Guillaume Touya

What Is the Level of Detail of OpenStreetMap?

1. Level of Detail Problems in OpenStreetMap Maps

The cartographic representation of geographical phenomena includes a level of detail (LOD) that can be considered as the scale to display the representation or the scale at which the phenomena can be meaningfully analyzed (Mackaness, 2007). Volunteered geographic information (VGI) is the result of the contributions of ordinary citizens (Goodchild, 2007), therefore, each contribution follows its own level of detail standards. For instance, very precise fountain features coexist with built-up area features in the same dataset. In OpenStreetMap (OSM), the most successful VGI project, level of detail heterogeneity both results from contributor will and from capture tools. Indeed, a contribution may be added using precise GPS tracks, vectorization on scanned maps or satellite imagery, or from open existing vector data (e.g. land use data in France is extracted from CorineLandCover® European data). The sources variety leads to heterogeneous data quality (Girres & Touya, 2010; Haklay, 2010).

This level of detail heterogeneity has a major drawback: whatever the scale, it is very difficult to make legible maps with OSM data. When scale is small, very detailed features like roads or buildings need cartographic generalization to be legible (Figure 1A & B). Indeed, (Gaffuri, 2011) claims generalization is a problem for most web maps. On the other hand, when scale is large, small scale features may be inconsistent with very detailed features (Figure 1C & D).

Figure 1 Four examples of maps (©OpenStreetMap) with level of detail problems. (a) The mountain road (orange) is too detailed compared with land use and coalesces at this scale. (b) Buildings are too small to be displayed at this scale. (c) The forest is much less detailed than buildings causing overlaps. (d) Industrial (purple) and commercial (grey) parcels overlap buildings as their level of detail is not compatible with maps with buildings at this scale.

2. Research Topics to Overcome OSM Limitations

VGI is very rich, sometimes richer than traditional datasets and it would be a pity not to overcome the above mentioned limitation to produce legible maps. Several research topics to tackle to overcome the limitations are described below.

The first research topic could be to define VGI-specific metadata, including LOD information, or to develop formal specifications for VGI (Brando, 2012). However, interviewed OSM users refuse to add heavy constraints to the contribution. Indeed, even simple attribute tags are theoretically foreseen but actually sparsely used in OSM (Girres & Touya, 2010).

Multiple representation databases (MRDB) are a research field that tries to design systems and models to store and represent multiple representations for geographic phenomena through scale and time (Balley et al, 2004). Converting VGI into MRDB is an interesting challenge, and it could be achieved using the linked data framework (Hahmann & Burghardt, 2010).

Besides MRDB, mapping OpenStreetMap data at different scales requires to harmonize the level of detail of mapped features. It means that spatial analysis techniques should be developed to automatically determine the level of detail from shape, accuracy or spatial relations of features. For instance, a forest area that has geometry with few vertices and overlaps several buildings (Figure 1C) should be considered as poorly detailed. Then, map generalization techniques could be used to harmonize features to the same level of detail.

Finally, a solution to overcome the level of detail heterogeneity could be to provide reconciliation tools when data is captured to avoid the inconsistencies created by LOD differences. (Brando, 2012) proposes such techniques, for inconsistencies more general than LOD differences, to allow VGI following classical integrity constraint by comparison with reference data. Figure 2 shows reconciliation example that helps overcome level of details problems.

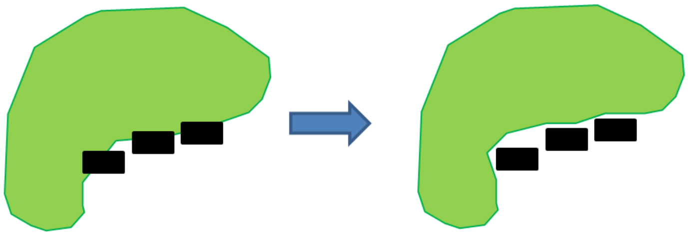

Figure 2 Reconciliation example: the forest should not overlap buildings so its outline is distorted.

References

Balley, S., C. Parent, and S. Spaccapietra (2004). Modelling geographic data with multiple representations. *International Journal of Geographical Information Science 18* (4), 327-352.

Brando, C. (2012). Edition collaborative d'un contenu géographique : une approche fondée sur des spécifications formelles et des relations spatiales explicites. In Journées de la Recherche IGN, Saint-Mandé, France.

Gaffuri, J. (2011). Improving web mapping with generalization. *Cartographica: The International Journal for Geographic Information and Geovisualization 46* (2), 83-91.

Girres, J.-F. and G. Touya (2010). Quality assessment of the french OpenStreetMap dataset. *Transactions in GIS 14* (4), 435-459.

Goodchild, M. F. (2007). Citizens as voluntary sensors: spatial data infrastructure in the world of web 2.0. *International Journal of Spatial Data Infrastructures Research 2*, 24-32.

Haklay, M. (2010). How good is volunteered geographical information? a comparative study of OpenStreetMap and ordnance survey datasets. *Environment and Planning B: Planning and Design 37* (4), 682-703.

Hahmann, S. and D. Burghardt (2010). Linked data – a multiple representation database at web scale? In *Proceedings of the 13th ICA Workshop on Generalisation and Multiple Representation*, Zurich, Switzerland.

Mackaness, W. A. (2007). Understanding geographic space. In W. A. Mackaness, A. Ruas, and T. Sarjakoski (Eds.), *The Generalisation of Geographic Information: Models and Applications*, Chapter 1, pp. 1-10. Elsevier.