Estimation of Building Types on OpenStreetMap Based on Urban Morphology Analysis

Conference Paper · January 2014

DOI: 10.1007/978-3-319-03611-3_2

3 authors, including:

Hongchao Fan
Norwegian University of Science and Technology

Alexander Zipf
Universität Heidelberg

Estimation of building types on OpenStreetMap based on urban morphology analysis

Hongchao Fan and Alexander Zipf

Chair of GIScience, Heidelberg University, Berlinerstr.48, 69120 Heidelberg, {hongchao, zipf} @uni-hd.de

Abstract: Buildings are man-made structures and serve several needs of society. Hence, they have a significant socio-economic relevance. From this point of view, building types should be strongly correlated to the shape and sized of their footprints on the one hand. On the other hand, building types are very impact by the contextual configuration among building footprints. Based on this hypothesis, a novel approach is introduced to estimate building types of building footprints data on OpenStreetMap. The proposed approach has been tested for the building footprints data on OSM in Heidelberg, Germany. An overall accuracy of 85.77% can be achieved. Residential buildings can be labeled with accuracy of more than 90%. Besides, the proposed approach can distinguish industrial buildings and accessory buildings for storage with high accuracies. However, public buildings and commercial buildings are difficult to be estimated, since their footprints reveal a large diversity in shape and size.

Keyword: data enrichment, building footprints, OpenStreetMap, urban morphology

1. Introduction

Nowadays, OpenstreetMap (OSM) is considered as one of the most successful and popular VGI (Volunteered Geographic Information) projects, with a global cast of volunteers. Currently, there are more than 1.4 million registered members (OSM, 2013) and OSM is growing rapidly. Sparked by the availability of high-resolution imagery from Bing since 2010, there has been an increase in building information in OSM, proving that volunteers do not only contribute roads or POIs to the database. According to the statistics (the values are derived from our internal OSM database which is updated daily), on November 20, 2013, the number of buildings in OSM was over 77 million. In Germany, there are almost 9 million objects with “building=yes” to the same time point.

Currently, building footprints data in OpenStreetMap (OSM) is mainly used for reconstructing 3D buildings. At present there are several projects which generate and visualize 3D buildings from OSM: OSM-3D, OSM Buildings, Glosm , OSM2World , etc. And applications based on these projects i.e. 3D navigation on mobile devices, web-based visualization, and simulation etc. are getting increased. The most of 3D buildings in these projects are rendered as polyhedral, extruded footprints with flat roofs, whereby the height information of a number of buildings are directly taken from the attribute of building footprints or converted from the number of stories, while the majority of 3D buildings own random heights. In OSM-3D, many buildings are modeled in LoD2 (Level of Detail according to CityGML) in case there are indications for their roof types (Goetz and Zipf, 2012). In further, Goetz (2013) proposed a conception to generate buildings in LoD3 and LoD4 in CityGML. Besides, buildings in different LoDs from other sources can be uploaded via OpenBuildingModels and visualized in OSM-3D. But the buildings for uploading have to be adapted with the corresponding building footprints in OSM (Uden and Zipf, 2012).

The applications based on the abovementioned projects i.e. 3D navigation on mobile devices (Li et al. 2012), web-based visualization, and simulation etc. are getting increased. 3D buildings could play important roles not only for the 3D visualization but also for other interoperability like spatial analyses and/or queries in the 3D environment. For this purpose, it is crucial to know the types of buildings. According to the latest research on quality assessment of OSM building footprints data, OSM building footprints data has a high completeness in terms of area covered, while there is limited attributive information such as building types (Fan et al. 2013).

Unfortunately, there are very few researches on building footprints data enrichment. To the best of the authors’ knowledge, only one detailed study has been conducted by Huang et al. (2013). In their work, an automatic method is developed by using the geometric and topological features in footprint data, in order to enhance the maps with the building usage information. The work utilized the knowledge that building types are strongly characterized by the geometric features of footprints. In addition, simple neighborhood relation was taken into account to improve the classification. However, size/area of building footprints was not considered. Another related work is proposed by Henn et al. (2012) to derive the architectural types of buildings based on coarse LoD1 block models with vertical walls and flat roofs by employing Support Vector Machines (SVMs), whereby, geometric features such as length, width, area, and degree of perpendicularity of building footprints, types of buildings, as well as height information of buildings are required for the classification process.

In this work, a novel approach is introduced to estimate building types of building footprints data in OpenStreetMap. The approach is proposed based on urban morphology analysis by using ATKIS building footprints data in five city districts in Heidelberg. First of all, the correlations among geometries (mainly area and rectangularity) of building footprints and their types are derived. Then it investigates the similarity in building types for the buildings whose footprints are similar in shape and size. Based on the results of this urban morphology analysis, a set of rules are established to estimate buildings types in OpenStreetMap. Prior to the rule-based estimation of building types, the existing records of building types in OSM are classified into six types: (i) residential buildings, (ii) industrial buildings, (iii) commercial buildings, (iv) public buildings, (v) accessory buildings for storage and (vi) accessory buildings for supply. The building footprints on OSM are then labeled to one of these six types using the proposed method.

The reminder of this paper is structured as follows. Section 2 describes the algorithm for detecting similar building footprints in urban area. Section 3 presents the urban morphology analysis. Section 4 describes the algorithm for estimation of building types on OSM. Section 5 presents the experimental results for the OSM data set in Heidelberg and evaluates the results by using ATKIS data. Finally, Section 6 concludes the whole work and gives some works in the future.

2. Finding similar building footprints in urban area

As indicated in the introduction, we would like to investigate if the building types are similar when these buildings are similar in shape and size. For this purpose, building footprints have to be clustered into a number of groups at first, whereas building footprints in each cluster have similar shape and size.

Most of the existing research works about finding similar building footprints focus on urban building clustering, because building footprints with similar shapes and size form patterns according to Gestalt theory. The most common approaches to find similar building footprints are to define polygon of building footprints using a set of parameters, i.e. minimum distance, area of visible scope, area ratio, edge number ratio, smallest minimum bounding rectangle (SMBR), directional Voronoi diagram (DVD) and so on (Yan et al. 2008). Similar footprints then have similar parameters, for instance, Qi and Li (2008) calculated the similarity of two footprints by comparing their intersected and united areas. These approaches, however, can only find approximately similar building footprints. Therefore, they are more suitable for building footprints with less detailed geometries. In further, most of these approaches are very sensitive to orientations.

In the field of computer vision, polygon curve representation is used to measure similarity of polygons with high accuracy. Arkin et al. (1991) introduced turning function to represent polygons. As shown on the left of Figure 1, let C be the polygon. The tangent angle at the starting vertex is $\theta_1 = \varphi_1$. Then θ_i can be calculated as $\theta_i = \theta_{i-1} + \varphi_i$. The right of Figure 1 shows the change of tangent angles (y-axis) along the normalized accumulated length of the polygon sides (x-axis). From this point of view, the tangent angle can be treated as a function of the normalized accumulated length $T_C(l)$. It can be called tangent function or turning function.

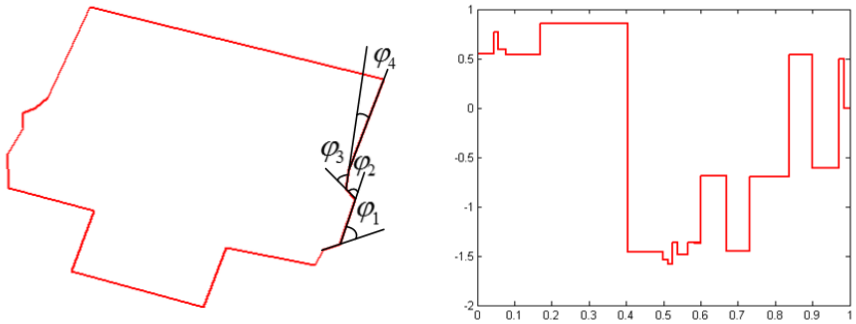


Fig.1. Turning function representation of polygon

The turning function $T_C(l)$ measures the angle of the counter-clockwise tangent as a function of the normalized accumulated length l. The cumulative angle increases with left hand turns and decreases with right hand turns. This kind of representation is invariant to rotation, because it contains no orientation information. Furthermore, it is invariant to scaling, since the normalized length makes it independent to the polygon size.

Then similarity of two polygons can be derived based on the $\mathrm{L}_2-\mathrm{norm}$ of their turning functions.

$$S(A,B) = d(A,B) = \|T_A - T_B\|_2 = \left(\int_0^1 \left(T_A(l) - T_A(l)\right)\right)^{\frac{1}{2}} \qquad (1)$$

Note that $S(A,B)$ denotes actually the dissimilarity between A and B. The smaller $S(A,B)$ is, the more similar are the two polygons. In the case A is identical to B, there is $S(A,B) = 0$.

However, the similarity measurement above is strongly affected by the starting point shift of the curve, because there is translation of the turning function when shifting the starting points. In the presented work, this problem is solved by resampling turning functions into power spectrum. Then similarity can be measured by comparing Fourier Descriptors of the power spectrum, as suggested by Lee et al. (2003).

3. Impact of urban morphology on types of buildings

Buildings are man-made structures and serve several needs of society. Hence, buildings have a significant socio-economic relevance. Their dimensions and architectures are strongly characterized by the purposes for that buildings are used, namely, the types of buildings. On the 2D map, these impacts are reflected by the shapes and sizes of building footprints. For instance, footprints of residential buildings normally have small area and relative simple shape – most of them are in form of rectangle, while public buildings normally have large and complicated footprints. In this work, the morphology of building footprints are analyzed using the ATKIS building footprints data in five city districts of the German city Heidelberg, namely, Altstadt, Bergheim, Boxberg, Emmertsgrund and Südstadt (Figure 2). ATKIS stands for Amtliches Topographisch-Kartographisches Informationsystem -- Authorative Topographic-Cartographic Information System. It is a common project of the Working Committees of the Survey Administrations of the States of the Federal Republic of Germany (AdV) (Grünreich, 2000). It contains information on settlements, roads, railways, vegetation, waterways, and more. Building footprints data in ATKIS is represented by polylines with building height, types, address and other attribute information.

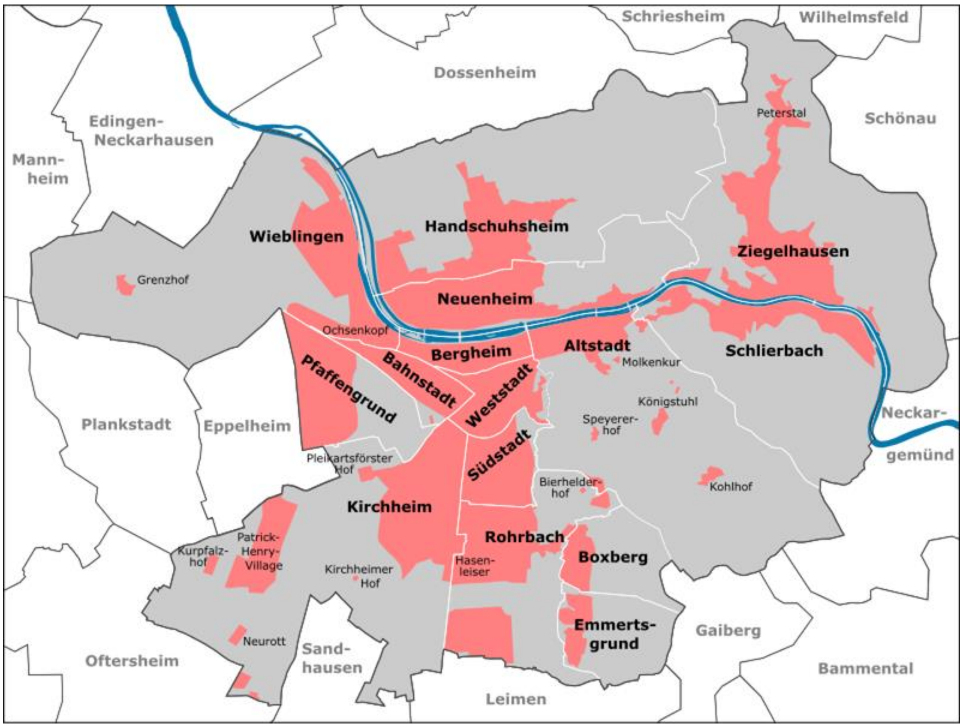


Figure 2. Administrative districts in Heidelberg

The morphological analyses are categorized in two classes. Firstly, it investigated whether different types of buildings in general differ from size and shape of building footprints, whereby size is represented by area and shape is measured by the rectangularity of the footprint polygon. The standard approach to measuring rectangularity is to use the ratio of the footprint area to the area of its minimum bounding rectangle (MBR), which is calculated as

$$rectangularity = \frac{area(polygon)}{area(MBR)} \tag{2}$$

The larger the *rectangularity* is, the simpler the building is in architectural style. Oppositely, the smaller the rectangularity is, the more complicated the building is constructed in architecture.

Table 1 summarizes the results of the statistical analysis on building footprints of six types of buildings respectively. It shows that building types are characterized by the area and rectangularity of building footprints, although the built area of a certain type of buildings varies as much as the average area of this type of buildings. In terms of built area, industrial buildings and accessory buildings for storage can easily be differentiated from other types of buildings, because industrial buildings are normally very large and accessory buildings for storages are normally very small. Residential buildings are majority in a city. The most of residential buildings have built area of 100 to 200 m^2. And their footprints are relatively simple with respect to architectural style. In the contrast, public buildings are normally large and the rectangularity of their footprints is low, because many public buildings are treated as landmarks due to their extraordinary dimension and architectural styles in their local environment.

Table 1. Geometrical analysis of building footprints for different types of buildings

	Number of buildings	Area		Rectangularity	
		Mean value (m^2)	Standard Deviation	Mean value	Standard Deviation
Residential building	3055	154.26	101.04	0.88	0.11
Industrial building	24	1912.55	1978.13	0.92	0.12
Commercial building	405	479.10	660.71	0.83	0.15
Public building	211	875.17	1240.04	0.72	0.17
Accessory building - storage	1390	38.21	33.24	0.95	0.09
Accessory building - supply	103	809.04	851.72	0.83	0.18

The second measurement of the morphological analysis is proposed according to the first law of geography (Tobler, 1970). It says: "Everything is related to everything else, but near things are more related than distant things”. In this work, the analysis is conducted based on the following hypothesis:

- Buildings share attributes i.e. types, heights, structure of roof and facades, if their footprints are similar in shape and size.

- For buildings with similar footprints, the closer they are located to each other, the more likely they share attributes such as type.

In the second morphological analysis, building footprints in a city district are classified into a number of clusters according to their shapes and sizes, as described in Section 2.2. As shown in Table 2, a cluster in city center (Altstadt) contains 4 buildings in average, while in the suburban region (Emmertsgrund and Südstadt) a cluster contains approximately 10 buildings in average. This means that building footprints in city center reveal a larger diversity than those in suburban regions. In each cluster, the similarity of building types is calculated based on Eq. (2).

$$Similarity_{type} = \frac{the\ number\ of\ the\ most\ frequency\ building\ type}{the\ number\ of\ building\ footprints\ in\ the\ cluster} \quad (3)$$

Comparing the average similarities in the five city districts, the types of buildings are about 80% similar in city center while they are more than 90% similar in suburban regions, if the building footprints are similar in shape and size.

Table 2. Context analysis on types of building footprints in city district

	Number of buildings	Number of clusters	Average similarity of building type in clusters
Altstadt	1978	534	79.17%
Bergheim	877	282	79.36%
Boxberg	545	111	94.18%
Emmertsgrund	333	41	97.50%
Südstadt	1189	133	89.88%

The similarity of building type in a cluster is getting higher, when decreasing the search area from city district to urban block which is defined as the smallest area that is surrounded by streets. Table 3 indicates that the possibility that buildings have same types is more than 95%, if they are located in the same urban block and their footprints are similar in shape and size.

Table 3. Context analysis on types of building footprints in urban block

	Number of urban blocks	Average similarity of attribute in clusters
Altstadt	97	95.42%
Bergheim	39	95.96%
Boxberg	32	95.86%
Emmertsgrund	12	100%
Südstadt	47	94.53%

4. Estimation of building types by rule-based approach

Based on the morphological analysis in Section 3, for a building footprint $foot_A$ on OpenStreetMap, its type can be estimated as follows:

Step 1: start with a building footprint $foot_A$ on OpenStreetMap. Keep the building type of $foot_A$, if there is an attribute for the building type.

Step 2: if there is no information about building type for $foot_A$

(a) In the same urban block, find building footprints $foot_{1,2...N}$ with similar shapes and sizes to $foot_A$. If there is information of building type in these building, take the majority type as the type of $foot_A$. Otherwise,

(b) In the whole area, find building footprints $foot_{1,2...N}$ with similar shapes and sizes to $foot_A$. If there is information of building type in these building, take the majority type as the type of $foot_A$. Otherwise,

(c) Estimate the building type using the results of the statistical analysis of building footprints in Table 1.

- If $|Area_{foot_A} - 150| \geq 50m^2$, $foot_A$ is a residential building.
- If $Area_{foot_A} \geq 2000m^2$ and $AI_{foot_A} \geq 0.9$, $foot_A$ is an industrial building.
- If $|Area_{foot_A} - 450| \geq 50m^2$, $foot_A$ is a commercial building.
- If $Area_{foot_A} \geq 750m^2$ and $AI_{foot_A} < 0.78$, $foot_A$ is a public building.

- If $Area_{foot_A} \leq 50m^2$, $foot_A$ is an accessory building for storage.
- If $|Area_{foot_A} - 800| \geq 50m^2$and $AI_{foot_A} \geq 0.83$, $foot_A$ is a an accessory building for supply.

(d) In an urban block, if there is an industrial building with $Area_{foot_A} \geq 2000m^2$ and $AI_{foot_A} \geq 0.9$, then change all the buildings in the urban block as industrial buildings.

5. Experiments and evaluation

The proposed approach is implemented and tested for the OSM data set of Heidelberg in Germany. The test area covers 46.38 km^2. The whole area is divided into 14 administrative districts which are called as city districts in this work, as shown in Figure 2. There are 32836 buildings in ATKIS while 14335 buildings in OSM, because many buildings in OSM are recorded as blocks of buildings in the real world on the one hand. On the other hand, a number of buildings are difficult to be recognized and mapped on OSM due to occlusion by vegetation and other buildings surrounding them. The ATKIS footprints in five city districts are used for morphology analysis, as described in Section 3, while the rest 9 city districts are used for the experiments and evaluation.

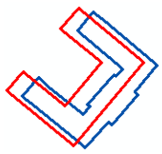
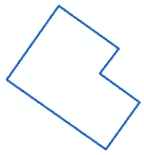
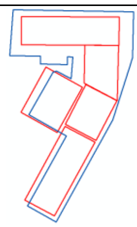
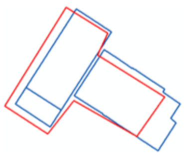
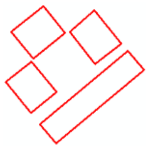
In this section, a preprocessing step is introduced to filter the building footprints which are recorded as block of buildings in OSM. Then the existing building types are classified into six types and used as input data for the process of estimation building types based on urban morphology analysis. The results are presented and evaluated by using authority data of ATKIS building footprints.

5.1 Pre-processing

- Finding semantically correctly recorded buildings in OSM

There might be 1:1, 1:n, 1:0, 0:1, n:1, and n:m relations between OSM building footprints and those in reference data, as shown in Table 4, whereby footprints in two data sets are distinguished in red and blue colors. While footprints in OSM are visualized in red color, footprints in ATKIS data are in blue.

Table 1. Possible relations between building footprints in two data sets

Relation	1:1	1:0	1:n
Illustration			
Relation	n:1	0:1	n:m
Illustration			

According to the OGC standard of CityGML building models (Groeger et al. 2008), semantic hierarchy and geometrical level of details (LoD) relate them inherently. Hence, only 1:1 buildings are recorded semantically correctly in OSM.

In this work, building footprints are selected for the test of type estimation, when they have 1:1 relation to ATKIS building footprints. Since the most of building footprints in OSM have been digitalized according to the Bing Map images (http://www.bing.com/maps) (Goetz and Zipf 2012; OSM 2013b, 2013c), there is normally offset between footprints in OSM and the reference data due to the distortion caused by oblique view of the used sensors. Considering this factor, large buildings in OSM have larger percentage of area overlap with their correspondence in the reference data, while small and high buildings might have smaller percentage of area overlap with their correspondence. The threshold of the judgment depends actually strongly upon the parameters of the Bing map images used for digitalization in OSM. In their work, Rutzinger et al. (2009) found out that the correspondence might be caused by their neighboring building if the overlapped area is less than 30%. Therefore, the threshold of the overlapping is set as 30%. If

$$\frac{Area_{overlap}}{\min\left(\mathrm{Area}\left(\mathrm{foot}_{osm\_i}\right),\mathrm{Area}\left(\mathrm{foot}_{ATKIS\_j}\right)\right)} > 30\% \qquad (4)$$

then the footprints foot_{osm_i} and foot_{ATKIS_j} are matched. A 1:1 relation is identified when a footprint in OSM can only be matched to one footprint in ATKIS.

- Classification of OSM building types into the pre-defined types

Because users are allowed to define building types by themselves (OSM, 2013c), there arise a large number of building types with high duplication and ambiguity. In this work, these building types are classified into six types (see Figure 3.): (i) residential buildings, (ii) industrial buildings, (iii) commercial buildings, (iv) public buildings, (v) accessory buildings for storage and (vi) accessory buildings for supply.

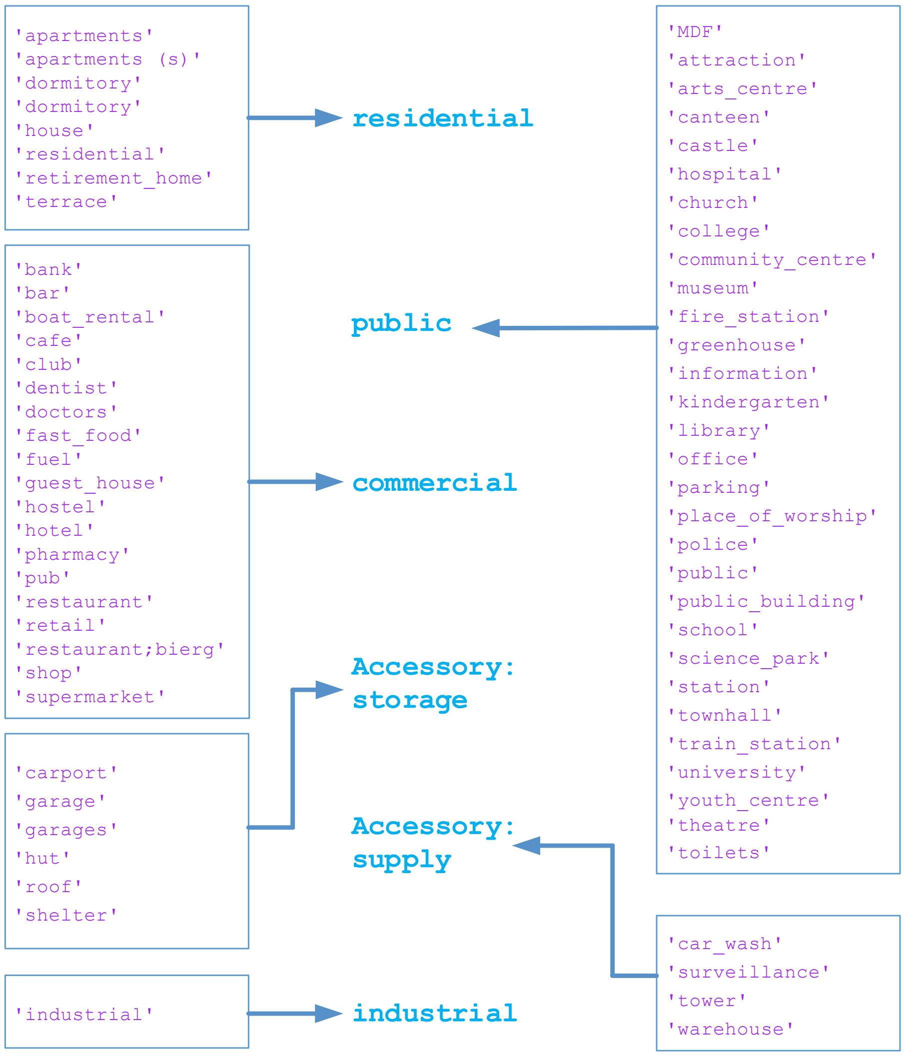


Figure 3. classification of building types in OSM

5.2 Experimental results and evaluation

In our test field, 12382 of 14335 building footprints in OSM were selected using the method in Section 5.1, because they have 1:1 relation to ATKIS footprints and are regarded as semantically correctly mapped on OSM. Among these selected building footprints, 2027 buildings are recorded with types of buildings on OSM. The building types of these 2027 building footprints are compared with those of ATKIS data. The results are listed in Table 4. It shows that: (i) the overall accuracy of the building type recording in Heidelberg on OSM is 88.86%; (ii) approximately 90% residential buildings are mapped with correct types, as well as industrial, public and storage buildings; (iii) commercial buildings and accessory building for supply are badly recorded with types, because they are normally difficult to be differentiated from residential and public buildings.

Table 4. Confusion matrix of the type record in Heidelberg on OSM

		Building types in ATKIS data: true data					
		Residential buildings	Industrial buildings	Commercial buildings	Public buildings	Accessory building - storage	Accessory building - supply
Estimated building types	Residential buildings	1303	1	20	5	13	17
	Industrial buildings	4	21	0	0	0	3
	Commercial buildings	25	0	35	6	5	6
	Public buildings	47	0	16	207	8	19
	Accessory building storage	20	1	0	0	229	4
	Accessory building supply	0	0	3	0	2	7
Total in column		1399	24	74	218	257	56
Producer accuracy		93.13%	87.50%	47.30%	94.95%	89.11%	12.50%
Overall accuracy		88.86%					

Although the types of 2027 buildings in Heidelberg are recorded with errors (with an accuracy of 88.86%), they are treated as correct and used as seeds in the process of type estimation, as described in Section 4. Table 5 summarizes the results of type estimation for the Heidelberg data on OSM. Residential buildings are estimated with high quality (with accuracy of 92.21%). There are several reasons. First of all, residential buildings are majority in the city. Even all the buildings are labeled as residential buildings; accuracy over 60% can be achieved. Secondly, in the input OSM data

(i) the majority of the buildings recorded with types on OSM are residential buildings; (ii) the residential buildings are recorded with high accuracy (93.13%); and (iii) the buildings recorded as residential building are distributed almost everywhere in Heidelberg. Therefore, they contribute much with respect of context effect during the process of type estimation.

Note: in the evaluation of type estimation in Heidelberg, the accuracy of type recording in the input OSM data is not considered.

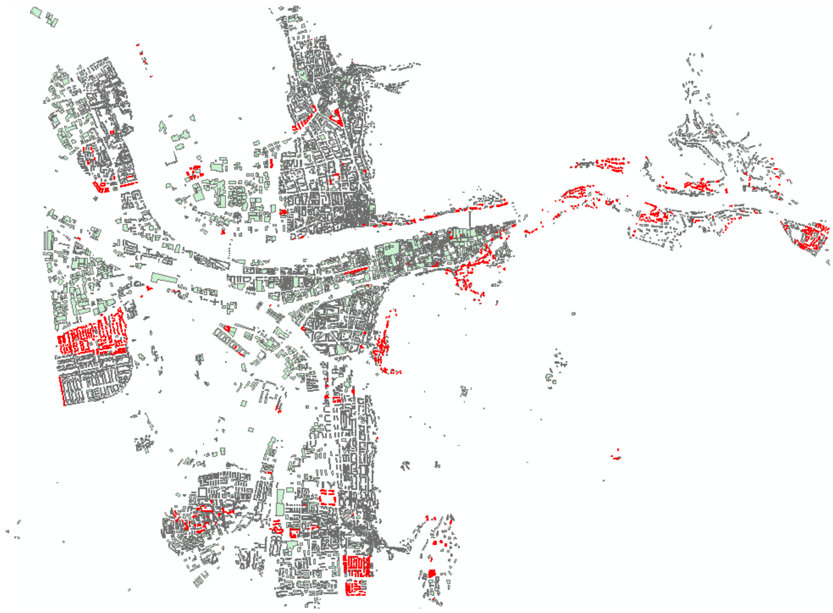


Figure 4. Residential buildings on OSM before type estimation

Despite of residential buildings, the accessory buildings for storage are estimated with high accuracy, too, because they can be easily distinguished from other buildings due to their characteristics in shape and size, namely, almost rectangular and in small size of area. Similarly, around 70% of industrial buildings are correctly estimated because of the simple architectural shape and large size in footprints. However, commercial buildings, public buildings and accessory building for supply are not so good estimated, because they reveal huge diversity in shape and size and therefore can be easily confused with each other. Many of these buildings are estimated as residential buildings, because buildings similar to residential buildings can also be used for commercial, public or (energy) supply buildings.

Table 5. Confusion matrix of the type estimation for OSM data in Heidelberg

		Building types in ATKIS data: true data					
		Residential buildings	Industrial buildings	Commercial buildings	Public buildings	Accessory building - storage	Accessory building - supply
Estimated building types	Residential buildings	9060	17	198	116	6	53
	Industrial buildings	0	111	7	7	0	8
	Commercial buildings	371	9	316	109	16	25
	Public buildings	254	12	126	304	5	3
	Accessory building storage	4	0	6	6	1242	54
	Accessory building supply	136	11	5	52	239	152
Total in column		9825	160	658	594	1508	295
Producer accuracy		92.21%	69.38%	48.02%	51.18%	82.36%	51.53%
Overall accuracy		85.77%					

For the nine city districts in Table 6, the process of type estimation is conducted again, and respectively. In most of city districts, better results are obtained than using the whole Heidelberg data as input data, since not only the overall accuracies are higher than that of using the whole Heidelberg data as input data, but also the producer's accuracies are higher. The reason is that the search area is reduced from the whole city to city district in the step 2b of the process of type estimation (Section 4). This verifies the hypothesis in Section 3: buildings with similar footprints might share attributive information, and the closer the similar footprints are located, the more likely they have same attributes.

Table 6. The accuracy of type estimation in the nine city districts in Heidelberg

	Producer accuracy						Overall accuracy
	residential	industrial	commercial	public	storage	supply	
Handschuhsheim	94.12%	100%	52.63%	57.14%	91.48%	71.42%	89.27%
Kirchheim	97.22%	79.31%	45.22%	78.26%	85.12%	92.98%	88.61%
Neuenheim	96.62%	60%	66.67%	48.89%	81.55%	37.04%	89.08%
Paffaffengrund	92.56%	100%	43.55%	41.67%	88.10%	52.22%	83.84%
Rohrbach	98.30%	-	41.66%	27.65%	78.66%	27.53%	90.85%
Schlierbach	95.30%	0%	0%	38.46%	78.62%	32.14%	86.41%
Weststadt	95.47%	0%	77.12%	58.82%	87.12%	63.93%	91.71%
Wieblingen	95.99%	95.12%	8.60%	25.93%	91.86%	48.70%	81.38%
Ziegelhausen	93.98%	0%	82.61%	66.67%	95.23%	72.46%	90.66%

6. Conclusion and future works

In this work, a rule-based approach for building type estimation is proposed with the attempt of data enrichment for OpenStreetMap footprints data. The rules are derived based on two hypotheses. First of all, building types are very impacted by the shapes and sizes of building footprints. Secondly, for buildings with similar footprints in shape and size, the more closely they are located to each other, the more likely they have the same building type. These two hypotheses are proofed by using authority footprints data in four city districts in Heidelberg by means of urban morphology analysis. The proposed approach is tested for OSM building footprints data in Heidelberg at first. The overall accuracy for the type estimation is 85.77%. With respect to the individual type of buildings, residential buildings are estimated with high accuracy, as well as industrial and accessory buildings for storage. Only about 50% of commercial buildings, public buildings and accessory building for supply are correctly estimated, because they reveal huge diversity in shape and size and therefore can be easily confused with each other. Many of these buildings are estimated as residential buildings, because buildings similar to residential buildings can also be used for commercial, public or (energy) supply buildings. When reducing the input area to city district level, better results can be yielded. Building types are estimated with both higher overall accuracy and producer's accuracies. This proofs our second hypothesis as well.

In comparison with the original accuracy of building type recording on the input data of OSM, the proposed approach achieve higher accuracy, in terms of both the overall accuracy and producer's accuracies. Besides, in the evaluation of type estimation in Heidelberg, the accuracy of type recording in the input OSM data is not considered. If this kind of impact could be calculated, even higher accuracy can be achieved by the proposed approach.

There are still many error estimations, especially for the commercial buildings, public buildings and accessory buildings for storage, because only 50% of these kinds of buildings can be correctly estimated with type. As mentioned above, one reason is that these kinds of buildings reveal huge diversity in shape and size. Other reasons could be that: (i) the types in the input OSM data are labeled with relative low accuracies, (ii) there are ambiguous classifications when classifying a lot of building types on OSM into six building types, and (iii) similarly, there are confusion classifications when classifying a lot of building types on authority data into six buildings types.

In the future, the abovementioned building type classification will be investigated for better correspondences in the two databases. Secondly, the proposed approach will be tested for larger cities in Germany. Thirdly, the algorithm of Support Vector Machine (SVM) will be used for the type estimation, whereby the authority data and buildings with known information of types will be used as training data respectively. In further, the results by using SVM will be compared with the proposed approach presented in this work.

7. References:

Arkin E.M., Chew L.P., Huttenlocher D.P., Kedem K. and Mitchell J.S.B. 1991. An Efficiently Computable Metric for Comparing Polygonal Shapes. In: IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 13, No. 3, March 1991.

Fan H., Zipf A., Fu Q. and Neis P. (2013): Quality assessment for building footprints data on OpenStreetMap. International Journal of Geographical Information Science. DOI: 10.1080/13658816.2013.867495

Goetz, M. (2013): Towards generating highly detailed 3D CityGML models from OpenStreetMap, International Journal of Geographical Information Science, DOI:10.1080/13658816.2012.721552

Goetz, M. and Zipf, A. (2012): OpenStreetMap in 3D – Detailed Insights on the Current Situation in Germany. AGILE 2012. Avignon, France.

Gröger, G., Kolbe, T. H., Czerwinski, A. and Nagel, C. 2008. OpenGIS City Geography Markup Language (CityGML) Encoding Standard – Version 1.0.0. OGC Doc. No. 08-007r1.

Henn A., Roemer C. and Groeger G. (2012): Automtic classification of building types in 3D city models using SVMs for semantic enrichment of low resolution building data. In: Geoinformatica, (2012) 16: 281-306.

Huang H., Kieler B. and Sester M. (2013): Urban building usage labeling by geometric and context analyses of the footprint data. In proceeding of: 26th International Cartographic Conference (ICC), 08/2013, Dresden, Germany.

Lee, D. J., Antani, S., & Long, L. R. (2003). Similarity measurement using polygon curve representation and fourier descriptors for shape-based vertebral image retrieval. In Medical Imaging 2003 (pp. 1283-1291). International Society for Optics and Photonics.

Li Q. and Li Z. (2008). An approach to building grouping based on hierarchical constraints. In: the International Archieves of the Photogrammetry, Remote Sensing and Spatial Information Science. Vol. XXXVII. Part B2. Beijing 2008.

OSM. 2013. Stats - OpenStreetMap Wiki. Retrieved 29/11/2013, from http://wiki.openstreetmap.org/wiki/Statistics

OSM 2013b. Bing - OpenStreetMap Wiki. Retrieved 29/11/2013, from http://wiki.openstreetmap.org/wiki/Bing

OSM 2013c. Buildings - OpenStreetMap Wiki. Retrieved 29/11/2013, from http://wiki.openstreetmap.org/wiki/Buildings

Rutzinger, M., Rottensteiner, F. and Pfeifer, N. 2009. A Comparison of Evaluation Techniques for Building Extraction From Airborne Laser Scanning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2(1): 11-20.

Tobler W., (1970) "A computer movie simulating urban growth in the Detroit region". Economic Geography, 46(2): 234-240.

Uden, M. and Zipf, A. (2012): OpenBuildingModels - Towards a platform for crowdsourcing virtual 3D cities. 7th 3D GeoInfo Conference. Quebec City, QC, Canada.

Yan H., Weibel R., and Yang B. 2008. A Multi-parameter Approach to Automated Building Grouping and Generalization. Geoinformatica 12, 1 (March 2008), 73-89. DOI=10.1007/s10707-007-0020-5 http://dx.doi.org/10.1007/s10707-007-0020-5