Urban Building Usage Labeling by Geometric and Context Analyses of the Footprint Data

Hai Huang, Birgit Kieler, Monika Sester

Institute of Cartography and Geoinformatics, Leibniz Universität Hannover

Abstract. We present an automatic building type (usage) labeling based on the footprint data. The usage information of buildings is of great interest for many applications, e.g., navigation, city planning and emergency management. This attribute, however, is generally not provided in the volunteered data sources like OpenStreetMap and is often incomplete even in the official cadastral maps. In this paper, we propose a method to enhance the maps with the building usage information exclusively using the geometric and topological features in the footprint data. A general category is predefined with four classes: residential, commercial, industrial and public. A novel inference framework is proposed using two new high-level (composite) geometric characteristics for the local description of the individual buildings and the Markov Random Field model to incorporate the contextual constraints of the neighborhood. Experiments are performed on both OpenStreetMap and cadastral data showing the potential of the proposed method.

Keywords: Building, Labeling, Usage, Footprint, OSM, MRF

1. Introduction

The usage (use and occupancy) information of buildings is of great interest for many applications, e.g., navigation, city planning, emergency management, etc. This attribute, however, is generally not provided to the buildings in a consistent way in the volunteered data sources like OpenStreetMap (OSM). Although a map feature catalogue exists, the volunteers are not obliged which attributes they set for the usage information. Even in the official cadastral maps, the building usage information is not always available. An approach to enhance OSM data is presented by Werder et al. (2010), in which an unsupervised classification of spatial data solely based on the geometric and topological characteristics is proposed. Both building outlines and road network information are employed. Lüscher et al. (2009) present a classification of buildings also based on topographic vector data by means of an ontology-driven approach. Supervised Bayesian inference is used to deal with the vagueness in definitions of spatial phenomena.

In this work, we classify the usage type of individual buildings in the urban area according also exclusively to their geometric and topological features derived from the given footprint data. A general category is predefined with four types of usage: (1) residential, e.g., single and multiple family houses, apartment buildings; (2) commercial, e.g., office buildings, supermarkets, shopping malls; (3) industrial, e.g., factory buildings, warehouses, and (4) public, e.g., museums, memorials, hospitals, theaters, stadia, universities/schools. We propose two new high-level geometric features: "effective width" and "branching degree", which are designed to quantify the living space and the structural complexity of the buildings, respectively. A novel inference framework is presented using (composite) geometric characteristics for the local description of the individual buildings and the Markov Random Field (MRF) model to incorporate the contextual constraints of the neighborhood. MRF (Kindermann & Snell, 1980), also known as Markov network, is an undirected graph model, in which the random variables hold Markov properties (cf. Section 3). It is widely used in image processing and computer vision (Li, 2009) for the labeling/segmentation of the image pixels or sub-regions and other applications like point cloud grouping, economics and sociology. In this work we use the vertices of the MRF to represent the individual buildings and the edges between vertices to encode the neighborhood constraints. By these means the buildings of dense urban areas can be classified and labeled with the above defined usage attributes more reasonably considering the geometric features and the neighborhood constraints.

The paper is organized as follows. In Section 2 we introduce the two new high-level composite geometric features, i.e., the effective width and the branching degree, and the definition of the local (unary) energy based on them. Section 3 presents the modeling of the building network and their neighborhood relationships via MRF, the definition of the contextual (binary) energy, and the optimization of the overall energy function. Experiment results and evaluation are demonstrated in Section 4. The paper ends up with conclusions in Section 5.

2. Building Attributes

First, we study the contribution of the local geometric features to the identification of the building types. One basic attribute that can be easily derived from the footprint data is the building area. It can somehow reflect the building usage, e.g., one building smaller than 200 square meters may be a single-family house and that of over 20000 square meters will very likely be a factory or warehouse. The problems of using such simple measure, however, is also clear: e.g., complex buildings such as apartment buildings may also have large footprint area and therefore cannot be distinguished from industrial or public building without considering the shape characteristics.

Considering the shape factor, a simple one can be defined as the ratio of building length and width. This L/W ratio helps to differentiate bar-like shape (often for residential or industrial buildings) and square-like shape (often for public or commercial buildings). But it works only well with rectangular buildings. For complex buildings, although the bounding box can be used to calculate L/W ratio, the values cannot reflect their real shape any more.

High-level attributes are therefore required to integrate multiple geometric attributes and provide more precise description to the building shape. In this work two new composite measures: "effective width" and "branching degree" are proposed specifically for the purpose of building usage classification.

2.1. Effective width

The effective width (EW) is an estimated width of buildings with arbitrary shapes. It is defined as the average width of the footprint along the centerline. For this we have to determine the centerline of the building skeleton, which describe the approximate length of the building.

Haunert & Sester (2008) compare different types of skeletons that are commonly used in geographic information systems for deriving polygon centerlines. As the basis of our work, we select a simple skeleton, the "straight skeleton", which only comprises straight lines in contrast to the "medial axis". The latter comprises also second-order lines, that would cause computational overhead. The straight skeleton is presented by Aichholzer et al. (1995) and is exemplarily shown in Figure 1 (a).

Please note, as shown in Figure 1 (building 1), the length of the centerline (L_C=1.50 m) can differ from the actual building length (L_B=7.50 m). To derive a reliable $\widehat{L_C}$ to approximate the L_B, we modify the original straight skeleton by extending the derived centerlines to the building boundary (red line in Figure 1, b). The effective width is practically calculated as the ratio of the building area (A_B) to the building length:

$$EW = \frac{A_B}{L_B} = \frac{A_B}{\widehat{L_C}} .$$

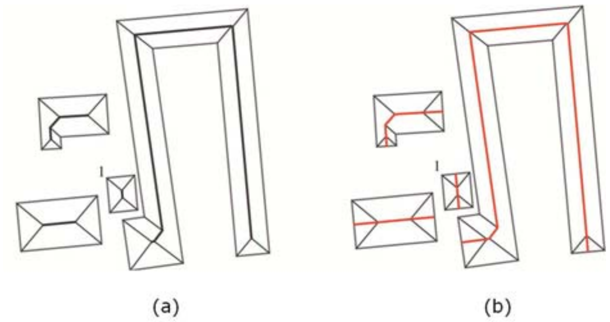


Figure 1. Building skeletons: (a) Straight skeleton of buildings (centerlines bold) and (b) modified centerlines (red lines).

The effective width is of interest in the usage classification because it actually implies the general living/movement space inside the building. By this means the residential buildings can be well distinguished from the industrial or public ones. In many building category definitions the single-family houses and multi-family houses (e.g., apartment buildings) must be defined as two separate classes because their area and complexity are remarkably different. Using EW, as demonstrated in Figure 2, the values of these two types of residential buildings show consistency, although the building areas and the complexities (calculated by "branching degree", cf. Section 2.2) are not close to each other.

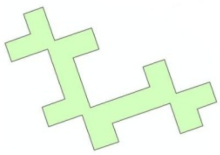
Footprint	Area (m^2)	Complexity	**EW**
	1144.75	2.85	**8.89**
	180.52	1.00	**7.78**

Figure 2. Effective width shows value consistency for the apartment building (top) and the single-family house (bottom) by indicating the living space of the buildings.

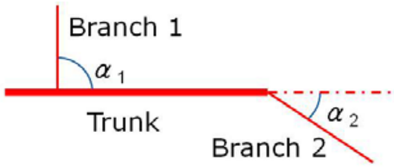
2.2. Branching degree

Another novel high-level attribute proposed in this work is the branching degree (BD). It scores the number and distribution of the building segments (called "branches") derived from the skeleton centerline to measure the structural complexity of the building. Please note that in this case the conventional straight skeleton as shown in Figure 1 (a) is employed for better overall structural analysis.

First, we define the longest linear segment of the skeleton centerline as the "trunk" and then the other segments as the "branches". The number, size, and branching angle of the branches are integrated as:

$$BD = \sum_{i=0}^{m} w_i \cdot \frac{L_i}{L_{trunk}} \ ,$$

where m is the total number of branches and L indicate the length of the i-th branch or the trunk (also the 0-th segment, $L_0 = L_{trunk}$). The weight of each branch or trunk is $w_i = 2 \cdot \alpha_i / \pi$ with α_i the intersection angle (in radians, $\alpha_i \in (0, \pi/2]$) of the current branch to its parent (as shown in Figure 3). The higher order branches have the same weight as that of the first order. A set of BD examples are given in Figure 3. Generally, we can imagine that the residential (except the apartment complex, cf. Section 2.1) and industrial buildings may have lower BD while the public buildings have normally higher complexity. One advantage of giving the weight w_i is the simple building with slight curve shape (Figure 3, building 5) can be better scored. The curved centerline in the straight skeleton is represented by a chain of linear segment and the weights can help to prevent the BD value being too high with the large number of "branches".



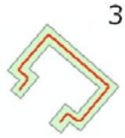
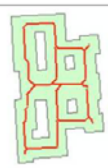
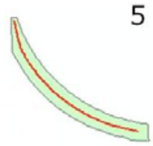
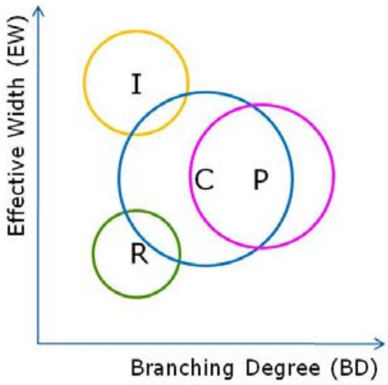
Footprint / straight skeleton centerline	1 	2 	3 	4 	5 
BD	1.00	1.19	2.81	4.12	1.46

Figure 3. Definition of trunk/branches and the intersection angle (top) and example values of the branching degree (bottom).

2.3. The local energy

Figure 4 shows a rough sketch to summarize the distribution of building clusters with different usage types in the space of EW-BD. Each building can be represented as a point in this 2D parameter space and the probability this building belongs to one of the classes is inversely proportional to its (standardized) distance to the centroid of the class, which is empirically given with generic values.



Usage	BD (range/centroid)	EW (range/centroid)
Residential (R)	Low / 1.5	Low / 12
Commercial (C)	Medium / 2.5	Medium / 25
Industrial (I)	Low / 1.5	High / 40
Public (P)	High / 3.2	Medium / 25

Figure 4. Distribution of buildings with different usage types in the parameter space of EW-BD.

The local potential of a building is then defined as a quaternary value of the probabilities:

$$p_{local} = \{p_R, p_C, p_I, p_P\}$$

that this building should be labeled. The probabilities are standardized with a sum of 1. E.g., if one building with a probability distribution of $\{p_R = 0.2, p_C = 0.1, p_I = 0.6, p_P = 0.1\}$ is given a label R, the current local energy of this vertex is 0.2, if label I then 0.6.

3. Context model

In this work, we use the MRF to model the buildings in the dense urban area and their neighborhood relationships. We define the graph model G as:

$$G = \{V, E\},$$

where the individual buildings are represented as vertices, $v = v_i, i \in V$, and the edges, $e = e(i, j), \{i, j\} \in E$, connecting pairs of vertices. Any pair of non-neighbor vertices is conditionally independent given all other vertices.

3.1. The definition of neighborhood

As shown in Figure 5, the neighborhood of buildings is defined based on the determination of Voronoi cells of the buildings centroids (distance-based approach). That is, as illustrated in Figure 5 (left), the polygons divide the whole area into seamless cells. For each centroid there will be a corresponding region consisting of all points closer to that centroid than to any other (Euclidean distance). And thus, all cells which share an edge are called neighbors. With the determination of neighbors, MRF holds the Markov properties: (1) pairwise Markov property: any two non-neighbor vertices are independent; (2) local Markov property: a vertex is conditionally independent to all other vertices given its neighbors, and (3) global Markov property: any two non-adjacent subsets are conditionally independent given a separating subset.

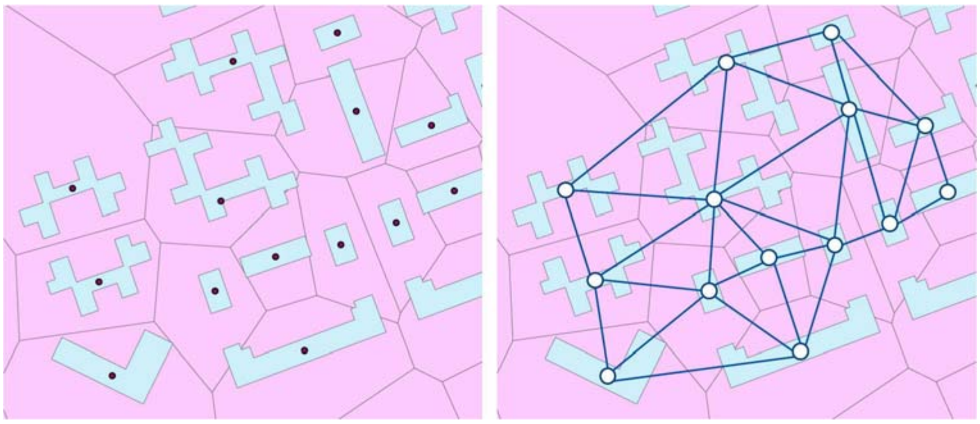


Figure 5. Definition of neighborhood of buildings: (left) polygons of buildings and their centroids marked with red points and their Voronoi cells; (right) the MRF model with the edges connecting neighbor buildings.

3.2. The overall energy

The overall energy function of the MRF consists of two components: the unary and the binary terms:

$$\mathcal{H} = \sum_{i} u(x_i, c_x) + \sum_{i,j} b(x_i, x_j) \ .$$

The unary energy $u(x_i, c_x)$ summarizes the local features of the individual buildings. It is calculated with the local potential described in Section 2.3:

$$u(x_i, c_x) = p_{local}(x_i)$$

with $x_i \in \{R, C, I, P\}$ being the random labeling assignment. It indicates the likelihood of the labeling.

The neighborhood inferences are encoded into the binary term, which implies the neighborhood plausibility of usage in the pairwise cliques. The plausibility is evaluated in two aspects:

1. Type consistency: neighbor buildings are inclined to have the same type;
2. Logical neighborhood: it reflects reasonable city planning for adjacent areas, e.g., residential buildings are more likely be found near public building instead of industrial zone.

The rewards as well as penalties of neighborhood proposals are embedded in a symmetric matrix N:

N	R	C	I	P
R	1	0	-1	0.5
C	.	1	0	0.5
I	.	.	1	-1
P	.	.	.	1

The binary energy of each clique is directly calculated based on this matrix as:

$$b(x_i, x_j) = N(i, j)\,.$$

The goal is to find the maximum $\mathcal{H}$ of the graph model. We use $\mathcal{K}$ to represent the configuration, i.e., a set of label assignments to all the vertices. The optimization task can be expressed as:

$$\widehat{\mathcal{K}} = \underset{\mathcal{K}}{argmax}\{\mathcal{H}\} = \underset{\mathcal{K}}{argmax}\left\{\sum_i u(x_i, c_x) + \sum_{i,j} b(x_i, x_j)\right\}$$

with $x_i, x_j \in \mathcal{K}$.

3.3. Stochastic sampling

Dealing with the data for dense urban area, the established MRF model is highly connected. In the optimization process the labels of all the vertices can be altered and lead to different configurations. These make the optimization task being a extremely high-dimensional problem and computational intractable for direct solution. In this work we employ statistical approximation by means of a Gibbs sampler (Geman & Geman, 1984) for this task. A Gibbs sampler is one Markov Chain Monte Carlo (MCMC) algorithm. It performs the random sampling specifically from the potentially complicated multivariate probability distribution with a large set of variables.

Let s be the step number and $\mathcal{M}$ the corresponding state of the model, the sampling process can be summarized as follows:

1. Initialization: $\mathcal{M}^{s=0}$, $\mathcal{K}^{s=0}$(give x_i based on unary likelihood only)

2. Propose a new state $\mathcal{M}'$ with the corresponding configuration $\mathcal{K}'$.

   2.1 Sample new label for the first building

   $$x_1' \sim p\left(x_1 \middle| x_2^{(s-1)}, \dots, x_n^{(s-1)}\right) = p\left(x_1 \middle| \tilde{x}_1^{(s-1)}\right)$$

   with n the total number of buildings and $\tilde{x}_1 \subset \{x_1, \dots, x_n\}$ the neighbors of building 1. As mentioned before, the current vertex is conditionally independent to the non-neighbor vertices. The conditional probability of the current vertex is defined following Bayesian inference:

   $$p(x|\tilde{x}) = \frac{p_{local}(x) \cdot p(\tilde{x}|x)}{p(\tilde{x})} ,$$

   where the discrete likelihood $p(\tilde{x}|x)$ can be directly derived from the Matrix N given different labels to x and update the $p_{local}(x)$. The resulting quaternary distribution is in practice directly normalized without calculating the margin probability of the neighbor buildings $p(\tilde{x})$.

   2.2 Sample new labels for the further buildings

   $$x_i' \sim p\left(x_1 \middle| x_1', \dots, x_{i-1}', x_{i+1}^{(s-1)}, \dots, x_n^{(s-1)}\right) = p(x_i|\tilde{x}_v', \tilde{x}_{\bar{v}}^{(s-1)})$$

   with $\tilde{x}_v$ the previously labeled neighbors and $\tilde{x}_{\bar{v}}$ the other neighbors.

   2.3 Calculate the overall energy $\mathcal{H}'$ (cf. Section 3.2) according to $\mathcal{K}'$.

3. Accept the new proposal with Metropolis-Hastings criterion

$$A(\mathcal{M}^{(s)}, \mathcal{M}') = min\left\{1, \frac{p(\mathcal{M}'|D)}{p(\mathcal{M}^{(s)}|D)} = \frac{\mathcal{H}'}{\mathcal{H}^{(s)}}\right\}$$

with $p(\mathcal{M}|D)$ the likelihood that the current model fitting the data D, whose ratio can be represented by that of the overall energy $\mathcal{H}$.

4. $\mathcal{M}^{(s+1)} = \mathcal{M}'$ if accepted, otherwise $\mathcal{M}^{(s+1)} = \mathcal{M}^s$.

The search stops when there is no more $\mathcal{H}$ improvement in the last 1000 iterations with the assumption that the overall energy converges.

4. Experiments

Experiments are performed on data-sets of urban areas from the OSM and the official cadastral maps. Figure 6 shows the OSM data of one part of Boston, USA, with 94 buildings. The manually labeled ground truth is given in Figure 1 (b). Figure 6 (c) presents a temporary labeling result based only on local geometric features: 68 out of 94 buildings (72.3%) are correctly identified. Figure 1 (d) shows the final labeling result considering both the local features and the contextual constraints. The classification accuracy is improved to 97.8%.

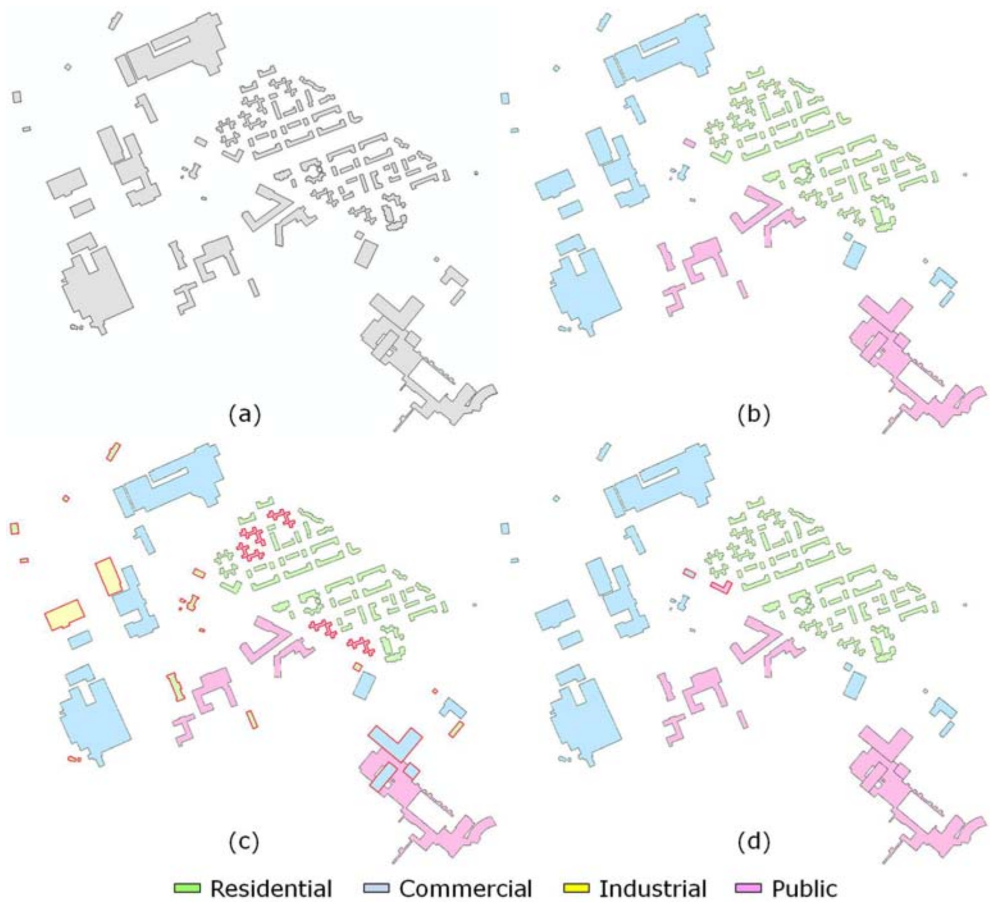


Figure 6. Example data of Boston, USA: (a) OSM data, (b) ground truth labeled manually, (c) labeling based on local features (unary energy) only, and (d) final labeling result considering both unary and binary terms. The incorrectly labeled buildings are highlighted with bold red contours in the results (c) and (d).

Another example for the cadastral map can be found in Figure 7. There are 456 buildings in total in this section of Hanover, Germany. Figure 7 (left) presents the final classification result with the accuracy of 89.7%. The majority of the buildings are correctly labeled with the proposed method. The errors often happen in the classification of the commercial buildings, which is in many cases tricky as they have less distinct geometric characteristics (moderate values of EW and BD, cf. also Figure 4) than the other types. Commercial buildings have, therefore, more possibility to be mislabeled to the other buildings and vice versa.

Please note that the classification is solely based on geometric and topologic criteria of the building footprint, i.e. shape and geographic context of the objects. The use of neighborhood information includes additional knowledge, which improves the classification based only on the characteristics of individual buildings – as shown in Figure 6.

As a matter of fact there is an inherent uncertainty in the classification of building usage, which sometimes makes it difficult or even impossible also for a human to decide on the correct classification – even when additional knowledge is taken into account. E.g., are residential buildings with shops in the ground floor residential or commercial buildings? Is a train station with shops still a public or a commercial building?

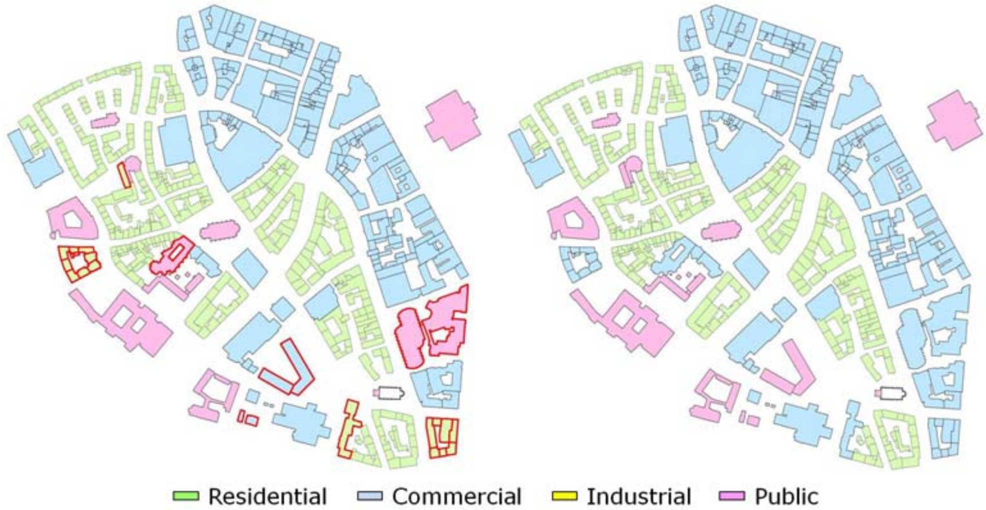


Figure 7. Example data (cadastral map) of Hanover, Germany: the labeling result (left) and the ground truth (right). The incorrectly labeled buildings are highlighted with bold red contours in the result.

5. Conclusion

This paper presents an automatic labeling of building type (use and occupancy) solely based on the building footprint data. A category is predefined with four classes: residential, commercial, industrial and public. We propose two new high-level geometric features: effective width and branching degree, which are designed to quantify the average living space and the structural complexity of the buildings, respectively. MRF is employed to model the network of buildings, in which the local geometric features are given to the vertices that represent the individual buildings while the contextual constraints are embedded in the edges that model the neighborhood relationship. The optimized labeling configuration is statistically searched by means of the Gibbs sampler. The OSM or cadastral maps can thereby be enhanced with the predicted building usage information, which is derived from the existing geometric and topologic features.

In this work we have proposed a general and rather rough category with only four types as our main goal is to explore the potential of using only the building footprint data. Please note that there is actually a wide variety of definitions for the building usage. More concrete classification or definition for specific purposes, e.g., the ten-classes building occupancy classification (International code council, 2006) primarily for the fire code enforcement, could be used. However, then more non-geometric building attributes are required, because of the class definition like "educational" (schools up to the 12th grade) or "high-hazard" (places that product and store flammable or toxic materials). If additional knowledge is available, it can be included, e.g. one could start off with a more elaborate (supervised) classification, and include more contextual knowledge, such as knowledge about the vicinity to other features, which might give an indication concerning the usage. An example would be the inclusion of the knowledge of a market square, which would increase the likelihood that the surrounding buildings are commercial.

For the future work we first consider a new definition of neighborhood, which is essential to the MRF model. In this work we simply use the Voronoi cells to defined the neighbors based only on the centroids of the buildings (cf. Section 3.1). More sophisticated methods like the "constrained Delaunay triangulation" considering the object points and lines (Sester, 2005) can be employed to determine more reasonable neighborhood and thereby improve the labeling result. More experiments can be performed on larger urban areas/whole cities with appropriate pre-partitioning.

Furthermore, the introduced framework for building classification is easily extendable and adaptable: new attributes/measures can be added into both the unary and binary terms to improve the labeling performance or the user can select certain attribute(s) corresponding to their specific task and definition of building types.

References

Aichholzer, O, Aurenhammer, F, Alberts, D, and Gärtner, B (1995) A novel type of skeleton for polygons. J.UCS: Journal of Universal Computer Science, 1(12): 752–761

Haunert, J-H, and Sester, M (2008) Area Collapse and Road Centerlines based on Straight Skeletons. GeoInformtica 12(2): 169-191

Geman, S and Geman, D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6(1): 721–741

International Code Council (2006) International Building Code

Kindermann, R. and Snell J. L. (1980) Markov Random Fields and Their Applications. American Mathematical Society

Li S Z (2009) Markov Random Field Modeling in Image Analysis, Springer.

Lüscher P, Weibel R, Burghardt D (2009) Integrating ontological modelling and Bayesian inference for pattern classification in topographic vector data. Computers, Environment and Urban Systems, 33(5): 363-374

Sester, M. (2005) Optimization approaches for generalization and data abstraction. International Journal of Geographical Information Science Vol. 19, No. 8–9, September–October 2005, 871–897

Werder S, Kieler B, Sester M (2010) Semi-Automatic Interpretation of Buildings and Settlement Areas in User-Generated Spatial Data, ACM-GIS'10, 330-339