A NEW APPROACH TOWARDS QUANTITATIVE QUALITY EVALUATION OF 3D BUILDING MODELS

Hanns-Florian SCHUSTER, Uwe WEIDNER

Institute for Photogrammetry, University Bonn, Nußallee 15, 53115 Bonn, Germany
schuster@ipb.uni-bonn.de

Institute for Photogrammetry and Remote Sensing, University Karlsruhe, Englerstraße 7, 76128 Karlsruhe, Germany
weidner@ipf.uni-karlsruhe.de

Working Group III/8

KEY WORDS: building models, quality measures, quantitative evaluation

ABSTRACT

The need of describing the quality of data ranges from data acquisition to the use of the data in geoinformation systems. The contractor should verify that the data he captured suffices the specifications and the end user wants to know, if the data is suited for a special task at hand. Both are interested in quantifying the quality, possibly by simple and meaningful measures, which can be easily computed without much further efforts – prohibitive with respect to involved labour and related costs. Much work has been already done on the standardization of principles of quality evaluation, reports and metadata (c.f. ISO standards 19113, 19114 and 19115), but only few contributions deal with the question of defining quality measures for a specific application, which possibly may be generalized for others as well.
A recent project in cooperation with the Surveying Office of North Rhine-Westphalia investigates the topic of quality evaluation of photogrammetrically captured building models with the aim to identify useful quality measures which can be used for contract specificatios and to implement an approach for automated quality control based on a comparision of measurement and reference data. This paper presents the concept of the approach and first results.

1 INTRODUCTION

In contrast to the huge amount of references dealing with semi-automatic or automatic approaches for building reconstruction from aerial images or laser scanner data (c.f. (Baltsavias et al., 2001) and previous proceedings of the Ascona-Workshop), there are only a few references focussing on quality evaluation of the extracted building models. Quality evaluation is important due to several reasons. First, it may give important information about deficiencies of an approach and may thereby help to focus further research activities. Second, quality evaluation is needed in order to compare the results of the different approaches and to convince a user, that an approach can be used in an operational workflow. Besides these reasons, the most important reason arise from the practical requirement, that a contractor should check his measurements and that a customer has to check the quality of the delivered data with respect to the specifications of the contract. For this purpose, quality evaluation should not only be based on visual inspection and thereby subjective control, but on quantitative quality measures. Therefore, a recent project in cooperation with the Surveying Office of North Rhine-Westphalia aims at indentifying useful quality measures and implementing an approach for automated quality control based on a comparision between measurement and reference data within representative areas of the entire area.

According to ISO 19113 different aspects of quality have to be taken into account: logical consistency, temporal accuracy, thematic accuracy, positional accuracy and completeness. In this contribution we focus on geometric aspects, thus dealing with the positional accuracy and – because of its importance for a surveying authority – also with the aspect of completeness. Geometry of an object may be defined as its topology and the metric, where topology may carry thematic aspects and therefore is correlated with the thematic aspect of quality. An example is a building consisting of a main building and an annex. As already mentioned above, we are up to now only interested in geometric aspects, thus only dealing with one object class *building* and do not distinguish between different parts of buildings. Therefore, we subsume building parts to one building and do not perform a more specific analysis of its aggregation and the related semantics. Nevertheless, the mere evaluation of different building data sets describing the same scene owns its own complexity due to the complexity of buildings.

This contribution gives a short introduction in the topic of quality evaluation. We will discuss – besides others – mainly two approaches for quality evaluation of building models in more detail, namely (McKeown et al., 2000) and (Ragia, 2001). Based on their principle ideas, a new approach is developed, taking the good aspects and combining them, avoiding – in our opinion – the single approaches' deficiencies, by using a combination of strategies and introducing alterated quality measures. Results will be shown for real data sets based on a 2D analysis – the full 3D implementation for complex building structures is still ongoing work and not yet finished – and for the simple scene of the test data set FLAT (c.f. (Sester et al., 1996)), although other quality measures, like RMS of point coordinates, may be more comprehensive on first sight for such simple buildings, but cause problems applied to complex building structures.

2 APPROACHES FOR QUALITY EVALUATION

In this section published approaches for quality evaluation of building models are discussed (c.f. Table 1). These approaches may be classified into the following categories:

- evaluation of building detection
- evaluation of building reconstruction
- evaluation of building detection and reconstruction in one step
- evaluation of building detection and reconstruction separately

Approach	Quality measures
(McGlone and Shufelt, 1994)	ρ_d, ρ_b, ρ_m, ρ_q
(Jamet et al., 1995)	R
(Sester et al., 1996)	R
(Weidner, 1997)	β, ρ_q, ρ_q^*, R
(Henricsson and Baltsavias, 1997)	β, ρ_s, R
(Brenner, 2000)	R
(McKeown et al., 2000)	ρ_d, ρ_b, ρ_q
(Ragia, 2001)	ρ_q, ρ_f, D
(Hanson et al., 2001)	R

R/D	RMS / distances	β	type 2 error
ρ_d	detection rate	ρ_b	branch factor
ρ_m	miss factor	ρ_q	quality rate
ρ_f	false alarm rate	ρ_s	shape dissimilarity
ρ_q^*	weighted quality rate		

Table 1: Approaches and their quality measures

For the purpose of building detection evaluation, mainly interested in completeness, well known quantities, like e.g. *type 1* and *type 2 error* or the *quality rate*, are used in almost all approaches, taking buildings as objects. The selection of the quantity depends on the aim of the evaluation: some of the quantities are symmetric and thus indicating missing and false buildings, some of them are unsymmetric, indicating either missing or false buildings. Furthermore, the quantities are not independent from each other (c.f. (Weidner, 1997), (Winter, 2000)).

Approaches which only evaluate the geometry (e.g. (Jamet et al., 1995), (Sester et al., 1996)) use RMS values for point coordinates as quality measures. In their investigations they only deal with simple building models like gable roof buildings with fixed structure/topology (c.f. also (Hanson et al., 2001)). For complex buildings this may hardly apply, because of matching problems (point to point), unless integral values are taken into account. For this purpose (Brenner, 2000) computes the RMS values for the heights on the basis of extracted building models and the used DSM. The building outlines are taken from existing GIS data set.

The focus in the reminder of this section is on two approaches, namely the approaches of (McKeown et al., 2000) and (Ragia, 2001), because both have impact on the new approach. The first is an example for approaches evaluating the data of building models in one step, the second is an example for approaches which split the evaluation - evaluation of building detection and building reconstruction

2.1 Approach of McKeown et al. 2000

(McKeown et al., 2000) is an extension of the the work already presented in (McGlone and Shufelt, 1994) dealing not only with the evaluation of building extraction, but also of other topographic objects like roads (c.f. (Heipke et al., 1997)). The principle idea of their approach is to transform the evaluation of building models into the evaluation of a classification by discretising the space either in pixel (2D) or voxel (3D) and to use well known quantities as already mentioned above, namely the *detection rate*

$$\rho_d = \frac{|\mathcal{S} \cap \hat{\mathcal{S}}|}{|\mathcal{S} \cap \hat{\mathcal{S}}| + |\mathcal{S} \setminus \hat{\mathcal{S}}|} = \frac{|\mathcal{S} \cap \hat{\mathcal{S}}|}{|\mathcal{S}|}, \quad \rho_d \in [0,1] \tag{1}$$

the *branch factor*

$$\rho_b = \frac{|\hat{\mathcal{S}} \setminus \mathcal{S}|}{|\mathcal{S} \cap \hat{\mathcal{S}}|}, \quad \rho_b \geq 0 \tag{2}$$

and the *quality rate*

$$\rho_q = \frac{|\mathcal{S} \cap \hat{\mathcal{S}}|}{|\mathcal{S} \cup \hat{\mathcal{S}}|} = 1 - \frac{|(\hat{\mathcal{S}} \setminus \mathcal{S}) \cup (\mathcal{S} \setminus \hat{\mathcal{S}})|}{|\mathcal{S} \cup \hat{\mathcal{S}}|}, \rho_q \in [0,1] \tag{3}$$

with $\mathcal{S}$ the building elements of the reference and $\hat{\mathcal{S}}$ the building elements of the captured data. In (McGlone and Shufelt, 1994) they also include the *miss factor*

$$\rho_m = \frac{|\mathcal{S} \setminus \hat{\mathcal{S}}|}{|\mathcal{S} \cap \hat{\mathcal{S}}|}, \quad \rho_m \geq 0 \tag{4}$$

in their investigation. Their approach is fully 3D and – if necessary, because only 2D reference data is available – can be applied for 2D building outlines as well. The evaluation is done in one step, thus the results of building detection and reconstruction influence the quantities. In our opinion, both steps should be evaluated separately, providing more and easier interpretable information. (McKeown et al., 2000) provide the quantities for the entire data set only, but point out the necessity to compute the quantities for each building. For this purpose, building objects of the captured and the reference data have to be matched. The authors propose a 1:1 match, only the largest matched overlapping building objects should be used. This strategy does not seem to be suitable in down town areas with complex building blocks including gaps. The last point to be mentioned is the fact that the used measures do not take the magnitude of the deviation into account. A pixel or voxel respectively in the captured data set far away from the reference data contributes to the quantities with the same factor than a pixel/voxel close to the reference. Both aspects will be considered in the new approach.

2.2 Approach of Ragia 2001

The approach of (Ragia, 2001) evaluates the success of building detection, the topology and the metric of building models. The success is measured by the *quality rate* as given in Eq. 3 and the *false alarm rate*

$$\rho_f = \frac{|\hat{\mathcal{S}} \setminus \mathcal{S}|}{|\mathcal{S}|}, \quad \rho_f \geq 0 \tag{5}$$

thus a symmetric and an unsymmetric quantity. The topology is used for matching of building parts given by 2D polygons for the aggregation of buildings in each data set and for matching buildings in the captured and reference data. The inherent determination of topological relations of the primitives takes the uncertainty of the polygons into account (Ragia and Winter, 1998). The resulting graphs are used for the matching and for the evaluation of the topology. The evaluation of the topology may be ambiguous, because the topology of two building models may differ due to the procedure during data acquisition, although the metric - position and aggregated outline - is the same. Thus, an evaluation of the topology may only indicate the structural complexity of a building. The metric deviation of the matched buildings is evaluated by distance histograms and profiles, where the distances are computed along the zonal skeleton.

For the heights statistics of point heights distinguishing between roof point types are used without taking the roof structure in detail into account. Therefore, the approach mainly evaluates the ground plan of buildings. Although it may be extended to 3D, the main drawback is the fact, that the approach is not able to handle all building configurations (e.g. court yards).

3 NEW APPROACH FOR QUALITY EVALUATION

This section presents the new approach for quantitative quality evaluation of 3D building models. The evaluation consists of an evaluation of the building detection, i.e. the completeness, and an evaluation of the building reconstruction. The aim is a general framework for the evaluation of the 3D geometry. Dependent on the requirements or the available data, the user should be able to decide whether to evaluate the full 3D geometry, the 2D positional geometry or the height. In this respect, the new approach extends the work of (McKeown et al., 2000) and (Ragia, 2001). The flow chart of the approach is given in Fig. 1. The basic strategy is presented in the following.

3.1 Building Detection

Preprocessing 2D The preprocessing consists of matching the primitives (building parts) including also their aggregation to buildings and the matching of buildings of the two data sets with each other. Both matchings are based on the ground plan information of the building parts or buildings respectively. A building in one data set may be matched to none or more than one building in the other. In the first case, the building is labelled as *unassigned*, in the second the buildings are regarded as one building object. An example is given in Fig. 2.

Determination of quality measures Based on the assigned and unassigned buildings in both data sets, the *quality rate* according to Eq. 3 and the *type 2 error*

$$\beta = \frac{|\mathcal{S} \setminus \hat{\mathcal{S}}|}{|\mathcal{S}|}, \quad \beta \in [0, 1] \tag{6}$$

are computed, thus like (Ragia, 2001), we use a symmetric and an unsymmetric quality measure. We preferr the *type 2 error*, because it directly indicates missing buildings. The other quantities may also be computed as additional information. Furthermore, histograms of the area of unassigned buildings will be provided giving an insight, if perhaps only small buildings are missing.

3.2 Building reconstruction

Preprocessing 3D The preprocessing consists of a conversion from vector to raster, i.e. voxel, for each assigned building or building block, followed by distance transformation for each data set and determination of connected components of the difference regions.

Determination of quality measures Based on the preprocessed data, the *weighted quality rate* (c.f. (Weidner, 1997))

$$\rho_q^* = 1 - \frac{\sum_{x \in (\mathcal{S} \cup \hat{\mathcal{S}})} (w(d(x, \mathcal{S})) + w(d(x, \hat{\mathcal{S}})))}{|\mathcal{S} \cap \hat{\mathcal{S}}| + \sum_{x \in (\mathcal{S} \cup \hat{\mathcal{S}})} (w(d(x, \mathcal{S})) + w(d(x, \hat{\mathcal{S}})))} \tag{7}$$

with

$$d(x, \mathcal{A}) = \inf\{\rho(x, a) : a \in \mathcal{A}\} \tag{8}$$

and

$$w(d(x, \mathcal{A})) = \frac{1}{\Delta_d} d(x, \mathcal{A}) \tag{9}$$

is computed . Δ_d may be choosen based on the resolution of the input data for building extraction or on the required resolution of the building models. Instead of the weight function in Eq. 9

$$w_T(d(x, \mathcal{A})) = \begin{cases} 0 & \text{if } d \leq d_T \\ \frac{1}{\Delta_d}(d(x, \mathcal{A}) - d_T) & \text{else} \end{cases} \tag{10}$$

may be used to take tolerance regions into account. Furthermore, histograms of the deviations, statistics and attributes of the connected components will be provided for further analysis.

4 RESULTS

Results of the new approach for quality evaluation of 3D building models are shown for the ISPRS data set *FLAT* (c.f. (Sester et al., 1996)) with simple buildings using a first prototype implementation. The building models are extracted by analysis of the given DSM using the approach of (Weidner, 1997). For all buildings the parametric model of a gable roof building was used. Furthermore, the approach for quality evaluation was applied to several real

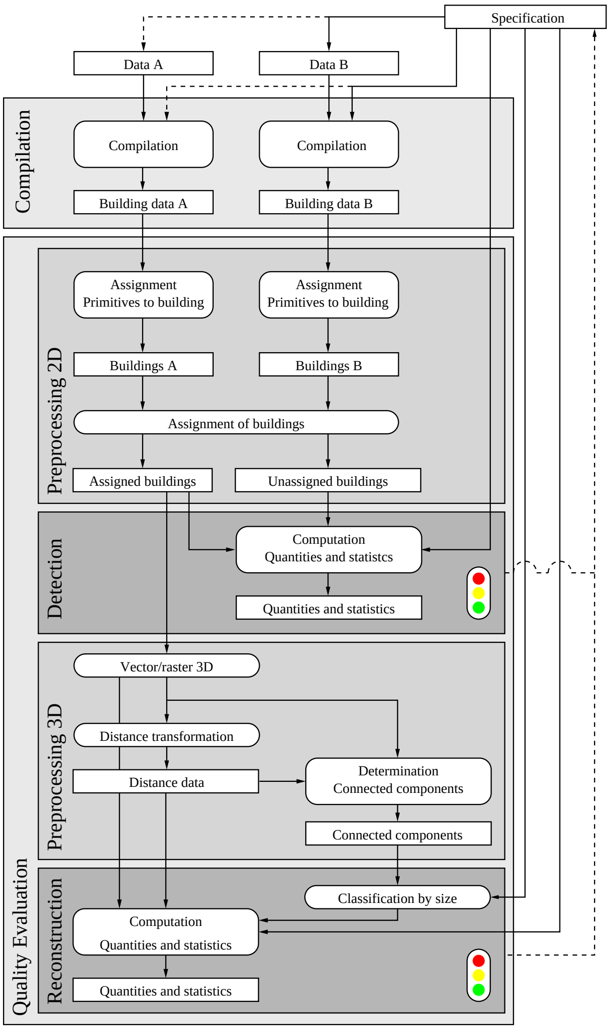

Figure 1: Flow chart of the approach

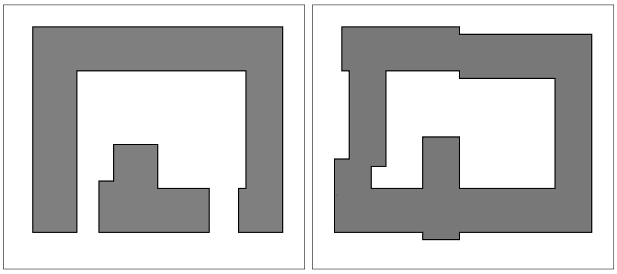

Figure 2: Example of building matching: Buildings (left) are regarded as one object and matched to the building of the other data set (right)

Resolution	ρ_d	ρ_b	ρ_m	ρ_q	ρ_q^*
$0.125\ m$	0.878	0.257	0.139	0.716	0.613
$0.250\ m$	0.877	0.260	0.140	0.714	0.592
$0.500\ m$	0.884	0.262	0.131	0.718	0.563
$0.750\ m$	0.864	0.250	0.158	0.710	0.507
$1.000\ m$	0.897	0.260	0.114	0.728	0.505

Table 2: Quality measures for test site *FLAT*

data sets: two test sites (1 and 2) of down town areas with complex building structures, two test sites (3 and 4) with cleary separated buildings with moderate complexity, and a test site (5) covering the area of the exhibition centre in Cologne with large connected buildings. The building models were compiled from aerial images (camera RMK TOP 30, image scale 1:13,000, scanned with $14\ \mu m$) by students with the programm package *inJECT* (c.f. (Gülch, 1997) for details). For each test site, both data sets – measurement and reference – were acquired independently, but using the same specifications. For these test sites the evaluation is based on 2D information only, because the implementation of the approach is still ongoing work.

Test Site *FLAT* Fig. 3 shows the reference and the extracted models (wire frames) of the test site *FLAT*. Table 2 summarizes some quantities like the *detection rate* ρ_d (Eq. 1), the *branch factor* ρ_b (Eq. 2), the *miss factor* ρ_m (Eq. 4), the *quality rate* ρ_q (Eq. 3) and the *weighted quality rate* ρ_q^* using the weight function of Eq. 9. For the computation of the quantities different discretisations are used. The systematic for the quantites of the *weighted quality rate* is in this case caused by the fact that voxel in a distance less than Δ_d have less impact on the quantity than larger deviations due to the weighing scheme. Such larger deviations are not present in the extracted models.

Test Sites 1 - 5 Fig. 4 shows a part of one of the aerial images used for the compilation of test site 1. Fig. 5 and 6 display the ground plans of the reference and the captured data set respectively. The first step of the analysis is the computation of the difference of the two data sets shown in Fig. 7. White areas indicate the set $\hat{\mathcal{S}} \setminus \mathcal{S}$, i.e. building parts which are present in the captured data $\hat{\mathcal{S}}$, but not in the reference data $\mathcal{S}$. Black areas indicate just the opposite – building parts present in the reference $\mathcal{S}$, but not in $\hat{\mathcal{S}}$. Table 3 summarizes the quantities and the statistics, where the *weighted quality rate* ρ_q^* is computed using the weight function given in Eq. 9 and ρ_{qT}^* using the weight function given in Eq. 10. The quantities are computed based on the entire data set and not for each building separately, using a discretisation of $0.5\ m$ and a tolerated deviation of $1.0\ m$.

Test site	β	ρ_q	ρ_q^*	ρ_{qT}^*
1	0.076	0.816	0.619	0.721
2	0.029	0.889	0.804	0.882
3	0.051	0.876	0.761	0.839
4	0.083	0.807	0.485	0.560
5	0.072	0.895	0.465	0.496

Test site	$\lvert\mathcal{S}\rvert$	$\lvert\hat{\mathcal{S}}\rvert$	$\lvert\mathcal{S} \setminus \hat{\mathcal{S}}\rvert$	$\lvert\hat{\mathcal{S}} \setminus \mathcal{S}\rvert$
1	129119	136295	9825	17001
2	562258	597776	16474	51992
3	67187	69392	3418	5623
4	37591	39567	3126	5102
5	1301677	1256124	93594	48041

Table 3: Quality measures and statistics for test sites 1 - 5

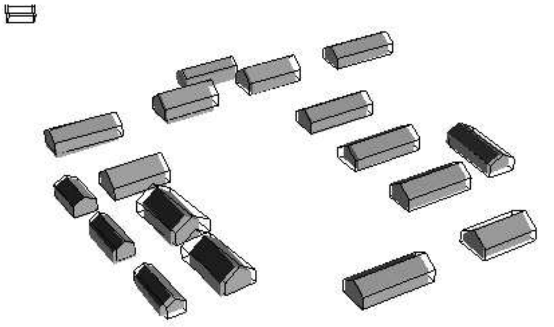

Figure 3: Overlay captured and reference data test site FLAT

Figure 7: Difference of captured and reference data

Figure 4: Aerial image of test site 1

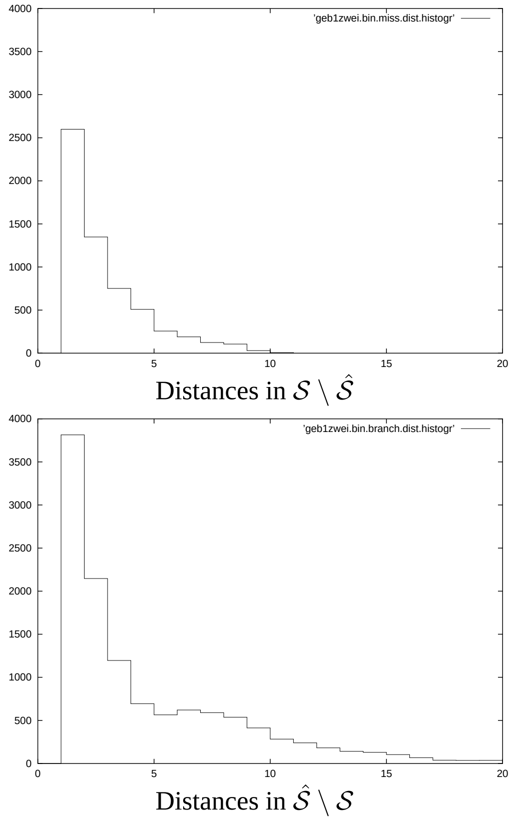

Figure 8: Histograms of distances (test site 1)

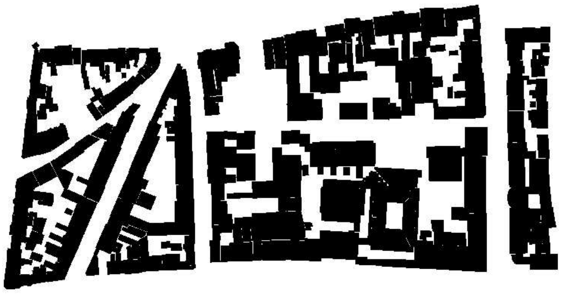

Figure 5: Outlines of reference data test site 1

Figure 6: Outlines of captured data test site 1

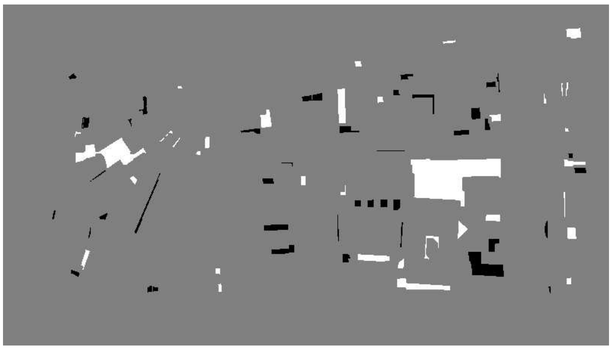

Figure 9: Difference of captured and reference data applying size and distance criteria

Comparing *quality rate* and *weighted quality rate* indicates the influence of the weighing scheme, when large differences and therefore large distances occur. It also emphasizes that each building should be checked separately. In case of test site 5, a few separated buildings in a larger distance from the exhibition center, which were not captured by one of the operators, reduce the weighted quality rate drastically, although the main part differs only a little.

Based on the segments displayed in Fig. 7 and the distance images further analysis can be performed, e.g. by an analysis of the distance histograms (c.f. Fig. 8) or taking the size of the segments into account. Fig. 9 shows all deviations of reference and extracted models of more than $1\ m$ and of a segment size with more than $4\ m^2$. Tests taking different discretisations for the evaluation show almost no differences in the quantities.

5 CONCLUSIONS

In this contribution we presented a new approach towards quantitative quality evaluation of 3D building models. The approach proposes a general framework for the evaluation of the 3D geometry and allows to treat the full 3D geometry, the 2D positional geometry or the height, dependent on the requirements of the user. Extracted building models are compared to reference data or building models of a second independent compilation. The approach is able to consider the accuracy of the data and the degree of generalisation during data compilation. The idea to have a general framework for the different possible evaluations is motivated not only by research interest, but also by the fact that 3D reference data is hardly available. Therefore, a full 3D evaluation may take place for representative areas based on a second independent compilation and a 2D evaluation for the entire area based on already available ground plan information provided e.g. by the land register.

The implementation of the approach is still ongoing work, which will also include thoroughly investigations of the applicability of the proposed quality measures in an operational workflow and the impact of the discretisation on the results.

6 ACKNOWLEDGEMENTS

The research is part of a project supported by the Surveying Office of North Rhine–Westphalia, who also provided the data of test sites 1 – 5.

REFERENCES

Baltsavias, E., Grün, A. and Gool, L. V. (eds), 2001. Automatic Extraction of Man-Made Objects from Aerial and Space Images (III). A. A. Balkema Publishers, Lisse / Abingdon / Exton (PA) / Tokyo.

Brenner, C., 2000. Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen. C, Vol. 530, Deutsche Geodätische Kommission, München.

Gülch, E., 1997. Application of semi-automatic building acquisition. In: A. Gruen, E. Baltsavias and O. Henricsson (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Birkhäuser, Basel, pp. 129–138.

Hanson, A., Marengoni, M., Schultz, H., Stolle, F., Riseman, E. and Jaynes, C., 2001. Ascender ii: A framework for reconstruction of scenes from aerial images. In: E. Baltsavias, A. Grün and L. V. Gool (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), A.A. Balkema Publishers, Lisse / Abingdon / Exton (PA) / Tokyo, pp. 25 – 34.

Heipke, C., Mayer, H., Wiedemann, C. and Jamet, O., 1997. Evaluation of automatic road extraction. In: IAPRS, Vol. 32, Part 3-2W3, pp. 47 – 56.

Henricsson, O. and Baltsavias, E., 1997. 3-d building reconstruction with aruba: A qualitative and quantitative evaluation. In: A. Grün, E. Baltsavias and O. Henricsson (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Birkhäuser, Basel, pp. 65–76.

Jamet, O., Dissard, O. and Airault, S., 1995. Building extracton from stereo pairs of aerial images: Accuracy and productivity constraint of a topographic production line. In: A. Grün, O. Kübler and P. Agouris (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images, Birkhäuser, Basel, pp. 231–240.

McGlone, J. and Shufelt, J., 1994. Projective and object space geometry for monocular building extraction. In: Proceedings Computer Vision and Pattern Recognition, pp. 54 – 61.

McKeown, D. M., Bulwinkle, T., Cochran, S., Harvey, W., McGlone, C. and Shufelt, J. A., 2000. Performance evaluation for automatic feature extraction. In: IAPRS, Vol. 33, Part B2, pp. 379 – 394.

Ragia, L., 2001. Ein Modell für die Qualität räumlicher Daten zur Bewertung der photogrammetrischen Gebäudeerfassung. Vol. 14, Gesellschaft für mathematische Datenverarbeitung - Forschungszentrum Informationstechnik GmbH.

Ragia, L. and Winter, S., 1998. Contributions to a quality description of areal objects in spatial data bases. In: IAPRS, Vol 32., Part 4, pp. 479 – 486.

Sester, M., Schneider, W. and Fritsch, D., 1996. Results of the test on image understanding of isprs working group iii/3. In: IAPRS, Vol. 31, Part B3, pp. 768–773.

Weidner, U., 1997. Gebäudeerfassung aus Digitalen Oberflächenmodellen. C, Vol. 474, Deutsche Geodätische Kommission, München.

Winter, S., 2000. Location similarity of regions. ISPRS Journal of Photogrammetry and Remote Sensing 55(3), pp. 189 – 200.