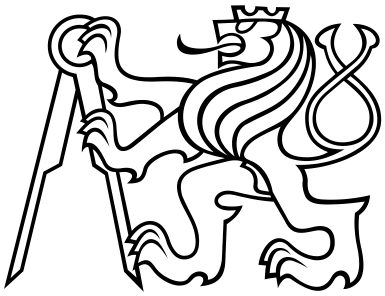
Czech Technical University in Prague
Faculty of Civil Engineering

Master's Thesis

Prague 2017

Bc. Žofie Cimburová

Czech Technical University in Prague
Faculty of Civil Engineering
Study programme: Geodesy and Cartography
Field of study: Geomatics



Master's Thesis

Urban morphology in Prague: automatic classification in GIS

Struktura městské zástavby v Praze: automatická klasifikace v GIS

Thesis supervisor: Doc. Ing. Jiří Cajthaml, Ph.D.
Department of Geomatics

Prague 2017

Bc. Žofie Cimburová

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta stavební
Thákurova 7, 166 29 Praha 6

ZADÁNÍ DIPLOMOVÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: Cimburová Jméno: Žofie Osobní číslo: 381446

Zadávající katedra: Katedra geomatiky

Studijní program: Geodézie a kartografie

Studijní obor: Geomatika

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce: Struktura městské zástavby v Praze: automatická klasifikace v GIS

Název diplomové práce anglicky: Urban morphology in Prague: automatic classification in GIS

Pokyny pro vypracování:

Cílem práce je vytvoření přehledu kvantitativních deskriptorů charakterizujících městskou zástavbu ve více měřítcích. Na podkladu těchto charakteristik bude bude vytvořena metodika pro automatickou klasifikaci struktury městské zástavby.
Pilotní projekt bude proveden na území hl. m. Prahy jako technická oponentura ke stávajícímu členění struktury městské zástavby.
K vytvoření metodiky budou použity nástroje geografických informačních systémů.

Seznam doporučené literatury:

Patrick M. SCHIRMER, Kay W. AXHAUSEN: A multiscale classification of urban morphology
Jorge GIL et al.: On the discovery of urban typologies: data mining the many dimensions of urban form

Jméno vedoucího diplomové práce: Doc. Ing. Jiří Cajthaml, Ph.D.

Datum zadání diplomové práce: 3.10.2016 Termín odevzdání diplomové práce: 8.1.2017

Údaj uveďte v souladu s datem v časovém plánu příslušného ak. roku

Podpis vedoucího práce Podpis vedoucího katedry

III. PŘEVZETÍ ZADÁNÍ

Beru na vědomí, že jsem povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je nutné uvést v diplomové práci a při citování postupovat v souladu s metodickou příručkou ČVUT „Jak psát vysokoškolské závěrečné práce" a metodickým pokynem ČVUT „O dodržování etických principů při přípravě vysokoškolských závěrečných prací".

3.10.2016

Datum převzetí zadání Podpis studenta(ky)

Abstract

The thesis proposes a methodology for automatic classification of urban morphology in GIS, which makes use of methods of unsupervised learning – principal component analysis and K-medoids clustering – and is based on morphological attributes of buildings and blocks. In theoretical part, an overview and description of morphological attributes is given and fundamentals of principal component analysis and clustering are explained. The case study is carried out in the study area of Prague, Czech Republic, resulting in detection and description of up to twelve various urban structures, which serve as an alternative to current urban structures definition used in the Metropolitan plan of Prague. The proposed methodology is fast, objective, easy to interpret and possible to extend and actualize.

Key Words

Urban morphology, GIS, urban structures, classification, principal component analysis, clustering

Abstrakt

Diplomová práce se zabývá návrhem metodiky pro automatickou klasifikaci městské zástavby v GIS. Klasifikace vychází z morfologických atributů budov a bloků a jsou pro ni použity metody strojového učení – analýza hlavních komponent a K-medoids clustrování. V teoretické části jsou představeny a popsány jednotlivé morfologické atributy budov a bloků a jsou vysvětleny základy analýzy hlavních komponent a clustrování. Pilotní projekt je proveden na území hl.m. Prahy. Výsledkem je detekování a popis až dvanácti různých struktur městské zástavby. Výsledek slouží jako alternativa k současné definici struktur zástavby používané v Metropolitním plánu Prahy. Navržená metodologie je rychlá, objektivní, jednoduše interpretovatelná a disponuje možností rozšíření a aktualizace.

Klíčová slova

Struktura zástavby, klasifikace, clustrování, analýza hlavních komponent, GIS

Declaration of authorship

I declare that I have completed the Master's thesis "Urban Morphology in Prague: automatic classification in GIS" by myself and I have cited all used literature according to the "Methodical Directive About Ethical Preparation of Final Academic Works" issued by the Czech Technical University in Prague.

Prague, 8th January 2017 .....

Acknowledgements

Many thanks belong to Mgr. Jiří Čtyroký, Ph.D., director of Section of Spatial Information, IPR Prague, and to Mgr. Eliška Bradová, IPR Prague, for giving me a chance to work on this inspirational topic.

I would also like to express my gratitude to the thesis supervisor, Doc. Ing. Jiří Cajthaml, Ph.D. for giving me a free hand, but still providing important feedback.

Contents

Introduction **1**

1 Attributes of urban morphology **6**
1.1 Buildings . . . 6
1.1.1 Characteristics derived from building geometry . . . 6
1.1.2 Characteristics derived from building courtyards . . . 11
1.1.3 Characteristics derived from building's bounding box . . . 12
1.1.4 Characteristics derived from building's enclosing circle . . . 14
1.1.5 Characteristics derived from building's convex hull . . . 15
1.1.6 Characteristics derived from building's influence zone . . . 17
1.1.7 Characteristics derived from building neighbourhood . . . 18
1.2 Blocks . . . 22
1.2.1 Characteristics derived from block geometry . . . 22
1.2.2 Characteristics derived from block defining streets . . . 23
1.2.3 Characteristics derived from arrangement of buildings in a block . . . 23
1.2.4 Characteristics derived from morphological attributes of buildings in a block . . . 28

2 Classification of urban morphology **32**
2.1 Principal component analysis . . . 33
2.2 Clustering . . . 35
2.3 Mode filtering . . . 39

3 Methodology **40**
3.1 Software . . . 41
3.2 Data acquisition and preprocessing . . . 41
3.2.1 Buildings . . . 42
3.2.2 Building compositions . . . 43
3.2.3 Blocks . . . 43
3.2.4 Street network . . . 46
3.3 Morphological attributes of buildings . . . 47
3.3.1 Characteristics derived from building geometry . . . 47
3.3.2 Characteristics derived from building courtyards . . . 48
3.3.3 Characteristics derived from minimum bounding geometry . . . 50
3.3.4 Characteristics derived from building's influence zone . . . 50
3.3.5 Characteristics derived from building neighbourhood . . . 51

3.4 Morphological attributes of blocks . . . 53
3.4.1 Characteristics derived from block geometry . . . 53
3.4.2 Characteristics derived from block defining streets . . . 53
3.4.3 Characteristics derived from arrangement of buildings in a block . . . 54
3.4.4 Characteristics derived from morphological attributes of buildings in a block . . . 58
3.5 Discovering urban structures . . . 59
3.5.1 Principal component analysis . . . 59
3.5.2 Clustering . . . 63

4 Results and conclusion **65**
4.1 Results . . . 65
4.1.1 Comparison of results . . . 65
4.1.2 Four urban structures . . . 67
4.1.3 Seven urban structures . . . 68
4.1.4 Eight urban structures . . . 70
4.1.5 Twelve urban structures . . . 72
4.2 Conclusion . . . 75

Bibliography **77**

Abbreviations **82**

List of figures **86**

List of tables **87**

Appendices **88**

A Contents of attached CD **88**

B Morphological attributes – summary **89**

C Morphological attributes – visualization **94**

D Principal components **99**

E Urban structures in Prague – visualization **103**

Introduction

What influences the perception of urban space? When walking through the streets or looking at a city plan, we are able to identify areas with similar character or distinguish historical centre from housing estate. What are the characteristics of cities which cause the feeling of spatial identity or dissimilarity, and how to identify such places? One way to answer this question is to search for similar patterns in the physical form of a city. The form of each structure – building, urban block or street – may be quantified through attributes of *urban morphology*. In general, morphology means the study of shape (Oxford Dictionaries [7]). Urban morphology, defined by Oliveira [27], "*means the study of urban forms, and of the agents and processes responsible for their transformation (...) urban form refers to the main physical elements that structure and shape the city—urban tissues, streets (and squares), urban plots, buildings.*" Evaluating the similarity of patterns in urban structures is based on an assumption that *if two places are perceived as similar, values of their morphological attributes are likely to correspond.* Urban structures represent a base stone for definition of various policies and regulations in city planning. They capture and define local character, which should be respected or cultivated, rather then disrupted, and unlike for example land use, they are persistent in time.

Urban morphology analysis in Prague – state of the art

The main principles stated in the proposal of the Metropolitan plan of Prague, set up by the Prague Institute for Planning and Development, are e.g. definitions of *public spaces*, *urban structures, forms and compositions* or *height regulations* (IPR Praha [16]). Urban structure, assessed through *structural diagnosis*, is considered a key concept for understanding and description of the character of localities, as well as explaining e.g. social stratification. It leads to a definition of basic structural elements of Metropolitan plan – *localities*, i.e. elementary city units of predominant urban structures with specific names. (IPR Praha [15])

In the up to date Metropolitan plan proposal, the structural diagnosis represents both *foundation* and *result* of a definition of localities. At first, localities were defined manually by architects, geographers and other specialists based on perceived urban structure, as well as local knowledge and terrain topography. The resulting maps of localities were overlaid and dissimilarities in locality borders were discussed. Furthermore, accuracy of borders definition was improved using borders of cadastral parcels. The identified localities then

became the spatial basis for an in-depth structural analysis. Four main structures are represented by *historical centre*, *town*, *suburbs* and *periphery* (Figure 1), distinguished by relation of private and public space and street and building line, character of public space or scale and typology of buildings. Each of the four zones comprises several finer defined urban structures (Figure 2). (IPR Praha [15])

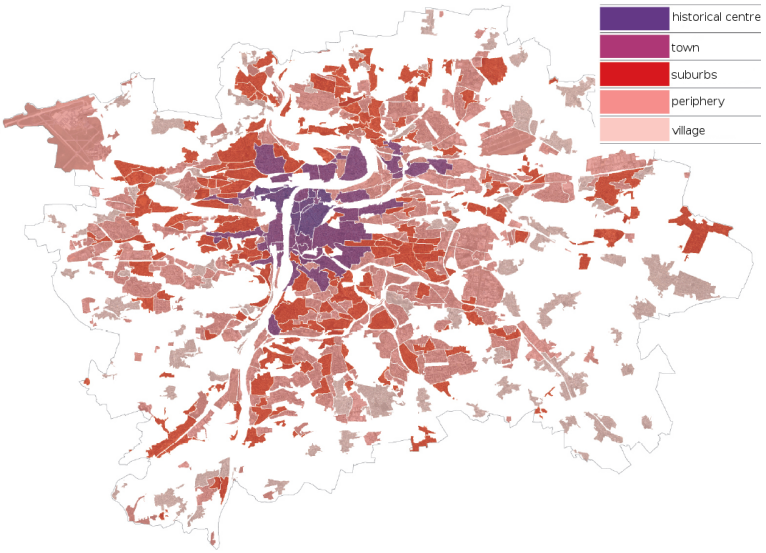


Figure 1: Map of structural diagnosis in Prague. Source: IPR Praha [15]

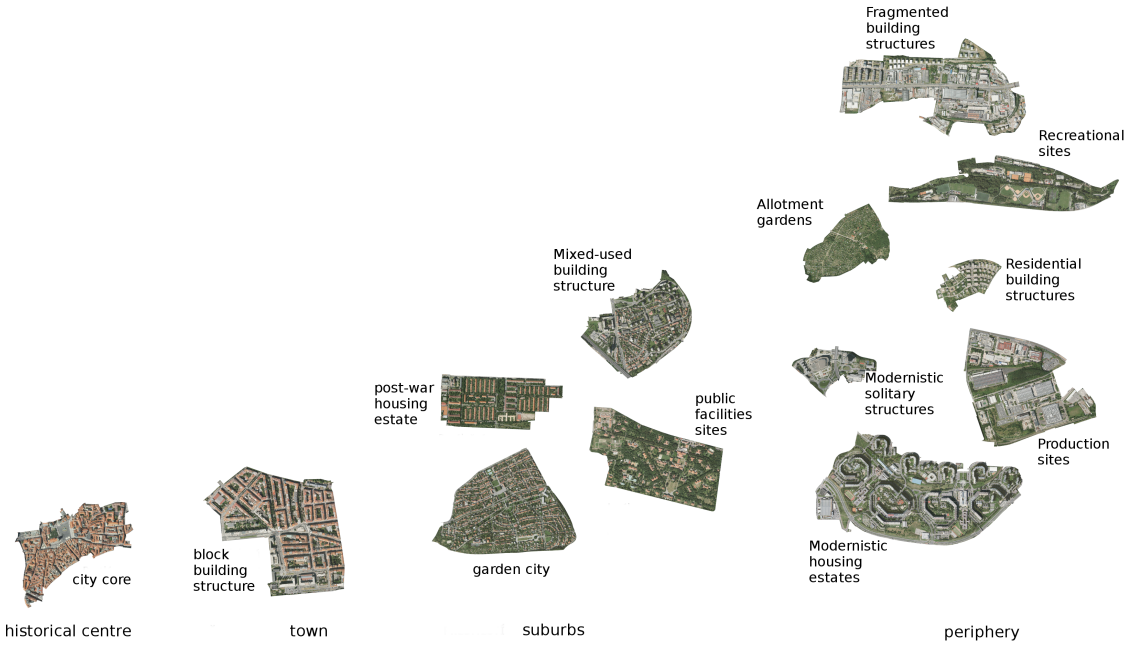


Figure 2: Urban sub-structures in Prague. Source: IPR Praha [15]

Objectives of the thesis

The aforementioned process utilized in assessing urban structures in Prague suffers from obvious drawbacks:

1. *Extension and actualization*: The criteria for distinguishing various urban structures are based on a relatively *small number of quantitative characteristics*. In case a new object is built, it might affect structural characteristics of not only this particular block, but also neighbouring blocks and structures. With current methodology, the process of actualization and extension is laborious and time consuming.

2. *Interpretability*: Structural diagnosis is difficult to interpret *beyond the borders of Prague*, which prevents from benchmarking and comparison of other cities.

3. *Objectivity*: Structural diagnosis is based on localities whose definition is mainly supported by *subjective decisions* and *prior knowledge*, and given another group of specialists performs the analysis, resulting borders might not be equally defined. Moreover, although most urban structures are easily differentiated from each other, *problematic areas* may occur and a solution for objective categorization may not exist.

The goal of this thesis is to address these problems with following solutions:

1. Provide a comprehensive *overview and description of morphological attributes* expressing various characteristics of urban structures (and hence reflecting the perception of urban space), which do not rely on any previous definition of locations and are therefore applicable regardless of the analysed city.

2. Present an *automatic classification technique* based on morphological attributes which is fast, and allows to easily utilize other morphological attributes or capture future changes in urban structure.

3. Using morphological attributes and their classification, suggest an objective *set of urban structures*, which are not biased by any previous knowledge about Prague or subjective perception.

Related work in automatic classification of urban morphology

Whilst a wide range of urban typologies has been explored and published (Oliveira [27]), the research carried out in its automatic classification is not very extensive and many authors call for a deeper investigation of this domain. Overall, the research may be classified from several points of view. Firstly, emphasis is not always put on classification itself, but rather the data acquisition. Only few researchers put stress on defining morphological attributes and automatic classification techniques. In most cases, the work deals with data acquisition and less attention is paid to following classification. Secondly, two ways of distinguishing urban types are generally presented – either a predefined set of classes or exemplified structures exist, and buildings or blocks are assigned into a group based on some decision criteria on their morphological attributes. The other approach is data-driven, when no previous knowledge about urban structures is available or utilized, and subgroups of structures are found through their similarity. Chapter 2 explains both approaches in detail.

The probably most similar research to the scope of this thesis in the domain of urban morphology classification was carried out recently by Gil et al. [10] and Schirmer and Axhausen [37]. Gil et al. [10] aim at presenting a "*method to support the description and prescription of urban form*" and suggest using K-means clustering in order to objectively classify urban environments, i.e. blocks and street types, and support larger urban studies. They stress the importance of K-means clustering as a method independent on previously defined urban types. Schirmer et al. [37] further develop the methodology suggested by Gil et al. [10] and claim to present a "*first attempt to define a set of attributes on urban morphology that allows for interpretation in different study areas*" as well as outline an analytical process for a quantitative description of urban morphology in four scales (i.e. object, composition, neighbourhood and municipality), and run a case study on the canton of Zurich. Different sets of morphological attributes representing the urban environment at different levels of perception are presented. These include form– and volume– based information about buildings, networks and street-blocks, such as height of building, length of façade that is perceived from the street or betweenness of road network. Subsequently, scale-dependent classification of locations is performed using K-means and K-medoids cluster analysis, resulting in so-called urban typologies. Each cluster consequently represents one typology. Thomas et al. [45] also propose to utilize K-medoids clustering, however they claim that clustering based on the curve of scaling behaviour is sufficient enough to explain the morphology of built-up area. Other authors do rather utilize the supervised leaning methods. Work by Steiniger et al. [41] was also carried out in Zurich. They present an approach for recognition of five predefined types of urban structures by means of different supervised learning classification methods (e.g. support vector machine) of morphological attributes of buildings derived from perceptual principles of Gestalt psychology (e.g. size, shape and built-up area). Colaninno et al. [5] first define seven classes of homogeneous structures in Barcelona and then perform clustering based on morphological indices, with cluster centres represented by statistics of representative samples of each class. Slightly different point of view on urban morphology is brought by Porta et al. [29] who point out an analogy of biological evolution in the built environment and observe correlation between age of urbanization and urban morphometrics of form-related properties

of streets, blocks, plots and buildings in Milan and Glasgow. A comprehensive research on the influence of quantifiable spatial attributes on spatial identity was done by Laskari [20]. By means of morphology measurements such as axial graph, fractal dimension, visibility etc. he illustrates degrees of identification and differentiation between spatial unities. Similar neighbourhoods in terms of morphological patterns of streets edges and buildings (footprint area, built front ratio, street width) are also searched by Venerandi et al. [48] in their study of the form of gentrification in London.

Other authors' objective is mainly to detect built-up areas by means of remote sensing and subsequent image processing and analyse these areas with spatial metrics (Taubenböck et al. [43], Taubenböck et al. [44], Banzhaf and Höfer [3], de Voorde et al. [6], Yu et al. [54]), or use these to improve the classification of land use and objects from 3D point clouds or remote sensing (Serna et al. [36], Herold et al. [14]).

Proposed methodology and thesis structure

Predefined objectives will be achieved through an automatic classification of urban structures using morphological attributes in the scale of buildings and blocks. The basic unit of classification will be an urban block, which encompasses characteristics of buildings as well as public spaces and neighbouring streets. Attributes will be derived from 2D, or alternatively 2.5D representation of spatial features, the latter providing information about object height. In order to utilize characteristics of buildings in the scale of blocks, such attributes will be summarized for appropriate blocks. Urban structures will be determined via unsupervised techniques of machine learning – principal component analysis (PCA) and clustering. An overview of workflow is illustrated in Figure 3.1.

The thesis is structured as follows: Chapter 1 presents an overview of morphological attributes for scale of buildings and urban blocks. Chapter 2 deals with techniques of urban morphology classification, i.e. PCA and clustering. In Chapter 3, implementation of morphological attributes and classification in Prague is described. Finally, the results are presented in Chapter 4.

Chapter 1

Attributes of urban morphology

Chapter 1 focuses on an overview of morphological attributes in scale of buildings and blocks. As already mentioned, attributes will be derived from 2.5D representation of buildings and 2D representation of urban blocks. Larger scales such as neighbourhoods are not processed separately, but rather incorporated in smaller scales through e.g. accessibility measures.

1.1 Buildings

Buildings represent the largest scale in urban morphology assessing. They are represented by a polygon of their footprint, which stores all necessary geometric information. In order to derive information about floorspace or building volume, it is useful to store an additional attribute of building height or number of floors (2.5D).

Three main groups of building morphology descriptors are presented in this section – *geometry-based* attributes, attributes based on *minimum bounding geometries* and finally attributes based on the *immediate neighbourhood of a building.*

Absolute *dimensions* of a building are evaluated using its polygon representation and size of minimum bounding geometries. Several *compactness* measures are presented in form of ratios between various combinations of building dimensions or dimensions of its bounding geometries. A good overview of compactness measures is given by MacEachren [22]. The relationship between *building and its surrounding space* is explored through its nearest neighbours and influence zone.

1.1.1 Characteristics derived from building geometry

Basic information about the shape of a building is derived from its geometrical information – area, perimeter and (if available) height or number of floors. Other characteristics in this chapter are derived from these basic attributes.

Area of building footprint

The actual space on the ground occupied by a building mass is the most common characteristic and is under different names used by Schirmer and Axhausen [37], Serna et al. [36], Steiniger et al. [41], Hartmann et al. [12], Colaninno et al. [5], Herold et al. [13] and Yu et al. [54]. Its value largely depends on the definition of building, which is set by data provider. For example, two touching buildings may be represented by one polygon, as well as by two, in case two address places exist, or by more, in case we distinguish between actual building and, let's say, appended porch.

Perimeter of building footprint

The total length of building façade is also widely used, e.g. by Schirmer and Axhausen [37], Serna et al. [36], Hartmann et al. [12] and Yu et al. [54]. Again, the value is influenced by the definition of building, and also by the scale and level of generalization in which data is provided.

Floorspace

Floorspace represents the potential usable area of a building. Two buildings might have equal floorspace, although the areas of their footprint differ. Schirmer and Axhausen [37] estimate floorspace as the area of building footprint multiplied by number of storeys.

Height

Height of a building as one of its characteristics is presented by Taubenböck et al. [43] or Yu et al. [54]. The precise value of height is not always available in the dataset, and its estimate may be derived from the number of storeys, assuming a standard height of one story.

Number of floors

Number of floors, used by Taubenböck et al. [43], as such already very well provides an image of urban morphology, with several storeys apartment buildings in the city centres, two-story family houses, multiple-storeys blocks of flats and high-rise buildings, and one-story bungalows at the suburbs. In case this attribute is not available in the dataset, it may be calculated from the building height, again assuming a standard height of one story.

Having both height and number of floors available allows to derive **average height of one story**. As high storeys are expected to occur in apartment houses in the city centres, and low storeys for example in blocks of flats, this characteristic is very useful as well.

Volume

The total amount of space occupied by a building is used by Serna et al. [36] and Yu et al. [54]. Building volume may be perceived and calculated in two different ways – either by simply multiplying the area of footprint by building height (representing for example the amount of heat necessary to heat up the building), or by adding area of building courtyards to the area of footprint and multiplying by height. This approach would capture the perception of a building from outside – whether the building is perceived as massive or subtle.

Fractal dimension

Colaninno et al. [5] aim at evaluating building efficiency using a ratio between its area and perimeter. This value also expresses the *fractal dimension* of an object (McGarigal [23]), i.e. complexity of a building derived from how much its perimeter is filling its area. It is worth noting that this attribute is scale dependent and reaches values from zero (small low complex buildings). Upper limit for values representing large buildings with large complexity is not set.

$$fractal\ dimension = \frac{footprint\ area}{footprint\ perimeter} = \frac{[m^2]}{[m]}$$

Ratio of building volume to façade area

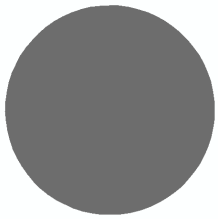
Similar attribute is proposed by Schirmer and Axhausen [37]. Compared values are one dimension higher than values used for *fractal dimension*, however the final value will be identical to *fractal dimension* as defined above.

$$fractal\ dimension = \frac{building\ volume}{facade\ area} = \frac{footprint\ area \cdot height}{footprint\ perimeter \cdot height} = \frac{[m^3]}{[m^2]}$$

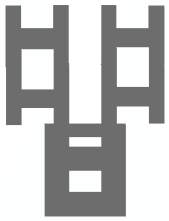
Gravelius index

Gravelius index aims at capturing compactness of a building through the relation between area and perimeter of building footprint (Colaninno et al. [5]).

$$gravelius\ i = \frac{building\ perimeter}{2\sqrt{\pi \cdot footprint\ area}} = \frac{[m]}{[m]}$$



(a) Gravelius index = 1.00



(b) Gravelius index = 3.65

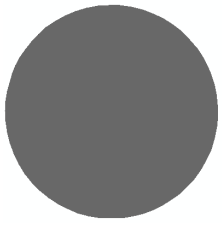
Figure 1.1: Gravelius index. Source: author

The lower border of Gravelius index is equal to one, i. e. the most compact buildings with circular footprint. High values are associated to complex buildings with perimeter much larger than area. No upper border for Gravelius index exists.

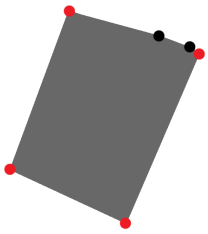
Number of building corners

Summing up the number of polygon points of exterior ring is proposed by Steiniger et al. [41] and Colaninno et al. [5]. This characteristic aims at distinguishing between low-variability houses of basic shapes, such as circle (0 corners), triangle (3 corners) or rectangle (4 corners), and complex structures, such as a church.

It is important to ensure that all walls are represented as polylines and not splines. The total number of corners largely depends on the level of detail of input data, as well as on the definition of a corner. In order to distinguish between significant (true corners) and insignificant corners (points in fact not representing corners, but rather not entirely straight walls) it is suggested to perform slight generalization.



(a) Zero corners – no corner is significant.



(b) Four corners – two corners are insignificant (in black).

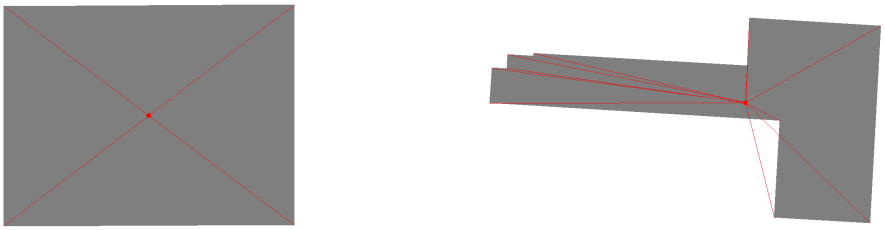
Figure 1.2: Number of building corners. Source: author

Distance of building corners to centroid

Schirmer and Axhausen [37] propose to use the average distance of building corners to centroid, however the standard deviation of this value is much more determining and therefore more interesting for urban morphology studies. If the individual distances in a building are scaled to 0-1 interval, the final standard deviation will not be affected by

the dimensions of the building, but only by its shape and will serve as a relative descriptor. The thesis also suggests to count in only distances from significant points, as presented in the section above.

Low values are associated to buildings with corners close to the border of circumscribed circle, whilst with high values the irregularity of a building increases.



(a) corner_dev = 0.01 **(b)** corner_dev = 7.63

Figure 1.3: Standard deviation of distance of building corners to centroid. Source: author

Squareness of building walls

Building squareness, expressed as a mean deviation of all building corners from a perpendicular angle is used by Steiniger et al. [41]. This thesis proposes to first classify the wall angle in one of five groups and explore its deviation from the group norm afterwards.

Low values are associated to buildings with walls at right angles. Such may be expected in villa quarters or housing estates. High values represent buildings with walls at either obtuse or acute angles, which may be found in historical centres, industrial zones, or modern solitaire buildings.

(a) corner_square = 0.00 **(b)** corner_square = 22.93

Figure 1.4: Building squareness. Source: author

Topological skeleton

To obtain a more detailed characteristics of building shape, Schirmer and Axhausen [37] suggest to construct a topological skeleton of building footprint polygon.

Topological skeleton (also *medial axis* or *centerline*) is a set of points inside a polygon that have more than one closest point among the points of border of the polygon

(O'Rourke [26]). These points form a boundary of Voronoi diagram of polygon edges. The advantage of topological skeleton representation is that it preserves all topological information of its polygon. However, it is very sensitive to the level of detail of polygon border. Topological skeleton is used e.g. in image processing or cartographic generalization (Eppstein [9]).

In work of Schirmer and Axhausen [37], four morphological attributes are based on topological skeleton. **Number of centerlines** provides information about building complexity. **Number of lines in skeleton that are only connected to one side** gives overview of number of building corners. **Total length of centerlines** assesses the overall size of building weighted by number of building corners. **Average width along centerline** provides information about average width of building. **Number of wings** is derived from the topological skeleton as well.

Another relevant usage of topological skeleton is determining the various attributes dealing with building orientation. Schirmer and Axhausen [37] observe for example **length of centerlines oriented at various cardinal directions** and **main orientation of centerline**. Describing building morphology in terms of orientation is also proposed by Steiniger et al. [41]. Here orientation is approximated by orientation of its major axis derived from bounding box. However, it is concluded that orientation is not a helpful descriptor of building morphology, as it "*cannot be expected that buildings of a particular structure type are aligned in the same direction*" (Steiniger et al. [41]).

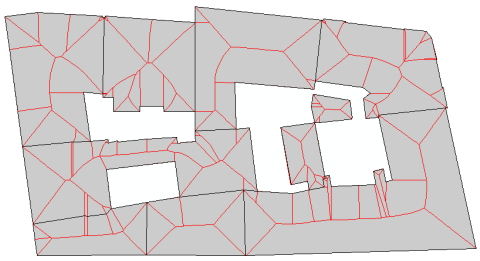


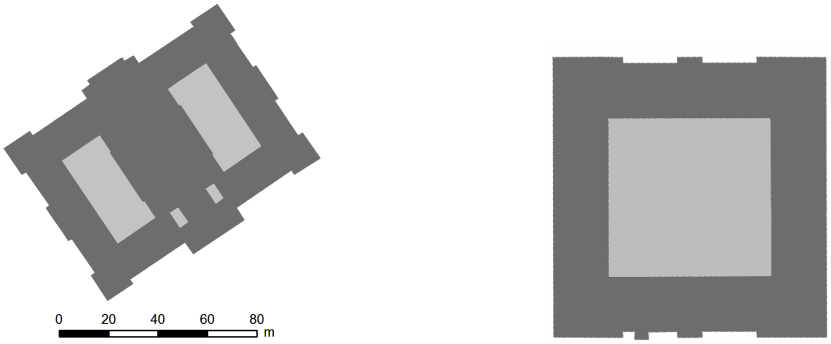
Figure 1.5: Topological skeleton of non-generalized buildings. Source: author

1.1.2 Characteristics derived from building courtyards

Courtyards are defined as places completely surrounded by building walls. According to Schirmer and Axhausen [37], *courtyards form a semipublic space that can substantially benefit a location.*

It is important to make difference between different types of "holes" in building footprint, e.g. ventilation shafts and courtyards which serve as places of people's encounters, free time activities etc. A proposed classification distinguishes between three types of courtyards. First group contains courtyards which are too small to serve any social activity, second comprises courtyards surrounded by too high walls, not allowing the sun to reach the ground. Finally, third group is formed by actual liveable courtyards.

Three measures of buildings characteristics are based on liveable courtyards. **Number of courtyards** is used by Schirmer and Axhausen [37] and Steiniger et al. [41]. **Total courtyard area**, used by Schirmer and Axhausen [37], distinguishes between e.g. small yards and footbal pitches in sport stadiums. Finally, **ratio of total courtyard area to building area** (Schirmer and Axhausen [37]) represents how big portion of building, as perceived from outside, is formed by its courtyards.



(a) Building with two liveable courtyards and two shafts

(b) Building with large courtyard ratio

Figure 1.6: Building courtyards. Source: author

1.1.3 Characteristics derived from building's bounding box

Bounding box is a rectangle with the smallest measure (e.g. area or perimeter) constructed around a set of planar features (points, lines, polygons), such that all of these features fall inside the bounding box (Page [28]). For the purpose of evaluating building footprint, bounding box minimizing the area is used.

Absolute values associated to bounding box are not of a big importance in describing the building structure. Much more valuable descriptors are derived from the relationship between shape of bounding box and actual building in form of ratios, allowing to compare buildings relatively, regardless their absolute size.

As for the absolute measures, Schirmer and Axhausen [37] and Serna et al. [36] use the **area** of bounding box, **perimeter** is used by Schirmer and Axhausen [37], and Schirmer and Axhausen [37] and Steiniger et al. [41] incorporate its **width** and **length** as a building characteristic.

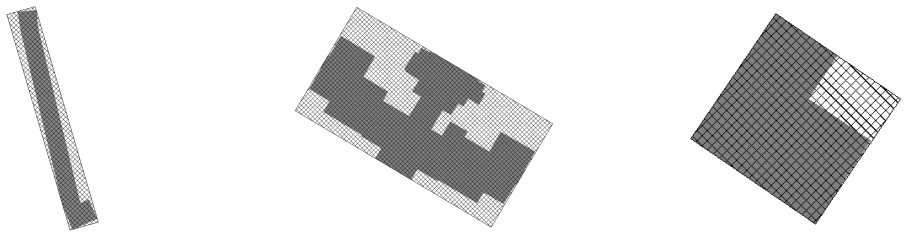
It would be interesting to explore the possibilities of three dimensional city models, and utilize a 3D bounding box for more precise description of building structure.

Elongation from bounding box

Elongation describes shape of a building in terms of ratio of its bounding box's width to length. Elongation can reach values from zero to one. The higher the ratio, the more

the bounding box resembles a square. With lower values the bounding box becomes thinner. Elongation should never reach zero, as this would represent a building with zero width. (Schirmer and Axhausen [37], Steiniger et al. [41])

$$elongation = \frac{BB\ width}{BB\ length} = \frac{[m]}{[m]}$$



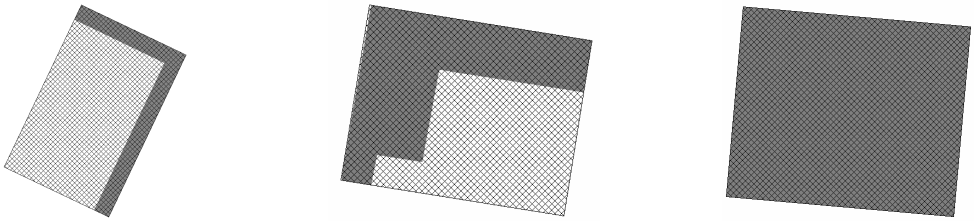
(a) $elongation = 0.13$ **(b)** $elongation = 0.53$ **(c)** $elongation = 1$

Figure 1.7: Elongation. Source: author

Ratio of building footprint area to bounding box area

To discover how much a footprint of a building resembles a rectangle, Schirmer and Axhausen [37] suggest to put footprint area in ratio with bounding box area. The result reaches values from zero to one. Low values are characteristic for thin buildings with many wings and courtyards. Values close to one are assigned to rectangular buildings.

$$BB\_ratio\_A = \frac{footprint\ area}{BB\ area} = \frac{[m^2]}{[m^2]}$$



(a) $BB\_ratio\_A = 0.22$ **(b)** $BB\_ratio\_A = 0.50$ **(c)** $BB\_ratio\_A = 1$

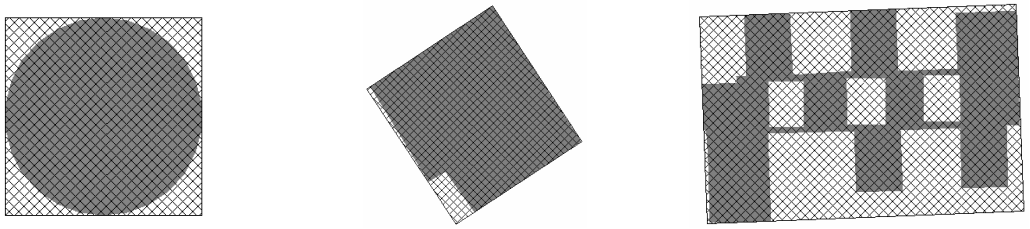
Figure 1.8: Ratio of building footprint area to bounding box area. Source: author

Ratio of building footprint perimeter to bounding box perimeter

Another measure of relative building compactness is proposed by Schirmer and Axhausen [37]. Ratio of building footprint perimeter to bounding box perimeter will result in values lower than one for buildings with smaller footprint perimeter than their bounding

box, i.e. compact buildings with triangular or circular footprint or round segments on otherwise rectangular façade. Values around one are typical for objects with rectangular footprint where perimeters of both building and its bounding box are similar. Values higher than one are characteristic for complex buildings – the higher the value, the longer perimeter of footprint compared to perimeter of bounding box.

$$BB\_ratio\_P = \frac{footprint\ perimeter}{BB\ perimeter} = \frac{[m]}{[m]}$$



(a) $BB\_ratio\_P = 0.79$ **(b)** $BB\_ratio\_P = 0.99$ **(c)** $BB\_ratio\_P = 1.93$

Figure 1.9: Ratio of building footprint perimeter to bounding box perimeter. Source: author

1.1.4 Characteristics derived from building's enclosing circle

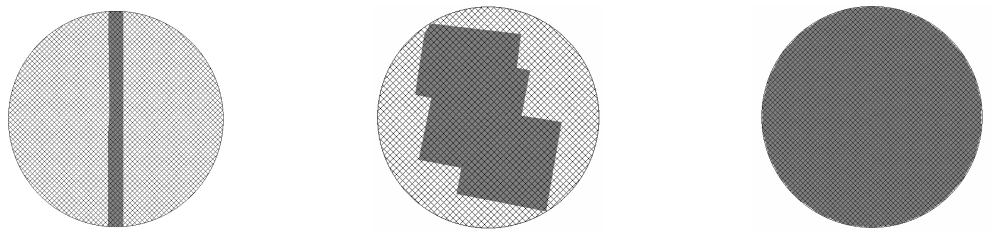
Morphological characteristics of buildings are also often evaluated by its *minimum enclosing circle* – the smallest circle enclosing an input planar feature. Both diameter and area of such a circle will be the smallest out of all enclosing circles.

Schirmer and Axhausen [37] propose to describe enclosing circle of each building in form of absolute values – **area**, **perimeter** or **diameter** (assess the size of building), as well as in form of ratios towards the building dimensions.

Ratio of building footprint area to enclosing circle area

Also called compactness by Porta et al. [29], this ratio characterizes narrow buildings with many wings and courtyards as less compact (values close to zero). The highest value of compactness (one) is assigned to buildings with circular footprint without holes.

$$EC\_ratio\_A = \frac{footprint\ area}{EC\ area} = \frac{[m^2]}{[m^2]}$$



(a) $EC\_ratio\_A = 0.09$ **(b)** $EC\_ratio\_A = 0.49$ **(c)** $EC\_ratio\_A = 0.89$

Figure 1.10: Ratio of building footprint area to enclosing circle area. Source: author

1.1.5 Characteristics derived from building's convex hull

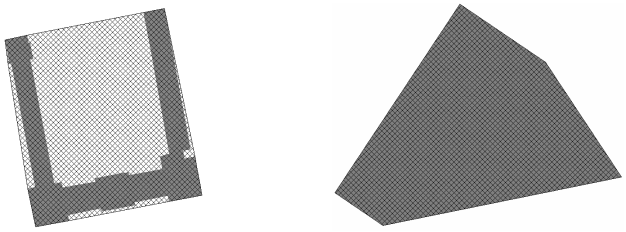
For any subset of the plane (set of points, rectangle, simple polygon), the convex hull is defined as the smallest convex set containing the subset. (Venkatasubramanian [49], van Kreveld [47])

Dimensions of convex hull are used by Schirmer and Axhausen [37] – **area** and **perimeter** describe the dimensions of a building. **Length** of convex hull expresses the distance between two most distant corners of a building.

Ratio of building footprint area to convex hull area

Schirmer and Axhausen [37] also suggest to assess the convexity of a building by comparing its footprint area to convex hull area. If all corners are convex and building contains no holes, both values will be identical and resulting value will be equal to one. Low values describe buildings with many concave angles between long walls.

$$CH\_ratio\_A = \frac{footprint\ area}{CH\ area} = \frac{[m^2]}{[m^2]}$$



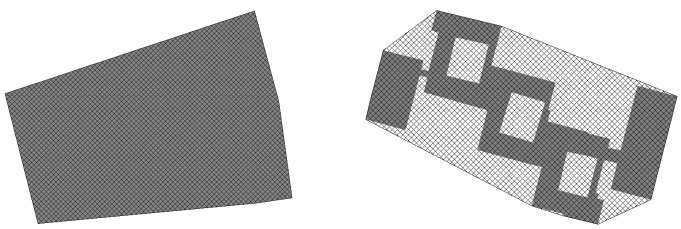
(a) $CH\_ratio\_A = 0.38$ **(b)** $CH\_ratio\_A = 1.0$

Figure 1.11: Ratio of building footprint area to convex hull area. Source: author

Ratio of building footprint perimeter to convex hull perimeter

Another ratio put forward by Schirmer and Axhausen [37] compares footprint perimeter to convex hull perimeter. Convex buildings will be characterized by values equal to one, whilst with rising values, the buildings are becoming more complex.

$$CH\_ratio\_P = \frac{footprint\ perimeter}{CH\ perimeter} = \frac{[m]}{[m]}$$



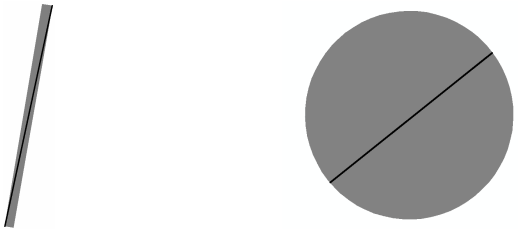
(a) $CH\_ratio\_P = 1.0$ **(b)** $CH\_ratio\_P = 2.16$

Figure 1.12: Ratio of building footprint perimeter to convex hull perimeter. Source: author

Schumm's longest axis to area ratio

The last ratio, suggested by Steiniger et al. [41] and Colaninno et al. [5], is another measure of compactness. Objects with circular footprint are evaluated as the most compact ones, with values of Schumm's ratio equal to one. Lowest values and lowest compactness is assigned to thin long buildings.

$$Shumm's ratio = \frac{2 \cdot \sqrt{\frac{footprint\ area}{\pi}}}{longest\ axis} = \frac{[m]}{[m]}$$



(a) Schumm's ratio 0.22 **(b)** Schumm's ratio 1.00

Figure 1.13: Schumm's longest axis to area ratio. Source: author

1.1.6 Characteristics derived from building's influence zone

Influence zone of a building is an important attribute which represents the area of attraction of each building in case no barriers exist in the open space. For each point in influence zone of a building, the distance to this building will be smaller than to any other building. Influence zones will therefore form a *Thiessen tessellation* of a city, computed from building polygons. Although the characteristic belongs to a building, it stresses the spatial arrangement of building neighbourhood. Schirmer and Axhausen [37] propose to utilize two attributes based on influence zone.

Area of influence zone

Small buildings in dense developments will have influence zones with small areas. The more compact housing, the smaller the areas of influence zones will be. Such values will apply for city centres as well as local centres. Influence zones with large areas are typical for remote objects and sparse development in the suburbs, as well as for buildings located at borders of concentrated development or buildings with large footprint areas.

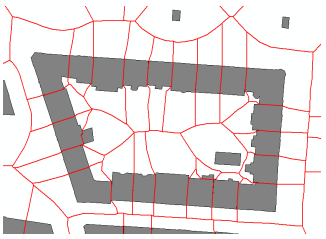
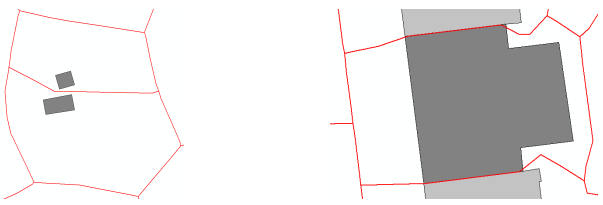


Figure 1.14: Influence zones. Source: author

Ratio of building footprint area to influence zone area

This attribute measures the proportion of influence zone occupied by a building. Similarly to *Area of influence zone*, this attributes aims at distinguishing between dense and sparse development, this time in a relative way. Low values close to zero will be assigned to remote buildings. High values characterize buildings in dense housing, with a limit value equal to one, indicating a building completely bounded by neighbouring buildings.



(a) IZ_ratio = 0.02 **(b)** IZ_ratio = 0.60

Figure 1.15: Ratio of building footprint area to influence zone area. Source: author

1.1.7 Characteristics derived from building neighbourhood

Spatial arrangement of building neighbourhood may be also described in form of relation between building and its nearest neighbour. An interesting derivative would be to observe neighbouring relationships formed by a street network, as it does not generally apply that the nearest neighbour is also the nearest to walk to.

Distance to closest building

Parameter proposed by Schirmer and Axhausen [37] expresses a relationship between each building and its nearest neighbour in terms of their distance. Zero distance stands for touching buildings. Small values may be expected in city centre, as well as in places with compact building groups. Large values characterize remote buildings.

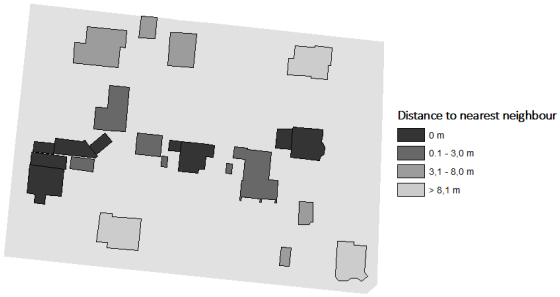
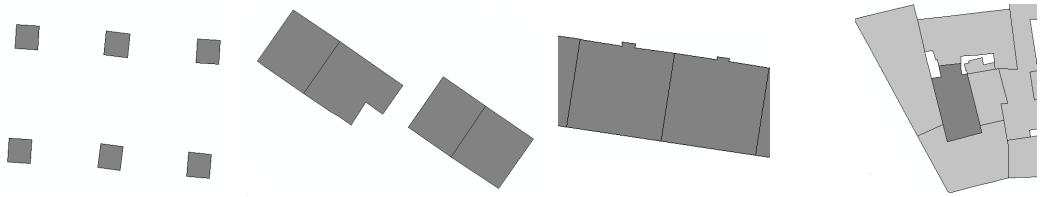


Figure 1.16: Distance to closest neighbouring building. Source: author

Number of neighbouring buildings, existence of neighbouring buildings

The thesis proposes to observe also the number of neighbouring buildings, as well as boolean variable expressing neighbour existence. Only those buildings sharing a line segment are deemed to be neighbours (i.e. not buildings touching at corners). The attribute aims at differentiation between different types of development (detached (no neighbours), semi-detached (one neighbour), terraced (two neighbours) and grouped (three and more neighbours)), typical for different urban structures.



(a) Detached **(b)** Semi-detached **(c)** Terraced **(d)** Groupped

Figure 1.17: Number of neighbouring buildings. Source: author

Distance to urban block border

The thesis suggests to observe also a relationship between a building and corresponding urban block. Distance to building from a street (a block border) describes not only the building itself, but also distribution of housing within the block. Low values apply for objects situated at block borders – such may be found in city centres or block building structure. Higher values are typical for villa quarters or housing estates. Very high values characterize buildings with bad accessibility, situated far from a street or road.

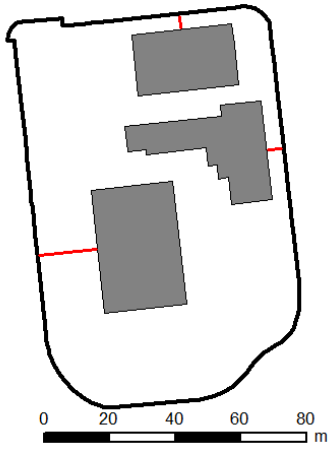


Figure 1.18: Distance to block border. Source: author

Characteristics based on street network

In the previous sections, the neighbourhood of a building was described through the building's relationship with its closest neighbours, and also its interaction with corresponding urban block. This section aims at describing a wider spatial relationships of each building through its neighbourhood. A neighbourhood in this context is not understood as a city part with fixed boundaries, but rather as walk-able surroundings of a building. Different results will therefore be reached if using an Euclidean distance or network distance. Schirmer and Axhausen [37] for example use three fixed Euclidean distances – 100 m, 300 m and 500 m in order to represent walk-able surroundings.

This definition of a neighbourhood permits not only for counting the simple number of buildings within a neighbourhood, but also much more complex *accessibility measures*. In the following sections, five network accessibility measures, presented in The Urban Network Analysis toolbox for ArcGIS (City Form Lab [46]), will be introduced. Unlike other accessibility measures, which are performed on the simple street network of a city, these incorporate buildings into the network representation, and a accessibility measure is performed separately for each building. This for example allows to take uneven building densities into account.

Reach A simple number of buildings j within a search radius of r from building i is represented by *Reach*. A shortest path distance $d[i,j]$ between building i and all other buildings j in graph G is computed, compared to r and if $d[i,j]$ is smaller than r, building j is added into the final *Reach* value. Additionally, each building may be weighted by $W[j]$, representing e.g. floorspace or volume, and enabling the *Reach* to represent e.g. the total floorspace or volume within search radius r.

$$Reach^r[i] = \sum_{j \in G - \{i\};\ d[i,j] \leq r} W[j]$$

Schirmer and Axhausen [37] observe built densities through the number of buildings, sum of footprints and floorspace in various search radii, however they use Euclidean distance, not network distance. Number of buildings in a neighbourhood is also observed by Steiniger et al. [41]. Additionally, Schirmer and Axhausen [37] also evaluate the network structure through the number of dead-ends, number of intersections and total length of network in the same radii.

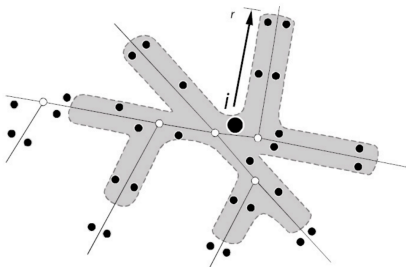


Figure 1.19: Reach. There are 20 buildings reachable from i. Source: [46]

Gravity In order to allow for the decreasing attractiveness of a building j with increasing distance from source i, *Gravity* is inversely proportional to the shortest path distance $d[i,j]$ between buildings i and j. Coefficient β controls how much $d[i,j]$ affects the result, and therefore it depends on the mode of travel modelled by the analysis. An empirical study of pedestrian trips to convenience stores in Oakland, CA by Handy and Niemeier [11] indicated that for walking distances measured in meters, β is approximately 0.00217.

$$Gravity^r[i] = \sum_{j \in G - \{i\};\ d[i,j] \leq r} \frac{W[j]}{e^{\beta \cdot d[i,j]}}$$

Betweenness The potential of passers-by of a building is represented by the *Betweenness* measure, and is significant for example in explaining the potential of retail distribution (Porta et al. [30]). *Betweenness* of building i calculates how many times i lies on a shortest path between all other buildings in search radius r in graph$G$, that is, the fraction of number of shortest paths $n_{j,k}[i]$ from j to k passign through i to the total number $n_{j,k}$ of shortest paths between j to k. Additional weighting by $W[j]$ is enabled.

$$Betweenness^r[i] = \sum_{j,k \in G-\{i\};\ d[j,k] \leq r} \frac{n_{j,k}[i]}{n_{j,k}} \cdot W[j]$$

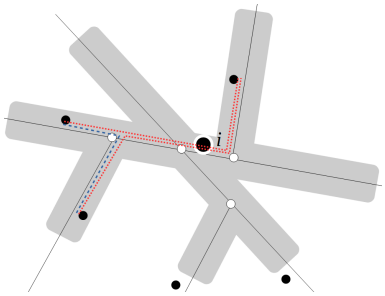


Figure 1.20: Betweenness. Two out of three shortest paths lead through i. Source: author

Closeness *Closeness* measure aims at reflecting how close in a given search radius r building i is located to other buildings j. It is therefore inverse of sum of distances $d[i,j]$ from building i to all other buildings j in a search radius. Additional weighting by $W[j]$ is enabled.

$$Closeness^r[i] = \frac{1}{\sum_{j \in G-\{i\};\ d[i,j] \leq r} d[i,j] \cdot W[j]}$$

Straightness Finally, Euclidean distances are compared to the network distance through *Straightness* measure. It calculates the ratio of straight (Euclidean) distance $\delta[i,j]$ between buildings i and j to the network distance $d[i,j]$ between the same buildings, for all buildings j in given search radius of r. This measure evaluates the overall network structure in given neighbourhood and is a useful measure of city permeability.

$$Straightness^r[i] = \sum_{j \in G-\{i\};\ d[i,j] \leq r} \frac{\delta[i,j]}{d[i,j]} \cdot W[j]$$

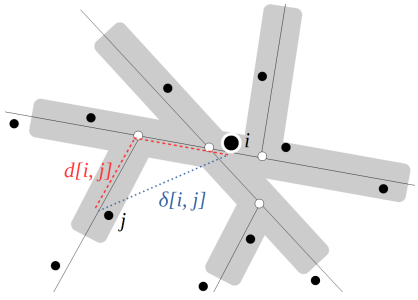


Figure 1.21: Straightness. Comparison of $\delta[i,j]$ and $d[i,j]$. Source: author

1.2 Blocks

Block is a complex two dimensional urban unit which encompasses built-up, as well as open space of various land use. Borders of a block are defined by surrounding street network. A block is characterized from various points of view – its *geometry* stored in a polygon representation describes the overall shape of a block as seen in a map or plan. Moreover, character of a block is also influenced by buildings which lay in it. From topological point of view as well as the definition, each building must be located in exactly one block. The *arrangement of buildings* influences the overall perception of a block, and so do *individual characteristics of comprised buildings* computed in the previous section, which may be expressed through summarization techniques such as sum, variance etc.

1.2.1 Characteristics derived from block geometry

Many morphological attributes presented in the previous section may be equally well utilized for description of blocks – namely attributes derived from building geometry and from minimum bounding geometries. Since the definition of block is based on fixed criteria, the results are expected to be less influenced by data granularity than in case of buildings, whose definition was set by the data provider. However, attribute values will still be affected by the level of detail of block enclosing streets.

Description of block geometry does not comprise the information about inner variability of block generated by appertaining buildings, but is just as important. The footprint of block varies with different urban structures and reflects arrangement of street network. Historical centres will consist of small irregular block, whilst regular rectangular blocks will occur in younger parts with block building structure, as well as in garden suburbs and villa quarters. The size of blocks is expected to grow with distance from centre, and will be reaching its limits in suburbs and sparsely inhabited areas, such as forests or fields. The shape of block is also largely influenced by the underlying topography. Whilst above mentioned examples will probably apply for flat parts of cities, steep terrain will cause the blocks to stretch in the direction of contour lines and narrow in the direction of gradient.

Characteristics applied for the scale of buildings are thoroughly described in Section 1.1 and their behaviour will apply equally for the scale of blocks. Namely, it is possible to make use of following descriptors (authors utilizing this attribute are mentioned in brackets):

- area of block footprint (Schirmer and Axhausen [37], Gil et al. [10], Porta et al. [29]),
- perimeter of block footprint (Gil et al. [10]),
- fractal dimension (Gil et al. [10]),
- Gravelius index,
- number of block corners,

- squareness of block borders,
- distance of block corners to centroid (Gil et al. [10]),
- all attributes derived from bounding box (Gil et al. [10]),
- all attributes derived from on enclosing circle (Porta et al. [29]),
- all attributes derived from on convex hull.

1.2.2 Characteristics derived from block defining streets

Perception of a block is indirectly influenced by neighbouring streets, especially by **street width**. Historical centres are often characterized by narrow streets, which, together with houses located directly at block borders, contribute to the feeling of closed space. On the other hand, other quarters might be perceived differently due to wide boulevards. In order to describe a block, it is necessary to *summarize* the width of all neighbouring streets and observe for example the average or minimum values.

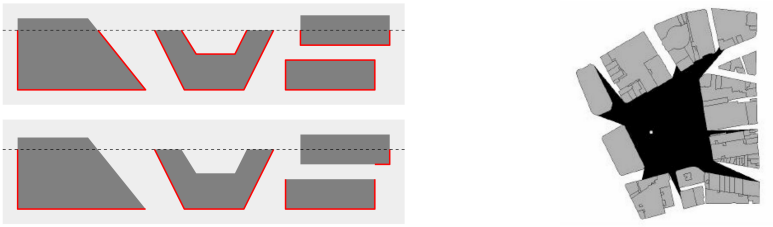
1.2.3 Characteristics derived from arrangement of buildings in a block

Perception of a block from the view of passers-by is influenced by the characteristics of comprised objects, as well as their arrangement and density. When walking on the street, we notice whether it is bordered by a wall or open space. We perceive how far objects are situated from the street and from each other, and if we can see between them or not. We also notice if the block gives impression of densely or sparsely built-up.

Façade which is perceived from the street

Blocks may be characterized by the percentage of building façade which is perceived from the street – some blocks are lined by building façades, some contain only buildings situated further inside the block. Whilst city centres generally consist of buildings directly neighbouring the streets, further away the buildings are moving inwards the block and the free space is often replaced by vegetation.

A façade will be considered as perceived from the street even if it does not directly border. Schirmer and Axhausen [37] for example define perceived façade as part of building façade contained by buffer of a certain distance from block border and Porta et al. [29] and Venerandi et al. [48] agree on 8 meters as limit distance for perception of façade. However, this approach brings several drawbacks – a façade may be detected as visible if it is situated in opposite side of a building or is at obtuse angle with block border. Also, façade hidden behind object situated closer to the street may be detected as well. A special care needs to be taken when creating a corresponding algorithm. It is also possible to utilize the *isovists* (field of vision) presented e.g. by Batty [4]. Difficulties and possible solutions are illustrated in Figure 1.22a.



(a) **(b)**

Figure 1.22: Issues in street facing façades detection – Figure **(a)** illustrates drawback of buffer method suggested e.g. by Schirmer and Axhausen [37]. In the upper Figure, façade at obtuse angle, façade situated on opposite side of building and façade hidden behind another object is detected. The corrected solution is presented in the bottom figure. Possible solution is utilization of isovists illustrated in Figure **(b)** . Source: **(a)** author, **(b)** Batty [4].

Three characteristics are based on described phenomenon – **total length of façade that is perceived from the street** (also called *built front* and used by Schirmer and Axhausen [37] and Porta et al. [29]), **percentage of façade that is perceived from the street** (used by Schirmer and Axhausen [37]) and **closeness of block**. Whilst first two values are associated with the influence of street on façades, the latter describes opposite effect, i.e. the influence of façade on street, and will be described separately.

$$Built\ front\ percentage = \frac{Built\ front}{Total\ facade\ perimeter} = \frac{[m]}{[m]}$$

Closeness of block

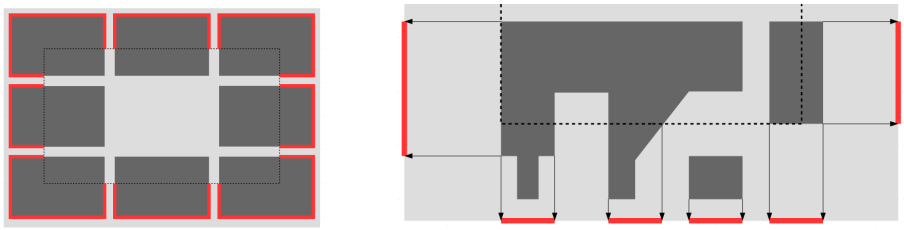
Closeness of block is defined in Schirmer and Axhausen [37] as *ratio of street facing façades to total length of perimeter of block*. Approach based on the aforementioned definition is very sensitive to the definition of *street facing façade* and it may eventually result in much higher values than one would expect (a completely closed block should probably be assigned closeness equal to one). Alternative definition is suggested by Porta et al. [29], Venerandi et al. [48] and Dibble et al. [8]. The term *closeness* is replaced by **built front ratio**. It is defined as *percentage of the street lined up by buildings within limit distance from the side-walk line* (Venerandi et al. [48]). Although the definition might seem similar to the one mentioned above, it addresses block borders, rather than façades. Built front ratio is obtained by orthogonal projection of street facing façades on the borders and comparing resulting length to the perimeter of whole block (see Figure 1.23b). This definition ensures that the highest value obtained will be equal to one.

$$BFR = \frac{Built\ front\ projection}{Block\ perimeter} = \frac{[m]}{[m]}$$

Built front ratio, as defined above, reaches values from zero to one. At zero, all buildings are further from block border than limit value. With rising values the blocks are

becoming bordered by buildings within the limit distance from block border, eventually reaching one, i.e. it is possible to perceive façades from each point of block border.

The value of built front ratio largely depends on chosen limit distance. Porta et al. [29] and Venerandi et al. [48] agree on 8 m limit distance, however, it is up to the data analysts and urban planners to decide at which distance the façade is no longer perceived.



(a) **(b)**

Figure 1.23: Two definitions of block closeness - 1.23a illustrates closeness of block as defined by Schirmer and Axhausen [37]. Length of street facing façades is 230 m, length of block border is 159 m, resulting in closeness value of 1.45 (limit distance is 5 m). 1.23b illustrates built front ratio as defined by Porta et al. [29], Venerandi et al. [48] and Dibble et al. [8]. Street facing façades in certain limit distance is projected on block border. Source: author

Number of closed liveable courtyards

The thesis proposes to use an identical attribute as in the scale of objects, i.e. *number of closed courtyards*. Similarly, only those courtyards evaluated as liveable are taken into the total count. Unlike in the scale of objects, closed courtyard of a block may be bounded by several different buildings, thus the final number of courtyards in a block might be larger than sum of courtyards in single buildings belonging to this block.

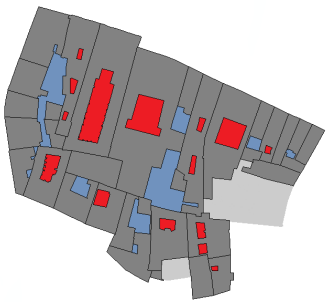


Figure 1.24: Closed courtyards. Red courtyards belong to single building, blue are bounded by several different buildings. Light gray spaces are not bounded and therefore do not represent courtyards. Source: author

High courtyard numbers are expected to occur in historical city centres with dense housing development or block development. On the other hand, housing estates and garden suburbs will probably have a very small number of courtyards.

Closeness of open space

When comparing the different effects on perception of open space, *closeness*, along with the total area might have the largest influence. This attribute shall not be confused with **Closeness of block** – whilst closeness of block focuses on its perception from the street, closeness of open space deals with perception when actually standing within the block. Consequently, the same block might have very high value of block closeness and low value of open space closeness, and vice versa (see Figure 1.25).

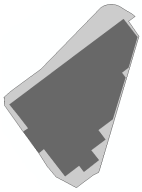
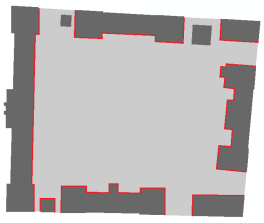
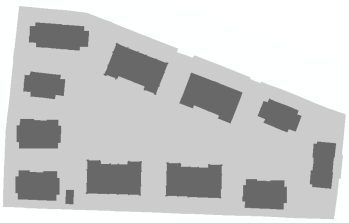


Figure 1.25: Difference between *Built front ratio* and *Closeness of open space*. Whilst BFR is relatively high (0.87) – building situated close to block border is making the block perceived closed from 87 % – OSC is lower (0.48), because open spaces are bounded only from inner side and remain open towards the street. Source: author

Schirmer and Axhausen [37] propose to model closeness using the number of perceived courtyards. Open space is perceived as courtyard if any gap between buildings is smaller than half of building height. In this thesis a different approach is suggested. Open space will be perceived as closed, if it is completely bound by buildings, i.e., when the length of border of open space is identical to length of border of enclosing buildings. These two values may be compared in a form of ratio, resulting in *Borders ratio*, which reaches value from zero (i.e. no shared border exists) to one (i.e. completely bound open space) (see Figure 1.26a). Buildings which are completely contained by open space (it is possible to "walk around"), are not considered as bounding, because they do not represent a barrier (see Figure 1.26b). Closeness of open space of a whole block will be produced as weighted average of borders ratio of all open spaces within block, with weight given by area of each open space. Closeness of open space aims mainly at distinguishing between block and detached development.



(a) Ratio of borders of open space (light gray) to borders shared with enclosing buildings (red) is 0.84. Source: author



(b) Buildings which are not forming barriers of open space. It is possible to "walk around". Source: author

Figure 1.26: Borders ratio. Source: author

Distance of buildings to block border

Whilst in the above mentioned characteristics an existence of buildings close to block border and following ratios are observed, it is possible to observe the single distance of each building to block border, as described in Section 1.1.7, and summarize this number for a whole block. Schirmer and Axhausen [37] suggest to consider *minimum* and *average*. The resulting value will not only help to stress the block types as defined in Section 1.2.3, but will also, in case that *standard deviation* is observed, provide information whether the development is situated in similar distance to block enclosing street (low values), or is evenly spread across the block (high values).

Number of touching buildings

Attribute proposed by Hartmann et al. [12] aims at differentiation between blocks with detached or group housing. Normalization by total number of buildings is essential, as only this number will provide the average number of building neighbours and will distinguish between detached (zero neighbours), semi-detached (one neighbour), terraced (two neighbours) and grouped (three and more neighbours)) development.

Nearest distance between buildings

Yu et al. [54] propose to summarize the distance of each building to its closest neighbour and observe the *average*. This attribute will however not provide much of useful information – take an example of zero average nearest distance, which may apply for both a block containing only grouped houses, as well as only semi-detached houses. *Standard deviation* of nearest distance between buildings would be of much higher information value, as it would help to distinguish between evenly or randomly spread buildings.

Permeability

The extension of previous phenomenon is the permeability, defined by Schirmer and Axhausen [37] as *number of views between buildings in ratio to perimeter of block*. Axial looks between buildings can be derived from Thiessen polygons of buildings within a block. A lower limit for distance between two buildings can be defined – this would express a gap too narrow to see or walk through, and such buildings would be considered neighbouring.

Permeability not only represents the visual aspect of perception, but also the possibility to walk through the block, in between the buildings (in case no barriers, e.g. fences, exist).

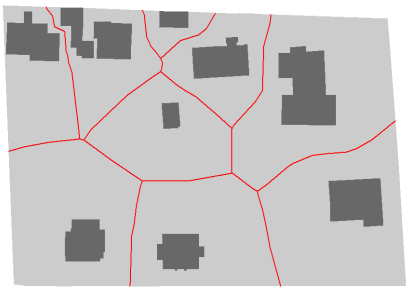


Figure 1.27: Permeability of urban block. 16 axial views are found by Thiessen tessellation. Gap between two buildings in upper left corner is not considered wide enough. Source: author

Number of buildings

Finally, perception of block is also influenced by number of buildings. However, as already mentioned, this value is largely affected by the granularity of input data and may therefore be biased. Number of buildings as block characteristic is used by Gil et al. [10], Porta et al. [29], Banzhaf and Höfer [3] and Schirmer and Axhausen [37].

1.2.4 Characteristics derived from morphological attributes of buildings in a block

In the previous section, block characteristics based on arrangement and density of comprised objects were presented. This section focuses on characteristics acquired by summarizing attributes of comprised objects, as presented in Section 1.1. This procedure will allow to transfer and utilize all attributes of scale of buildings in the scale of blocks and to describe its inner dimensions and variability, which co-determinates the perception of a block.

It is possible to observe many statistics of each summarized attribute, e.g. *minimum*, *maximum*, *average*, *sum* and *standard deviation* or *variance*. Each statistics will provide different information and it is crucial to determine which one will significantly contribute to the final classification procedure. *Minimum* and *maximum* may serve well when observing a lower or upper limit of certain feature, for example building height. *Average* reflects a value of certain attribute non-biased by dimensions of block or its number of buildings. Sometimes it is advantageous to use a *weighted average* – an example is given in Section 1.2.3. *Sum* captures the total of observed characteristic. It is important to decide whether this absolute value produces any contributing information – sometimes a relative value is more valuable. Finally, *standard deviation* and *variance* bring information about variability of observed phenomenon inside the block. Table 1.1 provides an overview of summarization of each attribute from Section 1.1. The statistics are proposed for the purpose of this thesis and in general it is possible to use any of the above mentioned for any attribute, in case a good reason exists. Some of these derived characteristics are also used in the literature and in such case authors are mentioned in brackets.

Building attribute	Min	Max	Sum	Average	St. dev.
Area of footprint	-	✓[54]	✓[37], [10]	✓[54]	✓[37], [54]
Perimeter of footprint	-	-	✓	✓	✓
Floorspace	-	-	✓[10], [29]	✓	-
Height	✓	✓[54]	-	✓[54], [29]	✓[37], [54]
Number of floors	✓	✓	✓[10]	✓[37]	✓[37]
Volume	-	✓[54]	-	✓[54]	✓[54]
Fractal dimension	-	-	-	✓	✓
Gravelius index	-	-	-	✓	✓
Number of corners	-	-	-	✓	✓
Squareness of walls	-	-	-	✓	✓
Distance of corners to centroid	-	-	-	✓	✓
Number of courtyards	-	-	-	✓	-
Total courtyard area	-	-	-	✓	-
R. of total courtyard area to footprint area	-	-	-	✓	-
Area of BB	-	-	-	✓	✓
Perimeter of BB	-	-	-	✓	✓
Width of BB	-	-	-	✓	✓
Length of BB	-	-	-	✓	✓
Elongation from BB	-	-	-	✓	✓
R. of footprint area to BB area	-	-	-	✓	✓
R. of footprint perimeter to BB perimeter	-	-	-	✓	✓
Area of EC	-	-	-	✓	✓
Perimeter of EC	-	-	-	✓	✓
R. of footprint area to EC area	-	-	-	✓	✓
Area of CH	-	-	-	✓	✓
Perimeter of CH	-	-	-	✓	✓
Length of CH	-	-	-	✓	✓
R. of footprint to CH area	-	-	-	✓	✓
R. of footprint perimeter to CH perimeter	-	-	-	✓	✓
Schumm's longest axis to area ratio	-	-	-	✓	✓
Area of influence zone	✓	✓	✓	✓	✓
R. of footprint area to influence zone area	-	-	-	✓	✓
Distance to closest neighbouring building	-	-	-	✓[54]	✓
Number of neighbouring buildings	-	-	-	✓[12]	✓
Existence of neighbouring buildings	-	-	-	✓[12]	✓
Distance to block border	✓[37]	-	-	✓[37]	✓

Table 1.1: Summarization of building attributes. (R. – ratio)

Several other characteristics are based on these summarized attributes, describing the influence of buildings on their block. These are based on the area of footprint of building, as well as the number of floors in a block.

Gross space index

Also called *built-up density* or *coverage*, gross space index represents a relationship between built and non-built space and is computed as a simple ratio of sum of building footprints to total area of block, reaching value from zero (low density) to one (high density). Blocks with dense development are found in city centres, while suburbs are often characterized by sparse built-up density. Gross space index is widely used as a block characteristics – e.g. in Schirmer and Axhausen [37], Gil et al. [10], Steiniger et al. [41], Taubenböck et al. [43], Yu et al. [54] and Urban Knowledge [39].

Subtracting the gross space index from one results in **ratio of open space to block area**, which describes reversed characteristic, i.e. the proportion of open space on the total area of block.

$$GSI = \frac{Built-up\ area}{Block\ area} = \frac{[\mathrm{m}^2]}{[\mathrm{m}^2]}$$

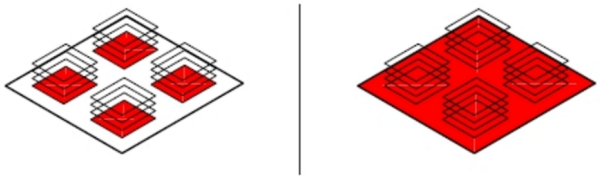


Figure 1.28: Principle of gross space index. Source: Urban Knowledge [39]

Floor space index

Also called *floor area ratio*, floor space index is a measure of the built intensity of a block expressed as a ratio of total floorspace to area of block. Low values (up to zero) represent low built intensities. The upper limit for floor space index is not set. Although two block might have a similar built-up density, they may differ in floor space index, because total floorspace in one may be different from total floorspace in the other. Analogously, two blocks may have equal floor space index although their built-up densities differ. Blocks with high built intensity are typical for city centres, as well as for administrative quarters with high-rise buildings. Floor space index is utilized e.g. by Yu et al. [54], Urban Knowledge [39] and Porta et al. [29].

$$FSI = \frac{Sum\ of\ floorspace}{Block\ area} = \frac{[\mathrm{m}^2]}{[\mathrm{m}^2]}$$

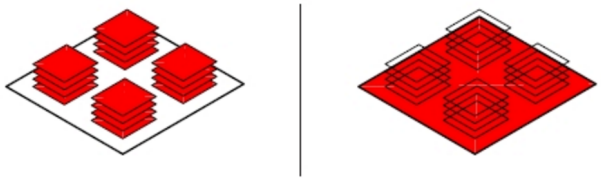


Figure 1.29: Principle of floor space index. Source: Urban Knowledge [39]

Open space ratio

Last attribute, utilized by Gil et al. [10] and Urban Knowledge [39], is also called *spaciousness*. Unlike *ratio of open space to block area* (1.2.4), it expresses how much open space exists per square metre of total floorspace. The result of spaciousness may be understood as the potential density of building inhabitants on the open space of a block.

$$OSR = \frac{1 - GSI}{FSI} = \frac{Open\ space}{Sum\ of\ floorspace} = \frac{[\mathrm{m}^2]}{[\mathrm{m}^2]}$$

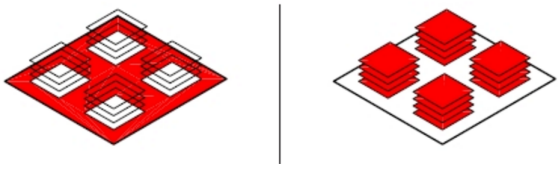


Figure 1.30: Principle of open space ratio. Source: Urban Knowledge [39]

Conclusion

No matter how much this chapter attempted to provide an overview of as many morphological attributes as possible, the list is in no ways exhausted. Various modification of proposed characteristics exist and many different attributes were probably overlooked in the literature. It is also worth noting that many of introduced building attributes might be equally well utilized in case building shape is represented in three dimension. Such representation would allow e.g. to assess not only the planar compactness, but also the three-dimensional, take into account style of roof or observe variances in building heights and many others (Serna et al. [36]).

Chapter 2

Classification of urban morphology

In the previous chapter an overview of various morphological attributes of buildings and urban blocks was presented. Each of these descriptors by itself already stores a piece of information about urban morphology – for example through Gravelius index the buildings may be classified into different categories of compactness and with gross space index it is possible to differentiate between several classes of built-up density. None of these attributes on its own can however explain the vast variability of urban morphology. In order to attempt to do so, the classification needs to take into account all these descriptors together.

Although each of presented attributes captures different characteristics of a building or urban block, a similar behaviour of some attributes will be most likely observed when comparing their values. For example, large area of a building will imply large area of its bounding box, as well as large area of its enclosing circle. Similarly, a building recognized as compact by Gravelius index will probably be identically characterized by ratio of building footprint area to enclosing circle area. This phenomenon, i.e. correlated attributes, could lead to excessive influence of such variables on the overall result. *Principal component analysis* is a method that helps to eliminate the correlation between variables and creates a smaller set of variables which together explain most of the variability in the dataset. (James et al. [17])

As already mentioned in the introduction, the approach of different authors towards the classification varies, but overall the used methods may be divided into two categories – *supervised methods* and *unsupervised methods*, both belonging to the field of *statistical machine learning*. With supervised classification it is possible to derive a probability of a response variable belonging to a particular category. In the urban morphology domain such approach is presented by defining zone-specific thresholds for observed attributes and assigning blocks to particular zone based on the values of these attributes. Supervised classification techniques comprise e.g. logistic regression, linear discriminant analysis or support vector machine. Unsupervised methods on the contrary aim at understanding the relationships amongst the measured values and observations without any prior knowledge. Such approach in the urban morphology field will find groups of blocks with similar attribute values and assign each of these groups a different category. Unsupervised

classification techniques comprise e.g. clustering. (James et al. [17])

The purpose of this thesis, i.e. an unbiased assess of the current classification of urban structures, implies the utilization of unsupervised learning. Values of set of attributes are observed for each urban block, but there is no prior knowledge about the resulting classes. The goal is then to discover the unknown subgroups of similar urban blocks in the city. *Clustering* is the method that allows to search for such patterns. Last but not least, in order to represent heterogeneous zone and remove outliers, *mode filtering* may be applied after clustering.

All the aforementioned methods, i.e. *principal component analysis*, *clustering* and *mode filtering* will be explained in the following sections.

2.1 Principal component analysis

Principal component analysis (PCA) is a process of reducing the dimensionality of data by finding so-called *principal components*. In scope of this thesis, a large set of p attributes is presented and many of them are likely to be correlated with others. This set of attributes defines a p-dimensional space in which n observations are measured. Through PCA these p attributes are summarized in a smaller number of representative variables – *principal components*. They represent the underlying structure in the data and together they capture as much as possible of information stored in the original set of attributes, i.e. explain as much as possible of its variability. Principal components are uncorrelated with each other and each of them is a linear combination of the original attributes. The direction of each principal component is the direction of highest variance in observations and also represents a line in p-dimensional space which is closest to the observations. (James et al. [17], The Pennsylvania State University [21], Analytics Vidhya [31])

First principal component Z_1 is a linear combination of original attributes which explains the maximum variance in the data set. The direction of Z_1 is the direction along which observations vary the most. If $\mathbb{A} = (A_1, A_2, ..., A_p)$ is a set of attributes, Z_1 is defined as follows:

$$Z_1 = \Phi_{11}A_1 + \Phi_{21}A_2 + ... + \Phi_{p1}A_p \tag{2.1}$$

Elements Φ_{11}, Φ_{21}, ..., Φ_{31} are called the loadings of Z_1 and together they form the loading vector Φ_1. The sum of square of all loadings must be equal to one in order to obtain a unique result (otherwise Z_1 could capture an arbitrary large variance). Finding the values of loadings of Z_1 is an optimization problem in which a linear combination of sample attributes with largest sample variance is searched.

$$\max_{\Phi_{11},\ldots,\Phi_{p1}} \left\{ \frac{1}{n} \sum_{i=1}^{n} (z_{ij})^2 \right\} \tag{2.2}$$

$$\text{where } z_{ij} = \sum_{j=1}^{p} \Phi_{j1} x_{ij} \tag{2.3}$$

$$\text{subject to } \sum_{j=1}^{p} \Phi_{j1}^2 = 1 \tag{2.4}$$

Values z_{11}, z_{2i}, ..., z_{n1} are called scores of Z_1. Problem 2.2 can be solved using linear algebra technique of eigen decomposition. It turns out, that, when computing a variance-covariance matrix $\mathbb{A}^T\mathbb{A}$ of analysed data set and ordering its eigenvalues from largest to smallest, the elements of corresponding eigenvectors are equal to loadings. (The Pennsylvania State University [21], James et al. [17], Analytics Vidhya [31])

Second principal component Z_2 is a linear combination of A_1, A_2, ..., A_p which has the largest variance of all linear combinations which are uncorrelated to Z_1. It can be demonstrated that the condition of zero correlation between Z_1 and Z_2 is achieved when loading vectors Φ_1 and Φ_2 are orthogonal. Therefore a similar problem to 2.2 is solved with Φ_1 being replaced by Φ_2 and an additional condition that Φ_1 and Φ_2 are orthogonal. Every other succeeding principal component is obtained by identical technique, i.e. it captures most of the remaining variation and is not correlated with the previous components. Altogether, $\min(n-1, p)$ of principal components can be constructed. (James et al. [17], The Pennsylvania State University [21], Analytics Vidhya [31])

The *proportion of variance explained* (PVE) by m-th principal component is computed as follows

$$PVE = \frac{\text{variance explained by m-th component}}{\text{total variance in data set}} = \tag{2.5}$$

$$\frac{\frac{1}{n}\sum_{i=1}^{n} z_{im}^2}{\sum_{j=1}^{p} \text{Var}(X_j)} = \tag{2.6}$$

$$\frac{\sum_{i=1}^{n} \left(\sum_{j=1}^{p} \Phi_{jm} x_{ij} \right)^2}{\sum_{j=1}^{p} \sum_{i=1}^{n} x_{ij}^2}. \tag{2.7}$$

In order to perform PCA, attributes should be normalized to zero mean and standard deviation equal to one, or be measured in the same units. If normalization is not performed prior to PCA, attributes with higher variance of observations will have larger loadings in the resulting principal components than attributes with smaller variance. This will cause the first principal component to be dependent on the attribute with high variance. Sum of PVE's of all $\min(n-1, p)$ principal components is equal to one.

After applying PCA a question remains how many of principal components to use. Generally, we are looking for the smallest number of first k principal components in order

to obtain the simplest interpretation of variance in data. However, we want the PVE of the first k principal components to explain a sizeable proportion of variance in data, ideally close to one. Literature proposes several methods of determining the ideal number of components. It is for example possible to observe the scree plot (dependence of PVE on number of principal components) and search for an "elbow", i.e. a point at which the proportion of variance explained by every other principal component drops off (James et al. [17], The Pennsylvania State University [21], Analytics Vidhya [31]). Another criterion is to use as many components so that *at least certain percentage* (80 %, 90 %) *of variance is explained* (Minitab Inc. [52]). Many authors (Minitab Inc. [52], Schirmer and Axhausen [37]) suggest a *Kaiser method*, i.e. standard deviation of principal component equal to et least one. However, none of these techniques is well-accepted to solve this problem and the solution depends largely on the specific case.

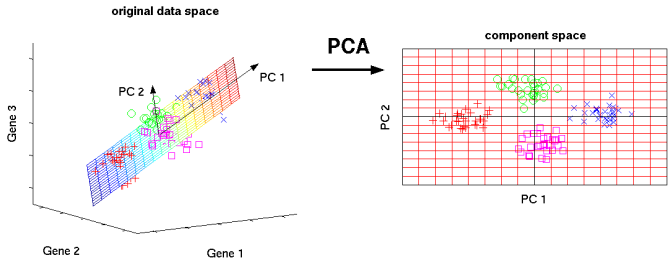


Figure 2.1: Transformation of a 3-dimensional data set to 2 dimensions using PCA. Source: Scholz [38]

2.2 Clustering

Clustering comprises a large amount of methods aiming at discovery of subgroups (clusters) in a dataset of n observations with p values. Observations belonging to one subgroup should be characterized by similar values, whilst the similarity in between clusters should be as small as possible.

Vast number of clustering methods exists. From the spatial point of view, clustering may be either performed on non-spatial data (*non-spatial clustering*), or spatial data (*spatial clustering*). Partitioning methods (such as K-means) and hierarchical methods (such as agglomerative clustering) represent two large groups of *non-spatial clustering* techniques. In the domain of this thesis, non-spatial clustering would be represented by grouping based on morphological attributes. In such case, no spatial relationship of blocks is considered, and both distant and close buildings may be grouped in the same cluster. On the other hand, *spatial clustering* takes into account the spatial proximity of objects, regardless the non-spatial attributes. Partitioning methods and hierarchical methods are also available for spatial clustering, accompanied by density-based, grid-based and constraint-based clustering. If blocks in this thesis were clustered spatially, several compact groups of neighbouring buildings would be detected regardless their attributes.

Only minor attention is paid to the combination of both, i.e. *clustering based on both non-spatial attributes and spatial proximity.* First approach is to handle both components separately. Two algorithms – spatially dominant SD_CLARANS and non-spatially dominant NSD_CLARANS were developed from a spatial clustering partitioning algorithm CLARANS (Ng and Han [25]). SD_CLARANS first clusters based on spatial component and afterwards "*performs attribute-oriented induction on the non-spatial description of points in each cluster. NSD_CLARANS first applies attribute-oriented generalization to the non-spatial attributes to produce a number of generalized tuples. Then, for each such generalized tuple, all spatial components are collected and clustered using CLARANS*". (Wang and Wang [51]) Second approach is to handle both spatial and non-spatial component together by combining their dissimilarities. Several algorithms capable of this approach were presented recently. GDBSCAN (Sander et al. [35]), derived from density-based DBSCAN, considers the non-spatial attribute as a weight when calculating the cardinality of neighbourhood of an object. Density based DBRS "*can take account of a property related to non-spatial attribute(s), by means of a purity threshold, when finding the matching neighborhood*". (Wang et al. [50]) Algorithm SOFM by Jiao and Liu [18] presents the concept of composite distance, which captures both spatial and non-spatial distance between objects. Zhang et al. [55] extend the technique and present a generalized distance function capturing distance between non-spatial attributes, boundary features and spatial events in order to cluster polygons.

Methods such as NSD_CLARANS, GDBSCAN and DBRS are currently not implemented in any major GIS or statistical software and their implementation is beyond the scope of this thesis. The technique of including spatial component suggested in this thesis is inspired by Jiao and Liu [18] and Zhang et al. [55]. It performs K-medoids clustering on composite distance and is easy to implement and reproduce.

Composite distance

Dissimilarity is a key concept of assessing the similarity between individual objects. It is expressed as distance in p-dimensional space, stored in $n \times n$ *dissimilarity matrix* (symmetrical with zeros at diagonal). There are various techniques of deriving a distance between observations, Euclidean distance or Manhattan distance being the most common ones, accompanied by others. In spatial domain, network distance should be considered as well, as it enables to define various spatial constraints etc. (James et al. [17], Jiao and Liu [18], Zhang et al. [55])

Dissimilarity matrix useful for combination of spatial and non-spatial attributes is computed from composite distance $D_{i,j}$, which is a weighted addition of *euclidean distance* $D^a_{i,j}$ *between objects in the attribute space* and *network or euclidean geometric distance* $D^n_{i,j}$.

$$D^a_{i,j} = \sqrt{\sum_{k=1}^{p}(z_{ik} - z_{jk})^2} \tag{2.8}$$

$$D^n_{i,j} = \begin{cases} network\ distance(i,j) \\ \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \end{cases} \tag{2.9}$$

$$D_{i,j} = w^a D^a_{i,j} + w^n D^n_{i,j} \tag{2.10}$$

z_{ik} and z_{jk} represent values of attribute k of observations i, j in p-dimensional attribute space. x_i (x_j) and y_i (y_j) are the coordinates of centroids of urban blocks. In case network distance is utilized, the point situated closest to the network is considered.

All attributes are considered to have equal weights. In order to be able to assess the influence of geometric and attribute distance, a condition is added:

$$w^a + w^n = 1 \tag{2.11}$$

With w^a equal to zero the clustering will result in spatial clustering. With w^n equal to zero the clustering will only be influenced by non-spatial attributes. In order to make the weighting easily interpretable, it is suggested to scale both attribute values and geometrical distances before the composite distance calculation. (James et al. [17], Jiao and Liu [18], Zhang et al. [55])

As mentioned in Section 2.1, it is possible to perform PCA before clustering and compute distances $D^a_{i,j}$ in low-dimensional space. It is in fact an appropriate solution to avoid taking into account too large number of attributes, as with very high number of attributes the distances $D^a_{i,j}$ become very similar and clustering might not perform as expected.

K-means and *K*-medoids clustering

K-means is a partitioning clustering technique which groups observations into predefined number K of clusters. Each observation will be assigned to at least one of K clusters, and no observation will belong to more than one cluster, i.e. clusters are not overlapping. The aim of K-means clustering is to minimize the total within cluster variation $W(C_k)$ summed over all K clusters C_k.

$$\min_{C_1,\dots,C_K} \left\{ \sum_{k=1}^{K} W(C_k) \right\} \tag{2.12}$$

Within-cluster variation of k-th cluster is expressed as sum of squared Euclidean distances of all observations belonging to this cluster, divided by total number of observations in this cluster. (James et al. [17])

Since there are almost K^n ways to partition n observations into K clusters, local optimum is searched in problem 2.12 rather than global. The algorithm may be expressed as follows:

Algorithm 1 K-means clustering (Source: Ng [24])

1: Randomly initialize K cluster centroids.
2: Iterate until the cluster assignments stop changing:
(a) Cluster assignment – each observation is assigned to closest cluster centroid (in terms of chosen distance – Euclidean, network etc.).
(b) Recompute centroids – move centroid to the mean of its cluster.

Solution provided by this algorithm is a local optimum which depends on the initial random assignment of cluster numbers. Therefore results of the algorithm will not be identical when computed twice for the same dataset with same number of clusters. It is possible to run the algorithm several times and select a solution with smallest value of objective 2.12. *Silhouette analysis* also helps to assess the overall strength of clustering. It compares how close each observation is to others in its own cluster and how close it is to observation in other clusters. Values close to one denote strong membership of this observation in a cluster, observations in between two clusters are assigned values close to zero and negative values indicate misclassified observations. The average silhouette width represents the overall strength of clustering. (Rousseeuw [33], Spector [40])

Other drawback of K-means, besides the predefined number of clusters, is that spatially, final clusters tend to be spherical and of similar size. Last but not least, K-means is very sensitive to outliers, because extreme values shift the computed mean significantly. This problem can be minimized by using ***K*-medoids** clustering, which is a technique similar to K-means. Unlike K-means, cluster center is not represented as mean value, but as medoid, which is an actual observation most centrally located in the cluster, with minimum sum of distances to other observations. K-medoids minimizes the sum of pairwise dissimilarities instead of a sum of squared Euclidean distances and compared to K-means it is more robust to noise and outliers (see Figure 2.2). (Sammut and Webb [34])

No best solution exist to choose the number K. It might be driven by the clustering purpose, when K is known. Other solution is to compare various clusterings using the *average silhouette width*. Large number of clustering indices, such as gamma index, silhouette width, Frey index and other 27 indices is offered by R package *NbClust*.

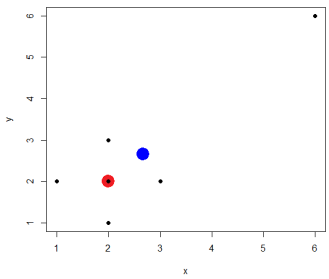


Figure 2.2: Data with an outlier and cluster centroids obtained by K-means (blue) and K-medoids (red). Source: author

2.3 Mode filtering

Even though performed as suggested above, the final zones of similar blocks might not be completely heterogeneous and spatially connected – a block of different group may for example lay amongst otherwise identically clustered blocks. This undesirable phenomenon, or outlier, may be suppressed using a *mode filter* on the resulting groups, as suggested e.g. by Steiniger et al. [41]. Inspired by image processing domain, each block is assigned the most common value (i.e. of cluster group) of blocks in its neighbourhood. The definition of neighbourhood is crucial in this technique. The size of neighbourhood defined by maximum distance will affect the size of resulting zones and the definition of distance will affect their shape – Euclidean distance will take into account all objects in a circular buffer, whilst network distance enables to model real accessibility and spatial obstacles.

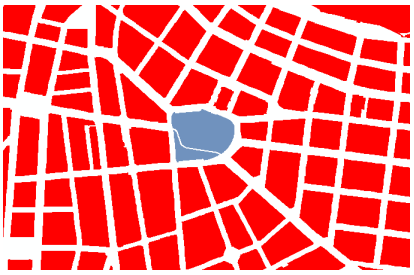


Figure 2.3: Outlier in otherwise heterogeneous zone. Source: author

Chapter 3

Methodology

In this section the complete methodology of urban morphology analysis in Prague is presented. The process begun with *acquiring necessary data* and considerable amount of time was spent by *data preprocessing* – topology validation and data cleaning, followed by *computing of morphological attributes*. Finally, a *classification process* was performed. The methodology workflow is illustrated in Figure 3.1. .

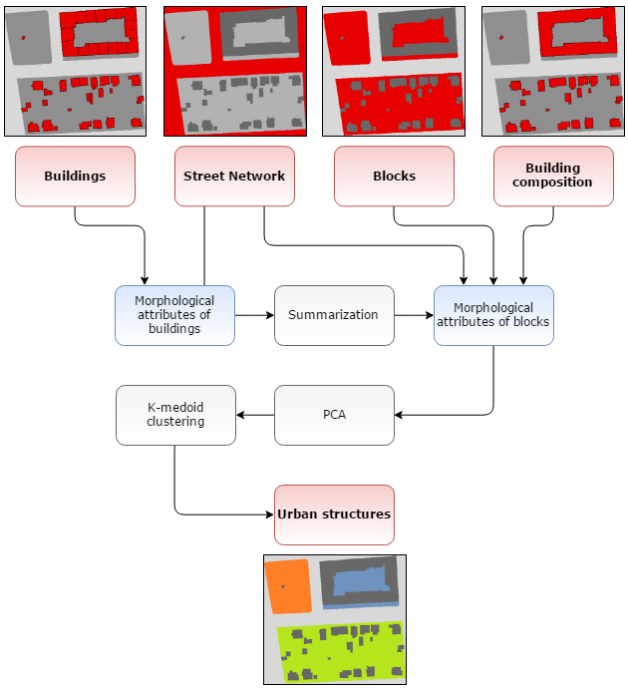


Figure 3.1: Schema of methodology workflow. Source: author

3.1 Software

Data preprocessing, computation of morphological attributes and mode filtering is performed using *ArcGIS*, versions 10.3 and 10.4 (ArcGIS Platform [2]). Scripts for computing individual morphological attributes and mode filtering are written in *ArcPy*, a site-package allowing to utilize the ArcGIS functionality together with Python (ArcGIS Platform [53]). Together they are stored in a toolbox, allowing for further usage, editing or extension. Individual morphological attributes as well as principal components and final urban structures are visualized as an ArcGIS Online map (ArcGIS Online [1]). PCA and clustering is computed via *R Project for Statistical Computing* (The R foundation [32]). However, the proposed methodology is not software-dependent and might be equally well performed using any other GIS or statistical programmes.

3.2 Data acquisition and preprocessing

Two main data sources are used to acquire spatial data necessary to perform the analysis – *Open Data* provided by *IPR Prague* and *StreetNet CZE* from *Central European Data Agency, a.s.* Four datasets were used from Prague Open Data – polygon feature class *Podlažnosti* (Floors), with spatial resolution 1:5 000, polygon feature class *Současný stav využití území* (Current Land Use), as well in spatial resolution 1:5 000, polygon feature class *Praha – hranice území* (Prague – borders of area), provided in spatial resolution 1:500 and finally polyline feature class of *Střední šířky ulic – linie* (Mean street widths – lines), provided in spatial resolution 1:5 000. Polyline feature class of routable streets is taken from StreetNet CZE in form of selected features from dataset *Úsek komunikace* (Road segment). Four feature classes, which serve as input into all following analyses, are derived from these data sets. The overview of data sources together with used attributes is given in Figure 3.2.

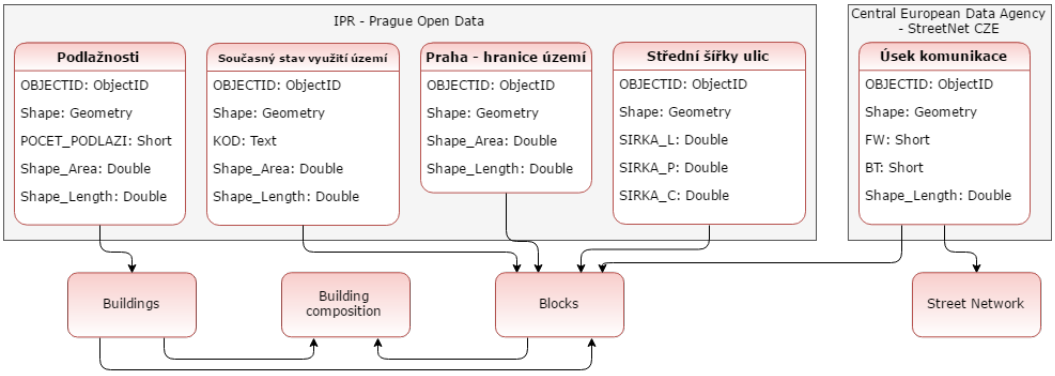


Figure 3.2: Schema of utilized data sources, datasets and derived feature classes. Source: author

3.2.1 Buildings

Polygon feature class of *Buildings* was acquired from dataset Podlažnosti (Floors). As already mentioned, the values of morphological attributes largely depend on the granularity of input data, i.e. on the definition of individual building. In case of Floors dataset, consistent scale and level of detail through the whole area of Prague was assumed, and therefore it was not necessary to create composites, like e.g. in Schirmer and Axhausen [37]. A thorough preprocessing was carried out before the data analysis.

Buildings represents the foundation stone of the entire study of urban morphology in Prague and the final precision of result derives from the consistency of buildings representation. Ortophoto raster layer of Prague was utilized to deal with disputable problems. Following issues were resolved in the preprocessing:

Topological inconsistencies First of all topology was checked using automatic topology control in ArcGIS. Violation of Must not overlap rule with cluster tolerance of 1 cm was observed. Small overlaps were assigned to bigger of involved buildings. Overlaps larger than one square metre were separated into unique building. Several entirely overlapping buildings (duplicates) were discovered and in such case only one building was kept.

Eliminating holes inside buildings Although holes in polygons of buildings usually represent shafts and courtyards, small holes may be considered a representation issue. All holes smaller than 0.5 m^2 were filled using Eliminate Polygon Part tool.

Figure 3.3: Eliminating holes inside buildings. Source: author

Eliminating holes in between buildings Thin or small gaps in between buildings, assumed a result of imprecise drawing, were eliminated by two methods – using Integrate tool with 3 cm tolerance, and by extracting only gaps and measuring their size. Gaps with area smaller than one square metre were manually checked and eventually filled. Gravelius index was computed in order to discover other inconsistencies.

Figure 3.4: Eliminating holes between buildings. Source: author

Eliminating small objects Polygons with area smaller than one square metre were considered not sufficiently large to represent a building (even a garden shed, for example) and were either deleted, merged to neighbouring building or expanded, based on the underlying ortophoto.

Eliminating atypical objects Other atypical polygons were discovered after calculating morphological attributes. Atypical objects and shapes were discover using Gravelius index, as well as setting limits for minimum width of bounding box. Such issues were manually resolved using the underlying ortophoto – deleted, merged with neighbouring buildings or expanded.

Polygon feature class of *Buildings* created by the aforementioned approach consists of 208 293 elements, each representing one building. The attribute table of *Buildings* contains following attributes: OBJECTID, a unique ID of each building, POCET_PODLAZI, an information about number of floors, and information about polygon perimeter and area, if feature class is stored in ArcGIS geodatabase.

3.2.2 Building compositions

An additional feature class of *Building compositions* was created, representing buildings dissolved by appropriate block. This feature class, apart from information about composition OBJECTID, area and perimeter, stores value of block OBJECTID.

3.2.3 Blocks

No available data source contained a feature class of urban blocks for the area of Prague, and therefore these had to be derived from available data. The initial definition of an urban block, i.e. area bordered by streets, was not sufficient for the purpose of this thesis. Large blocks were for example formed by sections of Vltava river and neighbouring areas. A decision was made to incorporate other spatial phenomena, not only communications, to represent block borders. These are *water bodies* due to the reason explained above, *railways*, because they are perceived as block barriers similarly to the roads, *paths* and *pathways* in order to divide large blocks in smaller ones, and finally *pedestrian areas* in district of Prague 1 – Václavské náměstí, Staroměstské náměstí, Náměstí republiky and Karlův most, because they are perceived similarly as roads. The last phenomenon was largely influenced by local knowledge. (Comparison of before / after situation is presented in Figure 3.5.)

(a) Before adding pedestrian area – all red buildings belong to one block, which does not represent the perception.

(b) After adding pedestrian area – eight blocks are created instead of one.

Figure 3.5: Reasons for incorporating pedestrian zones as block borders. Source: author

Three datasets were utilized in order to represent blocks in this form. Polygon features of appropriate land use attribute values from dataset Současný stav využití území (Current Land Use) were extracted first. The overview of utilized features is given in Table 3.1. Segments of railways in tunnels were deleted manually, because they do not form block barriers.

Land use	**Value of attribute KOD**
highways	DK
main roads	VN
streets and roads	VM
water	HY
railways	DZP
pathways	VPP
paths	VC
pedestrian areas (only in district of Prague 1)	VPN

Table 3.1: Utilized features of Current Land Use dataset

Because not all roads and streets are represented in this source, specific features from feature class Road segment of dataset StreetNet CZE were utilized as well. The overview of used features is given in Table 3.2. Additionally, features with values of attribute BT equal to 2 (ford) or 4 (tunnel) were excluded, because these do not form block barriers. Road segments are represented as polylines (two dimensional features), but by its nature, blocks should not directly border with each other. Road segments were therefore converted to polygon (three dimensional) representation by creating a buffer of half of street width around each segment. Width of road is not part of the attribute table of Road segment, and it was added manually based on street widths appropriate for road category, set by Czech technical standards (Kategorie komunikací [19]). This approach is rather questionable given the environment of a city, and would deserve a more detailed research or data acquisition if more precise results are required. Assigned street widths are presented in Table 3.3.

Communication type	Value of attribute FW
Highway	1
Motorway and other multi-lane roads	2
One-lane road	3
Roundabout	4
Slip-road	10
Service road	11
Parking lot entrance / exit	12
Service or shopping zone entrance / exit	13
Pedestrian zone	14

Table 3.2: Utilized features of Road segment feature class

Communication type	Communication class	Width [m]
Highway	-	27.5/2
Motorway and other multi-lane roads	highway	27.5/2
	1st class road	24.5/2
	2nd class road	9.5
	other communication	4.0
One-lane road	-	4.0
Roundabout	-	4.0
Parking lot	-	4.0
Slip-road	-	4.0
Service road	-	4.0
Parking lot entrance / exit	-	4.0
Service or shopping zone entrance / exit	-	4.0
Pedestrian zone	-	4.0

Table 3.3: Approximate street widths

Last data source for obtaining urban blocks was polygon dataset Praha – hranice území (Prague – area borders), which provides the extent of outer urban blocks. Communications from both data sources, as specified above, were erased from area of Prague in order to obtain individual blocks.

Resulting polygons represent urban blocks as defined in the first paragraph of this section. Two other adjustments were carried out in order to ensure the topological and morphological correctness of the data. Firstly, with respect to the definition, each building should be located inside exactly one block. Small slivers of building polygons were found to be located outside of blocks. This problem was solved by merging the blocks with buildings. Secondly, polygons which do not contain any buildings were deleted, because they do not represent urban blocks. Remaining problem is caused by buildings spanning over a road. Several ways exist how to resolve this case – it is possible to split the building and delete the segment which is situated above the road. It is also possible to consider these two connected blocks as one. The latter approach was utilized in this thesis.

Polygon feature class of Blocks created by the aforementioned approach consists of 10 439 elements, each representing one urban block. The attribute table of Blocks contains unique OBJECTID of each block, as well as information about its area and perimeter, if stored in ArcGIS geodatabase.

Algorithm 2 Blocks extraction

1: *Communications_LU* = Select by attribute from *Land_Use* based on 3.1
2: Delete segment of *Communications_LU* in railway tunnels
3: *Communications_ro* = Select by attribute from *Road segment* based on 3.2
4: Add field *street_width* in *Communications_ro*
5: Calculate field *street_width* using Czech technical standards [19]
6: *Communications_ro_b* = Buffer of (*street_width* / 2) around *Communications_ro*
7: *Communications* = *Communications_LU* + *Communications_ro_b*
8: *Blocks_temp_1* = Erase *Communications* from *Prague_area_borders*
9: *Blocks_temp_2* = Merge *Blocks_temp_1* and *Buildings*
10: *Blocks_temp_3* = Convert *Blocks_temp_2* to singlepart
11: *Blocks* = Select *Blocks_temp_3* containing *Buildings*

3.2.4 Street network

A detailed and topologically correct street network of Prague is essential for network analysis. It was provided by Central European Data Agency, a.s. in form of feature class Úsek komunikace (Road segment) of dataset StreetNet CZE. Due to restrictions on the amount of provided data, only features with specific Communication type attribute are available in the provided feature class. The overview is given in Table 3.4.

Communication type (FW)	**Participating in feature class**
Highway (1)	YES
Motorway and other multi-lane roads (2)	YES
One-lane road (3)	YES
Roundabout (4)	YES
Parking lot (6)	YES
Indoor parking lot (7)	YES
Slip-road (10)	YES
Service road (11)	YES
Parking lot entrance / exit (12)	YES
Service or shopping zone entrance / exit (13)	YES
Pedestrian zone (14)	YES
Pavement (15)	NO
Path (16)	NO
Stairs (20)	NO
Passage (21)	NO
Bicycle and pedestrian trail (24)	NO
Bicycle trail (25)	NO
Pathway (26)	NO
Ski trail (27)	NO

Table 3.4: Features of Road segment feature class

Besides, additional polyline feature class of *Street Widths* is acquired from dataset Střední šířky ulic – linie (Mean street widths – lines) and used to compute characteristics of blocks derived from block defining streets. The attribute table of *Street Widths* contains, besides others, following attributes: SIRKA_L, mean width of left side of street from a centerline in metres, SIRKA_P, mean width of right side of street from a centerline in metres and SIRKA_C, mean width of a street.

3.3 Morphological attributes of buildings

Altogether, 63 morphological attributes were computed for scale of buildings. Computational time varied based on the complexity of attribute – from tens of minutes up to tens of hours.

3.3.1 Characteristics derived from building geometry

Area and perimeter

Area and perimeter of building footprint are automatically computed from building geometry in case the data is stored in ArcGIS geodatabase. Otherwise they can be manually calculated from properties of the Geometry object.

Floorspace

Floorspace is calculated as number of floors multiplied by area of footprint. In case number of floors is zero, floorspace equals area of footprint.

Approximate height

Approximate height is derived from number of floors. A standard height 3.5 m of ground floor and 3 m of all other floors above ground is assumed. For side buildings the number of floors is not defined, and in such case this building is assigned height of 3 m (see Algorithm 3).

Algorithm 3 Approximate height

```
if floors_no = 0 then
    approx_height ← 3
else
    approx_height ← 3.5 + 3 · (floors_no − 1)
```

Number of floors

Number of floors is already present in the *Buildings* feature class in attribute *POCET_PODLAZI*.

Volume

Volume is computed from footprint with eliminated holes to capture the perception of building from outside. Holes in footprint are filled using *Eliminate Polygon Part* tool.

Compactness measures

Fractal dimension and *Gravelius index* are simply derived from values computed above. *Ratio of building volume to façade area* is not calculated, as it gives identical results as fractal dimension.

Number of building corners, squareness of building walls and distance of building corners to centroid

All these attributes are all calculated using the same algorithm. For each building, all vertices are stored in a list and an angle of each corner is computed using the coordinates of previous and following vertex. This angle is used to categorize the corner in a corresponding class and compare it to the class norm (see Table 3.5). An average deviation from norm for the whole building is derived. The value of angle is also used to decide whether the corner represents significant or insignificant point. Lines at angle between 170° and 190° are not considered to form a building corner. Number of significant corners is summed for each building and a distance of each significant corner to the building centroid is observed. Distances are scaled to 0-1 interval to allow for relative comparison and a standard deviation of scaled distances is calculated.

Angle value	Norm
$\alpha \leq 45°$	0°
$45° < \alpha \leq 135°$	90°
$135° < \alpha \leq 225°$	180°
$225° < \alpha \leq 315°$	270°
$315° < \alpha$	360°

Table 3.5: Angle classification

Topological skeleton

A decision was made not to include topological skeleton for evaluating buildings in Prague, as the algorithm is not part of any GIS software functionality. Additional implementation of the algorithm is beyond the scope of this thesis. However, it is possible to derive topological skeleton from Thiessen polygons of segments of building borders and following vectorization of resulting raster. This method is computationally very demanding, because the total number of building border segments is very high.

3.3.2 Characteristics derived from building courtyards

Courtyards are derived from polygon feature class of *Buildings*. Using *Eliminate Polygon Part*, all holes in *Buildings* are filled. Courtyards are obtained by subtracting original *Buildings* from resulting feature class using *Erase* tool.

Algorithm 4 Corner number, corner distance, squareness

```
1: for Building ∈ Buildings do
2:     Deviation = 0
3:     Corner_no = 0
4:     Distances = []
5:     for i := 1 → length(Building_corners) do
6:         Corner_angle = angle(Building corners[i-1], Building corners[i+1])
7:         if Corner_angle < 170 || Corner_angle > 190 then
8:             Distances.append(distance(Building_corners[i], Centroid))
9:             Corner_no = Corner_no + 1
10:        if Corner_angle ≤ 45 then
11:            Deviation = Deviation + |Corner_angle|
12:        else if Corner_angle > 45 & Corner_angle ≤ 135 then
13:            Deviation = Deviation + |90 − Corner_angle|
14:        else if Corner_angle > 135 & Corner_angle ≤ 225 then
15:            Deviation = Deviation + |180 − Corner_angle|
16:        else if Corner_angle > 225 & Corner_angle ≤ 315 then
17:            Deviation = Deviation + |270 − Corner_angle|
18:        else
19:            Deviation = Deviation + |360 − Corner_angle|
20:    for Distance ∈ Distances do
21:        Distance = Distance / max(Distances)
22:    Corner_square = Deviation / length(Building_corners)
23:    Corner_dist = stdev(Distances)
```

In order to classify courtyards into liveable and non-liveable, two assumptions were made – firstly that courtyards with area smaller than $4\,\mathrm{m}^2$ are not suitable for any outside activities (e.g. it is not convenient to place a table with two chairs in 2 by 2 metres space), secondly that to make any place liveable, the sun should reach the ground at least for a few hours in a year. Analysing courtyards in Prague, the maximum solar elevation angle on the summer solstice is 63° (Time and Date AS [42]), thus a courtyard bounded by building of height h should be at least $h/\tan(63°)$ long. A simplification is made here as the orientation of buildings towards sun is omitted. Comparing all possible pairs of vertices of courtyard border, the length of longest line contained by each courtyard is calculated and compared to the limit length.

Courtyards satisfying both conditions are summarized based on their building of origin using *Summary statistics* and *number of liveable courtyards* in each building as well as *total courtyard area* is computed. *Ratio of total courtyard area to building area* is derived from the footprint of filled buildings.

In total, only 627 buildings on the whole area of Prague were found to contain one or more liveable courtyards. These are mostly situated in city centre and nearby quarters. Several sport stadiums were detected as well – these are also the objects with highest values of courtyard area, whilst courtyards in historical development are significantly smaller. Values of courtyard ratio do not seem to follow any spatial pattern.

Algorithm 5 Closed liveable courtyards

```
Filled_buildings ← eliminate holes in Buildings
CO_multipart ← erase Buildings from Filled_buildings
CO_single ← split CO_multipart to singlepart
for Courtyard ∈ CO_single do
    max_length ← −1
    V_combination ← all combinations of Courtyard.Vertices
    for V_pair in V_combination do
        test_line ← Polyline(V_pair[1], V_pair[2])
        if test_line.length > max_length then
            if Courtyard contains test_line then
                max_length ← test_line.length
    CO_single.liveable_idx ← 0
    if Courtyard.area > 4 & max_length > Courtyard.height/1.96 then
        CO_single.liveable_idx ← 1
CO_single ← select CO_single with CO_single.liveable_idx = 1
CO_summary ← summarize CO_single by Blocks
Join Buildings and CO_summary
Buildings.CO_number ← CO_summary.count
Buildings.CO_area ← CO_summary.area
Buildings.CO_ratio ← Buildings.CO_area / (Buildings.area + Buildings.CO_area)
```

3.3.3 Characteristics derived from minimum bounding geometry

Bounding box, enclosing circle and enclosing convex area (convex hull) of each feature in polygon feature class of *Buildings* are computed using *Minimum Bounding Geometry* tool. All attributes are then derived using simple arithmetic operations.

3.3.4 Characteristics derived from building's influence zone

Influence zones of buildings form a Thiessen tessellation of space. Raster representation of Thiessen polygons is computed using *Euclidean Allocation* tool on polygon feature class of *Buildings*. Output cell size is set to 50 cm in order to reduce computational requirements. Resulting raster is then *converted to vector* representation, *dissolved* based on original OBJECTID, as sometimes a single cell is detected as individual polygon, and *clipped* with borders of municipality of Prague.

Thiessen polygons created by this method are not entirely precise due to cell size definition and raster vectorization and especially in case of neighbouring buildings the border of Thiessen polygon not always follows the shared border of buildings. This problem may be considered negligible regarding the scale of result.

Area of influence zone and *ratio of building footprint area to influence zone area* are easily derived from feature class of Thiessen polygons.

3.3.5 Characteristics derived from building neighbourhood

Distance to closest building

Distance to closest building is computed simply by *Near* tool with both input feature class and near feature class represented by *Buildings*.

Number and existence of neighbouring buildings

Number of neighbouring buildings is computed using *Polygon Neighbours* tool. Only those neighbours sharing boundary longer than 0 m (i.e. touching by walls, not by corners) are summarized and counted.

Distance to urban block border

Analogous approach as in previous case, i.e. *Near* tool with input feature class of *Buildings* and near feature class of *Blocks*, is applied to find the *distance to block border*.

Characteristics based on street network

The interaction of a building with its surroundings is evaluated using five accessibility measures computations, presented in Section 1.1.7, and an additional network evaluation. In all computations, accessibility is expressed through network distances rather than Euclidean. This approach better reflects the true arrangement of street network, affecting the morphology of Prague observed in this thesis. Street network is modelled by feature class *Road segment* of dataset StreetNET CZE.

Reach, *Gravity*, *Betweenness*, *Closeness* and *Straightness* are calculated for each building using three search radii – 100 m, 300 m and 500 m (Schirmer and Axhausen [37]), which represent three different walk-able surroundings of a building. In case of *Gravity*, coefficient β is set to 0.00217 (Handy and Niemeier [11]), assuming walking and impedance attribute in metres. Weighting is utilized in case of *Reach* – total number of floorspace and total area of building footprints were observed besides the total number of reachable buildings. In remaining four accessibility measure weighting is not included. All five indices are computed using *The Urban Network Analysis Toolbox* for ArcGIS, which is distributed and can be used under the Creative Commons Attribution – NonCommercial – ShareAlike 3.0 Unported License. (City Form Lab [46])

Additional network evaluation inspired by Schirmer and Axhausen [37] is carried out apart from the aforementioned accessibility measures. Number of reachable dead-ends and intersections and total length of reachable network is computed for each building, again using 100 m, 300 m and 500 m search radius. These measures are not computed using the UNA toolbox. For each building, approximated by its centroid, a service area of given radius is created and observed phenomena which fall inside this service area are

measured. Input features are polygon feature class of *Buildings*, polyline feature class of *Roads*, point feature class of *Juctions* and a *Network dataset* created from *Roads*. Two new attributes, *UNA_r_network_length* and *UNA_r_crossroads_number* are created (r stands for used radius, i.e. 100 m, 300 m or 500 m).

Algorithm 6 Accessible network

1: $radius = \{100, 300, 500\}$
2: *Service_area_layer* ← Make Service Area Layer(*Network_dataset*, *radius*)
3: **for** *Building* ∈ *Buildings* **do**
4: *Centroid* ← centroid(*Building*)
5: Add Location(*Service_area_layer*, *Centroid*)
6: Solve(*Service_area_layer*)
7: *Service_area_polygon* ← extract polygon of *Service_area_layer*
8: *Buildings.crossroads_number* ← Select *Junctions* intersecting *Service_area_polygon*
9: *Roads_clip* ← Clip(*Roads*, *Service_area_polygon*)
10: *Buildings.network_length* ← Summarize(*Roads_clip.Shape_Length*)

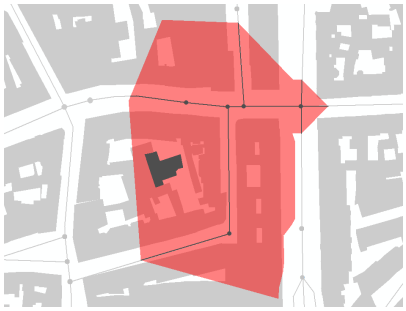


Figure 3.6: Number of junctions and network length with service are of a building. Source: author

3.4 Morphological attributes of blocks

Altogether, 144 morphological attributes were computed for scale of buildings. Computational time varied based on the complexity of attribute – from tens of minutes up to tens of hours.

3.4.1 Characteristics derived from block geometry

All attributes derived from block geometry were calculated analogously to attributes derived from building geometry in Section 3.3.1.

3.4.2 Characteristics derived from block defining streets

For each block, weighted *average* (by length of appropriate street segment) and *minimum width* of neighbouring streets is observed. As mentioned in Section 3.2.3, block borders are defined using several data sources, including two definitions of streets. Street width is stored in a polyline feature class *Střední šířky ulic* (Mean street width), which was not utilized in block definition, and therefore the detection of neighbouring streets is not straightforward. In order to detect such streets, a buffer of street width (separately for left and right side) is constructed around centerline and then those segments which intersect the block are selected and summarized. A search distance of 2 m is used as not always buffered street intersects block border. In case no neighbouring street segment is found, the same approach is repeated using approximate street width of polyline feature class *Úsek komunikace* (Road segment). In case no neighbouring street segment is found again, an approximation is made using the distance to nearest neighbouring block. This approach suffers from obvious disadvantages and the best solution would be to store the information about street width directly in block definition streets and utilize it in the block creation process.

Algorithm 7 Façade perceived from street

1: *Buffer_left* ← Buffer(*Street_width*, left_width)
2: *Buffer_right* ← Buffer(*Street_width*, right_width)
3: *Buffer_complete* ← Merge(*Buffer_left*, *Buffer_right*)
4: **for** *Block* ∈ *Blocks* **do**
5: Select *Buffer_complete* intersecting *Block* with 2 m tolerance
6: *Block*.StreetWidth_avg ← avg (left_width + right_width) weighted by *Street_width*.length
7: *Block*.StreetWidth_min ← minimum (left_width + right_width)

3.4.3 Characteristics derived from arrangement of buildings in a block

Façade which is perceived from the street

The algorithm is based on simple intersection of building façades with 8 m buffer (Venerandi et al. [48], Porta et al. [29]) within block borders. However, as already mentioned in Section 1.2.3, it is necessary to eliminate those parts of façade which are in fact not visible from the street. These may be sorted into two groups – segments situated on opposite side of a building or at obtuse angle with block border, and segments hidden behind another object (see Figure 1.22). A proposed solution solves the first issue by only taking into account those segments whose buffer on left topological side (i.e. outside of a building) intersects block border. This technique will eliminate any segments which are not facing the street with limit of 90°. Further research needs to be carried out in order to solve the second issue.

Two attributes are derived from built front – *length of façade which is perceived from the street* and *percentage of façade which is perceived from the street*.

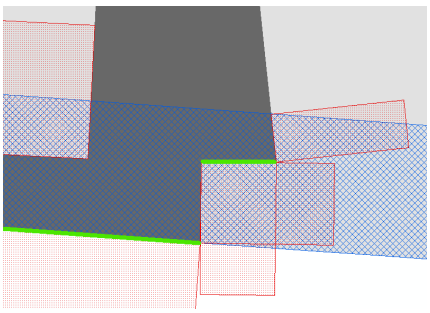


Figure 3.7: Façade perceived from a street is obtained by clipping all façade with inner block buffer (blue). Afterwards an outside buffer (red) is created for each suspect segment and only those segments with buffer intersecting block border are retained (green). Source: author

Algorithm 8 Façade perceived from street

1: *Building_borders* ← Polygon to Line(*Buildings*)
2: *Block_borders* ← Polygon to Line(*Blocks*)
3: *Block_buffer* ← Buffer(*Blocks*, −8 m)
4: *Buildings_clip* ← Clip(*Building_borders*, *Block_buffer*)
5: *Buildings_segments* ← Split at Vertices(*Buildings_clip*)
6: *Building_buffers* ← Buffer(*Buildings_segments*, 8 m, left)
7: Select *Building_buffers* intersecting *Block_borders*
8: Select *Buildings_segments* sharing a line segment with *Building_buffers*
9: *façade_length* ← Summarize length of *Buildings_segments* by *block id*

Closeness of block

Built front ratio of each block is computed as ratio of length of the portion of block border which is lined by buildings within 8 m (Porta et al. [29], Venerandi et al. [48]) from the border, to the total length of block border. Main steps of obtaining *BFR* are explained in Algorithm 9.

Input feature classes are polygon feature class of *Blocks* and polygon feature class of *Building compositions*. For each block, border of corresponding building composition is intersected with buffer of 8 m from the block border, in order to create all necessary vertices on façades. All vertices of façades are afterwards projected on block borders. Orthogonal projection is achieved using ArcGIS function *Polyline.queryPointAndDistance(in_point)*. Figure 3.8 illustrates the principle of this function. Since distance from first point of polyline is output of this function, it is easy to determine final length of block border which is in fact a projection of façade.

Algorithm 9 Built front ratio

distance_limit = 8 m
Building_borders ← Polygon to Line(*Buildings*)
Block_borders ← Polygon to Line(*Blocks*)
Block_borders_split ← Split at Vertices(*Block_border*)
Block_buffers ← Buffer(*Block_borders*, *distance_limit*)
for *Block* ∈ *Blocks* **do**
    Select *Building_borders*, *Block_buffers* and *Block_borders_split* belonging to this *Block*
    Create vertices at *Building_borders* on intersection with *Block_buffers*
    Line_segments ← empty polyline feature class
    for *Building_borders_segment* ∈ *Building_borders* **do**
        for *Block_borders_segment* ∈ *Block_borders_split* **do**
            distance_max ← maximum length of *Block borders*
            distance_min ← 0
            point_max ← empty point
            point_min ← empty point
            for *Corner* ∈ *Building_borders_segment* **do**
                Projected_point ← orthogonal projection of *Corner* to *Block_borders_segment*
                o ← distance from *Corner* to *Projected point*
                d ← distance from *Border segment* first point to *Projected point*
                if *d* < *minimum distance* **then**
                    distance_min ← *d*
                    point_min ← *Projected_point*
                if *d* > *maximum distance* **then**
                    maximum distance = *d*
                    point_max ← *Projected_point*
            Line_segments ← Insert(Polyline(*point_min*, *point_max*))
    Built_front_projection ← Dissolve(*Line_segment*)
    Built_front_projection_length ← Length(*Built_front_projection*)
    BFR ← *Built_front_projection_length* / Perimeter(*Block*)

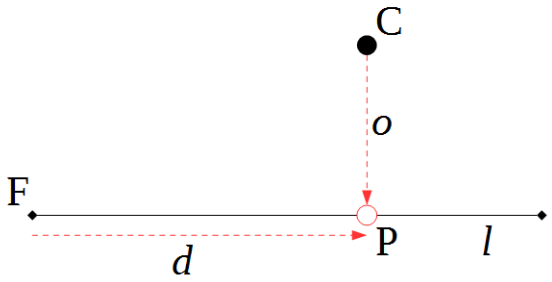


Figure 3.8: Principle of *Polyline.queryPointAndDistance(in_point)*: l – Polyline, C – in_point, P - orthogonal projection of C to l, F – l.firstPoint, d – distance from F to P, o – distance from C to P. Source: author

Number of closed liveable courtyards

The approach in calculating number of liveable is analogous to the one used for scale of *buildings* in Algorithm 5 with *Building* being replaced by *Building compositions*.

Closeness of open spaces

Open space is derived by subtracting *Building compositions* from *Blocks* using *Erase* tool. Afterwards, all holes in resulting open spaces are filled in order to avoid counting in buildings completely contained by open space. Similarly to the definition of *Courtyards*, areas smaller than $4\,\mathrm{m}^2$ are excluded from the computation and their *OSC* is NULL. Closeness of open space is computed as ratio between perimeter of shared border and perimeter of open space. The resulting value is weighted by open space area and is summarized in form of weighted average for each block.

As expected, OSC detected the difference between block and detached development. High values are found not only in the city centre, but also in local centres of town districts, whilst low values apply for garden suburbs, housing estates, industrial zones and blocks with sparse development (parks, forests, fields etc.).

Distance of buildings to block borders, number of touching buildings and nearest distance between buildings

Although these measures reflect the inner composition of a block, they are computed in a summarization process together with other building attributes (see Section 3.4.4).

Algorithm 10 Closeness of open spaces

1: *OS_multipart* ← erase *Buildings_composition* from *Blocks*
2: *OS_single* ← split *OS_multipart* to singlepart
3: *OS_filled* ← eliminate holes in *OS_single*
4: *Borders* ← intersect *OS_filled* and *Buildings*
5: **if** OS_single.area < 4 **then**
6: *OS_single.Border_Ratio* ← NULL
7: **else**
8: *OS_single.Border_Ratio* ← *Borders.length* / *OS_filled.length*
9: *OS_single.Weighted_Border_Ratio* ← *OS_single.Border Ratio* · *OS_single.Area*
10: *OS_summary* ← summarize *OS_single* by *Blocks*
11: *OS_summary.OSC* ← SUM(*OS_single.Weighted_Border_Ratio*) / SUM(*OS_single.Area*)
12: Join *Blocks* and *OS_summary*
13: *Blocks.OSC* ← *OS_summary.OSC*

Block permeability

Permeability expressed as number of axial looks between buildings is based on Thiessen polygons of buildings belonging to each block. The input to this algorithm is a polygon feature layer of *Blocks* and polygon feature layer of *Building compositions*. In order to avoid counting gaps too narrow to see or walk through, a limit distance of 1 m between any two objects is required and a buffer of 0.5 m is constructed for each building. Moreover, since we are not interested in permeability of courtyards, but only in permeability of open spaces accessible from the street, holes in building compositions are filled before the analysis. Each urban block is processed separately. Building composition belonging to this block is selected and Euclidean allocation (i.e. raster Thiessen polygons) is computed. Resulting raster is converted into polygons and afterwards into lines, which are clipped with the area of processed urban block. Permeability is expressed as number of these lines.

Algorithm 11 Permeability of urban blocks

1: *Filled_buildings* ← Eliminate polygon part(*Buildings_composition*)
2: *Buffered_buildings* ← Buffer(*Filled_buildings*, 0.5 m)
3: **for** *Block* ∈ *Blocks* **do**
4: Select *Buildings* belonging to this *Block*
5: *Permeability_raster* ← Eucliden Allocation(*Buildings*)
6: *Permeability_polygons* ← Raster to Polygon (*Permeability_raster*)
7: *Permeability_lines* ← Polygon to Line (*Permeability_polygons*)
8: *Permeability_clip* ← Clip(*Permeability_lines*, *Block*)
9: *Permeability* ← Get Count((*Permeability_clip*)

Number of buildings

Number of buildings is a product of building summarization process (see Section 3.4.4).

3.4.4 Characteristics derived from morphological attributes of buildings in a block

All attributes stated in Table 1.1 were computed. *Buildings* are intersected with *Blocks* in order to find out which block they belong to. Resulting feature class is summarized using *Summary statistics* tool based on feature id of corresponding block, and then joined to original attribute table of input *Blocks*. Additional attributes, i.e. *Gross space index*, *Floor space index* and *Open space ratio* are easily computed from the previously derived measures.

3.5 Discovering urban structures

Classification of urban structures is based on 144 morphological attributes computed for the scale of blocks. Closeness of open space was left out, as an error occurred in the processing – ArcGIS function Eliminate polygon part messes the output attribute table and after joining the result is not correct. This error was discovered after all subsequent processing was already carried out, and further incorporating of OSC would require several more tens of hours of computations. However, since OSC is probably highly correlated with other morphological attributes, the impact of its neglecting is not of high influence on the final results.

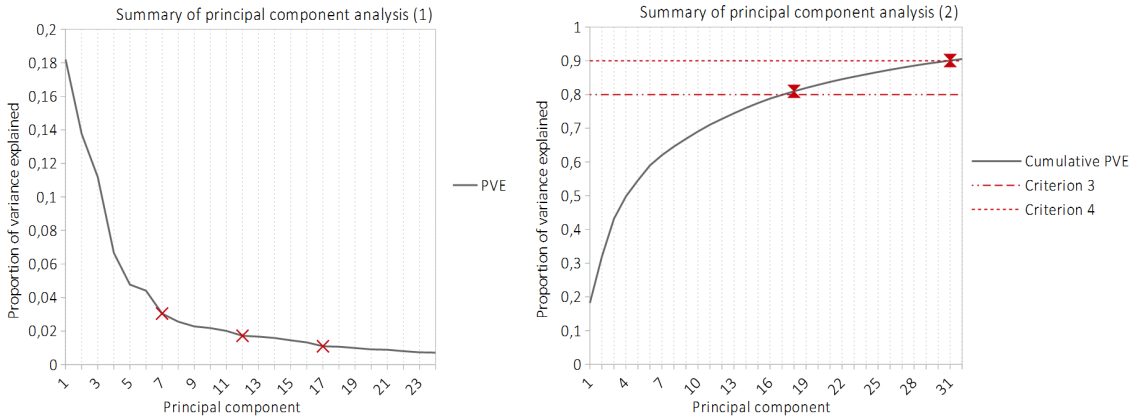
3.5.1 Principal component analysis

At first, all morphological attributes were scaled so that the mean value of each attribute is equal zero and standard deviation is equal to one. Principal component analysis was performed in *R* using function *prcomp()* from *pls* package. Computational time of PCA is up to few seconds. As explained in Section 2.1, each following component explains smaller proportion of variance than previous one. Due to large number of attributes, the growth of cumulative PVE is rather low – 99 % of cumulative PVE is reached at 73rd principal component.

In order to objectively decide how many principal components to use in further analysis, several criteria mentioned in Section 2.1 were taken into account. At first, a scree plot was explored. Three "elbows", **7**, **12** and **17** were picked up. Third criterion, i.e. use as many components so that at least 80 % or 90 % of variance is explained leads to 18 and 31 components, respectively. Kaiser method implies **24** components. At last, principal component were visualized in a map. It was found out that maps of **18**th and following principal components do not show any interesting pattern. An overview of all techniques and resulting components is given in Table 3.6. Table 3.7 and Figure 3.9 illustrate summary statistics of first 31 ordered components and highlight position of each limit principal component.

Criterion	**Explanation**	**Selected PC's**
Criterion 1	Scree plot – manual detection	1-7
		1-12
		1-17
Criterion 3	Cumulative PVE > 80%	1-18
Criterion 4	Cumulative PVE > 90%	1-31
Criterion 5	Kaiser method (Eigenvalue > 1)	1-24
Criterion 6	Visual checking of PC maps	1-18

Table 3.6: Various decision criteria for selecting an optimal number of principal components.



(a) Proportion of variance explained and results of scree test.

(b) Cumulative PVE and limit values for Criterion 3 and 4.

Figure 3.9: Summary of principal component analysis. Source: author

PC	Standard deviation	PVE	Cumulative PVE	PC	Standard deviation	PVE	Cumulative PVE
1	5.136	0.182	0.182	**17**	1.260	0.011	0.799
2	4.467	0.138	0.320	**18**	1.244	**0.011**	**0.810**
3	4.027	0.112	0.431	19	1.197	0.010	0.820
4	3.108	0.067	0.498	20	1.148	0.009	0.829
5	2.630	0.048	0.546	21	1.134	0.009	0.837
6	2.529	0.044	0.590	22	1.078	0.008	0.845
7	2.101	0.030	0.620	23	1.032	0.007	0.853
8	1.928	0.026	0.646	**24**	**1.019**	0.007	0.860
9	1.817	0.023	0.669	25	0.997	0.007	0.867
10	1.778	0.022	0.691	26	0.974	0.007	0.873
11	1.707	0.020	0.711	27	0.952	0.006	0.880
12	1.579	0.017	0.728	28	0.927	0.006	0.886
13	1.555	0.017	0.744	29	0.900	0.006	0.891
14	1.517	0.016	0.760	30	0.842	0.005	0.896
15	1.449	0.015	0.775	**31**	0.829	0.005	**0.901**
16	1.386	0.013	0.788	32	0.821	0.005	0.905

Table 3.7: Summary of principal component analysis. Bold principal components and statistics mark limit values for various criteria.

As it is not possible to decide which of the undertaken methods provides better results, all principal components selected by these methods (i.e. 1-31) were evaluated in further clustering.

Interpretation of principal components utilized in clustering

Each principal component is a linear combination of all 144 original attributes. However, the influence (loadings) of individual attributes vary and it is possible to understand the meaning of each principal component by looking at several first attributes sorted by highest absolute loadings. This way of interpretation does not mean that other attributes do not have any influence on value of principal component, their impact is just smaller. For each component used in final clustering (see following section), a subset from sorted attributes was extracted so that the sum of squares of their loadings is at least 0.5 (sum of squares of all loading is equal to one). An overview of extracted attributes is given in Table D.3 in appendix. Each concerned component is visualized in Figures D.1 and D.2 in appendix. The following paragraphs aim at stressing the meaning of each component utilized in final clusterings.

First principal component is influenced by characteristics of buildings: their size and variance in size. High values represent blocks with large buildings (area, volume, perimeter, length, width) and large variance in buildings size within block (perimeter, length, area, volume).

Second principal component is influenced by size of block, distribution of buildings in a block and neighbourhood density and accessibility. High values represent small blocks (area, perimeter) with large built-up density (GSI, FSI, IZ ratio), large built front ratio and dense neighbourhood with high accessibility measures (network length).

Third principal component is influenced by distribution of buildings in a block, neighbourhood structure and variance in shape of buildings. High values represent blocks with many buildings located both inside and close to block border, which vary in shape. Neighbourhood is characterized by large accessibility measures.

Fourth principal component is influenced by variance in both building size and shape, as well as by building height, its variance and accessible floorspace. High values represent blocks with large variance in building perimeter and shape, low buildings, low variance in height and low amount of accessible floorspace both in neighbourhood and block.

Fifth principal component is influenced by size of buildings and variance in their shape and height, as well as accessibility of neighbourhood and size of block. High values represent large blocks with large buildings with low variance in shape and height and good accessibility.

Sixth principal component is influenced by buildings, their shape and variance in size. High values represent blocks with complex buildings and low variance in size.

Seventh principal component is influenced by block, its size, shape and distribution of buildings. High values represent large blocks with small complexity of shape and small number of buildings having large influence zones and situated further from block border.

Eight principal component is influenced by block, its shape and accessibility of neighbourhood. High values represent circular or convex blocks with similar width and length, surrounded by sparse street network and containing buildings situated far from block border.

Ninth principal component is influenced by shape and height of buildings, surrounding network structure and distribution of buildings within a block. High values represent blocks with large built-up density and large built-front ratio, containing complex low buildings surrounded by sparse street network.

3.5.2 Clustering

Clustering arises from first 31 principal components computed in the previous section. K-me- doids clustering was performed in R using function *pam()* from package *cluster*. After a discussion with IPR, non-spatial clustering was utilized and no spatial component was taken into account.

In order to discover an optimal number of clusters K as well as number of principal components, clustering was computed several times for K running from 4 to 20 and 7 – 12, 17, 18, 24 and 31 first principal components. Lower values of K were not considered due to low predicative values of such clustering. Higher values of K were not considered due to descending trend of silhouette width for larger number of clusters. Additional numbers of principal components (8-11) were utilized in order to observe the behaviour of clustering in between values determined by different decision criteria. Computational time rises with number of clusters. Altogether, 170 clusterings were computed and evaluated, which took about half a day.

Assessing the right number of clusters is computationally very demanding task. For example, when computing all 30 indices from R package *NbClust*, the computational time for 10 439 blocks and K running from 4 to 20 is more than 10 hours. Due to lack of time, the optimal number of clusters was searched only using *average silhouette width*. A dependency of average silhouette width on number of clusters as well as number of principal components was observed and is illustrated in Figure 3.10.

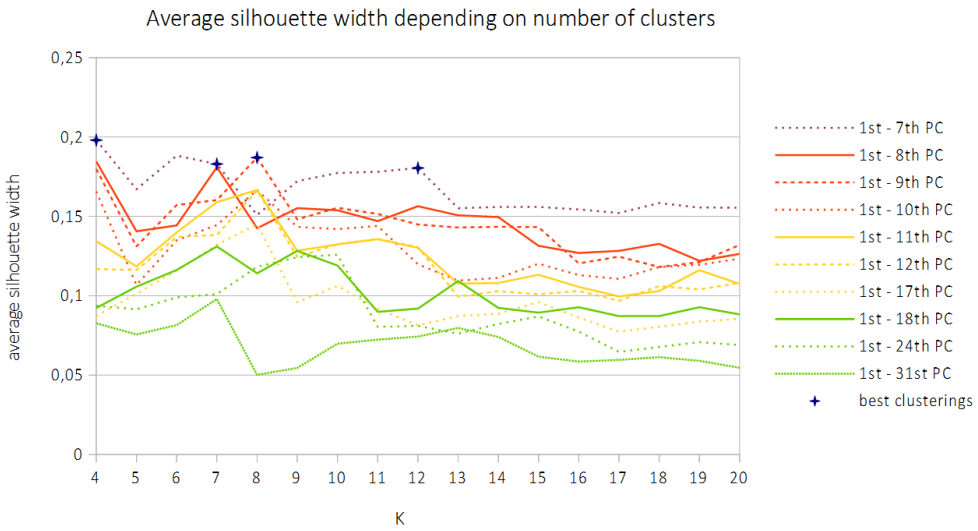


Figure 3.10: Dependency between silhouette width and number of clusters for various numbers of principal components. Source: author

Several conclusions may be inferred from the aforesaid plot. Smaller numbers of principal components generally tend to lead to more stable clustering (higher average silhouette widths). Similarities may be observed in the dependency of average silhouette widths on

K. Peaks in values are usually reached with K equal to 4, 7, 8 and 11, 12 or 13. Overall, the average silhouette widths are relatively small even for small numbers of clusters. This may imply a inappropriate clustering method, for example. Further research needs to be carried out in order to solve this problem.

As average silhouette width is only one of many indices evaluating the optimal number of clusters, and its values are very low, it is not straightforward to assess the optimal number of clusters and principal components. For this reason, four different clusterings were selected as a result and compared. Overall, the largest average silhouette width is the one of **4 clusters from 7 principal components**. As already mentioned, peaks appear at K equal to 7, 8 and 11, 12 or 13. Based again on largest average silhouette widths, **7 clusters from 7 principal components**, **8 clusters from 9 principal components** and **12 clusters from 7 principal components** were chosen as another possible results. Positions of best clusterings are illustrated in Figure 3.10 with a blue star. Clustering results will be discussed in the following chapter.

Chapter 4

Results and conclusion

4.1 Results

Four different divisions of urban structures were selected as an optimal result:

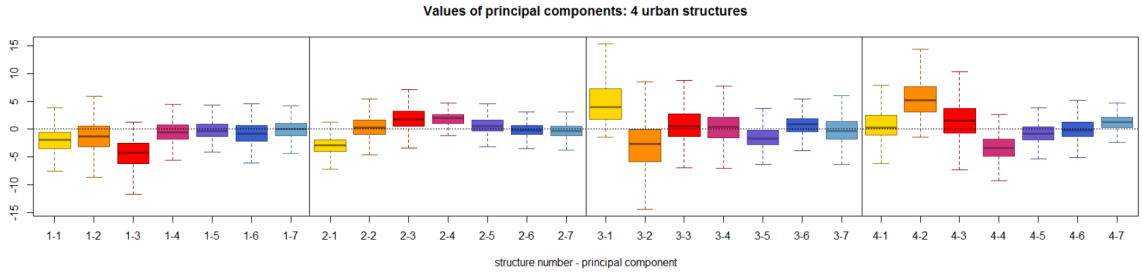
- 4 urban structures
- 7 urban structures
- 8 urban structures
- 12 urban structures

4.1.1 Comparison of results

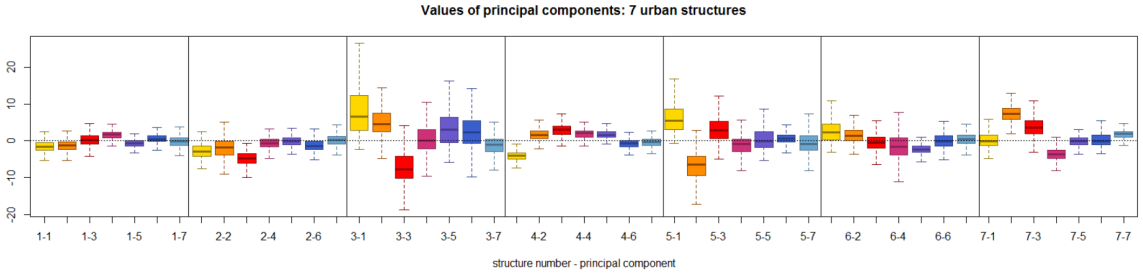
Figure 4.1 illustrates box-plots of individual urban structures, computed for each principal component. Length of box and whiskers illustrates the variability of values across all blocks belonging to particular structure. Low variability implies more similar blocks within structure. High variability implies larger differences amongst blocks within structure.

Decreasing importance of principal components may be observed in most groups. First principal components explain larger variance in the data sets and have usually the highest values, whilst last components have just a small differentiation significance.

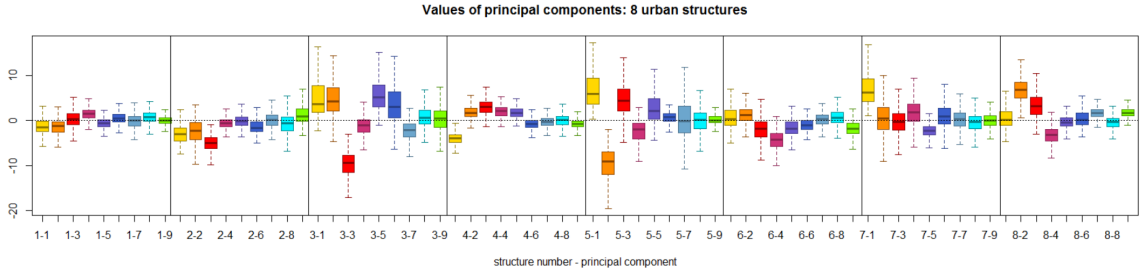
Similarities across results may be observed for some urban structures. Last group of box plots for example always characterizes block building structure. In visualization of result in the section below, similar structures are visualized in identical colors.



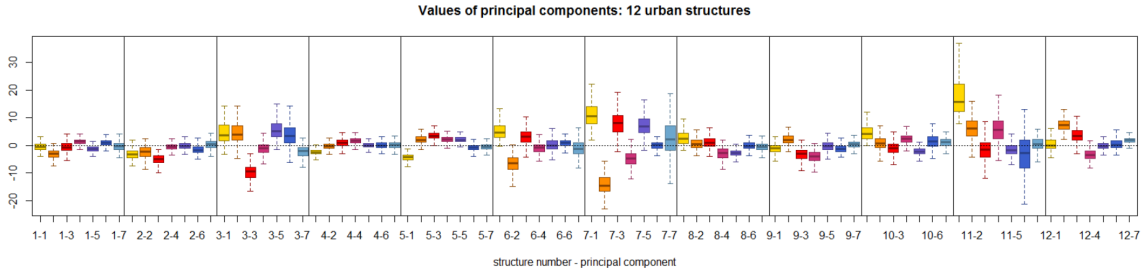
(a) 4 urban structures



(b) 7 urban structures



(c) 8 urban structures



(d) 12 urban structures

Figure 4.1: Summaries of results. Source: author

4.1.2 Four urban structures

In **first urban structure** (blue in Figures 4.2 and 4.3), medium to larger blocks of quite simple shapes contains smaller number of higher buildings of small to medium footprint with relatively simple shapes. Variance in both shape and size of buildings is small, however, variance in height is larger. Built up density of blocks is relatively low and buildings are located closer from block borders. Neighbourhood is sparse with lower accessibility. This structure often detects *housing estates*, as well as *blocks with one solitary building*.

Second urban structure (green in Figures 4.2 and 4.3) consists of medium size blocks of relatively simple shapes. Buildings tend to be low and have small footprint, but variance in their size and shape is considerable. Number of buildings in a block is higher than in first urban structure. Buildings are often separated from each other and are located relatively close to block borders. Neighbourhood is dense and has relatively large accessibility measures. *Garden suburbs and villa quarters* are often detected as this structure

Overall, **third urban structure** (light yellow in Figures 4.2 and 4.3) encompasses the largest and most complex and variable blocks. These contain larger number of low buildings with large complex footprint. Variability in building size is large. Blocks are characterized by low built-up densities and buildings are often located close to block borders, but built front ratios are low. Neighbourhood is sparse, with low accessibility measures. This structure contains various sparsely developed blocks, often *production sites*, *shopping zones*, or *peripheral blocks* of various other structures, where only small part of block is developed.

Fourth urban structure (red in Figures 4.2 and 4.3) detects typical *block building structure* as well as *city core*, as well as several housing estates with block character, i.e. high BFR. Blocks are small and of simple shapes and contain relatively large number of higher buildings with smaller relatively complex footprints. Variance in their shapes and heights is high, but size is rather uniform. Built up density, as well as built front ratio are high. Buildings are often neighbouring with each other and are located close to block borders. Neighbourhood is very dense and has high accessibility measures.

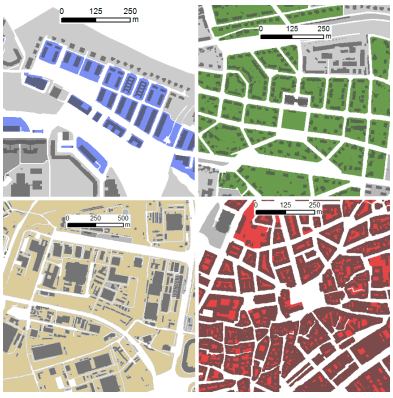


Figure 4.2: Overview of 4 urban structures. Source: author

Figure 4.3: Visualization of 4 urban structures. Source: author

4.1.3 Seven urban structures

First urban structure (light green in Figures 4.4 and 4.5) is characterised with medium sized, medium complex blocks. These contain low buildings with small medium complex footprint. Variance in both sizes and shapes is small. Number of buildings as well as built up density are relatively low. Buildings are separated from each other and are located close to block borders. However, BFR is low. Compared to fourth urban structure, which detects similar kind of urban structure, neighbourhood is sparser and accessibility lower. Peripheral parts of *garden suburbs and villa quarters* are often detected by this structure.

Second urban structure (beige in Figures 4.4 and 4.5) consists of blocks of various sizes and shapes, often larger and relatively complex. They usually contain only a small number of very small and low buildings, uniform in sizes and shapes. Built-up density is very low, and buildings tend to be situated further from block borders. Neighbourhood is sparser and is characterized by lower accessibility. This structure often detects *allotment gardens*, as well as suburban blocks with small number of very small buildings. However, it also often misdetects other structures.

Solitaire objects are detected by **third structure** (yellow in Figures 4.4 and 4.5). Blocks of medium sizes often contain only a small number of large complex buildings, sometimes accompanied by much smaller objects, causing the variability of both sizes and shapes to be very high. Built-up density of such block tends to be high. Neighbourhood is dense and has good accessibility. Third urban structure often comprises *churches* or *sport stadiums*, as well as other blocks with solitaire objects, e.g. a single apartment building.

Fourth urban structure (dark green in Figures 4.4 and 4.5) contains medium size blocks of relatively simple shapes. Buildings tend to be low and have small footprint and similar sizes, but variance in their shape is considerable. Number of buildings in a block is relatively high, as well as built-up density. Buildings are often separated from each other and are located relatively close to block borders. Neighbourhood is dense and has relatively large accessibility measures. Local centres of *garden suburbs and villa quarters* are often detected as this structure

Overall, **fifth urban structure** (light yellow in Figures 4.4 and 4.5) encompasses the largest and most complex and variable blocks. These contain large number of lower buildings with large more complex footprint. Variability in building sizes and shapes is

large. Blocks are characterized by very low built-up densities and buildings are often located close to block borders, but built front ratios are low. Neighbourhood is sparse, with low accessibility measures. This structure contains various sparsely developed blocks, often *production sites*, *shopping zones*, or *peripheral blocks* of various other structures, where only small part of block is developed.

In sixth urban structure (blue in Figures 4.4 and 4.5), smaller to medium blocks of quite complex shapes contain smaller number of higher buildings of comparatively large footprint with relatively complex shapes. Variance in both size and height of buildings is high, however, shape tends to be uniform. Built up density of blocks is relatively large and buildings are located at uniform distances from block borders. Neighbourhood is denser with larger accessibility. This structure very well detects *housing estates*, as well as *smaller production sites*.

Seventh urban structure (red in Figures 4.4 and 4.5) detects typical *block building structure* as well as *city core*. Blocks are small and of simple shapes and contain relatively large number of higher buildings with medium sized relatively complex footprints. Variance in their shapes and heights is high, but size is rather uniform. Built up density, as well as built front ratio are very high. Buildings are often neighbouring with each other and are located close to block borders. Neighbourhood is very dense and has good accessibility.

Figure 4.4: Overview of 7 urban structures. Source: author

Figure 4.5: Visualization of 7 urban structures. Source: author

4.1.4 Eight urban structures

First urban structure (light green in Figures 4.6 and 4.7) is characterised with medium sized, medium complex blocks, often of convex shapes. These contain low buildings with small, medium complex footprint. Variance in both sizes and shapes is small. Number of buildings as well as built up density are relatively low. Buildings are separated from each other and are located close to block borders. However, BFR is low. Compared to fourth urban structure, which detects similar kind of urban structure, neighbourhood is sparser and accessibility lower. Peripheral parts of *garden suburbs and villa quarters* are often detected by this structure.

Second urban structure (beige in Figures 4.6 and 4.7) consists of blocks of various sizes and shapes, often larger and relatively complex. They usually contain only a small number of very small and low buildings, uniform in sizes and shapes. Built-up density is very low, and buildings tend to be situated further from block borders. Neighbourhood is sparser and is characterized by lower accessibility. This structure often detects *allotment gardens*, as well as larger blocks with small number of very small buildings. However, it also sometimes misdetects other structures.

Solitaire objects are detected by **third structure** (yellow in Figures 4.6 and 4.7). Blocks of various sizes often contain only a small number of large complex buildings, sometimes accompanied by a smaller less significant object. Built-up density of such block tends to be high. Neighbourhood is dense and has good accessibility. Third urban structure often comprises *churches* or *museums*, as well as other blocks with solitaire objects, e.g. a single apartment building or office building.

Fourth urban structure (dark green in Figures 4.6 and 4.7) contains medium size blocks of relatively simple, convex or circular shapes. Buildings tend to be low and have small footprint and similar sizes, but variance in their shape is considerable. Number of buildings in a block is relatively high, as well as built-up density. Buildings are often separated from each other and are located relatively close to block borders. Neighbourhood is dense and has relatively large accessibility measures. Local centres of*garden suburbs and villa quarters* are often detected as this structure

Overall, **fifth urban structure** (light yellow in Figures 4.6 and 4.7) consists of the largest and most complex and variable blocks. These contain large number of lower buildings with large relatively complex footprint. Variability in building sizes and shapes is large. Blocks are characterized by very low built-up densities and buildings are often located close to block borders, but built front ratios are low. Neighbourhood is very sparse, with relatively low accessibility measures. This structure contains various sparsely developed blocks, often *production sites*, *shopping zones*, *the airport* or *peripheral blocks* of various other structures, where only small part of block is developed.

In sixth urban structure (dark blue in Figures 4.6 and 4.7), smaller to medium blocks of relatively complex shapes, often with circular convex footprints, contain smaller number of higher buildings with medium size footprint and relatively simple shapes. Variance in both size and height of buildings is larger, however, shape tends to be uniform. Built up density of blocks is relatively large and buildings are often located at uniform

distances from block borders. Neighbourhood is denser with long accessible network and large accessible floorspace. This structure very well detects *housing estates* with uniform elongated building compositions.

Seventh urban structure (light blue in Figures 4.6 and 4.7) comprises medium sized blocks of relatively complex shapes, which contain various numbers of relatively low buildings with large complex footprint. Variances in sizes and shapes are relatively high. Buildings are located closer to block borders and built-up density, as well as BFR, is relatively high. Neighbourhood is moderately dense and has relatively high accessibility. *Sport stadiums*, *shopping zones* and *industrial zones* are often detected as seventh urban structure.

Eight urban structure (red in Figures 4.6 and 4.7) detects typical *block building structure* as well as *city core*. Blocks are small and of simple convex, circular or rectangular shapes and contain relatively large number of higher buildings with medium sized relatively complex footprints. Variance in their shapes and heights is high, but size is rather uniform. Built up density, as well as built front ratio are very high. Buildings are often neighbouring with each other and are located close to block borders. Neighbourhood is very dense and has very good accessibility.

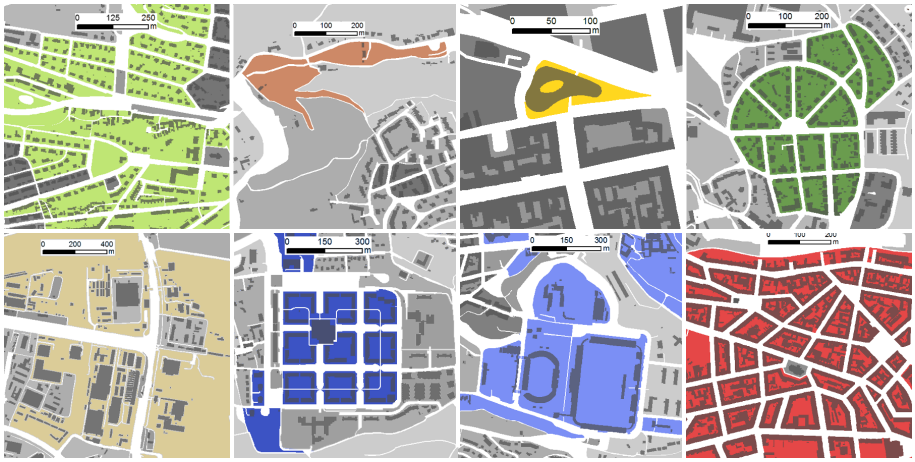


Figure 4.6: Overview of 8 urban structures. Source: author

Figure 4.7: Visualization of 8 urban structures. Source: author

4.1.5 Twelve urban structures

In **first urban structure** (green in Figures 4.8 and 4.9), medium to large blocks are comprised, containing various numbers of low buildings with small, relatively complex footprint. Variance in both sizes and shapes is moderate. Blocks are sparsely built-up and buildings are located further from block borders. Neighbourhood is sparse as well, and accessibility is low. First urban structure usually borders with fourth urban structure, detecting peripheral parts of *garden suburbs and villa quarters*, as well as *allotment gardens*.

Second urban structure (beige in Figures 4.8 and 4.9) encompasses blocks of various sizes and shapes, often larger and relatively complex. They usually contain a small number of very small and low buildings, uniform in sizes and shapes. Built-up density is very low, and buildings tend to be situated further from block borders. Neighbourhood is sparser and is characterized by lower accessibility. This structure often detects *allotment gardens*, as well as larger blocks with small number of very small buildings. However, it also sometimes misdetects other structures.

Solitaire objects are detected by **third structure** (yellow in Figures 4.8 and 4.9). Blocks of various sizes often contain only a small number of large complex buildings, sometimes accompanied by a smaller less significant object. Built-up density of such block tends to be high. Neighbourhood is dense and has good accessibility. Third urban structure often comprises *churches* or *museums*, as well as other blocks with solitaire objects, e.g. a single apartment building or office building.

Fourth urban structure (light green in Figures 4.8 and 4.9) is characterised with medium sized, medium complex blocks. These contain low buildings with small, medium complex footprint. Variance in both sizes and shapes is small. Number of buildings as well as built up density are relatively low. Buildings are separated from each other and are located close to block borders. However, BFR is low. Compared to fifth urban structure, which detects similar kind of urban structure, neighbourhood is sparser and accessibility lower. Less accessible parts of *garden suburbs and villa quarters* are often detected by this structure.

Fifth urban structure (dark green in Figures 4.8 and 4.9) contains medium size blocks of relatively simple shapes. Buildings tend to be low and have small footprint and similar sizes, but variance in their shape is considerable. Number of buildings in a block is relatively high, as well as built-up density. Buildings are often separated from each other and are located relatively close to block borders. Neighbourhood is dense and has relatively large accessibility measures. Local centres of*garden suburbs and villa quarters* are often detected as this structure

Overall, **sixth urban structure** (aquamarine in Figures 4.8 and 4.9) encompasses large and complex blocks, varying in sizes and shapes. These contain large number of higher buildings with relatively large and complex footprints. Variability in building sizes, shapes and heights is large. Blocks are characterized by low built-up densities and buildings are often located close to block borders, but built front ratios are low. Neighbour-

hood is sparse, with relatively low accessibility measures. This structure consists of various sparsely developed blocks, more centrally located than blocks of seventh structure. It often detects *production sites*, *shopping zones* or *peripheral blocks* of various other structures, where only small part of block is developed.

Compared to sixth urban structure, measures of **seventh urban structure** (light yellow in Figures 4.8 and 4.9) are more variable and overall more extreme. It consists of the largest blocks, which are however often more simple than blocks of sixth structure. Blocks contain large number of buildings of various shapes and sizes. Blocks are also characterized by very low built-up densities and buildings are often located close to block borders, but built front ratios are low. Neighbourhood is very sparse, but compared to sixth structure its accessibility is lower. This structure detects *peripheral blocks* with low densities, such as the airport, as well as several *industrial zones* and *peripheral blocks* of various other structures, where only small part of block is developed.

Eight building structure (light blue in Figures 4.8 and 4.9) consists of medium sized moderately complex blocks, containing high buildings with large, relatively simple footprints. Variance of building size and shapes is relatively high. Built-up density of blocks is higher, as well as number of buildings. Buildings are located close to block borders. Neighbourhood is slightly sparser, with lower accessibility measures. This structure detects mostly *housing estates with block character*, i.e. relatively closed block border. Sometimes, *block building structure* is detected as well.

In **ninth urban structure** (dark blue in Figures 4.8 and 4.9), smaller to medium blocks of relatively complex shapes, often with circular convex footprints, contain smaller number of high buildings with medium size footprint and relatively simple shapes. Both size and shape of buildings ten to be uniform, however, variance of height is larger. Built up density of blocks is relatively large and buildings are often located at uniform distances from block borders. Neighbourhood is denser with good accessibility values. This structure very well detects *housing estates* with high, uniform and elongated building compositions.

Tenth building structure (pink in Figures 4.8 and 4.9) comprises medium sized blocks of relatively simple shapes, which contain lower number of low buildings with large, complex footprints. Variance in size and shape is large. Buildings are located further from block borders and overall built-up density is relatively large. The accessibility is relatively good. Various complex structures are detected by tenth structure, such as parts of *housing estates* or *industrial zones*.

Eleventh urban structure (medium blue in Figures 4.8 and 4.9) comprises small to medium sized blocks of relatively complex shapes, which contain smaller numbers of relatively low buildings with large footprint. Variances in sizes and shapes are relatively high. Buildings are located closer to block borders and built-up density, as well as BFR, is relatively high. Neighbourhood is relatively dense and has high accessibility. *Sport stadiums*, *shopping malls and stores* or *industrial zones* are often detected as eleventh urban structure.

Twelfth urban structure (red in Figures 4.8 and 4.9) detects typical *block building structure* as well as *city core*. Blocks are small and of simple shapes and contain relatively

large number of higher buildings with medium sized relatively complex footprints. Variance in their shapes and heights is high, but size is rather uniform. Built up density, as well as built front ratio are very high. Buildings are often neighbouring with each other and are located close to block borders. Neighbourhood is very dense and has very good accessibility.

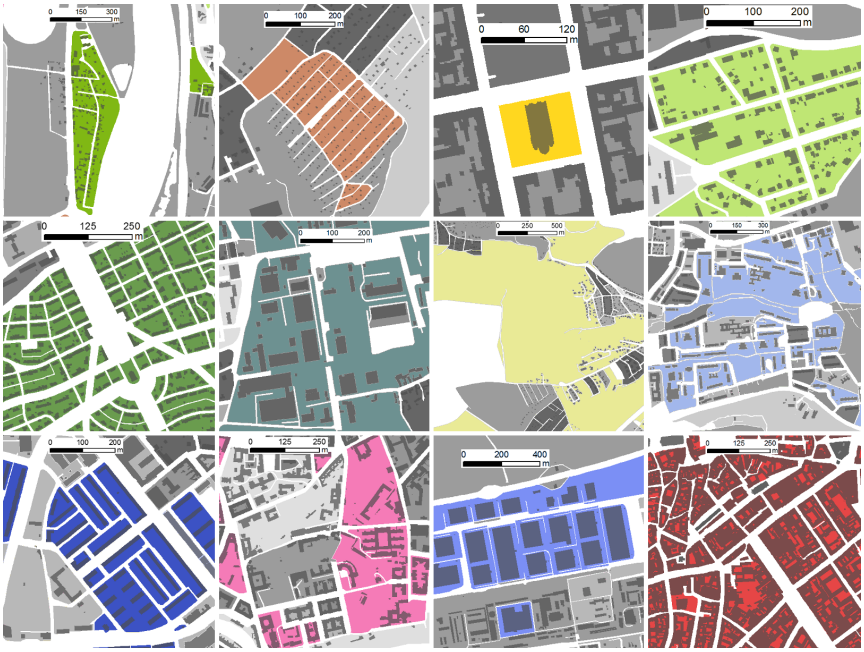


Figure 4.8: Overview of 12 urban structures. Source: author

Figure 4.9: Visualization of 12 urban structures. Source: author

4.2 Conclusion

This thesis proposed a methodology for assessing urban structures which is *fast*, *objective*, *easy to interpret*, *possible to extend and actualize* and is *not software dependent*. A pilot study of automatic classification of urban structures using the proposed methodology was carried out in the study area of Prague, Czech Republic.

An overview and description of *63 morphological attributes for scale of buildings and 144 morphological attributes for scale of blocks* was provided and computed in this study area. These attributes express various characteristics of urban structures, do not rely on any previous definition of locations and are therefore applicable regardless of the analysed city.

An *automatic classification technique* employing these morphological attributes was proposed, based on the methods of unsupervised machine learning – principal component analysis and clustering. A clustering method applying weighted sum of attributes in dissimilarity matrix was proposed for incorporating the spatial component together with non-spatial attributes. The proposed classification technique is faster compared to the previous deriving of spatial structures and allows to easily utilize other morphological attributes or capture future changes in urban structure.

Urban structures in Prague were assessed using principal component analysis and K-medoids clustering. The result of the proposed methodology is an objective *set of 4, 7, 8 and 12 distinct urban structures*, which are not biased by any previous knowledge about Prague or subjective perception.

The resulting clusterings succeed at differentiating between various urban structures, and equivalents may be found in the existing set of urban structures. However, misclassified blocks appear at times, and low average silhouette widths of individual clusterings prove this problem. This might for example imply an inappropriate clustering method. Overall, a small number of clusters leads to a more stable result, however, its contribution is rather low. A larger number of clusters on the other hand leads to more similar urban structures and various characteristics might overlap. A balance needs to be found between the stability of results and information value.

The results of this thesis will be beneficial as an alternative definition of urban structures for the Metropolitan plan of Prague. Urban structures represent a base stone for definition of various policies and regulations for urban development, because they capture and define local character, which should be respected or cultivated, rather then disrupted.

All morphological attributes, principal components and final urban structures are visualized as an ArcGIS Online map [1].

Although all objectives of the thesis were accomplished, the topic of automatic classification of urban morphology was by no means exhausted. A large space for further extension remains. During the processing of the thesis, various problems and questions arose that deserve an in-depth standalone research. Although the methodology aims to

[1] http://arcg.is/2i0ygLW

be as objective as possible, several steps are still influenced by human decision making. A sensitivity analysis should be performed in order to assess the influence of these input parameters on results. Namely, the author would like to explore the effect of changes in following the parameters and areas:

- granularity, scale and definition of buildings,
- definition of blocks,
- adding or deleting various morphological attributes,
- functions utilized in summarizing building attributes in scale of blocks,
- number of principal components used in clustering,
- assessing optimal number of clusters,
- various clustering techniques, including those without predefined number of clusters,
- including the spatial component in the case study of Prague.

The algorithms proposed for calculating more complex morphological attributes would deserve an optimization in order to be less computationally demanding. Last but not least, the suggested approach results in urban structures capturing solely the morphology of urban tissues. However, it can be easily extended with various other attributes, such as land use or socio-economic information, or even other morphological attributes.

Bibliography

[1] ArcGIS Online. *Esri*. [online]. [cit. 2017-01-07]. URL https://www.arcgis.com/home/index.html.

[2] ArcGIS Platform. *Esri*. [online]. [cit. 2016-12-26] URL http://www.esri.com/software/arcgis.

[3] BANZHAF E.; HÖFER R. Monitoring urban structure types as spatial indicators with CIR aerial photographs for a more effective urban environmental management. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. July 2008, vol. 1, no. 2, p. 129-138.

[4] BATTY, M. Exploring isovist fields: space and shape in architectural and urban morphology. *Environment and Planning B: Planning and Design*. 2001, vol. 28, p. 123-150.

[5] COLANINNO, N.; CLADERA, J.R.; PFEFFER, K. An automatic classification of urban texture: form and compactness of morphological homogeneous structures in Barcelona. *51st European Congress of the Regional Science Association International*. 2011.

[6] DE VOORDE, T. et al. Quantifying intra-urban morphology of the Greater Dublin area with spatial metrics derived from medium resolution remote sensing data. *Urban Remote Sensing Joint Event*. 2009.

[7] Definition of morphology in English. Oxford Dictionaries. [online]. Oxford University Press, © 2017. [cit. 2017-01-06]. URL https://en.oxforddictionaries.com/definition/morphology.

[8] DIBBLE, J.; PRELORENDJOS, A. et al. Urban Morphometrics: Towards a Science of Urban Evolution. *ISUF International Seminar of Urban Form, Rome, Italy*. September 2015.

[9] EPPSTEIN D. Geometry in Action: Medial axis. *Donald Bren School of Information and Computer Sciences*. [online]. The Regents of the University of California © 2005-2015 .[cit. 2016-12-26]. URL https://www.ics.uci.edu/ eppstein/gina/medial.html.

[10] GIL, J.; BEIRÃO, J.N.; MONTENEGRO, N. et al. On the discovery of urban typologies: data mining the many dimensions of urban form. *Urban Morphology*. 2012, vol. 16, no. 1, p. 27-40.

[11] HANDY, S.; NIEMEIER, A.D. Measuring Accessibility: an exploration of issues and alternatives. *Environment and Planning*. 1997, vol. 29, p. 1175-1194.

[12] HARTMANN, A.; MEINEL, G. HECHT, R. et al. A Workflow for Automatic Quantification of Structure and Dynamic of the German Building Stock Using Official Spatial Data. *International Journal of Geo-Information*. August 2016, vol. 5, no. 8, p. 142.

[13] HEROLD, M.; COUCLELIS, H.; CLARKE, K.C. The Role of Spatial Metrics in the Analysis and Modeling of Urban Land Use Change. *Computers, Environment and Urban Systems*. July 2005, vol. 29, p. 369-399.

[14] HEROLD, M. et al. Spatial Metrics and Image Texture for Mapping Urban Land Use. *Photogrammetric Engineering & Remote Sensing*. September 2003, vol. 69, no. 9, p. 991–1001.

[15] Institut plánování a rozvoje hlavního města Prahy. *Hranice města a lokality: Vymezení základní jednotky plánu, její popis a regulativy*. [online]. 2014. [cit. 2016-12-30]. ISBN 978-80-87931-08-0. URL http://plan.iprpraha.cz/cs/upp-hranice-mesta-a-lokality.

[16] IPR Praha. Územní plán hl. m. Prahy. *Institut plánování a rozvoje hl. m. Prahy*. [online]. [cit. 2016-12-26]. URL http://plan.iprpraha.cz/cs/metropolitni-plan.

[17] JAMES, G.; WITTEN, D. et al. *An Introduction to Statistical Learning: with Applications in R*. 1st. Springer-Verlag New York. 2013. XIV, 426 p. 978-1-4614-7137-0.

[18] JIAO, L.; LIU, Y. Knowledge discovery by spatial clustering based on self-organizing feature map and a composite distance measure. *The International Archieves of Photogrammetry, Remote sensing and Spatial Information Science*. 2008, vol.37, part B2.

[19] Kategorie komunikací. *České silnice a dálnice*.[online]. ceskedalnice.cz © 2002 – 2016. [cit. 2016-12-26]. URL http://www.ceskedalnice.cz/odborne-info/kategorie-komunikaci/.

[20] LASKARI, A. *Urban identity through quantifiable spatial attributes: Coherence and dispersion of local identity through the comparative analysis of building block plans*. London, 2007. Dissertation thesis. Bartlett School of Graduate Studies, University College London.

[21] Lesson 11: Principal Components Analysis (PCA). *STAT 505*. [online]. The Pennsylvania State University © 2016. [cit. 2016-12-26]. URL https://onlinecourses.science.psu.edu/stat505/node/49.

[22] MACEACHREN, A. M. Compactness of geographic shape: comparison and evaluation of measures. *Geografiska Annaler*. 1985, vol. 67B, p. 53-67.

[23] MCGARIGAL, K.; MARKS, B. J. 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Rep. PNW-GTR-351. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 122 p.

[24] NG, A. K-means algorithm. *Coursera*. [online]. Coursera Inc. © 2016 [cit. 2016-12-26]. URL https://www.coursera.org/learn/machine-learning/lecture/93VPG/k-means-algorithm.

[25] NG, R. T.; HAN, J. Efficient and Effective Clustering Methods for Spatial Data Mining. *Proceedings of the 20th VLDB Conference Santiago, Chile*. 1994.

[26] O'ROURKE, J. *Computational geometry in C*. Second edition. Cambridge University Press, 1998.

[27] OLIVEIRA V. *Urban Morphology: An Introduction to the Study of the Physical Form of Cities*. 1st. Springer International Publishing, 2016. XXIII, 192 p. 978-3-319-32081-6.

[28] PAGE, John D. Circumscribed rectangle, or bounding box. (Coordinate Geometry). *Math Open Reference*. [online]. 2011. [cit. 2016-11-23]. URL http://www.mathopenref.com/coordbounds.html.

[29] PORTA, S.; DA FONTOURA COSTA, L.; MORELLO, E. et al. Plot-based urbanism and urban morphometrics: measuring the evolution of blocks, street fronts and plots in cities. *University of Strathclyde*. 2011.

[30] PORTA, S.; STRANO, E. et al. Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and Design*. 2009, vol. 36, p. 450-465.

[31] Practical Guide to Principal Component Analysis (PCA) in R & Python. *Analytics Vidhya*. [online]. Analytics Vidhya © 2016. [cit. 2016-12-26]. URL https://www.analyticsvidhya.com/blog/2016/03/practical-guide-principal-component-analysis-python/.

[32] *R: The R Project for Statistical Computing*. [online]. The R Foundation © 2016. [cit. 2016-12-26]. URL https://www.r-project.org/.

[33] ROUSSEEUW, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*. 1987, vol. 20, p. 53-65.

[34] SAMMUT, C.; WEBB, G. I. (eds.) *Encyclopedia of Machine Learning*. 1st. Springer US, 2010. XXVI, 1031 p. 978-0-387-30164-8.

[35] SANDER J. et al. Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and its Applications. *Data Mining and Knowledge Discovery*. 1998.

[36] SERNA, A.; MARCOTEGUI, B. Detection, segmentation and classification of 3D urban objects using mathematical morphology and supervised learning. *ISPRS Journal of Photogrammetry and Remote Sensing*. July 2014, vol. 93, p. 243–255.

[37] SCHIRMER, P. M.; AXHAUSEN, K.W. A Multiscale Classification of Urban Morphology. *Journal of Transport and Land Use*. 2016, vol. 9, no.1.

[38] SCHOLZ M. PCA. Principal Component Analysis. *Nonlinear PCA*. [online]. [cit. 2016-12-26]. URL http://www.nlpca.org/pca_principal_component_analysis.html.

[39] Spacemate, Spacematrix. *Urban Knowledge*. [online]. Urban Knowledge © 2014. [cit. 2016-11-23]. URL http://www.urban-knowledge.nl/3/spacemate-spacematrix.

[40] SPECTOR P. Performing and Interpreting Cluster Analysis. *University of California, Berkeley Department of Statistics*. [online]. 2007. [cit. 2016-12-30]. http://www.stat.berkeley.edu/ spector/s133/Clus.html.

[41] STEINIGER, S.; LANGE, T. et al. An Approach for the Classification of Urban Building Structures Based on Discriminant Analysis Techniques. *Transactions in GIS*. 2008, vol. 12, no. 1, p. 31-59.

[42] Sunrise and sunset times in Prague, June 2016. *Timeanddate.com* [online]. Stavanger, Norway: Time and Date AS, © 1995-2016 [cit. 2016-11-23]. URL https://www.timeanddate.com/sun/czech-republic/prague?month=6&year=2016.

[43] TAUBENBÖCK, H.; WURM, M. et al. Integrating Remote Sensing and Social Science – The correlation of urban morphology with socioeconomic parameters. *Urban Remote Sensing Joint Event, Shanghai, China*. 2009.

[44] TAUBENBÖCK, H. et al. An Urban classification approach based on an object-oriented analysis of high resolution satellite imagery for a spatial structuring within urban areas. *1st EARSeL Workshop of the SIG Urban Remote Sensing Humboldt-Universität zu Berlin*. March 2006.

[45] THOMAS, I. et al. Clustering patterns urban built-up areas with curves of fractal scaling behavior. *Environment and Planning B: Planning and Design*. 2010, vol. 37, p. 942–954.

[46] Urban Network Analysis Toolbox for ArcGIS. *City Form Lab*. [online]. © City Form Lab [cit. 2016-12-26]. URL http://cityform.mit.edu/projects/urban-network-analysis.

[47] VAN KREVELD M. Computational geometry course. *Department of Information and Computing Sciences, University of Utrecht*. [online]. 2016. [cit. 2016-12-26]. URL http://www.cs.uu.nl/docs/vakken/ga/slides1.pdf.

[48] VENERANDI, A.; ZANELLA, M. et al. The Form of gentrification: Common morphological patterns in five gentrified areas of London, UK. *Working paper*. 2014.

[49] VENKATASUBRAMANIAN S. Convex hulls. *University of Utah, School of computing* [online]. 2013. [cit. 2016-12-26]. URL http://www.cs.utah.edu/ suresh/compgeom/convexhulls.pdf.

[50] WANG, X.; HAMILTON, H. J. DBRS: A Density-Based Spatial Clustering Method with Random Sampling. *7th Pacific-Asia Conference*. PAKDD 2003, Seoul, Korea, April 30 – May 2, 2003 Proceedings.

[51] WANG X.; WANG J. Using clustering methods in geospatial information systems. *GEOMATICA*. 2010, vol. 64, no. 3, p. 347-361.

[52] Ways to determine the number of principal components. *Minitab*. [online]. Minitab Inc. © 2016. [cit. 2016-12-30]. http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/multivariate/principal-components-and-factor-analysis/number-of-principal-components/.

[53] What is ArcPy? *ArcGIS for Desktop*. [online]. Environmental Systems Research Institute, Inc. © 2016. [cit. 2016-12-26]. URL http://pro.arcgis.com/en/pro-app/arcpy/get-started/what-is-arcpy-.htm.

[54] YU, B.; LIU H.; WU, J. et al. Automated derivation of urban building density information using airborne LIDAR data and object-based method. *Landscape and Urban Planning*. December 2010, vol. 98, no. 3, p. 210-219.

[55] ZHANG, J. et al. Polygon-Based Spatial Clustering. *Proceedings in 8th International Conference on GeoComputation*. August 2005.

Abbreviations

BB	bounding box
BFR	built front ratio
CH	convex hull
CO	courtyard
EC	enclosing circle
FSI	floor space index
GIS	geographic information system
GSI	gross space index
IPR	Prague Institute of Planning and Development
IZ	influence zone
OSC	open space closeness
OSR	open space ratio
PC	principal component
PCA	principal component analysis
PVE	proportion of variance explained
UNA	urban network analysis

List of Figures

1 Map of structural diagnosis in Prague. Source: IPR Praha [15] 2

2 Urban sub-structures in Prague. Source: IPR Praha [15] 2

1.1 Gravelius index. Source: author 9

1.2 Number of building corners. Source: author 9

1.3 Standard deviation of distance of building corners to centroid. Source: author 10

1.4 Building squareness. Source: author 10

1.5 Topological skeleton of non-generalized buildings. Source: author 11

1.6 Building courtyards. Source: author 12

1.7 Elongation. Source: author 13

1.8 Ratio of building footprint area to bounding box area. Source: author . . . 13

1.9 Ratio of building footprint perimeter to bounding box perimeter. Source: author 14

1.10 Ratio of building footprint area to enclosing circle area. Source: author . . 15

1.11 Ratio of building footprint area to convex hull area. Source: author 15

1.12 Ratio of building footprint perimeter to convex hull perimeter. Source: author 16

1.13 Schumm's longest axis to area ratio. Source: author 16

1.14 Influence zones. Source: author 17

1.15 Ratio of building footprint area to influence zone area. Source: author . . . 17

1.16 Distance to closest neighbouring building. Source: author 18

1.17 Number of neighbouring buildings. Source: author 18

1.18 Distance to block border. Source: author 19

1.19 Reach. There are 20 buildings reachable from i. Source: [46] 20

1.20 Betweenness. Two out of three shortest paths lead through i. Source: author 21

1.21 Straightness. Comparison of $\delta[i, j]$ and $d[i, j]$. Source: author 21

1.22 Issues in street facing façades detection – Figure **(a)** illustrates drawback of buffer method suggested e.g. by Schirmer and Axhausen [37]. In the upper Figure, façade at obtuse angle, façade situated on opposite side of building and façade hidden behind another object is detected. The corrected solution is presented in the bottom figure. Possible solution is utilization of isovists illustrated in Figure **(b)** . Source: **(a)** author, **(b)** Batty [4]. 24

1.23 Two definitions of block closeness - 1.23a illustrates closeness of block as defined by Schirmer and Axhausen [37]. Length of street facing façades is 230 m, length of block border is 159 m, resulting in closeness value of 1.45 (limit distance is 5 m). 1.23b illustrates built front ratio as defined by Porta et al. [29], Venerandi et al. [48] and Dibble et al. [8]. Street facing façades in certain limit distance is projected on block border. Source: author 25

1.24 Closed courtyards. Red courtyards belong to single building, blue are bounded by several different buildings. Light gray spaces are not bounded and therefore do not represent courtyards. Source: author 25

1.25 Difference between *Built front ratio* and *Closeness of open space*. Whilst BFR is relatively high (0.87) – building situated close to block border is making the block perceived closed from 87 % – OSC is lower (0.48), because open spaces are bounded only from inner side and remain open towards the street. Source: author 26

1.26 Borders ratio. Source: author 26

1.27 Permeability of urban block. 16 axial views are found by Thiessen tessellation. Gap between two buildings in upper left corner is not considered wide enough. Source: author 28

1.28 Principle of gross space index. Source: Urban Knowledge [39] 30

1.29 Principle of floor space index. Source: Urban Knowledge [39] 31

1.30 Principle of open space ratio. Source: Urban Knowledge [39] 31

2.1 Transformation of a 3-dimensional data set to 2 dimensions using PCA. Source: Scholz [38] 35

2.2 Data with an outlier and cluster centroids obtained by K-means (blue) and K-medoids (red). Source: author 38

2.3 Outlier in otherwise heterogeneous zone. Source: author 39

3.1 Schema of methodology workflow. Source: author 40

3.2 Schema of utilized data sources, datasets and derived feature classes. Source: author 41

3.3 Eliminating holes inside buildings. Source: author 42

3.4 Eliminating holes between buildings. Source: author 42

3.5 Reasons for incorporating pedestrian zones as block borders. Source: author 44

3.6 Number of junctions and network length with service are of a building. Source: author 52

3.7 Façade perceived from a street is obtained by clipping all façade with inner block buffer (blue). Afterwards an outside buffer (red) is created for each suspect segment and only those segments with buffer intersecting block border are retained (green). Source: author 54

3.8 Principle of *Polyline.queryPointAndDistance(in_point)*: l – Polyline, C – in_point, P - orthogonal projection of C to l, F – l.firstPoint, d – distance from F to P, o – distance from C to P. Source: author 56

3.9 Summary of principal component analysis. Source: author 60

3.10 Dependency between silhouette width and number of clusters for various numbers of principal components. Source: author 63

4.1 Summaries of results. Source: author 66

4.2 Overview of 4 urban structures. Source: author 67

4.3 Visualization of 4 urban structures. Source: author 68

4.4 Overview of 7 urban structures. Source: author 69

4.5 Visualization of 7 urban structures. Source: author 69

4.6 Overview of 8 urban structures. Source: author 71

4.7 Visualization of 8 urban structures. Source: author 71

4.8 Overview of 12 urban structures. Source: author 74

4.9 Visualization of 12 urban structures. Source: author 74

C.1 BB_ratio_A of individual buildings and summarized for blocks. Source: author 94

C.2 Visualization selected of morphological attributes of blocks (1). Source: author 95

C.3 Visualization selected of morphological attributes of blocks (2). Source: author 96

C.4 Visualization selected of morphological attributes of blocks (3). Source: author 97

C.5 Visualization selected of morphological attributes of buildings. Source: author 98

D.1 Visualization of 1st to 4th principal component. Source: author 99

D.2 Visualization of 5th to 9th principal component. Source: author 100

D.3 Interpretation of first nine principal component. Attributes are grouped by positive (+) or negative (-) loading, ordered by loading values and coloured based on attribute character (see legend in bottom right corner). Source: author 101

List of Tables

1.1 Summarization of building attributes. (R. – ratio) 29

3.1 Utilized features of Current Land Use dataset 44

3.2 Utilized features of Road segment feature class 45

3.3 Approximate street widths 45

3.4 Features of Road segment feature class 46

3.5 Angle classification 48

3.6 Various decision criteria for selecting an optimal number of principal components. 59

3.7 Summary of principal component analysis. Bold principal components and statistics mark limit values for various criteria. 60

B.1 Morphological attributes of buildings (1) 89

B.2 Morphological attributes of buildings (2) 90

B.3 Morphological attributes of blocks (1) 91

B.4 Morphological attributes of blocks (2) 92

B.5 Morphological attributes of blocks (3) 93

Appendix A

Contents of attached CD

SRC	folder of ArcPy source codes for UrbanMorphology.tbx
UrbanMorphology.gdb	ArcGIS geodatabase
buildings	polygon feature class of Buildings
buldingCompositions	polygon feature class of Building compositions
blocks	polygon feature class of Blocks
UrbanMorphology.mxd	ArcMap project
UrbanMorphology.tbx	toolbox for computation of morphological attributes

Appendix B

Morphological attributes – summary

Name	Description	Unit	Min	Max	Average
Shape_Area	area of building footprint	[m^2]	1.2	89 290.3	184.6
Shape_Length	perimeter of building footprint	[m]	4.5	3 025.3	50.0
Floorspace	floorspace	[m^2]	1.2	117 250.3	489.2
Approx_height	approximate height	[m]	3.0	84.5	5.5
POCET_PODLAZI	number of floors	[-]	0	28	1
Volume	volume	[m^3]	3.7	454 344.0	1540.2
Fract_dim	fractal dimension	[m]	0.277	59.954	2.440
Gravelius_i	Gravelius index	[-]	1.000	4.581	1.248
Corner_no	number of corners	[-]	0	365	7
Corner_square	squareness of walls	[°]	0.000	44.924	2.466
Corner_dist	st. dev. of distance of corners to centroid	[m]	0.000	0.452	0.092
CO_number	number of courtyards	[-]	0	4	0.004
CO_area	total courtyard area	[m^2]	0.0	10 292.1	1.0
CO_ratio	ratio of CO_area to Shape_Area	[-]	0.000	0.480	0.000
BB_area	area of bounding box	[m^2]	1.3	329 819.9	228.2
BB_perimeter	perimeter of bounding box	[m]	4.6	2 326.1	49.4
BB_width	width of bounding box	[m]	1.0	490.2	9.5
BB_length	length of bounding box	[m]	1.2	672.9	15.2
BB_elongation	elongation	[-]	0.025	1.000	0.671
BB_ratio_A	ratio of BB_area to Shape_Area	[-]	0.122	1.000	0.903
BB_ratio_P	ratio of BB_perimeter to Shape_Length	[-]	0.786	2.502	0.998
EC_area	area of enclosing circle	[m^2]	2.1	471 514.0	427.4
EC_perimeter	perimeter of enclosing circle	[m]	5.1	2434.2	55.8
EC_ratio_A	ratio of EC_area to Shape_Area	[-]	0.027	0.997	0.523
CH_area	area of convex hull	[m^2]	1.2	244 611.8	206.9
CH_perimeter	perimeter of convex hull	[m]	4.5	1 988.0	47.1
CH_length	length of convex hull	[m]	1.6	768.7	17.7
CH_ratio_A	ratio of CH_area to Shape_Area	[-]	0.211	1.000	0.954
CH_ratio_P	ratio of CH_perimeter to Shape_Length	[-]	1.000	2.648	1.039
CH_schumm_ratio	Schumm's longest axis to area ratio	[-]	0.164	0.999	0.720
IZ_area	area of influence zone	[m^2]	2.0	1 066 023.8	2 382.1
IZ_ratio	ratio of Shape_area to IZ_area	[-]	0.000	1.137	0.199
neigh_dist	distance to closest building	[m]	0.0	1 797.1	3.4
neigh_count	number of neighbouring buildings	[-]	0	13	1
neigh_exists	existence of neighbouring buildings	[0-1]	0	1	0.6
border_dist	distance to urban block border	[m]	0.0	558.5	8.4
UNA_100_reach	reach in 100 m radius	[-]	0	137	24
UNA_300_reach	reach in 300 m radius	[-]	0	536	129
UNA_500_reach	reach in 500 m radius	[-]	0	1248	286
UNA_100_gravity	gravity in 100 m radius	[-]	0.000	122.923	21.181
UNA_300_gravity	gravity in 300 m radius	[-]	0.000	345.096	88.215
UNA_500_gravity	gravity in 500 m radius	[-]	0.000	605.001	154.049

Table B.1: Morphological attributes of buildings (1)

Name	Description	Unit	Min	Max	Average
UNA_100_betweenness	betweenness in 100 m radius	[-]	0	5 018	155
UNA_300_betweenness	betweenness in 300 m radius	[-]	0	26 394	2 332
UNA_500_betweenness	betweenness in 500 m radius	[-]	0	95 922	8 050
UNA_100_closeness	closeness in 100 m radius	[-]	0.000	334.447	0.006
UNA_300_closeness	closeness in 300 m radius	[-]	0.000	5.958	0.000
UNA_500_closeness	closeness in 500 m radius	[-]	0.000	334.447	0.005
UNA_100_straightness	straightness in 100 m radius	[-]	0.000	1 719.531	21.621
UNA_300_straightness	straightness in 300 m radius	[-]	0.000	1 778.828	99.850
UNA_500_straightness	straightness in 500 m radius	[-]	0.000	1 827.170	211.747
UNA_100_floorspace	floorspace in 100 m radius	[m^2]	0.0	111 004.3	6 845.5
UNA_300_floorspace	floorspace in 300 m radius	[m^2]	0.0	464 042.6	46 405.5
UNA_500_floorspace	floorspace in 500 m radius	[m^2]	0.0	1 165 298.2	189 896.8
UNA_100_footprint	footprint area in 100 m radius	[m^2]	0.0	90 381.5	2 836.6
UNA_300_footprint	footprint area in 300 m radius	[m^2]	0.0	110 911.9	18 079.6
UNA_500_footprint	footprint area in 500 m radius	[m^2]	0.0	289 621.7	45 732.2
UNA_100_cross_no	number of crossroads in 100 m radius	[-]	0	28	3
UNA_300_cross_no	number of crossroads in 300 m radius	[-]	0	85	22
UNA_500_cross_no	number of crossroads in 500 m radius	[-]	0	207	58
UNA_100_net_length	length of network in 100 m radius	[m]	0.0	1 403.9	296.4
UNA_300_net_length	length of network in 300 m radius	[m]	21.0	7 364.2	1 944.8
UNA_500_net_length	length of network in 500 m radius	[m]	21.0	17 537.7	5 001.7

Table B.2: Morphological attributes of buildings (2)

Name	Description	Unit	Min	Max	Avg.
Shape_Area	area of block footprint	[m^2]	5.0	7 652 789.5	36 398.7
Shape_Length	perimeter of block footprint	[m]	8.6	35 548.5	894.2
Fract_dim	fractal dimension	[m]	0.206	255.689	22.848
Gravelius_i	Gravelius index	[-]	1.002	7.904	1.587
Corner_no	number of corners	[-]	0	925	38
Corner_dist	st. dev. of distance of corners to centroid	[m]	0.000	0.433	0.163
Corner_square	squareness of borders	[°]	0.001	42.597	5.743
BB_area	area of bounding box	[m^2]	5.2	17 391 326.2	64 742.9
BB_perimeter	perimeter of bounding box	[m]	9.5	16 732.3	739.4
BB_width	width of bounding box	[m]	1.5	3 856.4	112.0
BB_length	length of bounding box	[m]	2.4	4 509.8	249.7
BB_elongation	elongation	[-]	0.020	1.000	0.541
BB_ratio_A	ratio of BB_area to Shape_Area	[-]	0.055	1.000	0.736
BB_ratio_P	ratio of BB_perimeter to Shape_Length	[-]	0.789	3.802	1.044
EC_area	area of enclosing circle	[m^2]	6.6	22 927 559.3	125 488.7
EC_perimeter	perimeter of enclosing circle	[m]	9.1	16 974.0	837.4
EC_ratio_A	ratio of EC_area to Shape_Area	[-]	0.007	0.977	0.403
CH_area	area of convex hull	[m^2]	5.0	11 692 127.5	47 648.2
CH_perimeter	perimeter of convex hull	[m]	8.6	13 797.5	651.3
CH_length	length of convex hull	[m]	2.9	5 368.1	265.5
CH_ratio_A	ratio of CH_area to Shape_Area	[-]	0.083	1.000	0.894
CH_ratio_P	ratio of CH_perimeter to Shape_Length	[-]	1.000	4.380	1.165
CH_schumm_ratio	Schumm's longest axis to area ratio	[m]	0.085	0.991	0.623
Facade_length	length of façade perceived from the street	[m]	0.0	3 885.0	221.1
Facade_percent	percentage of façade perceived from the street	[-]	0	1	0.353
BFR_length	length of border lined by buildings	[m]	0.0	3 925.1	207.9
BFR_ratio	built front ratio	[-]	0.000	1.000	0.389
CO_number	number of closed courtyards	[-]	0	55	0.3
Permeability	permeability	[-]	0	607	20
Build_no	number of buildings	[-]	1	388	20

Name	Description	Unit	Min	Max	Avg.
GSI	gross space index	[-]	0.000	1.000	0.279
FSI	floor space index	[-]	0.000	12.438	0.738
OSR	open space ratio	[-]	0.000	102.203	242 235.406
Street_width_avg	avg. width of neighbouring streets	[m]	0.0	188.0	12.8
Street_width_min	minimum width of neighbouring streets	[m]	0.0	188.0	8.4
Build_Area_MAX	maximum building Shape_Area	[m^2]	2.7	89 290.3	856.9
Build_Area_SUM	sum of building Shape_Area	[m^2]	2.7	246 473.3	3 682.6
Build_Area_AVG	avg. building Shape_Area	[m^2]	2.7	25 903.4	310.1
Build_Area_STD	st. dev. of building Shape_Area	[m^2]	0.0	36 608.1	224.8
Build_Length_SUM	sum of building Shape_Length	[m]	6.5	30 253.4	997.5
Build_Length_AVG	avg. building Shape_Length	[m]	6.5	881.2	61.9
Build_Length_STD	st. dev. of building Shape_Length	[m]	0.0	806.9	26.6
Floorspace_SUM	sum of building Floorspace	[m^2]	2.7	377 783.7	9 762.1
Floorspace_AVG	avg. building Floorspace	[m^2]	2.7	70 294.0	870.8
height_MAX	maximum building Approx_height	[m]	3.0	84.5	9.6
height_MIN	minimum building Approx_height	[m]	3.0	63.5	4.1
height_AVG	avg. building Approx_height	[m]	3.0	63.5	6.3
height_STD	st. dev. of building Approx_height	[m]	0.0	48.8	2.0
floor_no_MAX	maximum of floor_no	[-]	0	28	2.7
floor_no_MIN	minimum of floor_no	[-]	0	21	0.5
floor_no_SUM	sum of floor_no	[-]	0	1 538	22.4
floor_no_AVG	avg. floor_no	[-]	0.0	21.0	1.4
floor_no_STD	st. dev. of floor_no	[-]	0.0	16.3	0.8
volume_MAX	maximum building volume	[m^3]	8.0	454 344.0	7 637.8
volume_AVG	avg. building volume	[m^3]	8.0	215 275.2	2 762.1
volume_STD	st. dev. of building volume	[m^3]	8.0	250 390.6	2 061.8
fract_dim_AVG	avg. building fract_dim	[m]	0.407	28.611	2.877
fract_dim_STD	st. dev. of building fract_dim	[m]	0.000	30.900	0.989
gravelius_i_AVG	avg. building gravelius_i	[-]	1.003	3.088	1.261
gravelius_i_STD	st. dev. of building gravelius_i	[-]	0.000	1.282	0.108

Table B.3: Morphological attributes of blocks (1)

Name	Description	Unit	Min	Max	Avg.	Name	Description	Unit	Min	Max	Avg.
corner_no_AVG	avg. building corner_no	[-]	2.0	182.0	8.4	EC_ratio_A_AVG	avg. building EC_ratio	[-]	0.094	0.977	0.520
corner_no_STD	st. dev. of building corner_no	[-]	0.0	147.8	3.9	EC_ratio_A_STD	st. dev. of building EC_ratio	[-]	0.000	0.419	0.078
squareness_AVG	avg. building squareness	[°]	0.0	43.1	2.6	CH_area_AVG	avg. building CH_area	[m^2]	2.7	28 370.5	350.9
squareness_STD	st. dev. of building squareness	[°]	0.0	29.1	2.6	CH_area_STD	st. dev. of building CH_area	[m^2]	0.0	40 097.0	273.4
corner_dev_AVG	avg. building corner_dev	[m]	0.0	0.4	0.1	CH_perimeter_AVG	avg. building CH_perimeter	[m]	6.5	752.1	57.3
corner_dev_STD	st. dev. of building corner_dev	[m]	0.0	0.3	0.1	CH_perimeter_STD	st. dev. of building CH_perimeter	[m]	0.0	755.8	22.4
CO_number_AVG	avg. building CO_number	[-]	0.0	4.0	0.0	CH_length_AVG	avg. building CH_length	[m]	2.3	329.3	21.7
CO_area_AVG	avg. building CO_area	[m^2]	0.0	3 614.8	2.9	CH_length_STD	st. dev. of building CH_length	[m]	0.0	305.0	8.8
CO_ratio_AVG	avg. building CO_ratio	[-]	0.000	0.362	0.001	CH_ratio_A_AVG	avg. building CH_ratio_A	[-]	0.475	1.000	0.950
BB_area_AVG	avg. building BB_area	[m^2]	2.7	31 038.4	386.2	CH_ratio_A_STD	st. dev. of building CH_ratio_A	[-]	0.000	0.419	0.053
BB_area_STD	st. dev. of building BB_area	[m^2]	0.0	43 870.0	312.9	CH_ratio_P_AVG	avg. building CH_ratio_P	[-]	1.000	2.130	1.045
BB_perimeter_AVG	avg. building BB_perimeter	[m]	6.5	769.4	60.3	CH_ratio_P_STD	st. dev. of building CH_ratio_P	[-]	0.000	0.760	0.045
BB_perimeter_STD	st. dev. of building BB_perimeter	[m]	0.0	800.1	24.5	CH_schumm_ratio_AVG	avg. building CH_shumm_ratio	[-]	0.306	0.991	0.717
BB_width_AVG	avg. building BB_width	[m]	1.5	128.4	11.4	CH_schumm_ratio_STD	st. dev. of building CH_shumm_ratio	[-]	0.000	0.291	0.058
BB_width_STD	st. dev. of building BB_width	[m]	0.0	126.6	4.8	IZ_area_MIN	minimum building IZ_area	[m^2]	2.0	802 111.4	3 124.7
BB_length_AVG	avg. building BB_length	[m]	1.6	323.6	18.8	IZ_area_MAX	maximum building IZ_area	[m^2]	54.7	1 066 023.8	15 384.0
BB_length_STD	st. dev. of building BB_length	[m]	0.0	300.2	8.0	IZ_area_SUM	sum of building IZ_area	[m^2]	54.7	7 370 106.3	47 531.1
BB_elon_AVG	avg. building BB_elon	[-]	0.075	1.000	0.669	IZ_area_AVG	avg. building IZ_area	[m^2]	51.7	802 111.4	5 540.3
BB_elon_STD	st. dev. of building BB_elon	[-]	0.000	0.553	0.154	IZ_area_STD	st. dev. of building IZ_area	[m^2]	0.0	500 979.1	3 583.3
BB_ratio_A_AVG	avg. building BB_ratio_A	[-]	0.373	1.000	0.898	IZ_ratio_AVG	avg. building IZ_ratio	[-]	0.000	0.754	0.178
BB_ratio_A_STD	st. dev. of building BB_ratio_A	[-]	0.000	0.448	0.086	IZ_ratio_STD	st. dev. of building IZ_ratio	[-]	0.000	0.567	0.085
BB_ratio_P_AVG	avg. building BB_ratio_P	[-]	0.790	1.883	1.002	neigh_dist_AVG	avg. building neigh_dist	[m]	0.0	1 797.1	7.5
BB_ratio_P_STD	st. dev. of building BB_ratio_P	[-]	0.000	0.752	0.028	neigh_dist_STD	st. dev. of building neigh_dist	[m]	0.0	377.8	4.0
EC_area_AVG	avg. building EC_area	[m^2]	4.2	85 190.2	739.6	neigh_count_AVG	avg. building neigh_count	[-]	0	4	1
EC_area_STD	st. dev. of building EC_area	[m^2]	0.0	106 486.9	598.5	neigh_count_STD	st. dev. of building neigh_count	[-]	0	2	1
EC_perimeter_AVG	avg. building EC_perimeter	[m]	7.3	1 034.7	68.3	neigh_exists_AVG	avg. building neigh_exists	[0-1]	0.0	1.0	0.5
EC_perimeter_STD	st. dev. of building EC_perimeter	[m]	0.0	958.0	27.7	neigh_exists_STD	st. dev. of building neigh_exists	[0-1]	0.0	0.6	0.3

Table B.4: Morphological attributes of blocks (2)

Name	Description	Unit	Min	Max	Avg.
border_dist_MIN	minimum building border_dist	[m]	0.0	66.3	1.0
border_dist_AVG	avg. building border_dist	[m]	0.0	84.5	6.4
border_dist_STD	st. dev. of building border_dist	[m]	0.0	76.8	5.5
UNA_100_ reach_AVG	avg. reach in 100 m radius	[-]	0	111	19
UNA_300_ reach_AVG	avg. reach in 300 m radius	[-]	0	481	114
UNA_500_ reach_AVG	avg. reach in 500 m radius	[-]	0	1 228	255
UNA_100_ gravity_AVG	avg. gravity in 100 m radius		0.000	100.658	16.938
UNA_300_ gravity_AVG	avg. gravity in 300 m radius		0.000	310.448	76.922
UNA_500_ gravity_AVG	avg. gravity in 500 m radius		0.000	593.447	136.551
UNA_100_ betweenness_AVG	avg. betweenness in 100 m radius	[-]	0	3 508	98
UNA_300_ betweenness_AVG	avg. betweenness in 300 m radius	[-]	0	19 371	1 756
UNA_500_ betweenness_AVG	avg. betweenness in 500 m radius	[-]	0	55 390	6 260
UNA_100_ closeness_AVG	avg. closeness in 100 m radius		0.000	5.958	0.007
UNA_300_ closeness_AVG	avg. closeness in 300 m radius		0.000	5.958	0.002
UNA_500_ closeness_AVG	avg. closeness in 500 m radius		0.000	5.958	0.005
UNA_100_ straightness_AVG	avg. straightness in 100 m radius	[-]	0.000	216.826	17.082
UNA_300_ straightness_AVG	avg. straightness in 300 m radius	[-]	0.000	381.614	87.084
UNA_500_ straightness_AVG	avg. straightness in 500 m radius	[-]	0.000	955.682	188.953
UNA_100_ floorspace_AVG	avg. floorspace in 100 m radius	[m^2]	0.0	100 356.1	7 222.7
UNA_300_ floorspace_AVG	avg. floorspace in 300 m radius	[m^2]	0.0	444 663.1	50 579.3
UNA_500_ floorspace_AVG	avg. floorspace in 500 m radius	[m^2]	0.0	1 072 245.2	131 000.4
UNA_100_ footprint_AVG	avg. footprint area in 100 m radius	[m^2]	0.0	59 990.5	2 682.7
UNA_300_ footprint_AVG	avg. footprint area in 300 m radius	[m^2]	0.0	99 261.1	18 191.9
UNA_500_ footprint_AVG	avg. footprint area in 500 m radius	[m^2]	0.0	272 415.3	46 127.8
UNA_100_ cross_no_AVG	avg. crossroads no. in 100 m radius	[-]	0	21	4
UNA_300_ cross_no_AVG	avg. crossroads no. in 300 m radius	[-]	0	80	24
UNA_500_ cross_no_AVG	avg. crossroads no. in 500 m radius	[-]	2	198	61
UNA_100_ net_length_AVG	avg. net. length in 100 m radius	[m]	40.1	901.5	313.4
UNA_300_ net_length_AVG	avg. net. length in 300 m radius	[m]	40.1	6 896.7	2 020.5
UNA_500_ net_length_AVG	avg. net. length in 500 m radius	[m]	40.1	17 299.9	5 110.8

Table B.5: Morphological attributes of blocks (3)

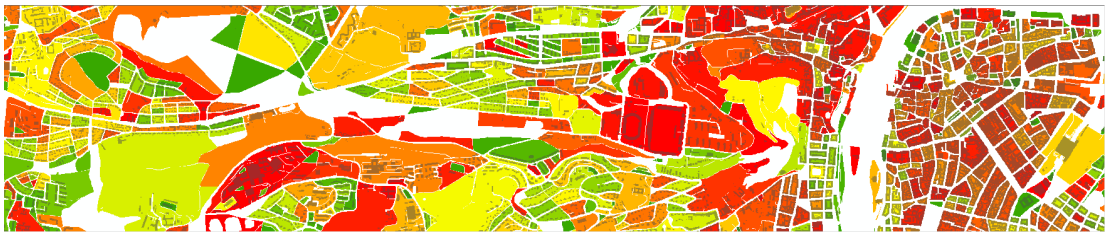
Appendix C

Morphological attributes – visualization

A subset of selected interesting morphological attributes of buildings and blocks is visualized on a section of study area of Prague. Values of attributes are classified into 32 classes using quantile method and visualized using green-yellow-red scale (highest values are assigned red color).

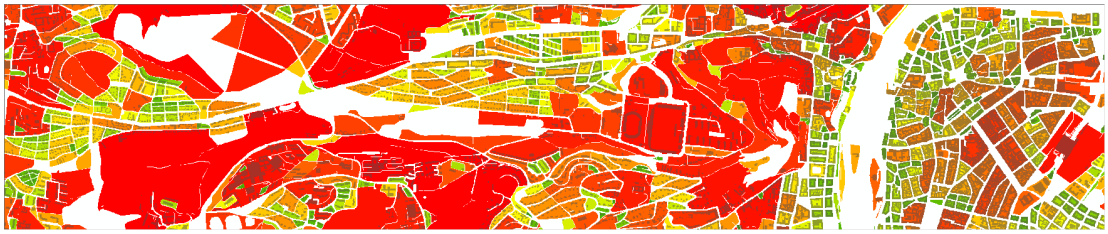
(a) buildings: BB_ratio_A

(b) blocks: BB_ratio_A_AVG



(c) blocks: BB_ratio_A_STD

Figure C.1: BB_ratio_A of individual buildings and summarized for blocks. Source: author

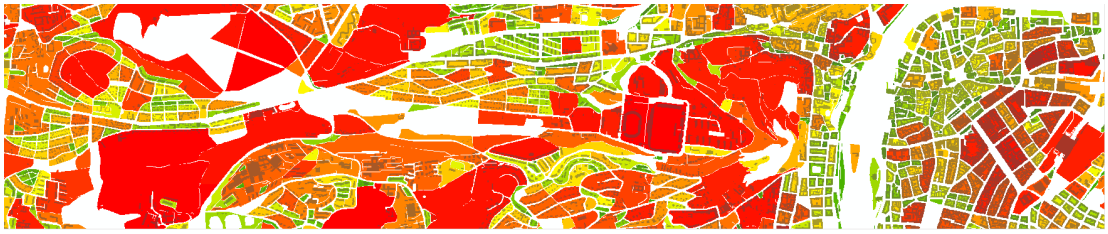


(a) Block area

(b) Elongation from bounding box

(c) Ratio of BB area to block area

(d) Schumm's ratio



(e) Fractal dimension

(f) Squareness of block borders

Figure C.2: Visualization selected of morphological attributes of blocks (1). Source: author

(a) Number of courtyards

(b) Ratio of perceived façade to total façade length

(c) Built front ratio

(d) Permeability

(e) Gross space index

(f) Floor space index

Figure C.3: Visualization selected of morphological attributes of blocks (2). Source: author

(a) Number of buildings

(b) Average distance of building to block border

(c) Average IZ ratio

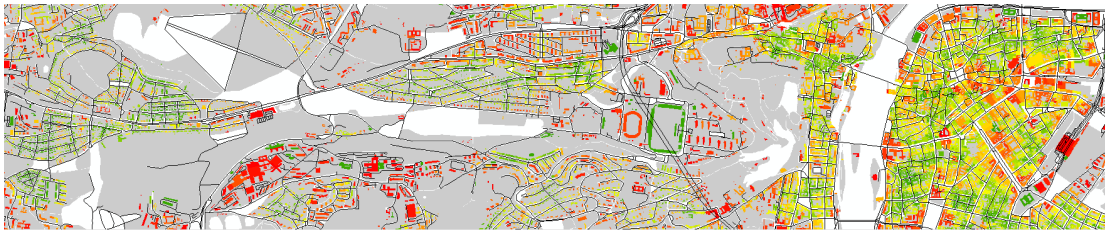
(d) Standard deviation of buildings height

(e) Average count of neighbouring buildings

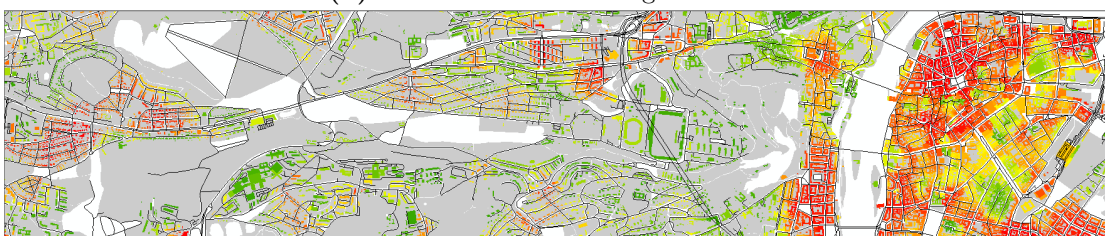
(f) Minimum street width

Figure C.4: Visualization selected of morphological attributes of blocks (3). Source: author

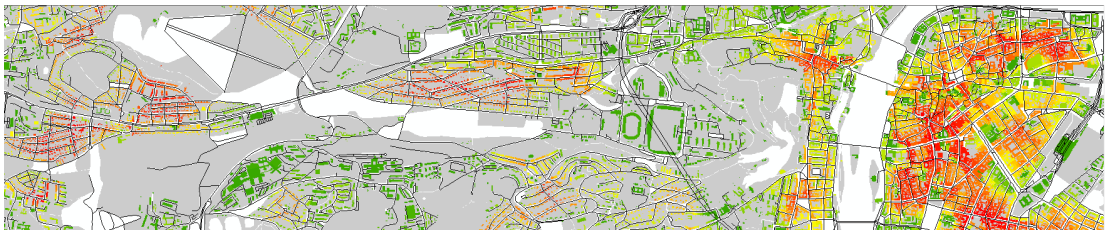
(a) Squareness of building walls



(b) Closeness in 100 m neighbourhood



(c) Network length in 300 m neighbourhood



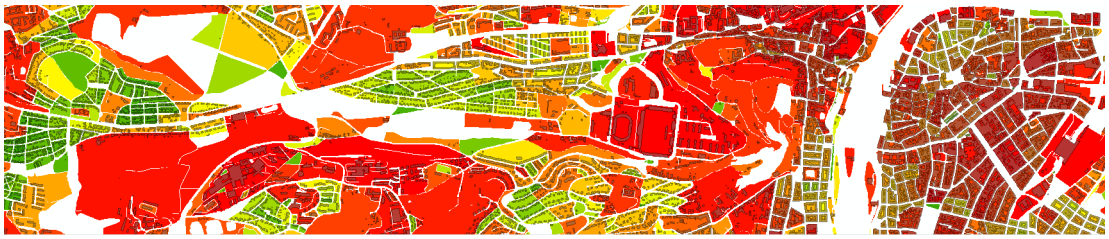
(d) Betweenness in 500 m neighbourhood

(e) Reach in 500 m neighbourhood

Figure C.5: Visualization selected of morphological attributes of buildings. Source: author

Appendix D

Principal components



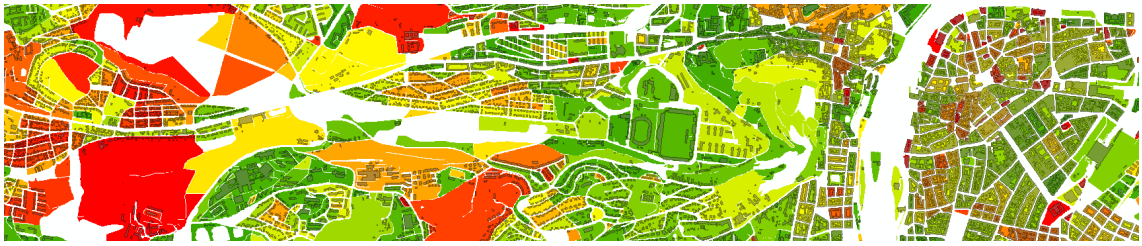
(a) 1st PC

(b) 2nd PC

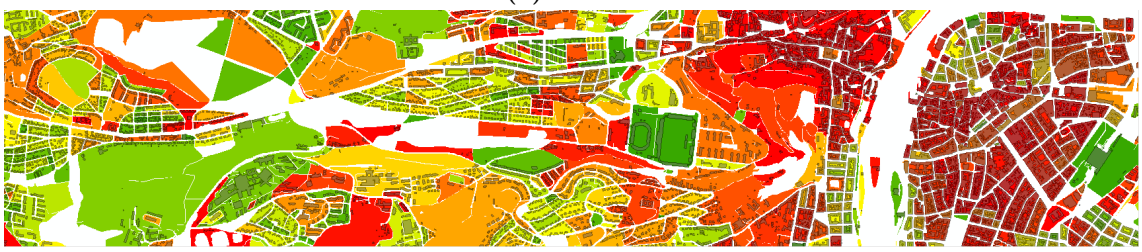
(c) 3rd PC

(d) 4th PC

Figure D.1: Visualization of 1st to 4th principal component. Source: author

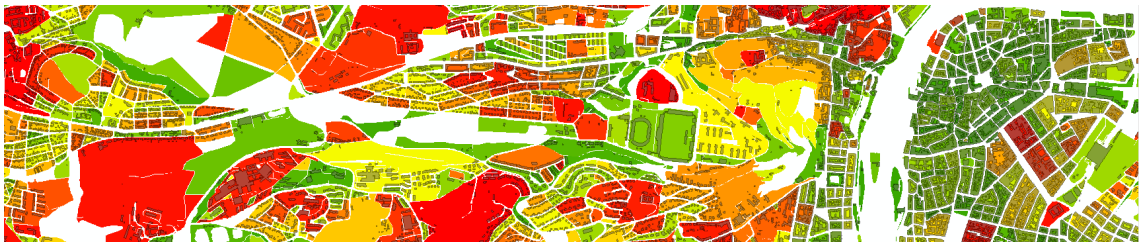


(a) 5th PC



(b) 6th PC

(c) 7th PC



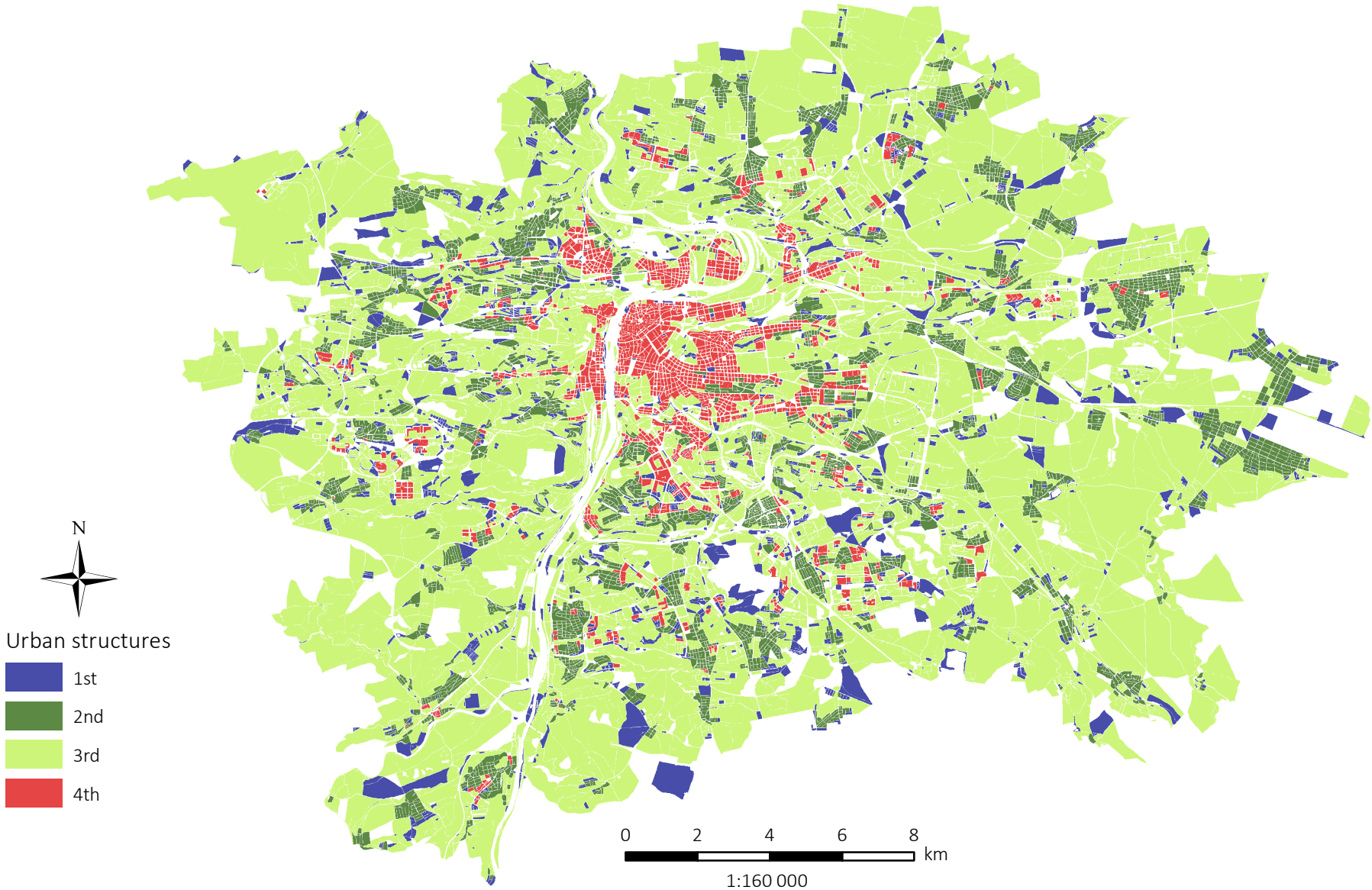
(d) 8th PC

(e) 9th PC

Figure D.2: Visualization of 5th to 9th principal component. Source: author

Figure D.3: Interpretation of first nine principal component. Attributes are grouped by positive (+) or negative (-) loading, ordered by loading values and coloured based on attribute character (see legend in bottom right corner). Source: author

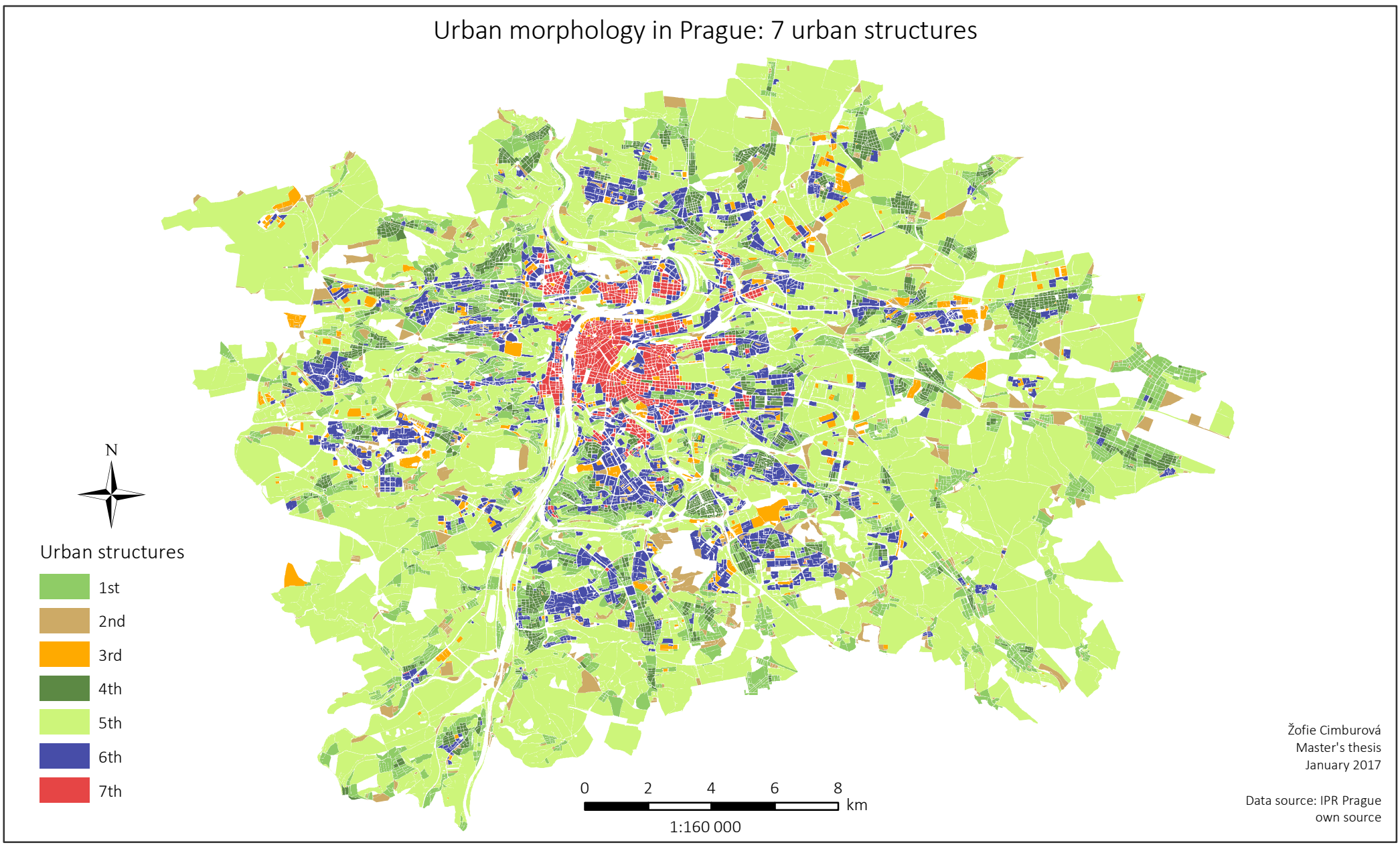
Urban morphology in Prague: 4 urban structures



Žofie Cimburová
Master's thesis
January 2017

Data source: IPR Prague
own source

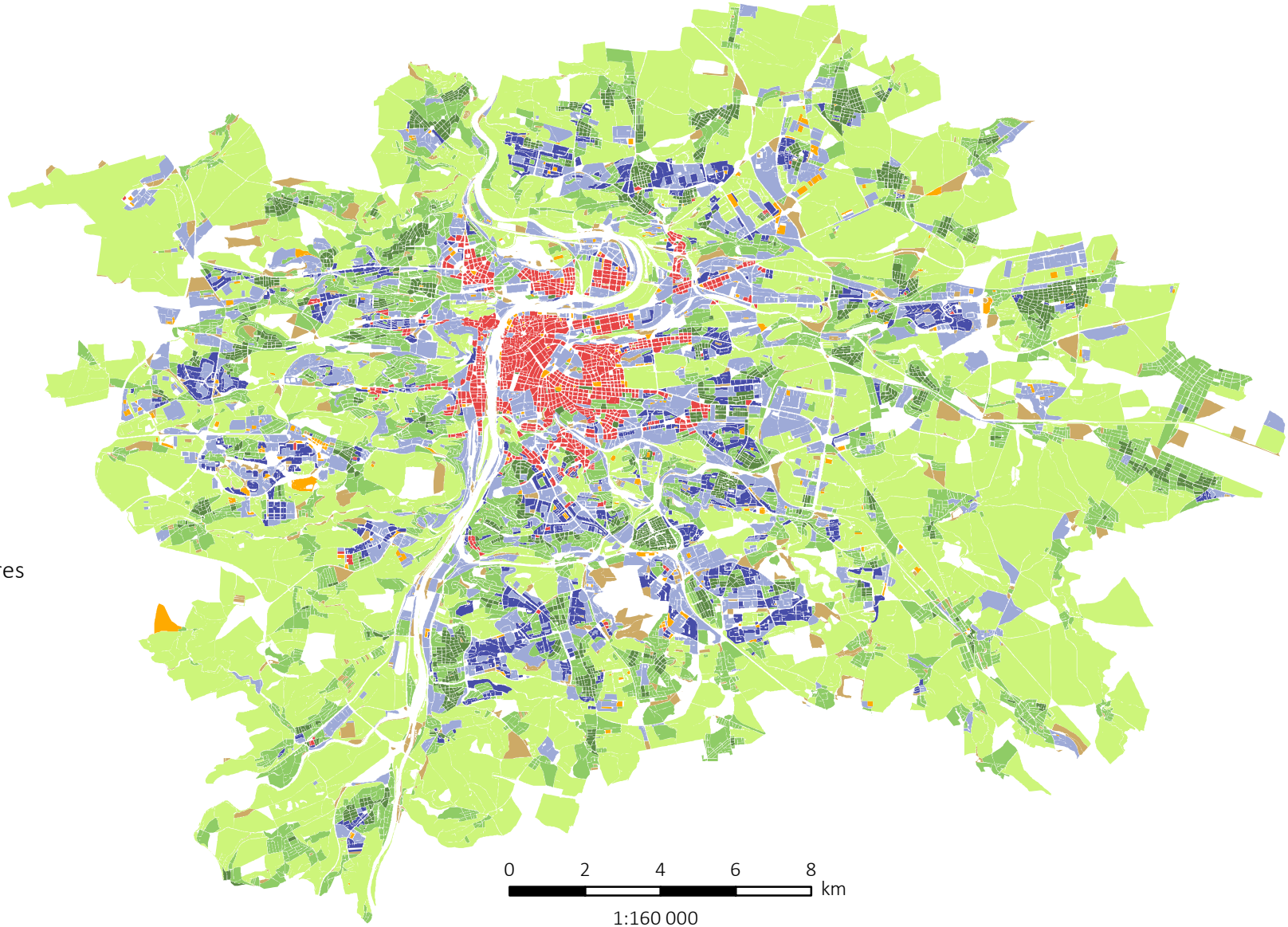
Urban morphology in Prague: 7 urban structures



Žofie Cimburová
Master's thesis
January 2017

Data source: IPR Prague
own source

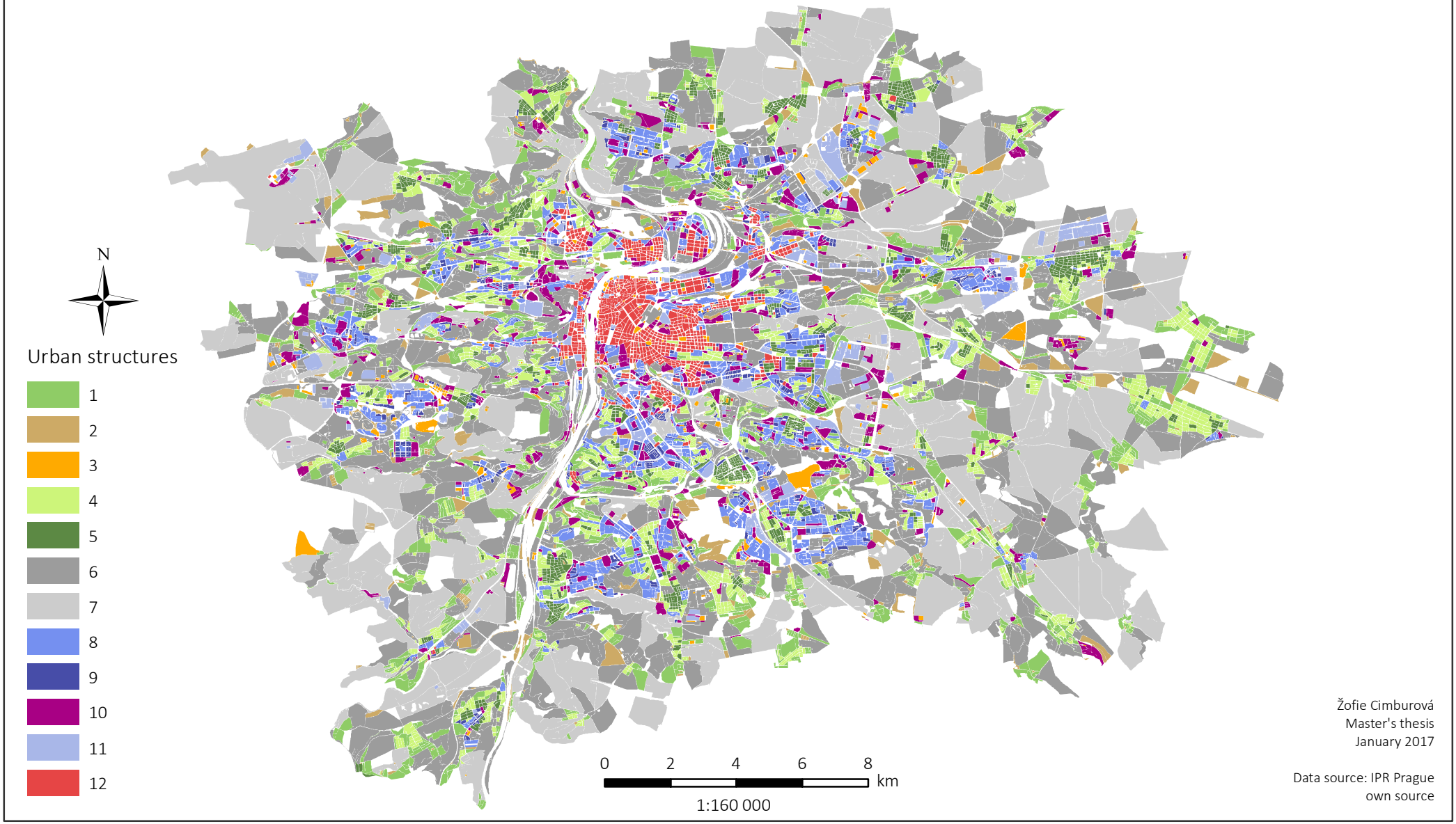
Urban morphology in Prague: 8 urban structures



Žofie Cimburová
Master's thesis
January 2017

Data source: IPR Prague
own source

Urban morphology in Prague: 12 urban structures



Žofie Cimburová
Master's thesis
January 2017

Data source: IPR Prague
own source