

RESEARCH PAPER

Road Geometry Estimation for Urban Semantic Maps using Open Data

Christian Landsiedel* and Dirk Wollherr

Institute for Automatic Control Engineering (LSR)
Technische Universität München

(August 27, 2016)

Complex robotic tasks require the use of knowledge that cannot be acquired with the sensor repertoire of a mobile, autonomous robot alone. For robots navigating in urban environments, geospatial open data repositories such as OpenStreetMap provide a source for such knowledge. We propose the integration of a 3D metric environment representation with the semantic knowledge from such a data base. The application we describe uses street network information from OpenStreetMap to improve street geometry information determined from laser data. This approach is evaluated on a challenging data set of the Munich inner city.

Keywords: spatial reasoning; hybrid maps; scene understanding

1. Introduction

As tasks devolved to robots become ever more complex and encompass more domains, also demands towards their understanding of relationships and autonomy are growing. Different sources of knowledge that can be tapped for a higher-level understanding of concepts and tasks, which is desirable for a more intuitive and user-friendly interaction with a robot, have been explored. Human interaction partners themselves have been used as a knowledge source for example in the IURO project [2], see Figure 1. Other approaches have considered the augmentation of robot knowledge using ontological models in databases that can be shared for learning and for usage by different robots [3, 4]. In this work, we consider OpenStreetMap (OSM), a community-driven online mapping framework, as a source for semantic information for robots moving autonomously in an urban environment. We propose the extension of a hybrid map, which includes a 3D occupancy grid as well as information about street and sidewalk objects in the environment, with semantic and topological information from this data base.

There are multiple reasons why a tighter integration between robot mapping frameworks with data repositories like OpenStreetMap is beneficial. For once, these repositories contain manually selected and curated information, which ensures that it is specified on a level that is understandable to humans and thus usable in interaction, for example for giving or receiving route instructions. Crowdsourcing the data means that additions and modifications to the database are possible for the general public. Thus, the data is updated continuously, and errors can generally be detected and corrected quickly. Furthermore, even state-of the art scene understanding

*Corresponding author. Email: christian.landsiedel@tum.de

⁰Part of the material covered in this article has been presented in workshop form in [1]. This article augments the previously published material by a more thorough description of the model used to infer street segment geometries, which has also been augmented by taking into account intersegment dependencies. Furthermore, the data set used for the evaluation in this paper is greatly extended in comparison to the prior version.



Figure 1. The robot IURO [2, 5] in an urban environment¹.

algorithms primarily rely on assigning labels on a per-pixel or per-region basis, and can have problems at determining distinctions between objects where this distinction happens primarily on a semantic level, i.e., two adjoining rooms with different functions in a space that is not clearly separated, or a building where different parts serve a different purpose. These will be hard to distinguish based on sensor data alone, but the information might be readily available as a bounding box in the OpenStreetMap annotation. On the other hand, the sensor repertoire used in robot mapping approaches will provide up-to-date metric spatial information in the near future, which can be uploaded to Open Data repositories for sharing with humans and other robots. Thus, the benefit of robots using open databases created by and for humans could be mutual.

This paper describes applications and possibilities offered by integrating 3D metric maps with rich semantic and geospatial Open Data repositories. An overview over related approaches in literature is given in Section 2. The data contained in OpenStreetMap that is relevant to this article is described in Section 3, and as an application scenario, it is described how street network information from OpenStreetMap can be used to improve understanding of street geometry based on 3D laser data in Section 4. The approach is evaluated on a challenging data set covering an area in downtown Munich. The results are presented in Section 5.3.

2. Related Work

Data retrieved from OpenStreetMap and similar information sources, in particular information about the topology and layout of the street network, has been used for multiple applications in robotics. An important requirement for the use of geospatial data is knowledge about the location of the robot on a global map, i.e., a solution of the localization problem. Hentschel and Wagner describe a localization method that uses building outlines from OpenStreetMap, which are matched to corresponding features in 3D laser scans [6]. Additionally, the work covers route planning on the OSM route network, and robot behavior control for the robot car's lights based

¹<http://www.iuro-project.eu>

on semantic attributes from OSM. Brubaker et. al. use the OSM route network to localize based on visual odometry data [7]. The localization problem is modeled as a dynamic system, where the state is the vehicle position related to the current route segment, and the visual SLAM trace is the input for filtering. An approach at localisation on the OSM route network with visual SLAM and an initial guess from GPS is presented by Floros et. al. [8]. The result of visual odometry is used as input to a particle filter, where the distribution is pruned based on comparison with the OSM street network. Recently, Ruchti et. al. described localization of a robot on the OpenStreetMap global map based on classification of 3D laser scan point clouds in street and non-street regions in a SLAM framework [9].

Geospatial data from open data repositories has also been used for the applications of place detection and image localization. 3D building geometries from sources similar to OpenStreetMap and vanishing point detection can be used to rectify and align training and query images for place recognition tasks [10]. Li et. al. [11] use 3D point clouds for global registration of images. The 3D point clouds are generated with Structure of Motion techniques from crowdsourced, geotagged monocular images.

The application considered in this paper, estimation of street geometry using information about the location of the street center from Open Data sources is also treated in [12], where aerial images are used as sensor data in combination with a street vector network, and in [13], where street geometry is inferred on the basis of high-resolution multispectral remote sensing satellite imagery. Geiger et. al. present an approach for urban scene understanding based on a generative model for street geometry and topographic information that is based on 3D data constructed from stereo vision recorded by a vehicle traveling on the street [14].

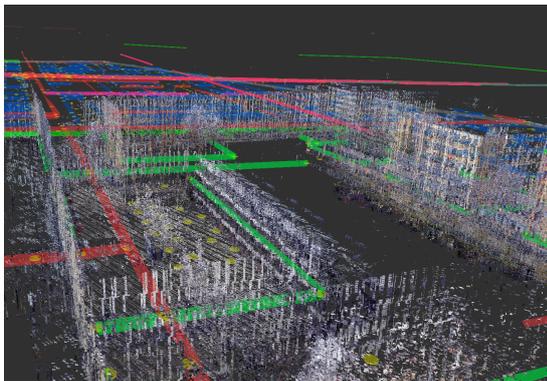
The key difference between these works and the approach presented here is in the data used for estimation. The approach presented here is designed for a static 3D point cloud, where no pose history or dynamics of other dynamic agents (e.g., cars) are available. Moreover, the data is recorded as if from a robot travelling on the sidewalk, such that large parts of the street may be occluded by parked cars or dynamic objects.

3. OpenStreetMap Data Model and Relevant Data

The data model of OpenStreetMap is a graph-like structure, where the basic building blocks are *nodes*, *ways* and *relations*. *Nodes* represent points on the map and are characterized by their latitude and longitude, as well as an optional elevation. *Ways* connect nodes to form open or closed paths and represent spatial entities like the path followed by railroad tracks, building outlines or the area covered by a football field. *Relations* describe higher-level characteristics of sets of nodes and ways, like all buildings belonging to an university campus, or the complete set of streets followed by a bus route. All instances of these three building blocks are identified by globally unique identifiers. Moreover, arbitrary tags can be applied to each instance of these data types, although there is an established set of tags and values that is largely adhered to, which can be used to automatically extract semantic information.

Many features from OpenStreetMap can be easily transferred to a metric map used for robot applications, provided that the transformation between the different global coordinate systems is known. Different localization approaches to address this problem have been proposed as summarized in Section 2, and this transformation is assumed to be known for the purposes of the work presented here. In this case, the mapping of spatial locations allows the transfer of features between the two maps, for example for route planning based on street addresses in an occupancy grid derived from sensor data, or for identifying all buildings belonging to a particular ensemble in a 3D map, as exemplified in Figure 2.

Since the positions of nodes in OSM are based on manual placement, which in turn is based on processed GPS data and aerial imagery, it is difficult to give an accuracy estimate. The errors depend on the accuracy of the recorded GPS data, the number of data points, where there usually



(a) 3D data set overlaid with OpenStreetMap street network



(b) Buildings on the TUM campus, extracted from OpenStreetMap building outlines

Figure 2. Examples for combinations of 3D laser data with additional RGB information and information from OpenStreetMap (best viewed in color). The visualizations of OpenStreetMap data in this paper are created with software based on the `open_street_map` ROS package².

is more data in cities and places of frequent travel, and the diligence and skill of the annotators. Hentschel and Wagner [6] give a visual comparison of a ground truth cadastral map with the building outlines extracted from OpenStreetMap in the context of robotic navigation. This can be interpreted to show errors in the building edges of a few metres. Fan et. al. [15] performed a quantitative analysis of positional accuracy of building outlines for the city of Munich, which, they state, is one of the most developed cities in OpenStreetMap. Their investigation showed an average positional error of 4 m with respect to administrative mapping data. Other research performed on road positions in other parts of the world give similar figures, e.g., [16] and [17]. This relative inaccuracy is one reason for fitting the geometry parameters based on sensor data, since this could be used to improve the positional accuracy of the OpenStreetMap data.

4. Street Geometry Estimation using Street Topology Information

The approach for street geometry information presented here is related to the work by Ruchti et. al. [9], where cells of a 3D laser-based map are classified point-by-point in order to enable localization of a robot in a street network like OpenStreetMap. In the work on semantic mapping presented here, additionally, the modelling imposes a strong geometric consistency constraint – street cells have to be adjacent and located in a strip around the street center. Depending on the intended robotic application, the term ‘street’ can be understood as either only the area of the street that is driven on, including parking spaces on the side of the road, or the combination of this drivable area with the sidewalk directly beside it. The simple model for street geometry used here incorporates both cases, but is not applicable if there is a larger spatial separation between drivable area and sidewalk.

For the work presented here, topological information about the street network is extracted from OpenStreetMap. The goal is to augment this graph with additional metric information in the form of street geometry, which is largely not existing as annotation in the OpenStreetMap database. This relies on the street network data being available and sufficiently accurate. This is the case for the regions considered in the evaluation of this paper, and has also been found to suffice for the different purposes of the other works that use street network data and perform evaluation on data from other parts of the world. However, street *width*, even though the infrastructure (an attribute tag defined for the purpose of annotating it) exists, is not annotated often. In the data set used for evaluation in this paper, only one street segment is annotated with a width tag in

⁰authored by Jack O’Quinn, https://github.com/ros-geographic-info/open_street_map.

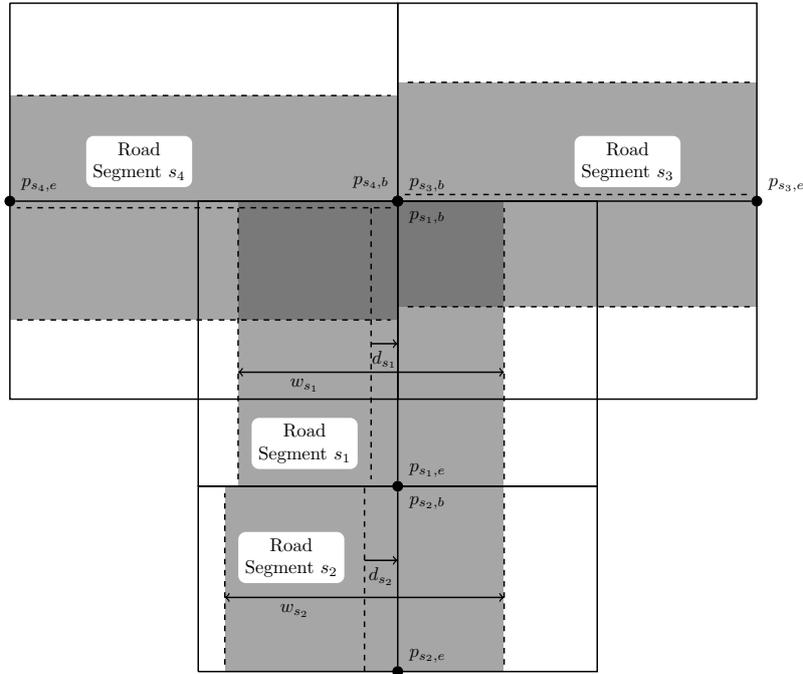


Figure 3. Illustration of street position and width model

the OpenStreetMap data base.

4.1 Modelling Street Geometry Information

Basis for the estimation is the street network from OpenStreetMap, which provides approximate street center lines subdivided into segments of varying length, within which the street is assumed to be straight. In order to reconcile this information with a metric 3D representation, two parameters need to be estimated for each street segment s : The vertical offset d_s of the actual street center from the vector connecting the OSM waypoints $p_{s,b}$ and $p_{s,e}$ defining the street segment in the street network, and the width w_s of the street around this actual center line. This model for the layout is displayed in Figure 3. Let the joint geometrical parameters for segment s be denoted by $\theta_s = (w_s, d_s)$, and the full set of parameters for all segments by Θ . The directions of the street segments from OpenStreetMap are assumed to be in keeping with the actual topology of the environment.

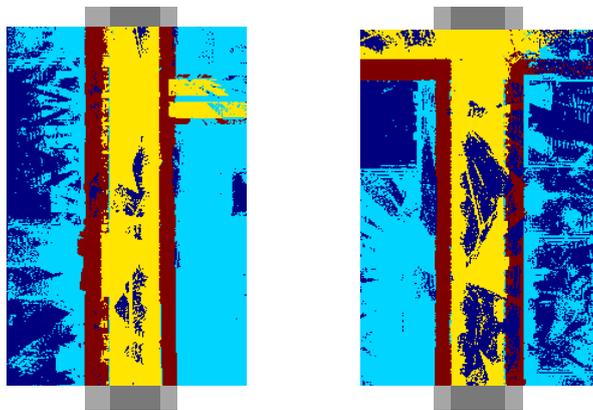
While this simple model of street geometry fits well with the interior of street segments, it does not cover intersection areas, as can be observed in Figure 3. In such areas, the vertical strip of the intersecting street does not conform well with this model. Since intersection points are known from the street network information from OSM, this information can be used to mask intersection areas for the purpose of street geometry estimation. The approach described below considers only areas that are at least 10m away from the middle of an intersection, so that the considered environment can be assumed to have the strip-like geometry expected under this model.

4.2 Inferring Street Geometry from 3D Laser Data and Street Network Information

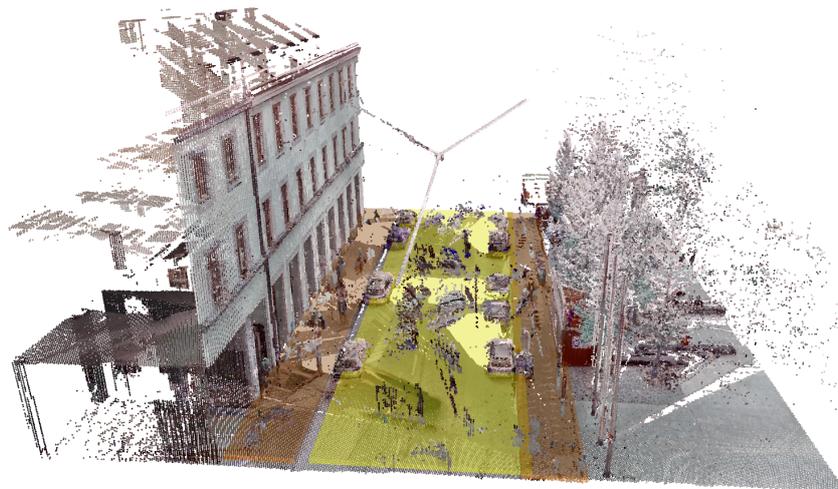
The approach for estimating street width from 3D laser data followed here is a two-step process. Firstly, based on the topology information from OpenStreetMap, areas of interest which contain the street segments are extracted from the point cloud of the covered area. In each of these areas, the information from the point cloud is condensed to a 2D grid, where features are computed for

each bin. A binary classifier provides an estimate about the assignment of each bin to the street or non-street class. Based on these estimates, the geometrical parameters for each street segment are determined by maximizing their probability determined through a Graphical Model.

The connectivity of the street network is encoded in a graph $\mathcal{G} = (\mathbf{P}, \mathbf{S})$. Its edges $s \in \mathbf{S}$ are the street segments in the relevant area extracted from OSM, and the nodes $p \in \mathbf{P}$ are the corresponding waypoints. Each segment connects two waypoints $p_{s,b}$ and $p_{s,e}$. This graph reflects the street topology and is used to model dependencies between the parameters of neighbouring street segments. From this graph, the set of pairs of neighboring segments \mathbf{N} can be derived as $\{\{s_1, s_2\} : s_1, s_2 \in \mathbf{S} \wedge p_{s_1,b} = p_{s_2,e}\}$.



(a) Ground truth annotation for the 2D grid representation – labeled bins (in color) and geometry parameters (grey bars top and bottom)



(b) Section of 3D map with overlaid ground truth annotation

Figure 4. Example annotated street segments. For (a), yellow areas belong to the drivable area, brown areas are on the sidewalk, light blue areas are non-street and for dark blue areas, no features are available because of occlusions. The grey blocks in the background show the annotation of the geometry parameters – light grey for the sidewalk; dark grey for the drivable area of the street. For (b), drivable area and sidewalk points are overlaid over the point cloud in yellow and brown, respectively (best viewed in color).

The positions of the start and end nodes of each segment also determine the area that is considered for estimating the street width. For this task, a candidate environment of a predefined width around each segment center line from the street network is retrieved from the 3D map. For the experiments reported in this paper, a total width of 40 m was chosen. This section of the map is then discretized in the ground plane, such that each segment is divided into a rectangular grid of bins of size $L \times N$, where L is the number of bins in the direction parallel to the street, and

N the number of bins in the considered area vertical to the direction of the street. The length of the sides of the square bins are chosen as 0.2m.

An illustration of this representation of the environment is given in Figure 4. Figure 4(a) shows two projected 2D segments with bins labeled according to their class membership. It also shows the street segment geometries for the drivable area and the sidewalk annotated for these segments. Figure 4(b) shows the full point cloud of one road segment.

For each resulting bin a set of local features is computed. The feature set contains standard geometric and appearance-based features. The geometric features are comprised of the mean, median, standard deviation and absolute range of the z -coordinates of all points projected to each bin as well as the polar angle of a normal vector computed for a small neighborhood around each point. Appearance-based features consist of the same statistics for the intensity values as well as histograms of the recorded color values in each bin in the RGB and HSV color spaces.

Using these features, a baseline classifier is trained to separate between street and non-street bins. For this binary classification problem, the labels are chosen as 0 for non-street bins and 1 for bins classified as belonging to a street. For the experiments described in this paper, a Support Vector Machine with radial basis function kernel is used for this purpose. The result of the classification for segment s is a matrix of labels $\mathbf{Z}_s \in \mathbb{B}^{L \times N}$, containing the classification result for each bin $z_s[i, j]$, $0 \leq i < L, 0 \leq j < N$. The entirety of estimates for all segments is denoted by \mathbf{Z} . Additionally, the confusion matrix \mathbf{C} of the classifier can be determined from the labeled data used for training the classifier.

The classification results then provide candidate information for the second step, which introduces the strong global geometric constraint described above on the inferred street geometry, i.e., that each street segment has straight parallel side lines.

These constraints are formalized in a probabilistic graphical model which encodes both the dependency of the geometrical parameters of a single segment on the estimate provided by the classifier, as well as dependencies of the geometrical parameters of neighbouring segments. The probability of a set of geometrical parameters is modeled by

$$P(\Theta|\mathbf{Z}) = \prod_{s_1, s_2 \in \mathbf{N}} P(\theta_{s_1}, \theta_{s_2}) \prod_{s \in \mathbf{S}} P(\theta_s | \mathbf{Z}_s). \quad (1)$$

The factors in the rightmost product of (1), the *segment geometry potentials*, describe the dependency of a segment geometry on the raw classification result. With the help of the definition of the matrix $\mathbf{X}_s(\theta_s) \in \mathbb{B}^{L \times N}$, which describes the labels assigned to each bin of the segment $x_s(\theta_s)[i, j]$ under a specific geometry θ_s , this can be further developed as

$$\begin{aligned} P(\theta_s | \mathbf{Z}_s) &\propto P(\theta_s) P(\mathbf{Z}_s | \mathbf{X}_s(\theta_s)) \approx P(\theta_s) \prod_{i, j} P(z_s[i, j] | x_s[i, j](\theta_s)) \\ &= P(\theta_s) \prod_{i, j} P(x_s[i, j](\theta_s), z_s[i, j]) / P(x_s[i, j](\theta_s)) \\ &= P(\theta_s) \prod_{i, j} \mathbf{C}[x_s[i, j](\theta_s), z_s[i, j]] / P(x_s[i, j](\theta_s)). \end{aligned} \quad (2)$$

The confusion matrix \mathbf{C} is used as an estimate of the classification error probability, and $P(x_s[i, j](\theta_s))$ denotes the class marginals. This distribution penalizes street segment geometries where many bins receive a label that is different from their initial classification result.

The prior distribution for the street geometry parameters $P(\theta_s)$ is chosen as a product of independent normal distributions for the segment offset and the logarithm of the street width as $P(\theta_s) = \mathcal{N}(d_s | \mu_d, \sigma_d) \mathcal{N}(\log(w_s) | \mu_w, \sigma_w)$, the mean and variance of which are estimated from a training set of segments.

The factors in the first product in (1), the *intersegment potentials*, serve the purpose of relat-

ing the geometrical parameters of neighbouring street segments. By transforming the geometry specified in offset from the segment center and the street width to the left and right boundary of the street according to

$$\begin{aligned} a_s &= d_s - w_s/2 \\ b_s &= d_s + w_s/2, \end{aligned}$$

it is possible to define a measure for the mismatch between the parameters of the two segments as

$$P(\theta_{s_1}, \theta_{s_2}) = \mathcal{N}(a_{s_1} - a_{s_2} | 0, \sigma) \mathcal{N}(b_{s_1} - b_{s_2} | 0, \sigma).$$

The value of σ is set to 0.5 for the experimental evaluation.

Given the model (1), street offset and width for all segments are determined as the parameters which maximise the probability density function as

$$\Theta^* = \underset{\Theta}{\operatorname{argmax}} P(\Theta | \mathbf{Z}). \quad (3)$$

4.3 Inference in the Graphical Model

Exact inference of (3) is intractable for street networks of general structure, since the street network may contain large loops. In order to obtain an approximate solution for the street geometry optimization problem, we use a Markov Chain Monte Carlo (MCMC) approach. For this computation, the values of the potential functions must be determined either analytically or in tabular form for all possible arguments.

While the intersegment potentials are Normal distributions with analytical expressions, the segment geometry potential (2) depend on the classification results in a more complex fashion. This table of values can be computed efficiently using a dynamic programming approach. This is best seen by transforming the probability into its logarithmic form so it can be expressed as a summation over the individual bins (i, j) of the segment grid. Since the computation can be done independently for each segment, the corresponding subscript s is dropped in the following.

$$\begin{aligned} \log(P(\mathbf{Z} | \mathbf{X}(\theta))) &= \log \left\{ \prod_{i,j} \mathbf{C}[x[i, j](\theta), z[i, j]] / P(x_s[i, j](\theta_s)) \right\} \\ &= \sum_{i,j} \log(\mathbf{C}[x[i, j](\theta), z[i, j]] / P(x_s[i, j](\theta_s))) \end{aligned}$$

Replacing the summands in this expression $\log(\mathbf{C}[x[i, j](\theta), z[i, j]] / P(x_s[i, j](\theta_s)))$ with the expression $\tilde{c}(x[i, j](\theta), z[i, j])$, one can use the fact that the labels assigned to each bin are

identical in the regions assigned as sidewalk and street, respectively, to simplify the expression:

$$\begin{aligned}
\log(P(\mathbf{Z}|\mathbf{X}(\theta))) &= \sum_{i,j} \tilde{c}(x[i,j](\theta), z[i,j]) \\
&= \sum_{j=0}^{a-1} \sum_{i=0}^L \tilde{c}(0, z[i,j]) + \sum_{j=a}^b \sum_{i=0}^L \tilde{c}(1, z[i,j]) + \sum_{j=b+1}^N \sum_{i=0}^L \tilde{c}(0, z[i,j]) \\
&= \sum_{j=0}^{a-1} \sum_{v \in \mathbb{B}} \tilde{c}(0, v) \sum_{i=0}^L \delta(z[i,j] = v) + \sum_{j=a}^b \sum_{v \in \mathbb{B}} \tilde{c}(1, v) \sum_{i=0}^L \delta(z[i,j] = v) \\
&\quad + \sum_{j=b+1}^N \sum_{v \in \mathbb{B}} \tilde{c}(0, v) \sum_{i=0}^L \delta(z[i,j] = v)
\end{aligned}$$

with the indicator function δ . Then, each of the inner summations does not depend on the geometry any more, but only on the number of bins that have been classified as sidewalk or street in each row of \mathbf{Z} . Thus, the value table for these factors can be efficiently computed for all possible segment geometries θ .

5. Experimental Evaluation

5.1 Munich Urban 3D Data Set

The data set that was used for experiments is in part overlapping with the one described in our earlier work on spatial relations in semantic maps [2]. It consists of 80 high-resolution laser range finder scans in 3D, acquired with an Z+F 5010C laser range finder, of an area in downtown Munich around the university campus. Additionally, laser intensity and RGB channels are recorded; GPS and odometry data are however not available. In this data set, object instances are manually segmented and annotated for object classes such as *building*, *street*, *sidewalk* or *car*, as well as for qualitative spatial relations, such as *left of* or *behind*, between objects. All OSM street segments covered by the point cloud are annotated with class labels for street, sidewalk or neither of the two on a per-point level, and ground truth street geometry parameters are determined, as shown in Figure 4. This street network comprises a set of 60 route segments with a total length of about 2 km.

The data set provides a challenging environment for scene understanding tasks, since it incorporates a considerable range of different environments, such as residential streets with parked and artefacts of moving cars, tunnels, and cobbled or gravelled streets closed for motor vehicles. Additionally, the laser scans are taken from positions on the sidewalk, such that in many cases the ground plane is not visible because of occlusions or dynamic objects blocking visibility at the time of registering the laser scan.

5.2 Registration of Point Cloud Data with OpenStreetMap

Since the data set is recorded sequentially with no ground truth information about the absolute robot position at the time of recording a scan, nor about the relative movement of the sensor between scans, a registration step is necessary to obtain a complete 3D representation of the area covered by the union of the different laser scans. To this end of estimating the relative transformations between the sensor positions for each recorded 3D scan, registration was carried out with multiple iterations of the 3D Iterative Closest Point algorithm [18], with the maximum allowed correspondence distance decreasing with each iteration, starting from a rough manually defined initial guess. Boundedness of the registration error was ensured by manually labeling key

points for pairs of scenes and monitoring the registration error, and by visual inspection of the registration result. Also for the lack of a global ground truth position data of the laser data, manual alignment of the 3D data with an export of OpenStreetMap data for the region covered by the laser data was carried out. This alignment was based on positioning building outlines in OSM in accordance with vertical surfaces in the point cloud, and the correctness of the alignment was determined by projecting the OpenStreetMap data into the pointcloud data. Note also that the accuracy of this alignment is not critical for the experimental evaluation as long as the road segments of interest in the point cloud are inside the regions of interest defined by the road network. This was verified visually.

Since handling the complete point cloud data for the combination of all laser scans is intractable, the data was filtered and downsampled using the RMAP algorithm [19]. This procedure produces a denoised occupancy grid at a variable resolution, where the grid size was chosen as 0.03 m for the experiments in this paper.

5.3 Experiments

The method for street geometry estimation described above was evaluated on this augmented Munich 3D Urban Data Set. In order to evaluate the benefits and limitations of the method as well as to gauge the influences of the different components of the model, a set of computational experiments with different settings was run.

As goals of the inference, two different applications were investigated: First, the target was to estimate the geometrical parameters of the drivable area of the street alone, counting all surrounding area as non-road. Secondly, the target area was defined to include both sidewalk and the drivable section of the street. The geometric model described in Section 4.1 can be used in both cases; the experiments are only distinguished by the choice of target class in the training of the baseline classifier.

The requirement of labeled training data for the training of the classifier, the calculation of the confidence matrix for the computation of the segment geometry potentials (2), and the parameters of the segment geometry prior requires splitting the data set into a training and a test set for evaluation. In order to be able to evaluate the full model, including the intersegment dependencies, on the complete available graph of street segments, a round robin scheme was adopted for the supervised training. For this, the data set was split into 5 folds, for each of which a classifier was trained on data from the 4 remaining ones. These were used to compute the potential value tables using (4). Then, inference was carried out with the full model including the intersegment dependencies on the full street network. For inference in the full model, 50 chains of MCMC inference were run for 1,000,000 iterations each.

A summary of the results in terms of per-bin retrieval of the correct labels, measured against the ground truth per-bin labelling. This metric, expressed as precision, recall and F_1 score, is given in Table 1. Additionally, a measure for the error of the estimated street widths against the manually labelled ground truth geometry parameters is given by the root mean squared error between the estimated segment width w_s^* and the true annotated width \hat{w}_s , weighted by the length of each segment l_s

$$RMSE_w = \sqrt{\frac{\sum_{s \in S} l_s (w_s^* - \hat{w}_s)^2}{\sum_{s \in S} l_s}}.$$

Different configurations of the method are evaluated. First, the baseline classifier by itself is evaluated (*raw*). Since its result do not include geometry information, no width error is given. Then, the geometries resulting from the probabilistic model are evaluated with (*full model*) and without (*single segment*) including the intersegment potential functions. As an upper bound to the achievable results, the geometric model is fitted to the per-bin ground truth class labels (*opt. on ground truth*). The final evaluation is the evaluation of bin-wise labels implied by the ground

Target Area	Method	P	R	F_1	$RMSE_w[m]$
driving area & sidewalk	raw	0.854	0.854	0.854	N. A.
	single segment	0.877	0.872	0.873	5.67
	full model	0.872	0.869	0.87	2.62
	opt. on ground truth	0.946	0.943	0.943	1.58
	ground truth geometry	0.943	0.937	0.938	0
driving area only	raw	0.815	0.808	0.81	N. A.
	single segment	0.856	0.851	0.852	5.61
	full model	0.876	0.876	0.876	2.65
	opt. on ground truth	0.955	0.953	0.953	1.54
	ground truth geometry	0.951	0.947	0.947	0

Table 1. Database recall metrics and root mean square error of the estimated street widths for the baseline classifier and for the solution including geometric constraints. The upper table contains result for the case where the sidewalk is included in the street; the lower doesn't include the sidewalk in the street area.

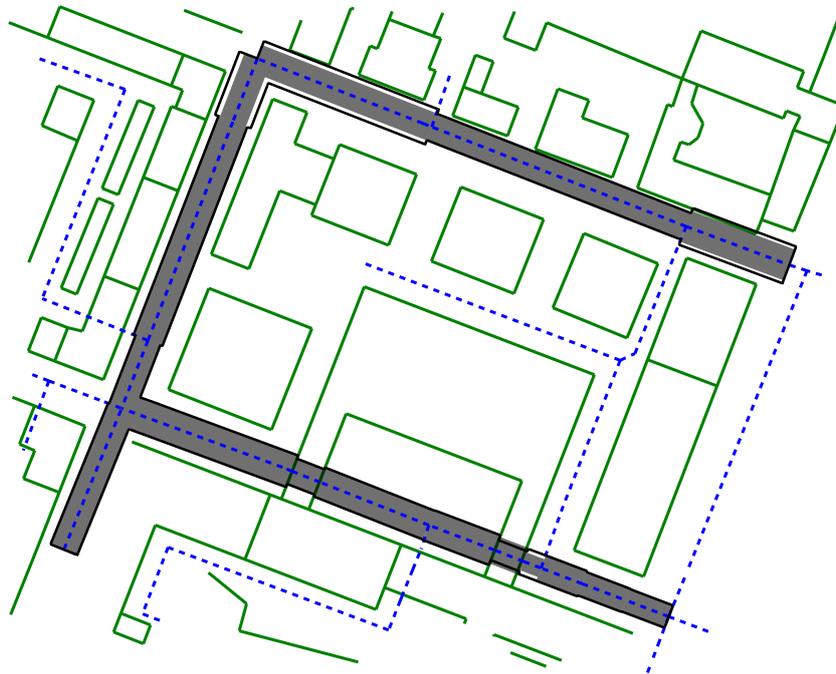


Figure 5. A part of the map with the estimated road geometries. Street center lines and building outlines are drawn as dashed blue and solid green lines, respectively. The estimated road geometries are shown as black outlines and ground truth geometries are shown in grey (image best viewed in color).

truth geometry (*ground truth geometry*). An illustration of the resulting geometries for a part of the data set in comparison with ground truth data is shown in Figure 5.

It can be seen in Table 1 that introducing the geometric constraints improves retrieval metrics of labels for individual regions, as well as it also decreases the error in the street width estimation. An analysis of the failure modes on segments where street geometry estimates exhibited larger errors showed that environments where the street area is directly adjacent to an open space with a surface very similar to the street were difficult to handle for the estimation procedure. Additionally, the data set also contains streets of different categories (i.e., residential urban streets as well as cobbled streets closed for general traffic and without sidewalks as well as tunnels), which again are quite different in nature from a generic scene.

6. Conclusion

In this paper, we have argued the benefits of including information from open geospatial repositories in hybrid maps. The application of street classification and street geometry estimation, parameters which are often missing in OpenStreetMap and could be added automatically from 3D maps, has shown that including a geometric constraint based on OpenStreetMap data provides an improvement in the geometry error over a baseline solution based on classification alone. Experiments have been carried out on a challenging data set, where laser scans have been recorded from the sidewalk, so that the full width of the street is often occluded, and which contains a widely varying array of street types, including tunnels. With the increase in mobile robot platforms navigating in urban scenarios that are equipped with a 3D laser scanners, it is to be expected that different avenues for use of additional information will be explored.

There are several directions in which the work presented here can be extended. Especially in the vein of improving urban scene interpretation by using mapping data from OpenStreetMap would be the use of information about additional properties of streets such as traversability and the existence of bike paths and sidewalks. Furthermore, it can be expected that knowledge about the type of street from the annotation as *residential*, *primary*, *secondary* etc. will be useful if separate models are built and conditioned on the different types of environment.

Acknowledgements

The authors would like to thank Mustafa Sezer, who worked on registration of the individual laser scans to a joint point cloud, and Sheraz Khan for providing assistance with the handling of the point clouds and the RMAP library.

References

- [1] Landsiedel C, Wollherr D. Augmenting and reasoning in semantically enriched maps using open data. In: Workshop on spatial reasoning and interaction for real-world robotics at IROS. 2015.
- [2] Wollherr D, Khan S, Landsiedel C, Buss M. The interactive urban robot IURO: Towards robot action in human environments. In: Hsieh AM, Khatib O, Kumar V, editors. Proc. int. symposium experimental robotics (ISER). Cham, Switzerland: Springer International Publishing. 2016. p. 277–291.
- [3] Tenorth M, Beetz M. Knowrob: A knowledge processing infrastructure for cognition-enabled robots. Int J of Robotics Research. 2013 Apr;32(5):566–590.
- [4] Riazuelo L, Tenorth M, Di Marco D, Salas M, Gálvez-López D, Mösenlechner L, Kunze L, Beetz M, Tardos JD, Montano L, et al.. RoboEarth semantic mapping: A cloud-enabled knowledge-based approach. IEEE Trans Automation Science and Engineering. 2015;12(2):432–443.
- [5] Buss M, Carton D, Khan S, Kühnlenz B, Kühnlenz K, de Nijs R, Turnwald A, Wollherr D. IURO – Soziale Mensch-Roboter-Interaktion in den Straßen von München [IURO – social human-robot interaction in the streets of Munich]. at – Automatisierungstechnik. 2015;German.
- [6] Hentschel M, Wagner B. Autonomous robot navigation based on OpenStreetMap geodata. In: Proc. IEEE int. conf. on intelligent transportation systems. IEEE. 2010. p. 1645–1650.
- [7] Brubaker MA, Geiger A, Urtasun R. Lost! Leveraging the crowd for probabilistic visual self-localization. In: IEEE conf. computer vision and pattern recognition. IEEE. 2013. p. 3057–3064.
- [8] Floros G, van der Zander B, Leibe B. OpenStreetSLAM: Global vehicle localization using OpenStreetMaps. In: Proc. IEEE int. conf. robotics and automation (ICRA). IEEE. 2013. p. 1054–1059.
- [9] Ruchti P, Steder B, Ruhnke M, Burgard W. Localization on OpenStreetMap data using a 3D laser scanner. In: Proc. IEEE int. conf. robotics and automation (ICRA). 2015 May.
- [10] Baatz G, Köser K, Chen D, Grzeszczuk R, Pollefeys M. Leveraging 3D city models for rotation invariant place-of-interest recognition. Int J Computer Vision. 2012;96(3):315–334.
- [11] Li Y, Snavely N, Huttenlocher D, Fua P. Worldwide pose estimation using 3D point clouds. In: Proc. european conf. computer vision (ECCV). Springer. 2012. p. 15–29.

- [12] Yuan J, Cheryadat AM. Road segmentation in aerial images by exploiting road vector data. In: Proc. IEEE int. conf. on computing for geospatial research and application. IEEE. 2013. p. 16–23.
- [13] Chen B, Sun W, Vodacek A. Improving image-based characterization of road junctions, widths, and connectivity by leveraging OpenStreetMap vector map. In: Proc. IEEE int. geoscience and remote sensing symposium. 2014 Jul. p. 4958–4961.
- [14] Geiger A, Lauer M, Urtasun R. A generative model for 3D urban scene understanding from movable platforms. In: IEEE conf. computer vision and pattern recognition. IEEE. 2011. p. 1945–1952.
- [15] Fan H, Zipf A, Fu Q, Neis P. Quality assessment for building footprints data on OpenStreetMap. *Int J Geographical Information Science*. 2014;28(4):700–719.
- [16] Haklay M. How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. *Environment and Planning B: Planning and Design*. 2010; 37(4):682–703.
- [17] Antunes F, Fonte CC, Brovelli MA, Minghini M, Molinari M, Mooney P. Assessing OSM road positional quality with authoritative data. In: *Confêrencia nacional de cartografia e geodesia (nat. conf. cartography and geodesy)*. 2015.
- [18] Besl P, McKay ND. A method for registration of 3-D shapes. *IEEE Trans Pattern Analysis and Machine Intelligence*. 1992 Feb;14(2):239–256.
- [19] Khan S, Dometios A, Verginis C, Tzafestas C, Wollherr D, Buss M. Rmap: a rectangular cuboid approximation framework for 3d environment mapping. *Autonomous Robots*. 2014;37(3):261–277. Available from: <http://dx.doi.org/10.1007/s10514-014-9387-y>.