Using shape, proximity and functionality to define neighborhoods with morphologically similar buildings

Raphaël Bubloz

Urban Patterns session
7 September 2019

UNIL | Université de Lausanne
Institut de géographie et durabilité

ECTQG 2019 Luxembourg

What, how, and why ?

- Trying to cluster morphologically similar buildings in an urban context
- Using features derived from simple building representations
- Clustering using a geo-constrained Self-Organizing Map (geo-SOM)
- Find (in-)homogeneous places in cities, compare internal morphologies of cities, etc.

Definitions and approaches

Literally the *study of urban form*, with several approaches (Kropf, 2017)

- Typo-morphological (evolution of building types)
- Configurational (topological relations, influence of spatial configuration)
- Historico-geographical (townplan study, hierarchy of elements)
- Spatial analytical (complex systems, physical and socio-economical dynamics)

Similar works

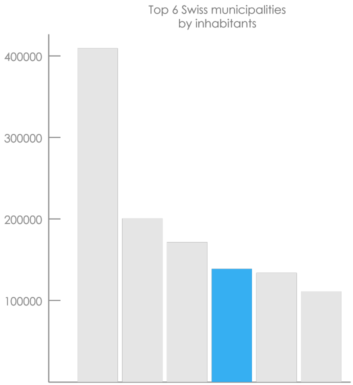
- Schirmer and Axhausen (2015;2019) proposed a set of features at multiple scales on which to apply clustering. They used it to quantitatively describe the urban morphology
- Steiniger *et al.* (2008) morphologically characterized building geometries and classified them for cartographic generalization
- Fan, Zipf, and Fu (2014)
- Dillenburger (2008)

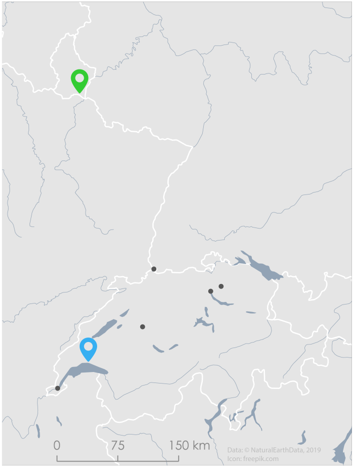
Similar works

- Schirmer and Axhausen (2015;2019)
- Steiniger *et al.* (2008)

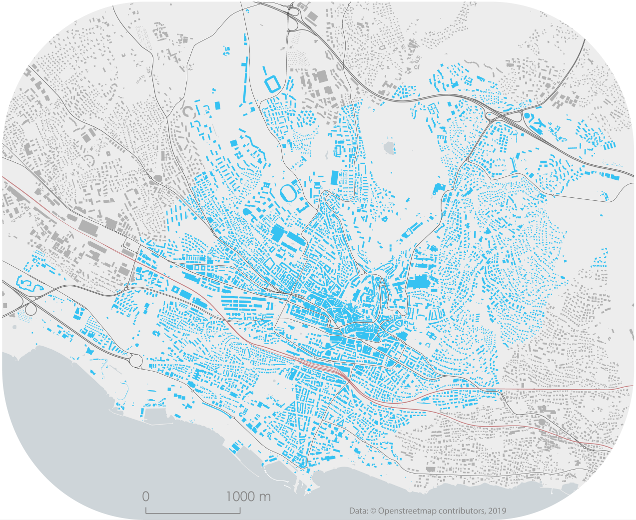
- Fan, Zipf, and Fu (2014) used a turning function to find similar footprints of buildings. They tried to classify the morphologies into (functional) types. They also developed a rule-based approach to estimate the type of a building according to its characteristics
- Dillenburger (2008) developed a bitmap-based index to retrieve parcels with similar buildings according to three characteristics: visibility, proximities of buildings, and orientation

Lausanne, Switzerland





Lausanne Municipality



OpenStreetMap data - Gathering footprints

- Open data generated by multiple contributors
- Available on a global scale but unevenly distributed
- Quality depending on several factors (available imagery, familiarity of the cartographer, etc.)

OpenStreetMap data - Gathering footprints

- Open data generated by multiple contributors
- Available on a global scale but unevenly distributed
- Quality depending on several factors (available imagery, familiarity of the cartographer, etc.)
- Limits: High geometrical variability, features of interest not sufficiently filled

Swisstopo data - Computing height

- Digital Elevation Model, 2016, 2m resolution, 0.5m accuracy
- Digital Surface Model, 2016, 2m resolution, 0.5m accuracy

- Compute the difference and assign mean value of cells intersecting each building
- Gives a sufficient approximation of the building height

Buildings and housing registry - Retrieving categories

- Central registry maintained by the Federal Statistical Office
- Data input by the municipalities
- General categories (6) and specific classes (26) adapted from EUROSTAT
- Quality assurance ?

Indicators derived - Feature creation

- Features based on the footprint geometry and the height
- Inspired by Schirmer and Axhausen (2015) with several additions (oriented envelope, compactness)
- Quite redundant and the question of which are the most suitable remains open

General intended pipeline

Contenu

Feature selection

NON

- Difficultés d'un choix non supervisé
- Réduction en fonction de ce qu'elles distinguent
- Distinction parmi les variables similaires

Clustering based on features

NON

- Pourquoi ?
- Qualité
- Interprétations

Model-based clustering

- Using the method implemented by Ceuleux *et al.* (2014)
- Simultaneously selects the number of clusters and a set of *relevant* features
- Needs to be run several times as results can vary

(Geo-)SOM

The Self-Organizing Map was proposed by Kohonen (2001)

- From a high-dimensional dataset to a simple (2D) representation
- Several other algorithm based on this method

Geo-SOM

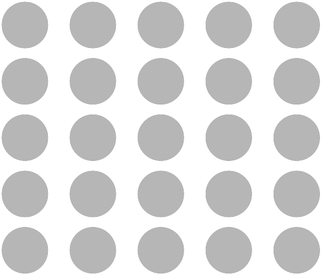
A Self-Organizing Map with a geographical constraint

Data normalized into range $[0, 1]$
Data as n-dimensional vectors d

$$d = [x_{coord}, y_{coord}, \xi_1, ..., \xi_n]^T$$

Initialisation of a map with
$l \times m$ neurons n

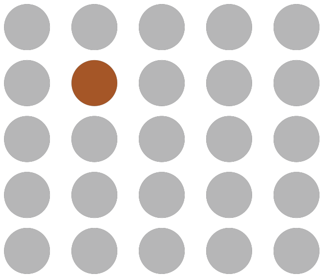
$$n = [x_{coord}, y_{coord}, \mu_1, ..., \mu_n]^T$$



Geo-SOM

For each d, find the most spatially similar neuron

$$Geo_{BM} = \min_i ||d_{xy} - n_{xy}||$$



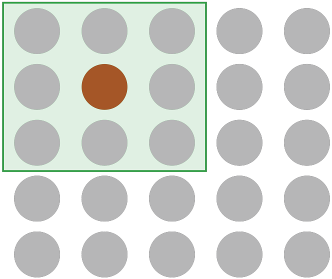
Geo-SOM

For each d, find the most spatially similar neuron

$$Geo_{BM} = \min_i ||d_{xy} - n_{xy}||$$

Then, find the most similar neuron in its neighbours

$$BMU = \min_j ||d - n||$$
$$with\ n \in Geo_{BM}\ neighbours$$



Frame Title

Contenu

Frame Title

Contenu

Future works

- Monitor the development of ML (Mapillary) and gamification solutions (StreetComplete) in the completion of height and number of stories

Future works

-

Thanks for your attention

Slides, routines (and more) are available on *GitHub*
`https://github.com/Raphbub/`

For any further questions, remarks or suggestions
`raphael.bubloz@unil.ch`

Frame Title

Contenu