Quality Evaluation of Ground Plan Generalization

Yevgeniya Filippovska, Martin Kada, Dieter Fritsch
Institute for Photogrammetry, Universitaet Stuttgart,
Geschwister-Scholl-Str. 24D, 70174 Stuttgart, Germany
firstname.lastname@ifp.uni-stuttgart.de

ABSTRACT

Unlike model-oriented generalization, which reduces the amount of information within a data base, cartographic generalization concerns the optimization of spatial data towards a certain scale with regard to visualization. As it occurs on the geometry level, the object's shape is altered in the process, thus introducing geometrical distortions. The assessment of the rate at which the objects change can be seen as a quality evaluation of the generalization process. This paper is particularly interested in the simplification of building ground plans and the achieved accuracy when comparing the results to their original appearance. A classification of possible changes is given, as well as quality characteristics and their graphical interpretation. The quality parameters are defined and normalized, so that they can be used separately, but also in combination with others, depending on the requirement of a certain application.

1. INTRODUCTION

Even though the generation of maps has considerably changed with the technological transition from paper to digital media, the transformation of spatial data from larger to smaller scale still remains very challenging today. It even obtained a new meaning, as particularly mobile and interactive products are an emerging market that puts additional requirements on the representation of data.

If the information density of a map is too high, the important aspects are hard to perceive. Instead of being helpful in the decision making process, the map becomes confusing and useless. Thus, the cartographic generalization aims to avoid information overloading by means of reducing the amount of presented details. This is accomplished by the use of generalization operators that work on single objects, object groups or even the whole map and e.g. select, emphasize or simplify the data in accordance to its importance. For a classification of generalization operators and their description on a more abstract level, see e.g. (Forberg, 2007; Hake et al., 2002; Sester, 2000).

However, on the implementation level, a multitude of alternatives are possible that each produces different results. If two or more operators are combined in the generalization process, the number of possible outcomes grows even larger. In order to choose the most appropriate result according to one's requirements, it is very important to have quality criteria that help to rate the different results. These criteria can also be defined as quality characteristics in conformity with the ISO 9000 standard, which identifies quality as a "degree to which a set of inherent characteristic fulfills requirements" (TC 176/SC, 2005).

The focus of this paper is on the quality assessment of generalized building ground plans. After a general discussion on the quality of cartographic data, where different quality aspects of the generalization process as well as related works are briefly described, new ideas concerning the quality characteristics of building ground plan generalization are proposed. They quantify the differences between the original and generalized object on the basis of their contours and areas. Further, some examples are provided and discussed in the analysis section. Finally, the main points of the paper are recapitulated and an outline on future work is presented.

2. QUALITY OF CARTOGRAPHIC DATA

In the age of paper maps, the generalization process was mainly done by hand. So its quality was highly dependent from the experience and creativity of the cartographer. As both the rate of abstraction and the geometric accuracy were determined by the given scale, the precision of the map coordinates was directly related to the graphical minimum dimensions (Stadler and Lechthaler, 2006). Depending on the minimal distance that is still recognizable on a certain scale, different results are to be expected from the generalization.

However, not only the coordinates of object vertices change by the generalization, but also some vertices might get deleted or even new ones added in the generalization process. Although these new vertices are inserted into the ground plan in order to preserve the orientation or the extension of the object, they have no resemblance to reality. They do not exist in the real world, which means the quality of the generalization cannot be measured by the accuracy of coordinates anymore. There are no original coordinates to compare them with. Instead, the difference between the original and the derived level of detail must be made on various other geometric aspects.

With respect to the change of geometric accuracy by the generalization, we classify quality metrics into four groups in compliance to the properties of map representations following (Hake et al., 2002). These classes called trueness of ground plan, trueness of location, trueness of extension and trueness of shape place emphasis on different geometric aspects.

(1) Trueness of ground plan is a boundary-based characteristic. It can e.g. be expressed by the largest deviation between the contour of the original and generalized ground plan measured by the Hausdorff distance (Hangouët, 2006; Schlüter, 2001). Another interesting aspect is the portion of the original outline that remains the same after generalization, either exactly or within a pre-defined tolerance (cp. section 3.1).

(2) Trueness of location describes the positional change that an object undergoes, e.g. expressed by the translation of the centroid before and after generalization. This can be quite interesting as the centroid changes with a geometric simplification, thus making a spatial situation different to perceive by a human viewer.

(3) Trueness of extension, which is the area distribution of an object in space, can be identified in various ways. Examples are the area difference (Podolskaya et al., 2007), eccentricity (aspect ratio) and orientation by means of central moments (Ballard and Brown, 1982) or oriented bounding rectangles (Hild, 2003), the latter two being invariant against translation. The extensional difference between two polygons can also be expressed by the sum of intrusions and extrusions measured relative to the original object (Schlüter, 2001), which is defined in set theory as the symmetric difference.

(4) Trueness of shape is a combined property which considers both the outline of a polygon and its area. In order to assess this characteristic, geometric properties such as compactness, roundness and convexity were suggested (Werff and Meer, 2008).

3. CHARACTERISTICS FOR QUALITY EVALUATION OF GENERALIZATION

As already mentioned, the number of vertices and edges of the original ground plan will be reduced by the generalization. Although some edges of the simplified ground plan polygon remain the same or at least partially the same, it is very difficult to establish a correspondence between the original and the generalized vertices and edges. A good solution could maybe be established if the procedural method of the applied generalization algorithm is known, which unfortunately in most cases isn't. So it is important to look at some other means of comparing two features using universal properties. Instead of considering an object to be a sequence of vertices and edges, it is regarded as a closed and bounded set of points in Euclidian space. As a consequence, set theory can be applied for comparing two ground plans, which is very convenient as it does not contradict the aforementioned classification. In the following, we will show how they can be used to evaluate trueness of ground plan (section 3.1) and trueness of extension (section 3.2).

3.1 Trueness of ground plan

The shape of the ground plan can be considered only by its boundary. After the notion of set theory, it can be defined as a set of points in its closure without the interior points. The difference between two sets of points can then be measured by means of Hausdorff distance. For each point of the first set, the distance to a closest point in the other set is computed and the maximum of all distances taken. This is the largest point distance between the two boundaries. In fact, this analysis regards two outlines by their isometry. It proved to be very meaningful in terms of detecting omitted object parts and can therefore be utilized as the prime control of the correctness of the generalization. If the Hausdorff distance is larger than the generalization distance, which is the minimum length that line segments should have after generalization, parts of the object were unnecessarily deleted. Figure 1 shows an example, where the distance from the front of the entrance porch to the main façade wall has been found to be below the generalization threshold. Therefore the porch is missing in the generalized building. Besides this example, the effect predominantly occurs to elongated parts of the ground plan with a large area, but one small side.

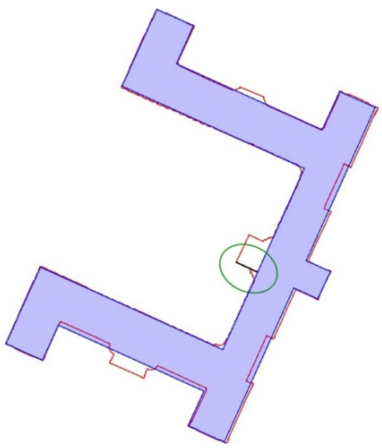


Figure 1: Hausdorff distance hinting towards an unnecessarily deleted building part.

Apart from the largest deviation between the boundaries of two objects, the percentage of the remaining part of the polygon is also of great interest. Again, the contours are considered as sets of points and the intersection operator known from set theory can be used. The result is the overlapping line segments of the two polygons.

Unfortunately, in spite of the simplicity of this solution, its implementation bears a few difficulties. They are related to the lack of strict geometric consistency in terms of small inaccuracies regarding the colinearity, parallelism and rectangularity of the line segments of the original building ground plan. It is inevitable that all measurements contain random errors, which cannot be completely excluded. One solution could be a preprocessing stage that adjusts the geometry to avoid such an effect. However, changing the original data is not always desired and the process might often be unambiguous. On the other side, generalization algorithms not only simplify polygons, but also enforce the aforementioned properties. This leads to small, but for our purpose significant differences that are sufficient to generate a very low overlap criterion of the two contours.

The problem can be solved by creating a buffer around the original polygon. This way, the matching of the two contours will not be based on strict geometric conditions, but rather tolerates small inaccuracies. However, the width of the tolerance region must be chosen with caution. If it is too large, the result will be overestimated. The size of the buffer reflects the strictness of conformity of the original ground plan to the main constraints.

Equation (1) represents the rate of boundary similarity of the original (O) and generalized (G) polygons or the outline intersection:

$$R_{boundary} = \frac{Buffer(O) \cap Perimeter(G)}{Perimeter(O)} \tag{1}$$

See Figure 2 for an example of this characteristic. The dashed contour parts represent the line segments that match the original shape.

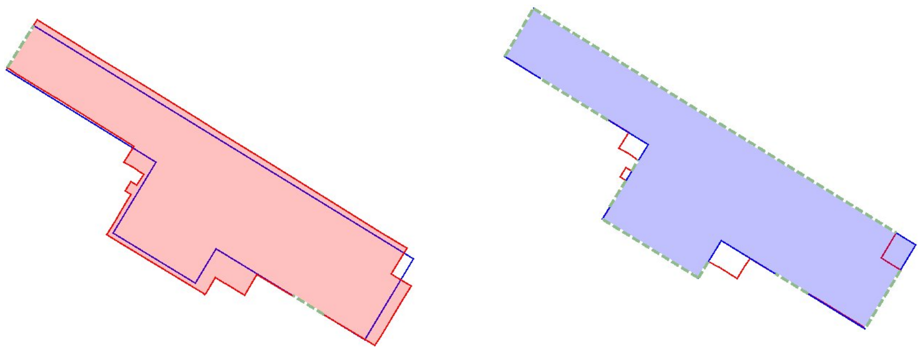


Figure 2: Matching contours (dashed) of two generalization results compared to their original shapes.

3.2 Trueness of extension

The modifications of the ground plan polygon that arises from the generalization can be characterized not only with the help of distance parameters. The change in the area of an object provides another very important aspect which can be used for the quality evaluation. The idea of such an approach is that the whole object including its boundary will be considered as a set of points. Although, the extension of two polygons will be compared instead of only their outlines, the analogy is the same as with the object boundary.

The unmodified area can be found as the intersection of the ground plan before and after generalization. The difference between this common area and the total area of the original polygon represents the area of the deleted parts, which is called intrusion. The added area, which was not part before the generalization, is called extrusion and can be computed as the difference between the area of the simplified object and its intersection with the original one. These rates are reflected by the equations (2) and (3).

$$R_{intrusion} = \frac{Area(O \cap G)}{Area(O)} \quad (2)$$

$$R_{extrusion} = \frac{Area(O \cap G)}{Area(G)} \quad (3)$$

In addition to the characteristics, which are based on the common part of two polygons, the extensions of these objects can also be compared as an area quotient between the generalized and original ground plans. The ratio allows the comparison of the polygon extensions even if they were noticeable displaced and, as a result of this, do not overlap. Otherwise, the areal ratio will reflect to what extent the intrusions and extrusions compensate each other. If the appearance of the distortions of an object caused by generalization is inevitable, they could be at least equalized.

4. ANALYSIS

In order to show the validity of the presented characteristics, they were tested on an example data set of 3D building models that were provided by the city surveying office of Stuttgart. However, even though the generalization of the data set was done in 3D, only the 2D ground plans were considered. The building models were simplified by the 3D generalization algorithms described in (Kada, 2007) which has a strong focus on generating a simplified topology, but does not explicitly generate coordinates that are true to the original ground plan. In order to have a data set to compare with, the generalized building models were adjusted by the approach of (Peter et al., 2008), which tries to best fit the ground plans to the original one.

The generalized and the adjusted polygons were both compared to the original ones using the quality characteristics presented in this paper. As these characteristics reflect different quality aspects, it is essential to have the ability to define the quality of the generalization process, taking into account all of the metrics simultaneously. This task is facilitated by the fact that the quality parameters are already normalized, where the best possible value equals 1 (apart from the Hausdorff distance which is to be considered separately). The smaller this value, the less similarity the original and generalized features possess. The only exception represents the area ratio, which exceeds 1 if the area of an object has increased after generalization. Equation (4) makes this quality parameter equal to the other characteristics. As for the tolerance buffer required for evaluating the rate of the boundary similarity, its width was determined empirical to be 0.15 m.

$$R_{area} = \begin{cases} \frac{Area(G)}{Area(O)} & if \quad Area(G) \leq Area(O) \\ \frac{2 * Area(O) - Area(G)}{Area(O)} = 2 - \frac{Area(G)}{A(reaO)} & if \quad Area(G) > Area(O) \end{cases} \quad (4)$$

For the purpose of aggregating the quality parameters (QP) to a total quality (TQ), the Euclidean distance was used, which ensures a better differentiation of the results than just a sum (see equation (5)). Finally, the outcomes were normalized so that a perfect match corresponds to 100. The quality characteristics calculated for the example data set (see Figure 3) are presented in Table 1.

$$TQ = \sqrt{\sum_{i=1}^{n} QP_i^2} \tag{5}$$

The general correctness of the generalization results were evaluated using the Hausdorff distance. For the 'New Palace', it turned out to be larger than the minimal generalization distance. Indeed it indicates the omission of a whole part of the ground plan. The pure adjustment of the outline cannot improve this characteristic as it does not change the topology. Altogether, the Hausdorff distance can worsen by the adjustment as it moves averaged edges away from small intrusions and extrusions to better fit an edge with a higher importance.

Building	Generalization algorithm	Min. gener distance	HD	Perimeter	Outline intersection	Original area	Generalized area	Intersection area		R_{area}	TQ	TQ (normalized) 0..100
								R_{intr}	R_{extr}			
		m				m²						
Hindenburgbau	Kada 2007	10	8,52	365,40	51,53	3975,09	3481,08	3424,10		0,77	1,58	79,0
					0,02			0,74	0,97			
	Peter 2008		7,36		311,57		3919,16	3851,92		0,97	1,90	94,9
					0,73			0,94	0,97			
Opera House	Kada 2007	5	4,32	314,08	107,85	4549,75	4691,08	4518,51		0,94	1,72	86,2
					0,12			0,99	0,93			
	Peter 2008		5,59		230,63		4644,69	4518,30		0,96	1,85	92,6
					0,54			0,99	0,95			
New Palace	Kada 2007	10	10,65	918,84	218,95	7676,60	7421,71	7179,16		0,93	1,67	83,7
					0,06			0,87	0,94			
	Peter 2008		10,65		503,02		7478,98	7284,41		0,95	1,76	88,0
					0,30			0,90	0,95			

Table 1: Quality comparison of two generalization algorithms by means of the individual and the aggregated parameters.

Regarding the total quality characteristic TQ, which aggregated all quality parameter like the rate of the boundary similarity, the rate of intrusions and extrusions, and the area ratio, were all improved after adjustment. As it was expected, it concerns especially the outline intersection of polygons where the changes of quality are drastic.

The rate of intrusions and extrusions was very effectively diminished by the generalization and makes over 0.90. Essentially this characteristic, as the area ratio, was only improved by the adjustment for the Hindenburgbau, where the area intrusion was significantly too large.

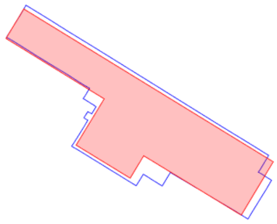
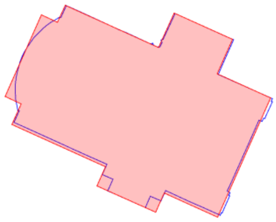
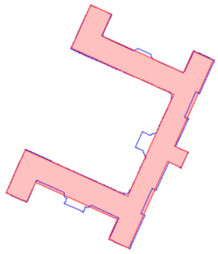
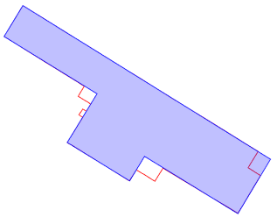
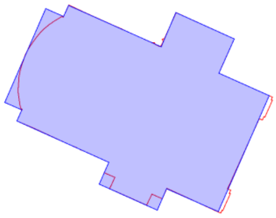
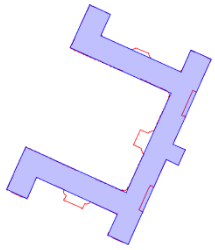
Generalization algorithm	Hindenburgbau	Opera House	New Palace
Kada 2007			
Peter 2008			

Figure 3: Example buildings used in the quality analysis.

Finally, the aggregated quality characteristic identifies the Hindenburgbau as the worst generalized ground plan which benefits most by the adjustment. The ground plans of the New Palace and the Opera House were also improved. However, due to its complexity, the ground plan of the 'New Palace' matches the original object less than the other two.

5. SUMMARY AND OUTLOOK

This paper is dedicated to the quality analysis of cartographic generalization. Specifically, it deals with single objects which are represented by building ground plans. Since there are plenty of generalization algorithms offering different results, it becomes important to be able to compare them in order to find the most suitable outcome. In many cases, it is almost impossible to identify a single best solution. The problem is that different properties of an object can be prioritized by the generalization process. Thus, some approaches aim to preserve the outline of a ground plan, the others tend to maintain its area. And these are not the only parameters which can be used to characterize similarity or, in other words, the quality of the polygons before and after generalization. This work suggests a classification of all quality parameters for properties an object can relate to. It includes trueness of ground plan, trueness of location, trueness of extension and trueness of shape.

According to the final total quality metric, which aggregates all the described characteristics except the Hausdorff distance, the ground plans generalized by the algorithm were evaluated to have been improved by means of the adjustment. Especially it concerns the outline intersection of polygons, which is the purpose of the adjustment and the focus of the paper. Other characteristics, especially the ones that have been mentioned in section 2, are subject of further research.

BIBLIOGRAPHY

Ballard, D.H. and Brown, C.M., 1982. Computer Vision. Prentice-Hall, London, 523 pp.

Forberg, A., 2007. Generalization of 3D Building Data Based on a Scale-Space Approach. ISPRS Journal of Photogrammetry and Remote Sensing, 62(2): 104-111.

Hake, G., Grünreich, D. and Meng, L., 2002. Kartographie. Walter de Gruyter Verlag, Berlin, 604 pp.

Hangouët, J.-F., 2006. Spatial Data Quality Assessment and Documentation. In: R. Devillers and R. Jeansoulin (Editors), Fundamentals of Spatial Data Quality. Geographical Information Systems Series. Fundamentals of Spatial Data Quality, pp. 211-235.

Hild, H., 2003. Automatische Georeferenzierung von Fernerkundungsdaten. Dissertation Thesis, Universität Stuttgart, Stuttgart.

Kada, M., 2007. Scale-Dependent Simplification of 3D Building Models Based on Cell Decomposition and Primitive Instancing, Proceedings of the International Conference on Spatial Information Theory: COSIT ’07, Melbourne, Australia.

Peter, M., Haala, N. and Fritsch, D., 2008. Preserving Ground Plan and Facade Lines for 3D Building Generalization, The XXI Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS 2008), Beijing, China, pp. 481-486.

Podolskaya, E.S., Anders, K.-H., J.-H.Haunert and Serester, M., 2007. Quality Assessment for Polygon Generalization, 5th International Symposium on Spatial Data Quality (SDQ 2007), ITC, Enschede, The Netherlands.

Schlüter, D., 2001. Hierarchisches Perzeptives Gruppieren mit Integration dualer Bildbeschreibungen, Bielefeld, pp. 222.

Sester, M., 2000. Maßstabsabhängige Darstellungen in Digitalen Räumlichen Datenbeständen. Professorial Dissertation Thesis, Universität Stuttgart, Stuttgart.

Stadler, A. and Lechthaler, M., 2006. Output Media Adapted Cartographic Visualisation, 10th International Conference on Information Visualization (IV'06). IEEE Computer Society, London, pp. 304-309.

TC 176/SC, 2005. Quality Management Systems - Fundamentals and Vocabulary, International Organization for Standardization.

Werff, H.M.A.v.d. and Meer, F.D.v.d., 2008. Shape-Based Classification of Spectrally Identical Objects. ISPRS Journal of Photogrammetry and Remote Sensing, 63(2).