and on the right the complete building is coloured red. These segments are left out from the automated building reconstruction algorithm as there was no relation found with the target graphs used in the target graph matching algorithm. The segments that are left out are transferred to the quality measures and shown to the user (as shown in red in the left part of Fig. 10). To judge whether the segments are correctly left out of the reconstruction method, the other roof segments of the building need to be analysed as well. That is why it would make sense to indicate that the whole building needs attention of the user (as shown in the right of Fig. 10).

Using the visual presentation of the quality measures, the user can have a first impression on the quality of the 3D models. Often statistics are used to judge the quality of the models. Similar to the visuals, statistics need to be presented and analysed with care as well. When is a 3D model considered to be correct?

From Table 5.1 it can be seen that the percentage of laser points that do not fit to the 3D models can be 1.8%, however the number of buildings that are affected by segments that were not used during the reconstruction is 20%. Even within the same category of quality measures, the percentage of correctness differ per unit that is analysed. It can be seen that Scene 2 (city centre) contains 'only' 7% laser points with more than 20 cm distance to the corresponding roof face, however more than 12% of the segments contain more than 20 laser points with high residuals and these affect about 35% of the buildings. Quality criteria should be derived from the user requirements in such a way that the models are correctly judged.

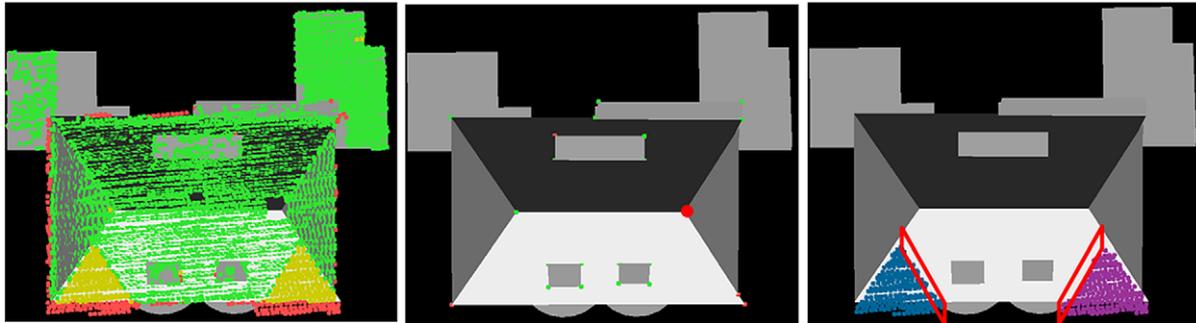


Fig. 9. Thresholds assigned to orthogonal distance (left), distance between object points and nearest laser point (middle) and segments not used in building reconstruction.

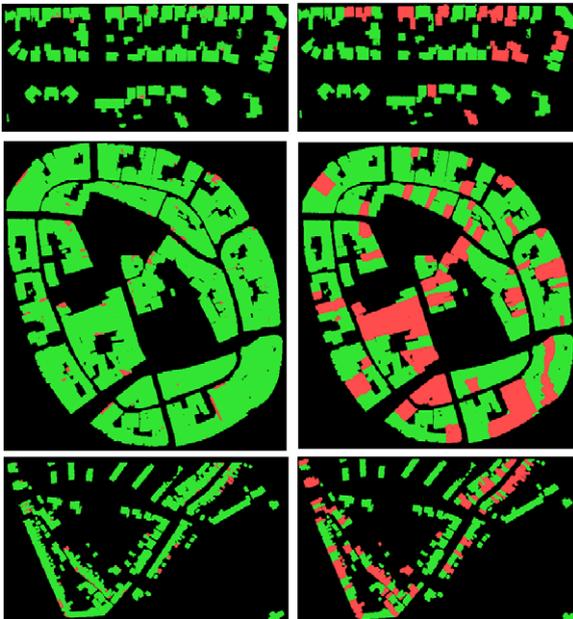


Fig. 10. Different visualisation of the same quality measure. Left: Segments not used are coloured red (dark grey); right: buildings that contain those segments are coloured red (dark grey).

6. Discussion

In our research no usage has been made of reference data and we did not include user requirements, although both of them are needed to answer the question whether the modelled data is suitable for a certain application. In this section we discuss the reason for focussing on only internal quality measures. Describing the quality of 3D building models only using internal measures does not give an absolute accuracy value. Using reference data is necessary for an independent and absolute accuracy description. However, before collecting reference data a deeper thought on the setup of the reference measurements is necessary. How many measurements are needed on what kind of object features? What do the differences tell between reference data and modelled data? To answer these questions we need to have insight in the behaviour of the quality of the reconstructed models. In this paper it was shown that some object features have higher accuracy than others. This insight is necessary when handling differences between reference data and modelled data: not all of the differences can be accounted for the inaccuracy of the reconstructed models. For example, if the ridge heights of the reconstructed buildings need to be checked, it is doubtful if suitable reference measurements can be found as most of the differences are due to the abstraction precision of the ridge itself plus the systematic errors in the laser data. So, for checking the accuracy of the ridge height, it is more

Table 5.1

Statistics on quality measures for 3 different scenes.

| Area ID | Scene 1 | Scene 2 | Scene 3 |
|---|-----------------------|-----------------------|----------------------|
| Average point density | 20 pts/m ² | 20 pts/m ² | 8 pts/m ² |
| # buildings | 61 | 324 | 250 |
| # laser points in roof segments | 176 k | 1232 k | 161 k |
| # roof segments | 460 | 1609 | 1081 |
| # laser points with residual > 20 cm | 3.1 k (1.8%) | 83 k (6.7%) | 5.8 k (3.6%) |
| # segments with more than 20 points with residual > 20 cm | 4 (0.9%) | 204 (12.7%) | 21 (1.9%) |
| # buildings affected by abovementioned segments | 3 (5%) | 114 (35%) | 13 (5%) |
| # 3D object points with distance > 1 m to laser data | 107 (11%) | 1107 (25%) | 1150 (26%) |
| # segments not used | 18 (4%) | 66 (4%) | 71 (7%) |
| # buildings affected by segments that have not been used | 12 (20%) | 48 (15%) | 55 (22%) |

appropriate to design the reference measurements such that they can detect and eliminate systematic errors than directly measuring the ridge height.

Important is the role of the user, and his user requirements, in defining criteria to indicate the quality of the automatic extracted building models. For volume calculations the exact 3D shape of the roof faces is less important than the ability to calculate an accurate mean height for a given ground plan. For automated roof texturing it is of high importance to work with 3D models that fit to the corresponding textures. Preferably, the user requirements are known in advance. Previous researches on user requirements focus on the semantics and the topological structures of 3D building models, e.g. in Nagel et al. (2009). So far, little is known on the geometric requirements of 3D building models, as most of requirements are limited to one planimetric and one vertical accuracy threshold for the whole dataset. It needs to be analysed whether our quality measures can be converted to such a simple threshold values, or that the list of geometric requirements needs to be extended and refined.

7. Conclusions

We have presented a method to evaluate the quality of 3D building models extracted from point clouds and demonstrated the usefulness of these methods in three different scenes. The models have been reconstructed using airborne laser scanner data with a point density of about 20 pts/m². The quality of the models depends on the input data and the processing algorithm in relation to the requirements and expectations of the user. Some features,

such as hip roof ridge lines, can be extracted with higher precision than other features, such as flat roof edges. An empirical quality check can be performed to analyse the correctness of assumptions made during the processing. The main assumptions are that roof planes are planar and that the roof topology corresponds with one or more graphs in the target database. Segments that topologically do not fit to one of the target graphs are left out from the automated building reconstruction. This assumption is not always correct, but the biggest advantage of this method is that these situations can easily be shown to the user as part of the quality description of the building models.

The complexity of quality parameters of 3D models show up in the quality criteria. Until now, there is little known on the user's requirements and acceptance criteria of 3D city models. When is a 3D model considered to be of good quality? Future customers of 3D city models, who are about to let data producers model 3D models, would be well-advised to carefully set up well defined quality criteria. Both visual and quantitative quality measures can be presented such that the model seems to be of better quality than it actually is.

If precise reference data were available in the future to determine the quality of a certain set of reconstructed buildings, it is necessary to correctly analyse differences between the two datasets. The method presented in this paper gives insight in the construction of the errors in models constructed by airborne laser scanner data. Not only will reference data face the fact that the 3D models are not perfect. Aerial images are often used for texturing the 3D models. Even if these images have been captured at the same time as the airborne laser data, it cannot be expected that the object edges exactly fit to the image features. It is of high importance to analyse whether the textures should be adapted to the shape of the model, or that the model should be adapted to the texture. This analysis can only be done after a thorough quality analysis of both of the datasets.

At this moment, usage has been made of fixed threshold values to analyse the quality measures. Future work includes the integration of the expected quality to judge whether the empirical quality as measured from the internal check between laser data and 3D building models is significant or not. It includes making the evaluation criteria dependent on the expectations of the quality. Additionally, the integration of realistic user requirements is planned to actually check whether the reconstructed models fit the use.

References

- Becker, S., 2009. Generation and application of rules for quality dependent façade reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (6), 640–653.
- Brenner, C., 2000. Towards fully automatic generation of city models. *International Archives of Photogrammetry and Remote Sensing* 33 (Part B3), 85–92.
- Crombaghs, M., Oude Elberink, S., Brugelmann, R., de Min, E., 2002. Assessing height precision of laser altimetry DEMs. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 34 (Part 3a), 85–90.
- 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.
- Hofmann, A., 2004. Analysis of TIN-structure parameter spaces in airborne laser scanner data for 3-D building model generation. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 35 (Part B3), 302–307.
- Jochem, A., Hoefle, B., Rutzinger, M., Pfeifer, N., 2009. Automatic roof plane detection and analysis in airborne lidar point clouds for solar potential assessment. *Sensors* 9 (7), 5241–5262.
- Kaartinen, H., et al., 2005. Accuracy of 3D city models: euroSDR comparison. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part 3/W19), 227–232.
- 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.
- Karel, W., Pfeifer, N., Briese, C., 2006. DTM quality assessment. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part 2), 7–12.
- Kolbe, T., Groeger, G., Pluemer, L., 2005. CityGML – interoperable access to 3D City models. In: Oosterom, P.V., Zlatanova, S., Fendel, E. (Eds.), *Int. Symposium on Geo-Information for Disaster Management*. Springer, Berlin, Heidelberg, pp. 883–899, Delft, The Netherlands.
- Milde, J., Brenner, C., 2009. Graph-based modeling of building roofs. In: *AGILE Conference on GIScience, Hannover, Germany (on CD-ROM)*.
- Nagel, C., Stadler, A., Kolbe, T., 2009. Conceptual requirements for the automatic reconstruction of building information models from uninterpreted 3D models. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 38 (Part 3-4/C3) (on CD-ROM).
- 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.
- Oude Elberink, S., 2010. Quality measures for building reconstruction from airborne laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 38 (Part 3A), 198–203.
- Oude Elberink, S., Vosselman, G., 2009a. Building reconstruction by target based graph matching on incomplete laser data: analysis and limitations. *Sensors* 9 (8), 6101–6118.
- Oude Elberink, S.J., Vosselman, G., 2009b. 3D information extraction from laser point clouds covering complex road junctions. *The Photogrammetric Record* 24 (125), 23–36.
- Pfeifer, N., Oude Elberink, S., Filin, S., 2005. Automatic tie elements detection for laser scanner strip adjustment. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part 3/W19), 174–179.
- 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.
- Rentsch, M., Krzystek, P., 2009. Automatically reconstructed roof shapes for lidar strip adjustment and quality control. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 38 (Part 3/W8), 117–122.
- Ripperda, N., Brenner, C., 2009. Evaluation of structure recognition using labelled facade images. In: *Pattern Recognition. In: Lecture Notes in Computer Science*, vol. 5448/2009, pp. 532–541.
- 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), 13–18.
- 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) (on CD-ROM).
- Vosselman, G., 2008. Analysis of planimetric accuracy of airborne laser scanning surveys. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 37 (Part 3A), 99–104.
- Vosselman, G., Kessels, P., Gorte, B.G.H., 2005. The utilisation of airborne laser scanning for three-dimensional mapping. *International Journal of Applied Earth Observation and Geoinformation* 6 (3–4), 177–186.
- Vosselman, G., Maas, H.-G., 2001. Adjustment and filtering of raw laser altimetry data. In: *Proceedings of OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Terrain Models*, OEEPE Publication No. 40, 11 p. (on CD-ROM).