HAL open science

Generalising Unusual Map Themes from OpenStreetMap

Guillaume Touya, Jean-François Girres

To cite this version:

Guillaume Touya, Jean-François Girres. Generalising Unusual Map Themes from OpenStreetMap. 17th ICA Workshop on Map Generalisation and Multiple Representation, Sep 2014, Vienna, Austria. hal-02280105

HAL Id: hal-02280105

https://hal.science/hal-02280105v1

Submitted on 6 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Generalising Unusual Map Themes from OpenStreetMap

Guillaume Touya[1], Jean-François Girres[1,2]
1 COGIT - IGN France, 73 avenue de Paris 94165 Saint-Mandé France
2 UMR GRED - Université Paul-Valéry Montpellier III, Route de Mende 34199 Montpellier Cedex 5

1. Introduction

Research on generalisation algorithms mainly focused on the most prominent themes of a topographic map, namely roads, buildings, hydrography, terrain and land use. Now that many efficient algorithms exist for such prominent themes (Stanislawski et al. 2014), minor themes like ditches (Savino et al. 2011) or swamp patches (Anderson-Tarver et al. 2011) gain interest from the research community. In OpenStreetMap (OSM), the loose specifications resulted in a large range of unusual themes, e.g. electric lines (Figure 1c), and generalising maps out of OSM data requires to handle such unusual themes. Moreover, another particularity of OSM is the very high level of detail that can be found for some features, compared to traditional geographical databases, e.g. airports, port sites (Figure 1a) or parks (Figure 1b). The unusual details contained in such areas also require specific algorithms to preserve the structure and the relations between the components. Generalising OSM data is a challenging task (Sester et al. 2014) but this paper only focuses on the generalisation of those map themes or areas exhibited in OSM.

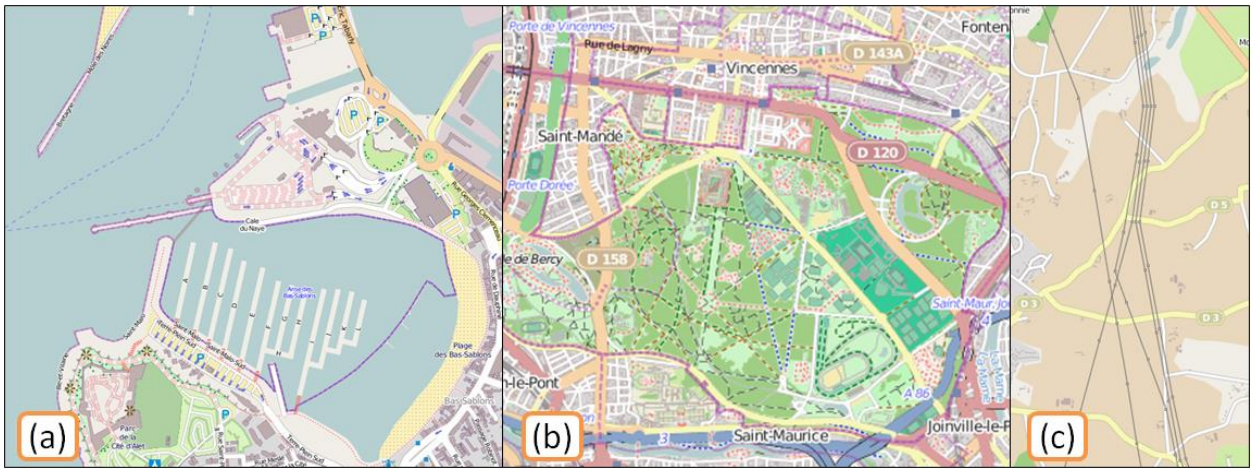

Figure 1. Sites with unusual detail ((a) a port and (b) a park), and unusual map theme ((c) electric lines) found in OpenStreetMap. ©OpenStreetMap

The second part of the paper describes algorithms to generalise airports and their components. The third part focuses on railways generalisation. The last part concludes and presents some perspectives.

2. Generalising Airports

2.1. Modelling Airports

Airports are composed of specific features that are generally well captured in OSM: runways, taxiways, aprons and terminals, for instance (Figure 2). Additional features like helipads, hangars or windsocks are omitted here to focus on the main features. These features can either be polygons or lines, and when scale decreases, generalisation is clearly required to typify these features (Figure 2). In order to generalise airports properly, airports are modelled as complex objects composed of these features, that sometimes can either be lines or polygons (Figure 3).

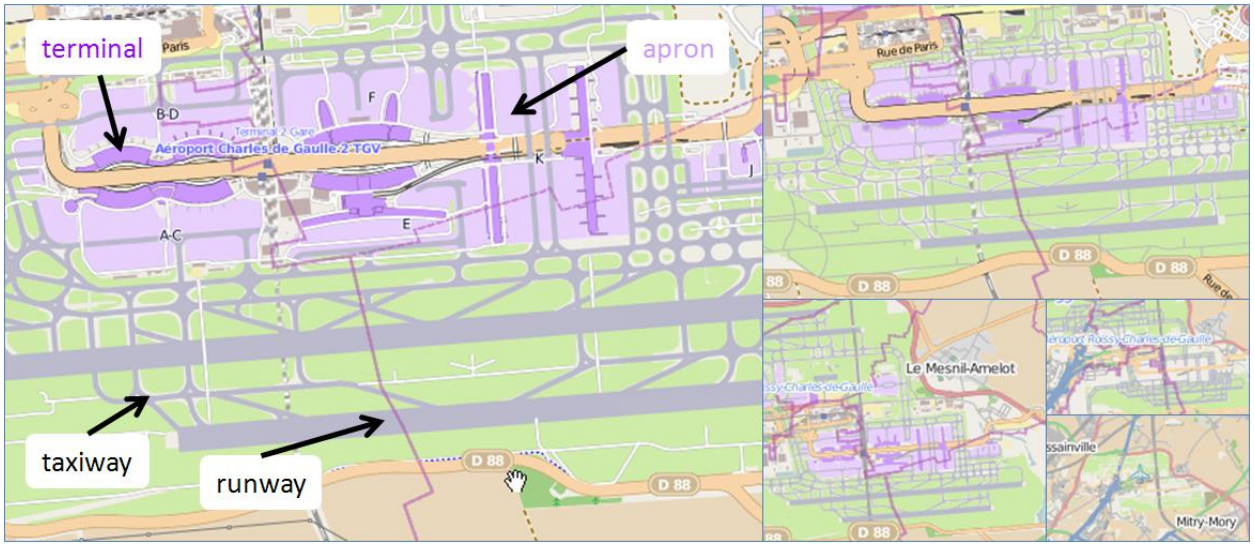

Figure 2. Highly detailed airport: when zooming out (pictures on the right), such details require generalisation. ©OpenStreetMap

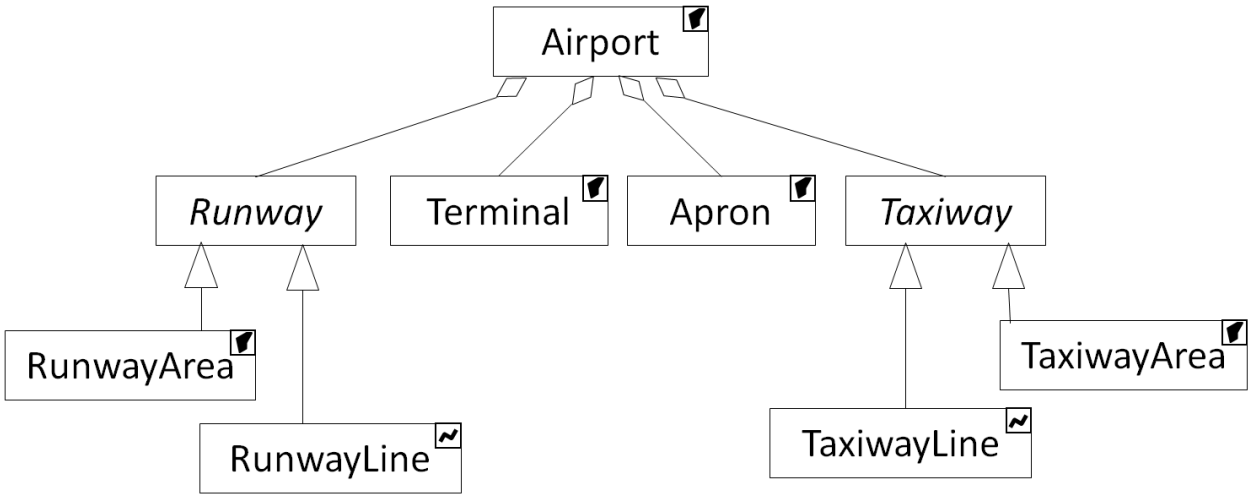

Figure 3. UML class model of airports used to generalise OSM airports.

Several algorithms are proposed to generalise these features: runway collapse, taxiway collapse and typification, apron and terminal amalgamation, and algorithms change whether the

taxiways/runways are lines or polygons. Figure 4 show a simplified version of the proposed generalisation guidelines for OSM airports, into a ScaleMaster (Brewer & Buttenfield 2007). Airport features are mainly large scale features, as all should be removed at the 1:200k for instance (only the extent with a ponctual symbol should be represented from these scales).

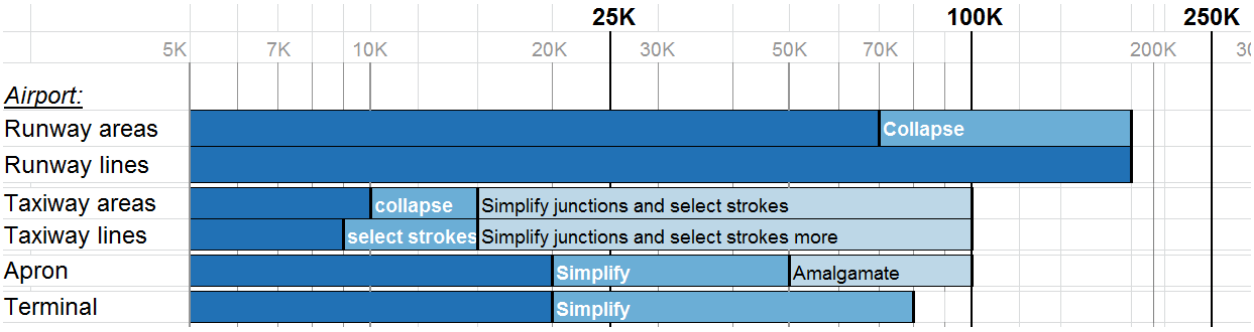

Figure 4. ScaleMaster (Brewer & Buttenfield 2007) modelling of airports generalisation.

The algorithms described in the next subsections can be used gradually to increase the generalisation level as scale decreases. Only runway and taxiway algorithms are presented here for the sake of brevity.

2.2 Runway Collapse

Runways are long straight polygonal lanes that need to be collapsed into straight lines when scale decreases. Rather than complex skeletonizing algorithms, we use the longest segment that can be drawn inside the runway polygon that is parallel to the orientation of the polygon (Duchêne et al. 2003) (Figure 5). As the general orientation algorithm is based on the minimum bounding rectangle of the runway polygon, it is not so sensible to the granularity of the geometry. As a consequence, the algorithm obtains good straight generalised runways for airports found across the world in OSM.

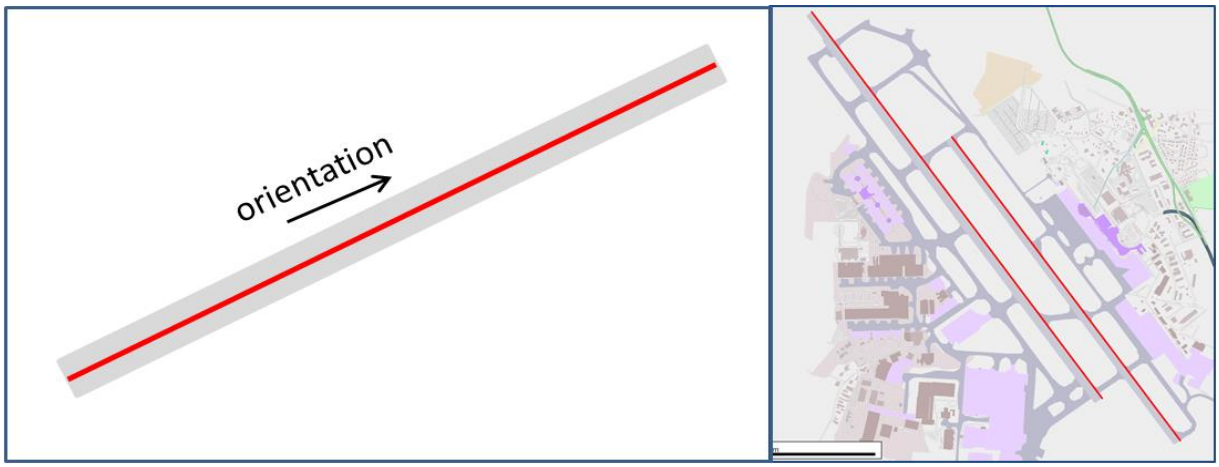

Figure 5. Principles of the runway collapse algorithm and results on Blagnac airport on the right. ©OpenStreetMap

2.3 Taxiways Collapse and Typification

Taxiways can have either polygonal geometries or linear geometries. When geometries are polygons, taxiways have to be collapsed, beginning by the thin parts, and then including all taxiways as scale decreases. A complete collapse algorithm like the one by Szombara (2013) could be used, but we chose a two-steps process: first thin parts are identified by erosion, and then the thin parts are skeletonized.

To skeletonize taxiways, a three-step process is proposed. Polygonal taxiways in OSM are generally represented through a complex assembly of connected polygons. Thus, adjacent borders of neighboring polygons (of taxiways and runways) are firstly detected (Figure 6a) in order to preserve the structure of the taxiway network. To collapse taxiways, the skeleton is obtained using Voronoi diagrams extracted from a triangulation (Figure 6b) in each polygon. Diagrams are then connected through adjacent borders detected in the first step, and finally pruned and filtered in order to generate linear taxiways, as exposed in Figure 6c. For instance, in the example of the generalisation of taxiways on Blagnac airport in Figure 6, a minimal length of 100 meters is needed to preserve linear taxiways.

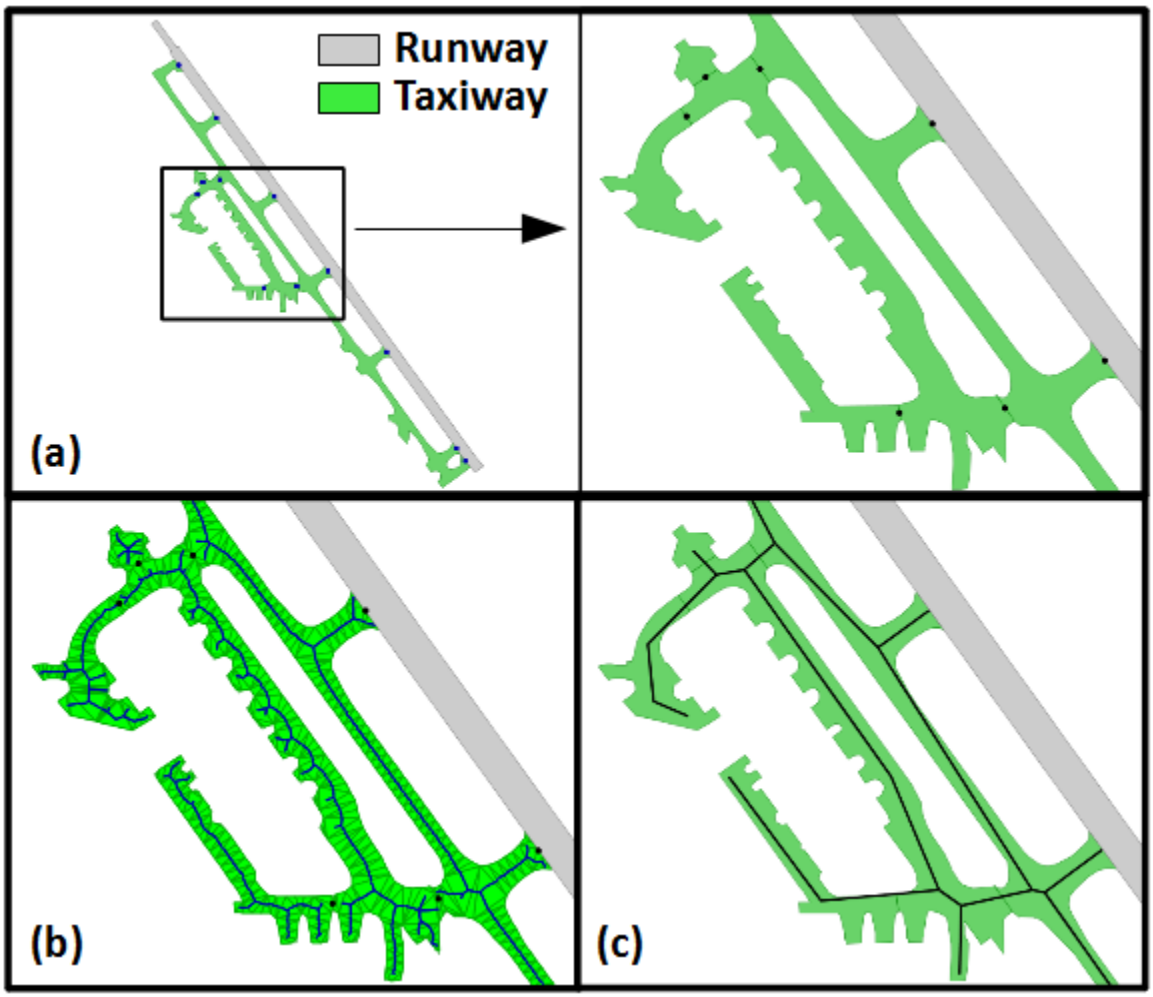

Figure 6. Three-step process to collapse taxiways on Blagnac airport. ©OpenStreetMap

When taxiways are lines, two algorithms are proposed: crossroad typification and stroke-based selection. Complex crossroads are identified using graph and shape measures (Figure 7) and collapsed by creating a simple connection, like branching crossroads can be collapsed in a road network. More precisely, three types of complex crossroads are identified in all studied airports with linear taxiways: simple branchings (in red in Figure 7), branchings with a central connection (in pink), and double branchings (two connected simple branchings, in blue). For instance, the simple branchings are the taxiway graph faces with three edges and three degree-3 nodes, that have a triangular shape. Then, the edge to keep is identified by angle considerations, and the node that is not connected to this edge to keep is extended to reach the edge (Figure 7). Similar methods are applied to both other types of complex crossroads.

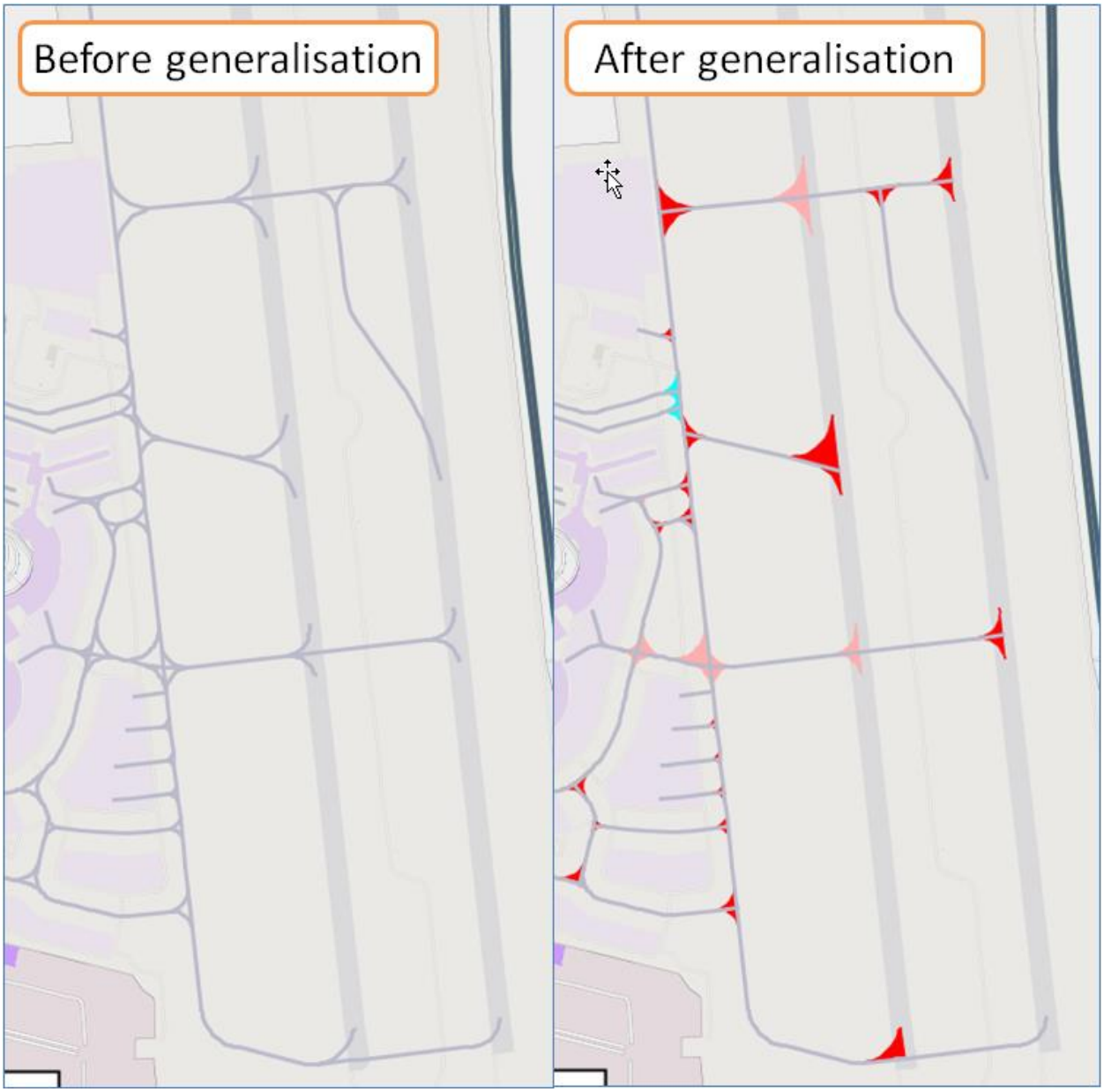

Figure 7. Linear taxiway complex crossroads are identified (coloured areas) and collapsed.

Then, strokes are computed and only the longest ones are selected. Length threshold are related to the scale of the generalised map. For instance, taxiway strokes smaller than 0.4 map mm are removed. If a stroke connects two other strokes that are above the threshold, it is kept to preserve connectivity.

2.4 Results

Figure 8 shows some 1:25k generalisation results in three airports in different countries (Lyon, Vienna and Rio) to show robustness across countries. In this case, all three airports had linear runways and taxiways so generalisation algorithms were triggered on taxiway lines, aprons and terminals.

Figure 8. 1:25k generalisation of three airports in Lyon, Vienna and Rio.

3. Generalising Railways

Railway networks are composed of main railway lines and sidetracks, isolated or grouped in triage areas. Main tracks can simply be generalised by a stroke-based selection (Thomson & Brooks 2007). But, due to OSM heterogeneity, main tracks are sometimes divided in two or more parallel lines corresponding to the rail lanes (Figure 9). In this case, a collapse into a single lane is necessary before stroke-based selection can be applied.

Figure 9. Three, then four parallel railway lines. ©OpenStreetMap

3.1. Collapsing Main Railways

To collapse such parallel railway lines, the first step is the identification of the railway portions where such a structure appears. Using the spatial relation model from (Touya et al 2012), the sections where two railways are close and parallel are identified, and then collapsed using a computed average line between both sections (Figure 10). When more than two sections are involved (Figure 9), the average is computed between all the sections.

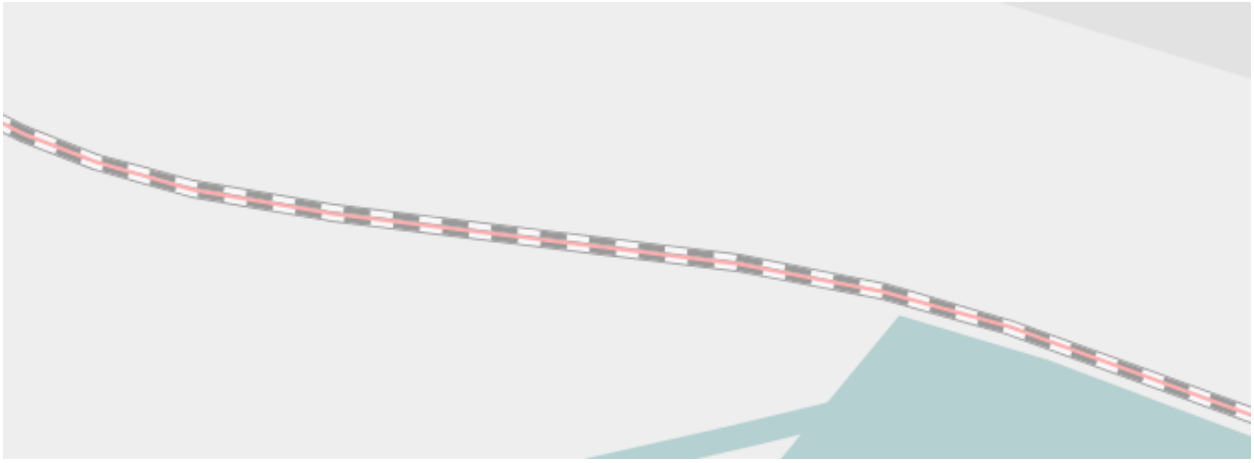

Figure 10. Results of main railways collapse (collapsed line in pink).

Two types of junctions can occur for parallel railway lines: parallelism ends with one or more lines turning in a different direction, or the lines intersect (Figure 11). The intersection case is easy to handle with the intersection point used as the final point of the collapsed line. When some lines change directions, the lines are reconnected at the point with parallelism ends (Figure 11).

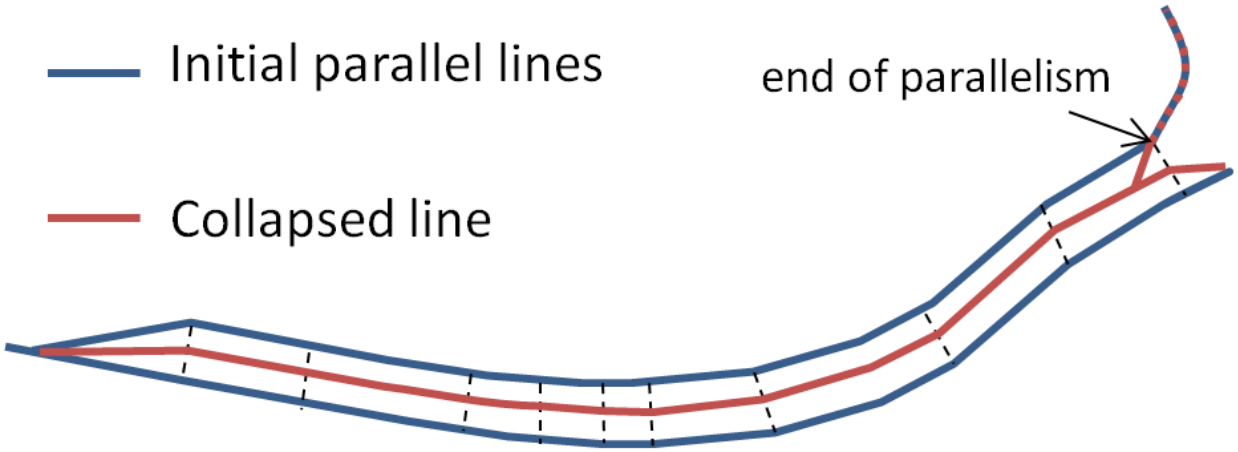

Figure 11. Results of main railways collapse (collapsed line in pink).

3.2. Sidetracks Typification

When grouped in a triage area or a train station, sidetracks are grouped in a particular structure that needs to be preserved by typification as scale decreases (Figure 12). Sidetracks are identified when a line is tagged by both “railway” and “service” tags. Triage areas are then identified by grouping sidetracks by connexity. Strokes (Thomson & Brooks 2007) are computed inside each triage areas and sorted according to length by descending order. As long as it is bigger than a minimal length, long strokes are selected and the surrounding strokes (that are shorter) are removed from the list to avoid symbol overlap. Finally, if the selected strokes are not connected to the main railways, features enabling connection are selected. Results can be seen in Figure 12 for a generalisation to 1:25000 scale.

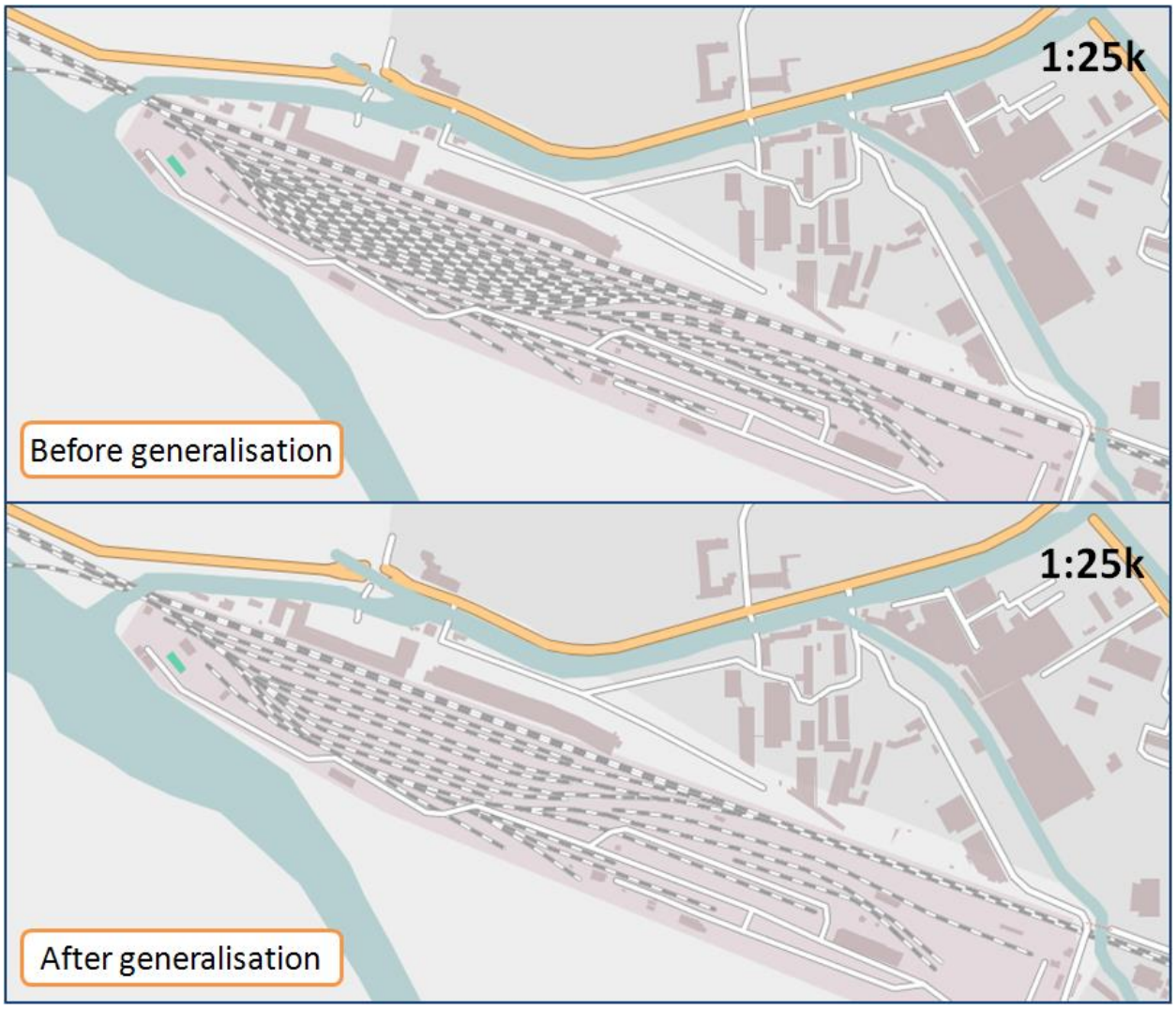

Figure 12. Results of sidetracks typification around the main station of Pau. ©OpenStreetMap

4. Discussion

Due to the varying expertness of contributors, and the imprecision of some capturing methods (e.g. recognition on satellite imagery), the quality of OSM data is very heterogeneous (Girres & Touya 2010, Haklay, 2010). This heterogeneity has an impact on generalisation processes that are mostly designed for good and homogeneous data. Regarding, the unusual themes presented in this paper, some railroad sections might be disconnected to their neighbours, for instance, which makes the proposed algorithms useless. We believe that one way to tackle the heterogeneity is to provide correcting pre-processes with each algorithm, which correct the data regarding the algorithm requirements. For instance, our railway algorithms require a planar graph (the OSM specification recommend to capture the railways as a planar graph). So the algorithms are combined with a pre-process that makes the railways network a planar graph.

Despite this quality heterogeneity, most airports are similarly captured across countries (Figure 13), because of the unique set of specifications, and airports can be captured from aerial photos

by european contributors all over the the world. So a generalisation process set up with european airports is quite applicable elsewhere in the world.

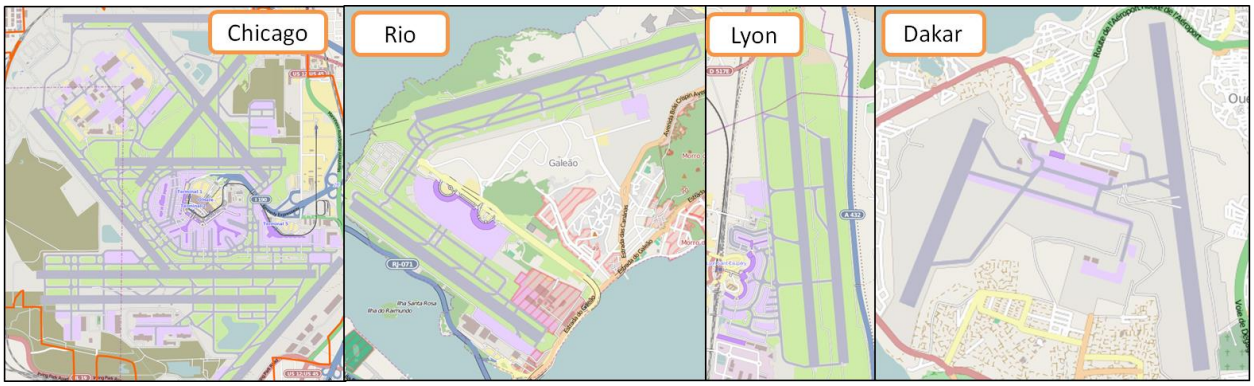

Figure 12. Airports in different parts of the world: they are captured similarly. ©OpenStreetMap

(Touya & Brando 2013) shows that there is also a level of detail heterogeneity in OSM. Geographical features are not always captured at the same level of detail for different themes (e.g. cadastral buildings vs. forests from Corine LandCover dataset), and even in the same theme, as objects may be captured with sources (GPS tracks, satellite imagery) at different levels of detail. Taxiways are a good example of this level of detail heterogeneity as it is sometimes captured with details at complex junctions (e.g. Chicago in Figure 12), and sometimes with simplified representations (e.g. Dakar in Figure 12). As a consequence, similar objects do not require the same degree of generalisation for a same final scale, as the initial level of detail may differ (i.e. there no need for complex junctions collapse in Dakar airport).

Traditional generalisation orchestration seeks to analyse the geographical context of an object (or a group of objects), in order to find the best suited algorithms and parameters for this object regarding its context. With OSM, we believe that orchestration also has to analyse the quality and level of detail of the objects to generalise, in order to choose the best suited algorithms.

5. Conclusion and Further Work

To conclude, making maps out of OSM data requires the development of new algorithms either because new type of geographical features are included in the dataset, or because new geographical structures to generalise emerge from the high details OSM data can have in some active areas. In this paper, algorithms are proposed and tested in OSM data, for airport features and railways.

Map generalisation is not only about algorithms, and the main difficulty is to orchestrate the use of the algorithms in an automatic process. So, to go further, the proposed algorithms have to be included in orchestration processes like CollaGen (Touya & Duchêne 2011) or the ScaleMaster2.0 (Touya & Girres 2013). Moreover, other types of unusual map themes, such as the ones presented in Figure 1, have to be tackled, in order to improve automatic mapping from OSM data.

References

Anderson-Tarver, C., S. Leyk, and B. P. Buttenfield (2011). Fuzzy modeling of geometric textures for identifying archipelagos in Area-Patch generalization. *Cartography and Geographic Information Science 38* (2), 137-145.

Brewer, C. A. and B. Buttenfield (2007). Framing Guidelines for Multi-Scale Map Design Using Databases at Multiple Resolutions. *Cartography and Geographic Information Science 34* (1), 3-15.

Duchêne, C., S. Bard, X. Barillot, A. Ruas, J. Trevisan, and F. Holzapfel (2003). Quantitative and qualitative description of building orientation. In *fifth workshop on progress in automated map generalisation*, Paris, France.

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

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.

Haunert, J.-H. and M. Sester (2004). Using the straight skeleton for generalisation in a multiple representation environment. In *6th CA Workshop on progress in automated map generalisation*. ICA.

Regnauld, N. and P. Revell (2007). Automatic amalgamation of buildings for producing ordnance survey 1:50 000 scale maps. *The Cartographic Journal 44* (3), 239-250.

Savino, S., M. Rumor, and M. Zanon (2011). Pattern recognition and typification of ditches. In A. Ruas (Ed.), *Advances in Cartography and GIScience*, Volume 1 of *Lecture Notes in Geoinformation and Cartography*, pp. 425-437. Berlin, Heidelberg: Springer.

Sester, M, Jokar Arsanjani, J, Klammer, R., Burghardt, D. and Haunert, J-H (2014). Integrating and Generalising Volunteered Geographic Information. In: Burghardt, D., Duchêne, C. and Mackaness, W. (Eds.) Abstracting Geographic Data in a Data Rich World. Springer, Berlin, 119-155.

Szombara, S. (2013). Unambiguous collapse operator of digital cartographic generalisation. In *Proceedings of 16th ICA Workshop on Generalisation and Multiple Representation*, Dresden, Germany.

Thomson, R. C. and R. Brooks (2007). Generalisation of geographic networks. In W. A. Mackaness, A. Ruas, and L. T. Sarjakoski (Eds.), *Generalisation of Geographic Information: Cartographic Modelling and Applications.*, pp. 255-267. Elsevier.

Touya, G. and C. Brando-Escobar (2013). Detecting Level-of-Detail inconsistencies in volunteered geographic information data sets. *Cartographica: The International Journal for Geographic Information and Geovisualization 48* (2), 134-143.

Touya, G. and C. Duchêne (2011). CollaGen: Collaboration between automatic cartographic generalisation processes. In A. Ruas (Ed.), *Advances in Cartography and GIScience*, Volume 1 of *Lecture Notes in Geoinformation and Cartography*, pp. 541-558. Berlin, Heidelberg: Springer.

Touya, G. and J.-F. Girres (2013). ScaleMaster 2.0: a ScaleMaster extension to monitor automatic multi-scales generalizations. *Cartography and Geographic Information Science 40* (3), 192-200.

Touya, G., S. Balley, C. Duchêne, K. Jaara, N. Regnauld, and N. Gould (2012). Towards an ontology of generalisation constraints and spatial relations. In *Proceedings of 15th ICA workshop on generalisation and multiple representation*, Istanbul, Turkey.