Assessing global OpenStreetMap building completeness to generate large-scale 3D city models

Filip Biljecki[1,*] and Li Min Ang[1]

[1] Urban Analytics Lab, National University of Singapore, Singapore; filip@nus.edu.sg, limin.ang1@gmail.com

* Author to whom correspondence should be addressed.
This abstract was accepted to the Academic Track of the State of the Map 2020 Online Conference after peer-review.

Quality assessment of OpenStreetMap (OSM) data has been an important topic since the inception of the project [1]. Much research has been done on this topic by many research groups around the world, and it can mostly be seen as permutations of three aspects: (1) spatial data quality element(s) in focus (e.g. positional accuracy, completeness), (2) theme (e.g. amenities, buildings, roads), (3) geographical area (e.g. particular city or country); e.g. positional accuracy of cultural features in Italy [2].

Completeness is one of the principal quality aspects of geospatial data, and our research focuses on developing a method to assess the completeness of buildings in OSM on a large scale (spanning several countries). While there are many robust OpenStreetMap completeness techniques and studies developed, they mainly focus on limited areas, mostly developed countries with ground truth data at hand for comparisons.

Doing the same for less developed regions is rare as the lack of authoritative data inherently hampers it, and the methods hardly ever scale: such an analysis done simultaneously for more than one administrative region is seldom carried out because there are other research challenges such as disparate urban morphology, different data sources and standards to bridge in order to facilitate ground truth, and varying understanding of what a building is. Furthermore, the development of a method that would scale across dozens of countries is limited by computational resources.

We are currently developing a method that uses several indicators derived from remote sensing, which are available on the global level, that may hint at the building completeness and would scale across the world. A regression model to predict the approximate volume of buildings in a given area is trained in areas in which there is an indication of high completeness of buildings in OSM.

OSM building completeness is estimated by comparing the area of mapped buildings against their expected (predicted) coverage in reality. The method has the potential to scale worldwide, and completeness is estimated for a grid of resolution of approximately 1x1 square kilometres, and simultaneously for administrative regions to enable cross-country comparisons.

Biljecki, F., & Ang, L. M. (2020). Assessing Global OpenStreetMap building completeness to generate large-scale 3D city models In: Minghini, M., Coetzee, S., Juhász, L., Yeboah, G., Mooney, P., & Grinberger, A. Y. (Eds.). Proceedings of the Academic Track at the State of the Map 2020 Online Conference, July 4-5 2020. Available at https://zenodo.org/communities/sotm-2020
DOI: 10.5281/zenodo.3922285

The work is being implemented in Google Earth Engine, mostly relying on imagery and indicators such as normalised difference built-up index (NDBI) and normalised difference vegetation index (NDVI). Preliminary results suggest a substantial disparity in OSM building completeness around the world, with areas that are entirely complete to those with inadequate completeness.

This preliminary paper aims to report the progress and the encountered challenges in the ongoing project such as selecting indicators that are consistently available on the global level, applying the method widely, accounting for different mapping practices around the world, different notions of a building, and varying morphologies of cities and urban areas. We also investigate the relation between OSM building completeness and socio-economic parameters such as the gross domestic product (GDP) to understand their relationships to mapping intensity and quality. These may offer the potential to be used as additional predictors.

This work is part of a broader project conducted at the Urban Analytics Lab at the National University of Singapore on investigating the potential of generating 3D city models by extruding building footprints in OpenStreetMap to a building height that is predicted using artificial intelligence [3]. This ongoing portion of the project is the crucial first step in the project, as it will enable us to understand what is the completeness of building footprints in OpenStreetMap around the world and manage expectations about the potential coverage of 3D city models.

Acknowledgements

This research is part of the project Large-scale 3D Geospatial Data for Urban Analytics, which is supported by the National University of Singapore under the Start Up Grant R-295-000-171-133.

References

[1] Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C., & Haklay, M. M. (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1), 139-167.

[2] Balducci, F. (2019). Is OpenStreetMap a good source of information for cultural statistics? The case of Italian museums. *Environment and Planning B: Urban Analytics and City Science*.

[3] Biljecki, F., Ledoux, H., & Stoter, J. (2017). Generating 3D city models without elevation data. *Computers, Environment and Urban Systems*, 64, 1-18.