Conditional Random Fields for Urban Scene Classification with Full Waveform LiDAR Data

Joachim Niemeyer[1], Jan Dirk Wegner[1], Clément Mallet[2],
Franz Rottensteiner[1], and Uwe Soergel[1]

[1] Institute of Photogrammetry and GeoInformation, Leibniz University Hannover,
Nienburger Str. 1, 30167 Hannover, Germany
{niemeyer,wegner,rottensteiner,soergel}@ipi.uni-hannover.de
[2] Laboratoire MATIS, Institut Géographique National, Université Paris Est,
73 avenue de Paris, 94165 Saint-Mandé, France
clement.mallet@ign.fr

Abstract. We propose a context-based classification method for point clouds acquired by full waveform airborne laser scanners. As these devices provide a higher point density and additional information like echo width or type of return, an accurate distinction of several object classes is possible. However, especially in dense urban areas correct labelling is a challenging task. Therefore, we incorporate context knowledge by using Conditional Random Fields. Typical object structures are learned in a training step and improve the results of the point-based classification process. We validate our approach with two real-world datasets and by a comparison to Support Vector Machines and Markov Random Fields.

Keywords: Conditional Random Fields, 3D Point Cloud,
Full Waveform LiDAR, Urban, Classification

1 Introduction

Airborne LiDAR (Light Detection And Ranging) has become a standard technology for the acquisition of elevation data in remote sensing. A new generation of laser scanners captures the full waveform (FW) of the backscattered pulse and provides additional valuable hints for classification purposes, which is mainly used in forestry today. However, object detection in other types of complex scenes can also benefit from FW data. In urban areas the classification of a point cloud is an essential but challenging task, required to generate products such as three-dimensional (3D) city models. Urban scenes include many different objects like buildings of various shapes, hedges, single trees, grouped vegetation, fences, and ground in complex spatial arrangements (Fig. 1). Here, FW data can support the separation of objects that are close to each other. Further improvements can be achieved by considering spatial structures. Thus, the adjacency of some object classes can be modelled to be more likely than other combinations. For example, it is not likely that points labelled as buildings are surrounded by water. In this paper we present a powerful approach for contextual classification of a point cloud obtained with FW laser scanners.

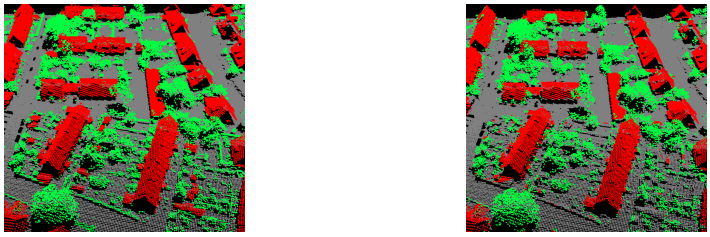

Fig. 1. 3D point cloud of an urban scene. Ground truth (left) and classified points (right). Classes *ground* (grey), *building* (red), and *vegetation* (green)

1.1 Related Work

Many approaches for object extraction from LiDAR data have been developed. Most of them focus on a certain object class like buildings or vegetation. In this case the classification task is reduced to a binary decision process. The classification of an entire point cloud into several object classes is more challenging. In some cases, non-probabilistic approaches like Support Vector Machines (SVM) are used. For instance, in our previous work (Mallet, 2010) a point-based multi-class SVM-classification was applied to FW data. Although SVM are known to work well even with high dimensional feature spaces, this type of approach labels each point without considering its neighbourhood. It may thus lead to inhomogeneous results in complex scenes such as urban areas. Other classifiers such as Random Forests exhibit similar characteristics (Chehata et al., 2009). An improvement can be achieved by incorporating contextual information. Common probabilistic approaches are based on graphical models such as Markov Random Fields (MRF) and Conditional Random Fields (CRF) (cf. Section 2). MRFs have become a standard technique for considering context in classification processes. As the application of graphical models to large LiDAR datasets leads to high computational costs, those context-based approaches became feasible only recently. First research on contextual point cloud labelling was carried out in the fields of robotics and mobile terrestrial laser scanning. Anguelov et al. (2005) proposed a classification of a terrestrial point cloud into four object classes with Associated Markov Networks (AMN), a subclass of MRF. This approach is based on a segmentation, assigning all points within a segment to the same class. Neighbouring points are assumed to belong to the same object with high probability. Thus, an adaptive smoothing of classification results is performed. In order to reduce the number of graph nodes, ground points are eliminated based on thresholds before the actual classification. Munoz et al. (2008) also used point-based AMNs, but they extended the original isotropic model to an anisotropic one, in order to emphasize certain orientations of edges. This directional information enables a more accurate classification of objects like power lines. Rusu et al. (2009) were interested in labelling an indoor robot environment described by point clouds. For object detection points are classified using CRFs according to the geometric surface they belong to, such as cylinders or planes.

They applied a point-wise classification method, representing every point as a node of the graphical model. Compared to our application they deal with very few points (≈ 80000), and they even reduce this dataset by about 70% before the classification based on some restrictions concerning the objects' positions. CRF were also used by Lim and Suter (2007) for point-wise classification of terrestrial LiDAR data. They limit the computational complexity by adaptive point reduction. The same authors improved their approach (2009) by segmenting the points in a previous step and classifying these 'super-pixels' afterwards. They also considered both a local and a regional neighbourhood. An approach utilizing CRFs for remotely sensed LiDAR point clouds is presented in (Shapovalov et al., 2010). They classify a point cloud obtained from airborne laser scanning. Similar to the work mentioned previously, a segmentation is carried out in a preliminary step by k-means clustering. This step helps to cope with noise. However, small objects are not represented in the segmentation because they are merged with larger ones. As a consequence, important object details might be lost. A second restriction of the approach is feature selection. The authors solely use geometric features such as angles, spin images, height distributions, and shape parameters. On the one hand, this method is flexible because these features can be calculated for every point of a cloud, but on the other hand, usually additional point attributes such as intensity values are available and can be incorporated to improve the classification. Today, more features can be derived by the FW laser scanning technique. A combination of geometrical features and those obtained by FW analysis is expected to result in a more robust classification.

1.2 Contribution

We propose a new probabilistic approach for classification of LiDAR data by applying CRFs to point clouds. It combines three main advantages to improve the classification accuracy. Firstly, we utilize point clouds acquired by FW sensors. The point attributes like the echo width provide additional features and enable a more reliable separation of object classes. Secondly, we do not perform a segmentation, which usually leads to smoothing effects and loss of information, but we apply a point-based classification to preserve even small objects. Thirdly, CRFs provide a flexible classification framework that is also capable to model the spatial structure of the data. The proposed supervised classifier is able to learn context to find the most likely label configurations. Due to these three characteristics, our framework yields good results even in urban areas, which will be shown by an evaluation of the method.

2 Conditional Random Fields

Conditional Random Fields (CRF) have originally been proposed by Lafferty et al. (2001) to label sequential data. They are probabilistic models and thus provide probabilities instead of crisp decisions, which is a major benefit in terms

of post-processing. CRFs belong to the family of graphical models as, for example, Markov Random Fields (MRF), and represent data as a graph $G(\boldsymbol{n}, \boldsymbol{e})$ consisting of nodes $\boldsymbol{n}$ and edges $\boldsymbol{e}$. The nodes $\boldsymbol{n}$ can represent spatial units such as pixels in images or points of point clouds. In our case, each node n_i corresponds to a single point i of the LiDAR point cloud. The edges $\boldsymbol{e}$ link pairs of adjacent nodes n_i and n_j in the graph; they can be used to represent contextual knowledge. The goal of our classification is to assign class labels $\mathbf{y}$ to all nodes $\boldsymbol{n}$ in $G(\boldsymbol{n}, \boldsymbol{e})$ based on data $\boldsymbol{x}$. Here, $\mathbf{y}$ is a vector whose element y_i corresponds to the class label of the node n_i. In the multi-class case, the labels can take more than two values, each corresponding to a single object class. Thus, the class labels of all nodes are determined simultaneously in this process, and multiple object classes are discerned.

In order to highlight the theoretical differences between MRFs and CRFs we will first give a short review of MRFs. A MRF is a *generative* approach that models the joint probabilities $P(\boldsymbol{x}, \boldsymbol{y})$ of the data and the labels, in a way similar to the Bayesian Theorem:

$$P(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{Z(\boldsymbol{x})} \exp\left(\sum_{n_i \in \boldsymbol{n}} \log P_i(x_i|y_i) + \sum_{n_i \in \boldsymbol{n}} \sum_{n_i \in N_i} \beta \cdot \delta(y_i = y_j)\right) \quad . \tag{1}$$

In (1), the joint distribution $P(\boldsymbol{x}, \boldsymbol{y})$ is decomposed into a product of factors $P(\boldsymbol{x}|\boldsymbol{y}) \cdot P(\boldsymbol{y})$, whereas the probability of the data is represented by the partition function $Z(\boldsymbol{x})$, which acts as normalisation constant turning potentials into probabilities. The difference to the Bayes Theorem is that the prior $P(\boldsymbol{y})$ is replaced by a random field over the labels. The data observed at node n_i are represented by x_i, and P_i is the likelihood function for node n_i. N_i is the set of all nodes n_j linked to a particular node n_i by an edge. The function $\delta(y_i = y_j)$ yields +1 if $y_i = y_j$ and -1 otherwise. Hence, the prior term $\beta \cdot \delta(y_i = y_j)$ will increase the posterior probability if neighbouring labels are identical by a weight $\boldsymbol{\beta} > 0$, whereas it will penalize the inequality of neighbouring labels. Both will lead to a smoothed classification (Kumar and Hebert, 2006).

CRFs, unlike MRFs, are *discriminative* models that model the posterior distribution $P(\boldsymbol{y}|\boldsymbol{x})$ directly, leading to a reduced model complexity compared to MRFs (Kumar and Hebert, 2006):

$$P(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{Z(\boldsymbol{x})} \exp\left(\sum_{n_i \in \boldsymbol{n}} A_i(\boldsymbol{x}, y_i) + \sum_{n_i \in \boldsymbol{n}} \sum_{n_j \in N_i} I_{ij}(\boldsymbol{x}, y_i, y_j)\right) \quad . \tag{2}$$

In (2), the same symbols are used as in (1). $A_i(\boldsymbol{x}, y_i)$ is called the association potential. It links the data to the class labels and can be interpreted in a similar way as the data term in the MRF model. The difference is that the association potential of a single node n_i may depend on all the data $\boldsymbol{x}$ rather than only on the data observed at the node n_i. The assumption of conditionally independent

features of different nodes inherent in the model described by (1) is relaxed for CRFs. $I_{ij}(\boldsymbol{x}, y_i, y_j)$ is called the pairwise interaction potential. Unlike the prior term in (1), it is not only dependent on the class labels at the two nodes linked by an edge, but also on potentially all the observed data $\boldsymbol{x}$. For instance, this property can be used to compensate for the smoothing effect of MRFs, if the observed data indicate that the class labels of two nodes are different. This along with the possibility to integrate the entire data $\boldsymbol{x}$ makes CRF a very flexible technique for contextual classification. The model in (2) only provides a framework that can be filled by various probabilistic classifiers for the potentials. We use generalized linear models (GLM) for both the association potential A_i and the interaction potential I_{ij}:

$$A_i(\boldsymbol{x}, y_i) = \log P'(y_i|\boldsymbol{x}) \text{ where } P'(y_i = l|\boldsymbol{x}) \propto \exp\left(\boldsymbol{w}_l^T \cdot \boldsymbol{h}_i(\boldsymbol{x})\right) \tag{3}$$

$$I_{ij}(\boldsymbol{x}, y_i, y_j) = \begin{cases} \boldsymbol{v}_l^T \cdot \boldsymbol{\mu}_{ij}(\boldsymbol{x}) & \text{if } y_i = y_j = l \\ 0 & \text{if } y_i \neq y_j \end{cases}. \tag{4}$$

In (3), $\boldsymbol{h}_i(\boldsymbol{x})$ is a feature vector of node n_i computed from data $\boldsymbol{x}$. As stated above, we are not restricted to data only of that node (as in the MRF case) but may also include data of other nodes. $P'(y_i = l|\boldsymbol{x})$ is the probability for the label of node n_i taking the value l given the data. $\boldsymbol{w}_l$ is a vector of feature weights for class l. There is one such vector per class, which has to be determined by training. In (4), $\boldsymbol{\mu}_{ij}(\boldsymbol{x})$ is an edge feature vector, determined for each edge in the graph and potentially depending on all observations, and $\boldsymbol{v}_l$ is an associated vector of feature weights that has to be learned from training data. Note that the weights are only required if the two labels are identical, so that there is again one such weight vector per class. As the MRF model in (1), the model for the interaction potentials (4) will have a smoothing effect, but the degree of smoothing is dependent on features $\boldsymbol{\mu}_{ij}(\boldsymbol{x})$: if the difference of the node features $\boldsymbol{h}_i(\boldsymbol{x})$ and $\boldsymbol{h}_j(\boldsymbol{x})$ is used to define $\boldsymbol{\mu}_{ij}(\boldsymbol{x})$, a high difference could suppress smoothing. This leads to small objects changes being better preserved if there is sufficient evidence in the data, which is a major advantage in complex urban areas.

3 CRFs applied to 3D-FW Point Clouds

The classification is performed on point clouds acquired by airborne FW laser scanners. We are interested in the object classes *ground*, *building* and *vegetation*. In Section 3.1 we discuss the advantages of FW data for classification. Due to the irregular distribution of the LiDAR points, an adaptive graph structure is required for the CRF, which is presented in Section 3.2. The features are described in Section 3.3. Finally, training and inference are discussed in Section 3.4.

3.1 Full waveform LiDAR

Classical laser scanners only provide the coordinates of first and last echoes. A new generation of LiDAR sensors is able to record the entire signal of the

backscattered energy. Thus, information about any illuminated objects in the path of the pulse is contained in this waveform. We operate on pre-processed data in terms of a point cloud derived from Gaussian decomposition of raw waveforms (Wagner et al., 2006). Apart from the position of the echoes (corresponding to peaks in the waveforms), additional indicators related to the geometry and the radiometric properties of the illuminated objects can be the amplitude A and the width σ of the Gaussians (Fig. 2). In particular, applications dealing with vegetation extraction benefit from this technology: the echo width and the fact that there are frequently multiple echoes have turned out to be discriminative features for vegetation (Chehata et al., 2009) and even enable single tree detection and tree species classification. Thus, most of the research in FW laser scanning has dealt with forestry applications. However, classifications of complex urban scenes can benefit from this additional information as well (Mallet, 2010). The echo width, for instance, is related to the slope of the illuminated object and to the local surface roughness because it is broadened on slanted or rough surfaces. Consequently it can help to separate oblique roofs and flat ground points or smooth objects from vegetation. For classification tasks, multiple echoes as well as the local distribution of points also provide hints for points belonging to a certain object type. For example, a volumetric distribution indicates trees, whereas points distributed on a plane may be associated to ground or roofs.

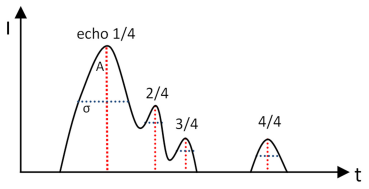

Fig. 2. Backscattered energy recorded by full waveform laser scanner. Gaussians are fitted to the sampled signal. Features amplitude A and width σ can be used to characterise echoes.

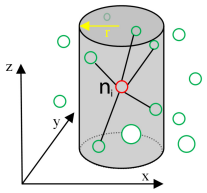

Fig. 3. Generation of graph structure. All points within a vertical cylinder with radius r are linked by edges to the currently investigated point n_i in the centre.

3.2 Graph architecture

Compared to images a point cloud is more complex, because points are irregularly distributed in 3D space. There is no straightforward definition of the neighbourhood that can be used to define the edges of the graph. In contrast, images are arranged in a lattice, and each pixel has a defined number of neighbours (usually four or eight). Thus, a graph structure with an adaptive number of edges per node is required for point clouds in order to capture locally varying point densities. In our application, each point (node) n_i is linked to the points located within a vertical cylinder of radius r whose axis passes through

n_i (Fig. 3). This radius, which has to be specified by the user, and the point density determine the number of neighbours of each node and thus the number of edges emanating from a node. The impact of varying radii r is analysed in Section 4.4. We chose a cylinder for defining the local neighbourhood in order to make sure that points on tree canopies and points near the roof edges will have links to ground points. This will be very helpful in classification because, for instance, the local height differences can be used as edge features, which can give valuable hints for the local configuration of classes. An alternative definition of the local neighbourhood, e.g. a spherical neighbourhood or k nearest neighbours, could not ensure this important information to be preserved. In any case, based on the cylindrical neighbourhood, we construct a graphical model with a varying number of edges per node according to the local point density. This results in a fairly complex graph structure.

3.3 Features

For each 3D point, a set of geometrical and FW features can be determined. In our previous work (Chehata et al., 2009), the contribution of these features to urban point cloud classification using Random Forests was analysed. Based on the experience gained in that work, we selected ten features for classification in our CRF-based method: amplitude, echo width, normalized echo number, height difference between first and last echo, distance to ground, variances of normal vectors and elevations, residuals of an estimated plane as well as the features omnivariance and planarity based on the eigenvalues of the local scatter matrix. A detailed description of the features can be found in (Chehata et al., 2009). In addition, a bias parameter of value one is added to each feature vector (Kumar and Hebert, 2006). These features are determined for each node n_i and represent the node feature vectors $\boldsymbol{h}_i(\boldsymbol{x})$ in the association potentials (3). The edge features $\boldsymbol{\mu}_{ij}(\boldsymbol{x})$ in the interaction potentials (4) are defined as the differences of the two adjacent node feature vectors, thus $\boldsymbol{\mu}_{ij}(\boldsymbol{x}) = \boldsymbol{h}_i(\boldsymbol{x}) - \boldsymbol{h}_j(\boldsymbol{x})$.

3.4 Training and Inference

Due to the fact that our graph $G(\boldsymbol{n}, \boldsymbol{e})$ contains cycles, exact training and inference is computationally intractable, so that approximative methods have to be applied in both cases.

The weight vectors $\boldsymbol{w}_l$ and $\boldsymbol{v}_l$ for the node and edge features in the association (3) and interaction potentials (4) are derived by training. For that purpose, a fully labelled point cloud is required. The training task is a non-linear numerical optimisation problem, in which a cost function is minimized. Following Vishwanathan et al. (2006), we use the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method for optimisation of the objective function $f = -\log(P(\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{y}))$, where $\boldsymbol{\theta}$ is a parameter vector containing the weight vectors $\boldsymbol{w}_l$ and $\boldsymbol{v}_l$ for all classes l. The L-BFGS method requires the computation of the objective function and its gradient as well as an estimation of the partition function $Z(\boldsymbol{x})$ (2) in each iteration; we use the method described in

(Vishwanathan et al., 2006) in the variant with Loopy Belief Propagation (LBP) (Frey and MacKay, 1998) for that purpose.

LBP, a standard iterative message passing algorithm for graphs with cycles, is also used for inference, i.e., for determining the optimum configuration of labels based on maximising $P(\boldsymbol{y}|\boldsymbol{x})$ (2). Although this technique does not ensure convergence to the global optimum, it has been shown to provide good results.

The result of inference is a probability value per class for each 3D point. Maximum a posteriori (MAP) selects the highest probability and assigns the corresponding object class label to the point.

4 Experiments

In this section we present experiments in order to evaluate the classification performance of our proposed method. First, we describe the datasets. Then, the point-wise CRF classification results are assessed on two datasets. For performance validation a comparison to state-of-the-art classifiers SVM and MRF is done. An analysis of the impact of varying neighbourhood size keeping in mind the computational costs is given afterwards. This section finishes with a discussion of our approach.

4.1 Datasets

We evaluate our algorithm on two airborne LiDAR datasets, acquired by a RIEGL LMS-Q560 FW laser scanner (under leaf-on conditions), depicted in Fig. 4. The first one, Bonnland, is a rural scene in Germany of about 2.02 ha, and consists of scattered buildings, streets, single trees, forested areas, and grassland. 759,790 points were recorded. The other dataset (553,033 points, 1.93 ha), was acquired in Kiel, also situated in Germany, and has a suburban character with regularly arranged buildings, streets, and lots of vegetation. The point density is about 3.7 $\frac{points}{m^2}$ and 2.9 $\frac{points}{m^2}$, respectively. We fixed the cylinder radius according to (5) (Section 4.4) to 0.75 m (Bonnland) and 1 m (Kiel).

Ground truth was generated by labelling both point clouds manually. This fully classified 3D data can be considered to be almost accurate and is used for training and validation.

4.2 Results of CRF-Classification

For evaluation purposes we divided each dataset into three parts of nearly the same sizes of about 250,000 points (Bonnland) and 184,330 points (Kiel). We performed three-fold cross-validation for both datasets. Parameters are trained on two parts of the data and tested on the third. This is done three times (i.e., folds), each time with another combination of the data subsets. Finally, we report the mean performance parameters of all three trials. Experiments show that our approach leads to good results. Two subsets of the classification results are illustrated in Fig. 4. In particular the overall accuracy for dataset Bonnland

is 94.3 % (Table 1) with a standard deviation of 0.2 % for the three trials. For all three classes completeness and correctness rates are above 88 %. The overall accuracy in Kiel amounts to 89.3 % (standard deviation 0.6 %).

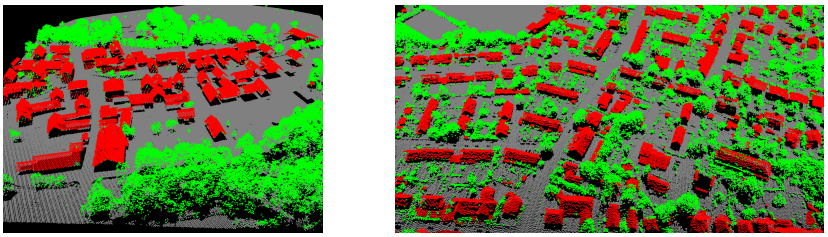

Fig. 4. Classification results of datasets Bonnland (*left*) and Kiel (*right*).

Table 1. Classification performance with completeness and correctness rates [%].

Bonnland		SVM	MRF	CRF
ground	corr.	97.4	94.5	95.5
	comp.	85.5	95.6	95.5
building	corr.	28.9	94.3	88.1
	comp.	82.0	81.7	88.2
vegetation	corr.	90.9	92.5	94.0
	comp.	71.6	93.4	93.9
overall accuracy		**80.3**	**93.8**	**94.3**

Kiel		SVM	MRF	CRF
ground	corr.	78.2	86.6	91.1
	comp.	93.1	94.4	93.5
building	corr.	52.5	93.1	85.7
	comp.	67.8	75.2	86.6
vegetation	corr.	81.6	85.2	88.2
	comp.	26.0	81.4	83.4
overall accuracy		**74.3**	**87.1**	**89.3**

4.3 Comparison to SVM and MRF

In order to evaluate the effectiveness of our CRF model, we compared the results of our method to those obtained by two state-of-the-art classifiers: SVM and MRF. SVM are a common approach for a large variety of classification tasks. This non-probabilistic method yields good performance, but a drawback is the very local functional principle. MRF overcome this disadvantage by incorporating contextual information into the classification process. We used GLM and Potts model as prior terms for MRF. Table 1 depicts the results of the comparison. As expected, the non-probabilistic classification based on SVM (with Gaussian kernel) leads to the worst results. Note that there are many class assignment errors because a considerable number of vegetation points are labelled as *buildings*. Thus, the correctness rate for building points is only 28 % in Bonnland. In addition, several errors occur on roofs. Due to the point-based classification, no context is incorporated (despite of feature calculation within a small local neighbourhood) and it is not possible to separate flat objects into

roof or ground. An improved classification performance is obtained with MRFs. In case of Bonnland, the overall accuracy increases by 13.5 percentage (93.8 % vs. 80.3 %). Although this rate is obviously better, especially the completeness rate of object class *building* is relatively low (81 % Bonnland and 75 % Kiel). CRFs further improve the results. Fig. 5 gives an example for the different performances. It shows a building with oblique and flat roof parts (highlighted with arrows). Especially the flat-roofed part is difficult to classify due to features similar to ground. SVM and MRF cannot detect it, whereas CRF extracts the roofs nearly without errors. It benefits from the more general model and enables a correct classification.

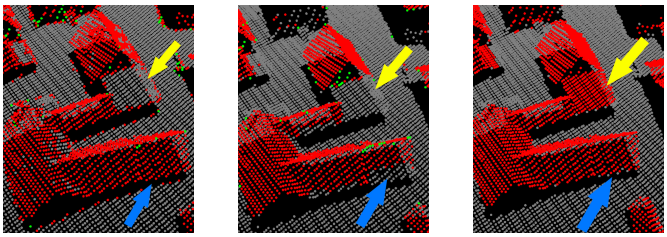

Fig. 5. Comparison of results obtained from SVM, MRF and CRF classification. Yellow arrows show a flat roof, whereas blue arrows mark an oblique roof.

4.4 Optimal neighbourhood size

A critical parameter is the radius r of the cylinder used for graph construction. It is the only one that has to be specified manually. As the number of neighbours within this radius is strongly correlated to the point density of the scene, we analysed this number directly. Fig. 6 depicts the relation of overall accuracy and the number of linked edges, which corresponds to the number of neighbours. We performed tests for 1-10, 15 and 20 neighbours per node. It can be seen, that increasing the number of edges leads to slightly increased overall accuracy. The best result is achieved for 20 edges. However, since the computational complexity increases considerably with higher number of edges, we decided to use the local maximum at 7 neighbours for the experiments. The relation between point density, number of edges and cylinder radius r is given by

$$r = \sqrt{\frac{\text{number of edges}}{\text{point density} \cdot \pi}} \quad . \tag{5}$$

4.5 Computational Costs

The high degree of obtained details with the CRF-classification is related to a higher computational burden. Consisting of many nodes and edges, CRF applied

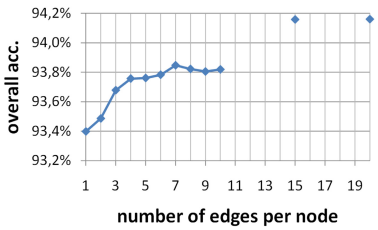

Fig. 6. Overall accuracy related to the number of edges per node.

to point clouds are computationally expensive. Especially the training is time intensive due to the iterative optimisation steps, whereas the final classification time is negligible. For a training area of Kiel consisting of 403,221 nodes and 1,775,582 edges with 11 features each, 253 min were required. The classification of the unlabelled point cloud with 177,343 nodes and 770,735 edges took 0.9 min. In contrast, parameter training for MRF of the same area with the same number of nodes and edges took 202 min. We processed the tests on a machine with a 2.8 GHz Quad-Core CPU and 16 GB RAM.

4.6 Discussion

Our experiments revealed that CRFs show the best performance for point cloud labelling. In contrast to non-probabilistic local classifiers like SVM the accuracy can be improved using graphical models that take into account context information. Building roofs, for example, are classified correctly (Fig. 5, yellow arrows). Compared to a standard MRF our approach further improved the overall accuracy slightly, in case of Kiel by 2.2 percentage. In particular, completeness rates of classes *building* and *vegetation* benefit from using CRFs (Table 1). Complex objects like the flat roof in Fig. 5 (yellow arrows) can be classified correctly. Although it is computationally costly in terms of time and memory, this approach provides a FW point cloud classification with high accuracies and preserves details due to the point based method. Thus it is suitable for advanced requirements in remote sensing. An additional advantage of CRFs are the directly resulting probabilities, which allow the introduction of a separate class for points with low posteriors $P(\boldsymbol{y}|\boldsymbol{x})$.

5 Conclusions and Outlook

In this paper we have addressed the task of 3D point cloud classification. We applied the probabilistic machine learning approach Conditional Random Fields (CRF) to label several object classes, namely *ground*, *buildings*, and *vegetation*. By incorporating local context and full waveform features classification performance is improved. Our method shows good results for our two datasets of Bonnland and Kiel. Overall accuracy rates yield 94.3 % and 89.3 %, respectively. It

is demonstrated that the results are more accurate compared to state-of-the-art classifiers Support Vector Machines (SVM) and Markov Random Fields (MRF). Future work will comprise a more sophisticated formulation of contextual knowledge. In urban areas, for instance, parcels are often bounded by hedges or fences. These typical arrangements can be used to improve classification results further. Moreover, we are going to validate other training and inference methods.

Acknowledgment We used the 'Matlab toolbox UGM', developed by Mark Schmidt, `www.cs.ubc.ca/~schmidtm/Software/UGM.html`. This work is funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG).

References

Anguelov, D., Taskar, B., Chatalbashev, V., Koller, D., Gupta, D., Heitz, G., Ng, A.: Discriminative learning of markov random fields for segmentation of 3d scan data. In: Proc. of IEEE CVPR. vol. 2, pp. 169–176 (2005)

Chehata, N., Guo, L., Mallet, C.: Airborne lidar feature selection for urban classification using random forests. In: IntArchPhRS. vol. 38-3, pp. 207–212 (2009)

Frey, B., MacKay, D.: A revolution: Belief propagation in graphs with cycles. In: Advances in Neural Information Processing Systems 10. pp. 479–485 (1998)

Kumar, S., Hebert, M.: Discriminative random fields. International Journal of Computer Vision 68(2), 179–201 (2006)

Lafferty, J., McCallum, A., Pereira, F.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Proc. 18^{th} Int. Conference on Machine Learning. pp. 282–289 (2001)

Lim, E., Suter, D.: Conditional random field for 3d point clouds with adaptive data reduction. In: Int. Conference on Cyberworlds. pp. 404–408 (2007)

Lim, E., Suter, D.: 3d terrestrial lidar classifications with super-voxels and multi-scale conditional random fields. Comp. Aided Design 41(10), 701–710 (2009)

Mallet, C.: Analysis of Full-Waveform Lidar Data for Urban Area Mapping. Ph.D. thesis, Télécom ParisTech (2010)

Munoz, D., Vandapel, N., Hebert, M.: Directional associative markov network for 3-d point cloud classification. In: Int. Symposium 3DPVT (2008)

Rusu, R., Holzbach, A., Blodow, N., Beetz, M.: Fast geometric point labeling using conditional random fields. In: IEEE IROS. pp. 7–12 (2009)

Shapovalov, R., Velizhev, A., Barinova, O.: Non-associative markov networks for 3d point cloud classification. In: PCV 2010. IntArchPhRS, vol. 38(A), pp. 103–108 (2010)

Vishwanathan, S., Schraudolph, N., Schmidt, M., Murphy, K.: Accelerated training of conditional random fields with stochastic gradient methods. In: 23^{rd} Int. Conference on Machine learning. pp. 969–976 (2006)

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., Studnicka, N.: Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry and Remote Sensing 60(2), 100–112 (2006)