Efficient Interpretation of 3D Point Clouds by Assessing Feature Relevance

Effiziente Interpretation von 3D-Punktwolken durch die Abschätzung der Relevanz von Merkmalen

Martin Weinmann, Clément Mallet, Stefan Hinz, Boris Jutzi

The semantic interpretation of 3D point cloud data acquired with mobile laser scanning (MLS) systems has become a topic of major interest for photogrammetry, remote sensing and computer vision. In this paper, we propose a methodology for the semantic interpretation of point cloud data in terms of assigning each 3D point a semantic label. Our methodology involves (1) individual neighborhoods of optimal size in order to provide distinctive geometric features for each 3D point and (2) feature relevance assessment in order to reduce the computational burden with respect to processing time and memory consumption. More specifically, our approach for feature relevance assessment relies on a general relevance metric composed of seven different, classifier-independent feature selection strategies and thus addresses different intrinsic properties of the given training data. The results derived for a labeled benchmark dataset with about 1.3 million 3D points reveal that, instead of including as many features as possible in order to compensate a lack of knowledge about scene and data, a crucial task such as the semantic scene interpretation can be carried out with only few relevant features without a significant loss in classification accuracy.

Keywords: Mobile laser scanning, point cloud, feature extraction, feature relevance, classification, interpretation

Die semantische Interpretation von 3D-Punktwolken, welche mithilfe von mobilen Laserscanning (MLS) Systemen aufgenommen werden, ist in den Bereichen der Photogrammetrie, Fernerkundung und Computer Vision zu einem wichtigen Thema geworden. In diesem Beitrag wird eine Methodik zur semantischen Interpretation von Punktwolken im Sinne der Zuordnung einer semantischen Klassenzugehörigkeit zu jedem 3D-Punkt präsentiert. Diese Methodik nutzt (1) individuelle Nachbarschaften von optimaler Größe, um für jeden 3D-Punkt aussagekräftige geometrische Merkmale zu erhalten, und (2) eine Abschätzung der Relevanz von Merkmalen, um den Rechenaufwand bezüglich Rechenzeit und Speicherbedarf zu reduzieren. Im Speziellen basiert der Ansatz zur Abschätzung der Relevanz von Merkmalen auf einer allgemeinen Relevanz-Metrik, welche sich aus sieben verschiedenen Klassifikator-unabhängigen Strategien zur Auswahl von Merkmalen zusammensetzt und dadurch verschiedene intrinsische Eigenschaften der vorliegenden Trainingsdaten erfasst. Die Ergebnisse für einen Standarddatensatz mit etwa 1,3 Millionen 3D-Punkten und entsprechenden Referenzwerten bezüglich der Klassenzugehörigkeit zeigen, dass statt der Nutzung von möglichst vielen Merkmalen, um fehlendes Wissen über Szene und Daten zu kompensieren, mit nur wenigen relevanten Merkmalen ein kritischer Schritt wie die semantische Szeneninterpretation durchgeführt werden kann, ohne die Klassifikationsgenauigkeit wesentlich zu verringern.

Schlüsselwörter: Mobiles Laserscanning, Punktwolke, Merkmalsextraktion, Relevanz von Merkmalen, Klassifikation, Interpretation

1 INTRODUCTION

Due to the technological advancements in the last decade, mobile laser scanning (MLS) systems are increasingly used for fast, dense and reliable 3D mapping. The acquired three-dimensional representation of a scene in the form of point cloud data with a relatively high spatial resolution, in turn, represents an important prerequisite for further tasks such as 3D scene analysis which typically relies on the automatic interpretation of the acquired data. Consequently, the automatic interpretation of 3D point cloud data has become a topic of major interest in photogrammetry, remote sensing and computer vision.

Recent research involving mobile laser scanning data for 3D scene analysis focuses on a variety of subtasks such as object detection (e. g. /Pu et al. 2011/, /Velizhev et al. 2012/ or /Zhou & Vosselman 2012/), urban accessibility analysis /Serna & Marcotegui 2013/ or the semantic perception for ground robotics /Hebert et al. 2012/. However, many of these subtasks rely on the results of point cloud classification which aims at assigning a (semantic) class label to each 3D point of the given point cloud and faces challenges arising from the complexity of 3D scenes caused by the irregular sampling of 3D points, varying point density and very different types of objects.

In this paper, we address point cloud classification and thereby involve approaches to quantify the importance of involved features in order to select only the few most relevant ones among them for classification. More specifically, we focus on the use of a standard set of geometric point cloud features, increase their distinctiveness by involving individual neighborhoods of optimal size and evaluate their relevance for classifying point cloud data acquired via mobile laser scanning. Since relevance can be quantified via different relevance metrics, we present a general relevance metric taking into account different criteria and thus providing an objective ranking of features with respect to their suitability.

In the following, we first reflect related work in Section 2. Subsequently, we present our methodology in detail in Section 3. By showing experimental results derived for a benchmark MLS dataset, we demonstrate the performance of our approach in Section 4. The derived results are discussed in Section 5, and, finally, concluding remarks as well as suggestions for future work are provided in Section 6.

2 RELATED WORK

For 3D scene analysis in terms of assigning each 3D point a (semantic) class label, two different strategies may be exploited. On the one hand, we may focus on the individual classification of each 3D point which only relies on the respective feature vector. For this purpose, standard classifiers such as Random Forests /Chehata et al. 2009/, Support Vector Machines /Mallet et al. 2011/ or Bayesian Discriminant Classifiers /Khoshelham & Oude Elberink 2012/ are commonly used. On the other hand, we may focus on contextual classification which also involves a modeling of relationships among 3D points in a local neighborhood and thus accounts for the fact that class labels of neighboring 3D points tend to be correlated. Respective classifiers are for instance represented by Associative Markov Networks /Munoz et al. 2009/, Non-Associative Markov Networks /Shapovalov et al. 2010/ or Conditional Random Fields /Niemeyer et al. 2012/.

Since we mainly address feature relevance assessment, we focus on the individual classification of each 3D point and thus only on the respective feature vectors. In general, a variety of features may be involved which describe radiometric or geometric properties, or properties assessed during data acquisition (e. g. full-waveform features). Respective investigations categorizing a variety of features into different feature types have for instance been presented in /Mallet et al. 2011/ or /Guo et al. 2014/. In order to provide a fundamental solution, we only exploit information which is shared by all available datasets, i. e. 3D geometry. Further features simply extend the feature vectors and do not change subsequent steps of the proposed methodology.

When using geometric features for describing the local 3D structure, these are typically derived at a single scale, where the scale parameter may be represented by (1) the radius of a spherical neighborhood, (2) the radius of a cylindrical neighborhood or (3) the number of points within the neighborhood. In order to obtain informative features, however, we have to consider that the choice of this scale parameter may strongly influence the feature representation, i. e. the feature vector, and thus also the classification results. Typically, the scale parameter is selected based on heuristic or empiric knowledge on the scene and defined to be identical for all 3D points of the considered point cloud. Consequently, this selection is specific for each dataset. In order to obtain a generic solution, the concepts of dimensionality-based scale selection /Demantké et al. 2011/ and eigenentropy-based scale selection /Weinmann et al. 2014/ have recently been proposed for obtaining individually optimized 3D neighborhoods. These concepts not only avoid the use of heuristic or empiric knowledge on the scene, but they also result in a significant improvement in classification accuracy /Weinmann et al. 2014/. As alternative to selecting optimal neighborhoods for each individual 3D point, it has been proposed to calculate features at different scales /Brodu & Lague 2012/ or even based on different entities such as points and regions (e. g. /Xiong et al. 2011/ or /Xu et al. 2012/). This, however, results in feature vectors of significantly higher dimension.

Furthermore, it has to be taken into account that, due to a lack of knowledge about scene and data, often as many features as possible are extracted and used for classification. However, some features may be more relevant, whereas others may be less suitable or even irrelevant. This is of great importance since, in theory, many classifiers are considered to be insensitive to the given dimensionality, whereas redundant or irrelevant information has been proven to influence their performance in practice. Consequently, feature selection techniques have been proposed in order to gain predictive accuracy, improve computational efficiency with respect to both time and memory consumption, and retain meaningful features /Guyon & Elisseeff 2003/. As a crucial step, such techniques rely on quantifying feature relevance. Accordingly, the basic idea consists of defining a ranking procedure and selecting a subset of the best-

ranked features. In the context of lidar data processing, the ranking procedure often relies on the interaction with a classifier (e. g. /Mallet et al. 2011/ or /Khoshelham & Oude Elberink 2012/) and, consequently, the derived feature subsets are only optimized with respect to a specific classifier. In order to obtain a more general solution, other approaches exploit a classifier-independent ranking procedure (e. g. /Weinmann et al. 2013/ or /Weinmann et al. 2014/).

3 METHODOLOGY

As shown in *Fig. 1*, the proposed processing workflow consists of four components which are explained in detail in the following subsections.

3.1 Neighborhood Selection

For deriving informative features for a given 3D point $\boldsymbol{X}_0$, we consider its neighborhood consisting of the respective k closest neighbors /Linsen & Prautzsch 2001/. Such a neighborhood definition accounts for varying point density by preserving flexibility with respect to the absolute geometric size of the neighborhood. However, the selection of a suitable value for the scale parameter k still remains an important topic of recent research. Whereas selecting an identical value for k across all 3D points provides a straightforward solution, it has to be taken into account that this involves heuristic or empiric knowledge on the scene and, consequently, the derived value is specific for each dataset. In contrast, considering individual values for k across all 3D points provides a generic solution, and it also accounts for the fact that k rather depends on the local 3D structure as well as the local point density.

Among a variety of techniques for selecting the "optimal" scale parameter for each individual 3D point of a given point cloud, the concepts of dimensionality-based scale selection /Demantké et al. 2011/ and eigenentropy-based scale selection /Weinmann et al. 2014/ have proven to be suitable for mobile laser scanning data. Interestingly, a significant gain in classification accuracy (mainly resulting from a significant gain in the recall values) can be observed when involving such a technique instead of a fixed scale parameter across all 3D points of a given point cloud /Weinmann et al. 2014/. Consequently, locally optimized neighborhoods increase the distinctiveness of local 3D shape features, i. e. the resulting features are more informative than those features based on a fixed scale parameter across all 3D points.

Since respective investigations clearly reveal that eigenentropy-based scale selection provides the best solution in comparison to various other neighborhood definitions /Weinmann et al. 2014/, we focus on this approach for defining appropriate neighborhoods. Consequently, for varying values of the scale parameter k, we exploit the 3D coordinates of a given point $\boldsymbol{X}_0$ and its k closest neighbors in order to calculate the respective 3D covariance matrix which is commonly referred to as 3D structure tensor $\boldsymbol{S}$:

$$\boldsymbol{S} = \frac{1}{k+1} \sum_{j=0}^{k} \left(\boldsymbol{X}_j - \overline{\boldsymbol{X}} \right) \left(\boldsymbol{X}_j - \overline{\boldsymbol{X}} \right)^{\mathrm{T}} . \quad (1)$$

In this equation, $\overline{\boldsymbol{X}}$ denotes the center of gravity. The three eigenvalues λ_1, λ_2 and λ_3 of the 3D structure tensor $\boldsymbol{S}$ with $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ are normalized by their sum $\sum_\lambda$ which yields normalized eigenvalues e_1, e_2 and e_3 with $e_i = \lambda_i / \sum_\lambda$ for $i \in \{1,2,3\}$. Since the normalized eigenvalues sum up to 1, the measure E_λ of eigenentropy is defined as the Shannon entropy:

$$E_\lambda = -e_1 \ln(e_1) - e_2 \ln(e_2) - e_3 \ln(e_3) . \quad (2)$$

Minimizing the measure of eigenentropy across different values of k corresponds to minimizing the disorder of 3D points within the local 3D neighborhood and, consequently, the optimal neighborhood size for the point $\boldsymbol{X}_0$ corresponds to the respective k with the minimal eigenentropy.

3.2 Feature Extraction

Once a 3D point $\boldsymbol{X}_0$ has been assigned its corresponding neighborhood size, the next step consists of deriving respective features. Since many datasets only contain information in the form of spatial 3D coordinates, we focus on the use of geometric features which are derived from the spatial arrangement of all points within the local neighborhood. More specifically, we follow the strategy of deriving a variety of both 3D and 2D features /Weinmann et al. 2013/, but we increase their distinctiveness by taking into account the optimal neighborhood size of each individual 3D point as described in Section 3.1.

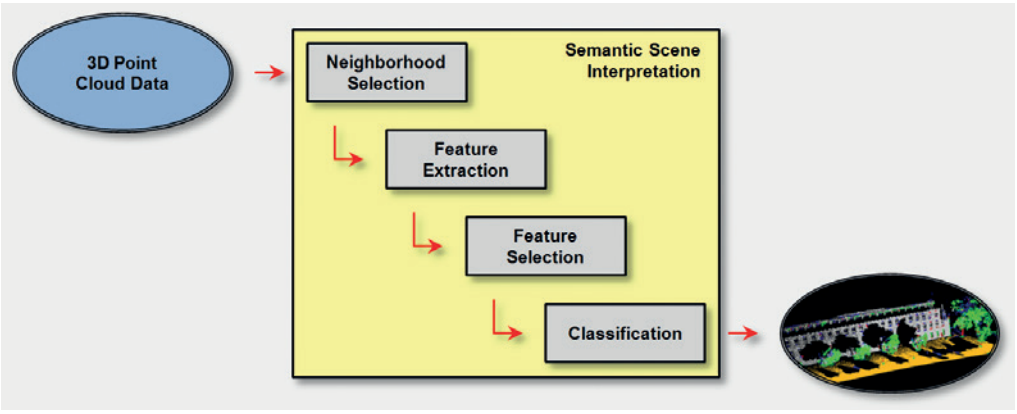


Fig. 1 | The proposed processing workflow to get from measured 3D point cloud data to semantic objects in the scene.

First, we focus on deriving 3D features. Whereas an analytical consideration of the eigenvalues λ_1, λ_2 and λ_3 of the 3D structure tensor $\boldsymbol{S}$ already allows a characterization of specific shape primitives /Jutzi & Gross 2009/, these eigenvalues are typically used to define more general local 3D shape features. Accordingly, we exploit the normalized eigenvalues e_1, e_2 and e_3 in order to extract a set of eight eigenvalue-based features according to /West et al. 2004/ and /Mallet et al. 2011/. These features are referred to as linearity L_λ, planarity P_λ, scattering S_λ, omnivariance O_λ, anisotropy A_λ, eigenentropy E_λ, sum $\sum_\lambda$ of eigenvalues and change of curvature C_λ. Further 3D features for characterizing the local neighborhood arise from basic geometric properties of the considered neighborhood, e. g. absolute height H, radius $r_{k-\mathrm{NN}}$ of the neighborhood, local point density D, verticality V which is derived from the vertical component of the normal vector, and maximum height difference $\Delta H_{k-\mathrm{NN}}$ as well as height variance $\sigma_{H,\,k-\mathrm{NN}}$ within the local neighborhood. Thus, we have defined a total number of 14 geometric 3D features.

Subsequently, we take into account that urban environments contain numerous man-made objects which, in turn, typically provide almost perfectly vertical structures (e. g. building façades, walls, traffic signs or utility poles). In order to exploit such information, the use of 2D features resulting from a 2D projection onto a horizontally oriented plane is advisable. Respective 2D features may describe basic geometric properties such as the radius $r_{k-\mathrm{NN},\,2D}$ of the neighborhood after the projection onto the horizontal plane or the local point density D_{2D} in 2D. Further 2D features arising from the 2D projection of all 3D points within the local neighborhood may be defined as the sum $\sum_{\lambda,\,2D}$ and the ratio $R_{\lambda,\,2D}$ of the eigenvalues of the 2D covariance matrix which is also known as the 2D structure tensor. Finally, we project all 3D points onto the horizontally oriented plane and construct a 2D accumulation map with discrete, quadratic bins (here with a side length of 0.25 m as proposed in /Weinmann et al. 2013/). For each bin, respective characteristics are described by the number M of points as well as the maximum difference ΔH and standard deviation σ_H of height values within that bin. Thus, we have defined a total number of seven geometric 2D features.

Combining all 3D and 2D features as summarized in *Tab. 1* thus yields a 21-dimensional feature vector for each 3D point, and since the features represent properties with different units, a normalization across all feature vectors is introduced in order to map the values of each dimension onto the interval [0,1]. This yields normalized feature vectors $[X_{1,\,\ldots,\,n}] = [X_1,\,\ldots,\,X_{21}]^{\mathrm{T}}$ characterizing the local 3D structure at the respective 3D points.

3D features		2D features		
Basic properties	Local shape features	Basic properties	Local shape features	Accumulation map
H, $r_{k-\mathrm{NN}}$, D, V, $\Delta H_{k-\mathrm{NN}}$, $\sigma_{H,k-\mathrm{NN}}$	L_λ, P_λ, S_λ, O_λ, A_λ, E_λ, $\sum_\lambda$, C_λ	$r_{k-\mathrm{NN},\,2D}$, D_{2D}	$\sum_{\lambda,\,2D}$, $R_{\lambda,\,2D}$	M, ΔH, σ_H

Tab. 1 | Categorization of the 21 involved 3D and 2D features.

3.3 Feature Selection

Although, in theory, many classifiers are considered to be insensitive to the given dimensionality, redundant or irrelevant information has been proven to influence their performance in practice. Consequently, techniques for finding compact and robust subsets of relevant and informative features have been investigated in order to gain predictive accuracy, improve computational efficiency with respect to both time and memory consumption, and retain meaningful features /Guyon & Elisseeff 2003/. In this context, a feature is defined to be statistically relevant if its removal from a feature set will reduce the prediction power. In the following, we will quantify feature relevance by assigning values in a specific interval.

Since we intend to select a generally versatile set of relevant features which is not optimized with respect to a specific classifier, we focus on a classifier-independent solution which, in turn, results in both simplicity and efficiency. Consequently, we apply a filter-based feature selection method. Such methods exploit a score function directly based on the training data. Note that the respective score function may address different intrinsic properties of the given training data such as distance, information, dependency or consistency. Since considering a single property may not be sufficient, we apply a general relevance metric which is based on several score functions. More specifically, we involve $N_s = 7$ different score functions which "evaluate" different intrinsic properties of the given training data /Weinmann et al. 2013/:

- The Pearson correlation coefficient $s_{\mathrm{Pearson}} = s_1$ indicates to which degree a feature is correlated with the class labels.
- The F-score or Fisher score $s_{\mathrm{Fisher}} = s_2$ represents the ratio between interclass and intraclass variance.
- The measure of Information Gain $s_{\mathrm{IG}} = s_3$ reveals the dependence between a feature and the class labels.
- The Gini index $s_{\mathrm{Gini}} = s_4$ provides a statistical measure of dispersion and thus an inequality measure which quantifies a feature's ability to distinguish between classes.
- The measure $s_\chi = s_5$ results from a χ^2-test which is used as a test of independence in order to assess whether a class label is independent of a particular feature.
- The measure $s_t = s_6$ results from applying a t-test on each feature and checking how effective it is for separating classes.
- The ReliefF measure $s_{\mathrm{ReliefF}} = s_7$ indicates the contribution of a feature to the separation of samples from different classes.

For each score function, we derive a separate ranking of all features and denote the rank of a specific feature X_i given the score function s_j as $r(X_i \mid s_j)$. The rank $r(X_i \mid s_j)$ is thus an integer value in the interval $[1,\,N_f]$, where N_f denotes the number of involved features (in our case $N_f = 21$). Smaller values reveal features with higher relevance when considering the respective score function s_j, whereas higher values reveal less suitable features. In order to obtain a general relevance metric $R(X_i)$ taking into account several score functions, we combine the separate ranking results across all score functions s_j by taking the mean rank $\overline{r}\,(X_i)$ of each feature X_i according to

$$\overline{r}\,(X_i) = \frac{1}{N_s} \sum_{j=1}^{N_s} r\,(X_i \mid s_j), \tag{3}$$

and we introduce a mapping to the interval [0, 1] in order to interpret the result as relevance $R(X_i)$ of the feature X_i:

$$R(X_i) = 1 - \frac{\overline{r}(X_i) - 1}{N_f - 1}. \quad (4)$$

3.4 Classification

For the sake of simplicity and applicability, we focus on individual point classification (i. e. the classification of a 3D point $\boldsymbol{X}_0$ by simply exploiting the respective feature vector), where a good trade-off between classification accuracy and computational effort can be achieved when using Random Forests /Breiman 2001/. Generally, Random Forests belong to the category of ensemble learning techniques which are based on the idea of strategically generating a set of weak learners (represented by decision trees) and combining them via bagging in order to create a single strong learner. More specifically, bagging is based on the strategy of using bootstrapped replica of the training data, i. e. subsets of the complete training data which are randomly drawn with replacement, in order to train several decision trees. The random sampling results in randomly different weak learners and thus in diversity in terms of de-correlated hypotheses across the weak learners. Consequently, taking the respective majority vote over all hypotheses can be expected to result in improved generalization and robustness, and thus to provide a suitable classifier.

4 EXPERIMENTAL RESULTS

For demonstrating the performance of our methodology, we involve a standard benchmark dataset (Section 4.1), where ground truth labels are available for each 3D point. In our experiments (Section 4.2), we intend to quantify feature relevance and define suitable feature subsets for interpreting 3D point cloud data. Consequently, when presenting the derived results (Section 4.3), we focus on the relevance of single features in order to select appropriate feature subsets and on the classification accuracy obtained when using the different feature sets for point cloud classification.

4.1 Dataset

For demonstrating the performance of the proposed methodology, we involve the Oakland 3D Point Cloud Dataset /Munoz et al. 2009/ which has been acquired in an urban environment with a mobile laser scanning system. This system allows to capture the local 3D geometry with side looking SICK LMS laser scanners used in push-broom mode. More specifically, the dataset contains spatial 3D coordinates of about 1.3 million 3D points as well as respective reference labels which have been obtained in a manually assisted way. The reference labels are assigned with respect to five semantic classes: wire, pole/trunk, façade, ground and vegetation. A separation of the whole dataset into training set and test set is available and, for both of them, the distribution of 3D points belonging to the different classes is very inhomogeneous. Consequently, in order to avoid a bias in feature selection as well as a detrimental effect on classification, we introduce a class re-balancing by reducing the training set (which comprises about 37 000 labeled 3D points) to a reduced training set encapsulating 1 000 training examples per class. The test set contains 1.3 million labeled 3D points, where 70.5 % represent ground, 20.2 % represent vegetation, 8.4 % represent façade, 0.6 % represent pole/trunk, and 0.3 % represent wire.

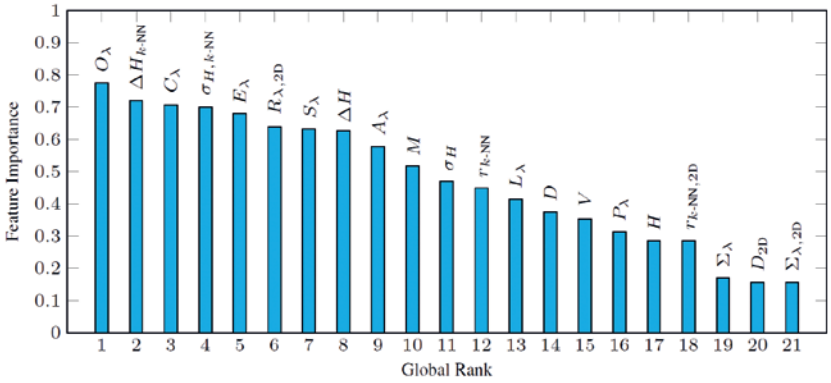


Fig. 2 | Ranking results obtained by applying the proposed general relevance metric. Relevant features are assigned a higher feature importance and hence a lower global rank.

4.2 Experiments

By exploiting the concept of eigenentropy-based scale selection /Weinmann et al. 2014/, we focus on automatically selecting an optimal scale parameter for each individual 3D point. Based on the derived individual neighborhoods of optimal size, the 21 defined low-level geometric 3D and 2D features are calculated. Subsequently, we consider the use of all features as well as the use of different feature subsets derived via the general relevance metric. For classification, we use a Random Forest classifier composed of $N_T = 100$ decision trees with a maximum tree depth of $d_{max} = 15$. In each tree, a node is only split if it is reached by at least $n_{min} = 20$ training samples, and the number of active variables (i. e. the number of features) to be used for the test in each tree node is set to the square root of the number of features (i. e. $n_a = \lfloor \sqrt{d} \rfloor$ where d represents the number of involved features).

For evaluation, we consider (1) recall which represents a measure of completeness or quantity, (2) precision which represents a measure of exactness or quality, (3) overall accuracy (OA) which reveals the overall performance of the involved classifier on the test set, and (4) mean class recall (MCR) which reveals the capability of the involved classifier to detect instances of different classes.

# Features	Recall					Precision					OA	MCR
	W	P/T	F	G	V	W	P/T	F	G	V		
5	72.69	72.08	63.28	98.61	79.49	10.32	23.41	75.15	96.35	95.03	91.55	77.23
10	78.91	75.89	67.02	98.69	79.82	8.01	18.90	87.15	97.24	96.31	92.03	80.07
15	85.40	73.30	67.87	98.34	82.38	7.81	32.24	84.19	97.24	95.53	92.38	81.46
21	85.56	79.54	67.88	98.36	78.24	8.30	23.47	82.72	96.77	96.10	91.59	81.91

Tab. 2 | Classification results (in %) obtained for the different feature sets comprising 5, 10, 15 and all 21 features (W: wire, P/T: pole/trunk, F: façade, G: ground, V: vegetation).

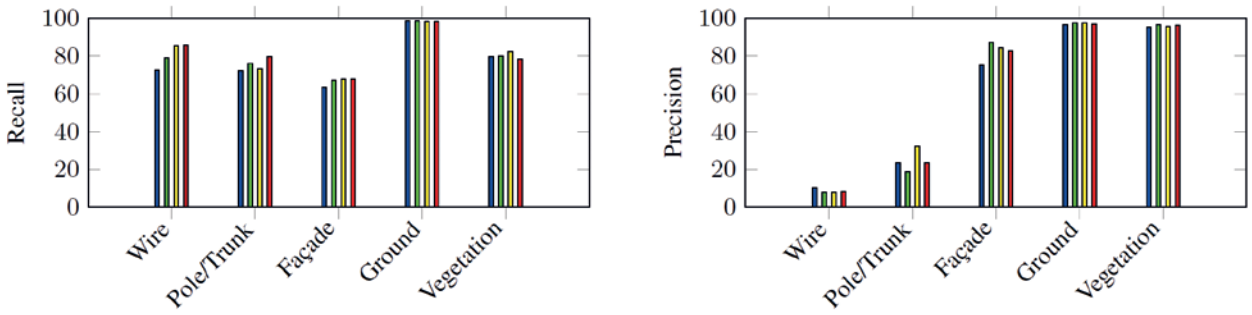


Fig. 3 | Recall and precision values (in %) for the different feature sets comprising 5 features (blue), 10 features (green), 15 features (yellow) and all 21 features (red).

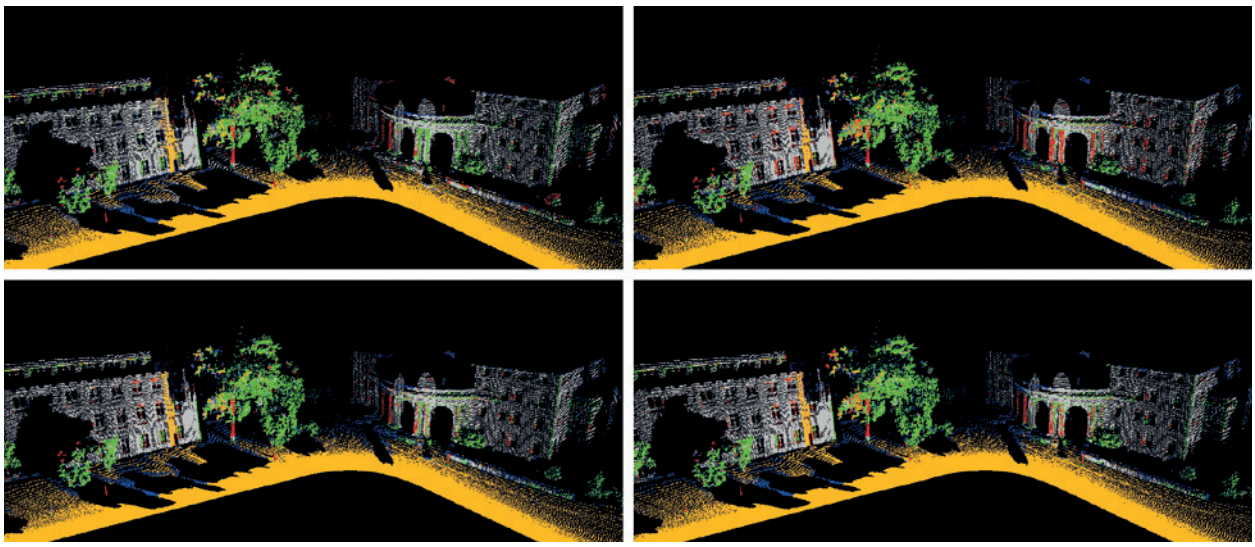


Fig. 4 | Classified point clouds obtained for the feature sets comprising 5 features (top left), 10 features (top right), 15 features (bottom left) and all 21 features (bottom right) when using a standard color encoding (blue: wire, red: pole/trunk, gray: façade, brown: ground, green: vegetation). The noisy appearance results from individual point classification.

4.3 Results

For the conducted experiments, the proposed measure for quantifying feature relevance and thus describing feature importance is visualized in *Fig. 2*. Thereby, the general relevance metric may result in a feature ranking which is quite different from a ranking involving neighborhoods with a fixed scale parameter across the whole dataset /Weinmann et al. 2013/, since the concept of using individual neighborhoods of optimal size has proven to yield an increased distinctiveness of the extracted features /Weinmann et al. 2014/ which, in turn, influences the importance of these features.

Based on the extracted features, we conduct feature selection by selecting a feature subset which consists of the few best-ranked features (*Fig. 2*). More specifically, we compare the classification results obtained for feature subsets of the 5, 10 and 15 best-ranked features as well as the full feature set. The respective classification results are provided in *Tab. 2* and, in order to ease comparison, the class-wise recall and precision values are visualized in *Fig. 3*. Furthermore, the classified point clouds are depicted in *Fig. 4*.

5 DISCUSSION

From the derived results, it becomes clearly visible that some features are more and others less relevant (*Fig. 2*). Thereby, the gap between highly relevant and less relevant features is quite considerable. In comparison to /Weinmann et al. 2013/, we may state a slight change of the ranking which is due to the fact that we increased the distinctiveness of the involved features by conducting eigenentropy-based scale selection. Consequently, the feature of eigenentropy E_λ (which is used for deriving the optimal scale and thus the optimal neighborhood size) is assigned a significantly higher relevance. Whereas the feature of omnivariance O_λ obtained the highest relevance and thus also reveals a significant gain in relevance compared to /Weinmann et al. 2013/, the maximum height difference $\Delta H_{k-\mathrm{NN}}$ as well as the height variance $\sigma_{H,\,k-\mathrm{NN}}$ within the local neighborhood and the change of curvature C_λ are again within the 5 best-ranked features. Furthermore, the ratio $R_{\lambda,\,2\mathrm{D}}$ of the eigenvalues of the 2D covariance matrix and the maximum difference ΔH of height values within a bin are among the most relevant features which is in accordance with comparable investigations /Weinmann et al. 2013/.

Whereas the proposed general relevance metric directly allows to quantify feature relevance, the impact of feature relevance assessment on 3D scene analysis becomes visible after classification. The classification results provided in *Tab. 2* and *Fig. 3* reveal that the use of only the 5 best-ranked features (which corresponds to 23.8 % of the required memory for storing extracted features), the 10 best-ranked features (47.6 %) and the 15 best-ranked features (71.4 %) still leads to reasonable results, which are partially even better than when using all features, while simultaneously reducing memory consumption significantly. More specifically, the overall accuracy (OA) is in the range of 91.5–92.4% for all four feature sets, and the respective mean class recall (MCR) values are in the range of 77.2–81.9%. For the class-wise considerations, only minor differences with respect to the recall and precision values may be stated when using the different feature sets. Consequently, the classified point clouds depicted in *Fig. 4* only reveal minor differences. However, in comparison to /Weinmann et al. 2013/, the class-wise recall values and thus also the mean class recall (MCR) values are significantly higher due to the consideration of individual neighborhoods of optimal size. Furthermore, a significant increase in the precision values may be stated for the smaller classes of wire and pole/trunk.

Since the conducted feature selection is only based on the training set, all features only have to be calculated for the training set which contains 5 000 points. In contrast, for the test set containing about 1.3 million points, only those features considered as relevant have to be calculated. Thus, feature selection via the proposed general relevance metric can be considered as an interesting option for large-scale 3D scene analysis, where efficiency with respect to the computational burden in terms of processing time and memory consumption represents an important prerequisite. In order to get an impression on the computational effort for our experiments on a high-performance computer (Intel Core i7-3820M, 3.6GHz, 64GB RAM), we may have a look at the processing time required for each subtask. Considering the training set with 5000 points, the required processing times are about 21s for neighborhood selection, about 4s for feature extraction, about 25s for feature selection and less than 1s for training. For the test set containing 1.3 million points, the maximal processing times are reached when considering the full feature set for classification, and these processing times are given by about 758s for neighborhood selection, about 2793s for feature extraction and about 6s for classification. Consequently, feature extraction is the most time-consuming subtask and a significant improvement in efficiency for this component directly corresponds to a significant improvement of the whole processing workflow.

6 CONCLUSIONS

In this paper, we focused on feature relevance assessment in order to increase efficiency for the semantic interpretation of mobile laser scanning data. In this context, we proposed a methodology for 3D scene analysis where efficiency is increased by (1) involving individual neighborhoods of optimal size which, in turn, increases the distinctiveness of the involved features and (2) selecting a subset of the few best-ranked features according to a general relevance metric which, in turn, reduces the computational burden with respect to both processing time and memory consumption. The derived results clearly demonstrate the feasibility of the proposed methodology with respect to both criteria. The increased distinctiveness of the involved features becomes visible when considering the significantly beneficial impact of individually optimized neighborhoods on the classification results and particularly on the mean class recall values. On the other hand, the derived results reveal that 3D scene analysis may be conducted with only very few, but suitable features without significantly reducing the quality of the classification results. For future work, we plan to address large-scale 3D scene analysis as well as a more detailed scene analysis up to object level. Furthermore, contextual learning approaches or spatial smoothing techniques could be applied in order to improve the classification results.

REFERENCES

Breiman, L. (2001): Random forests. In: Machine Learning, 45(2001)1, 5 – 32.

Brodu, N.; Lague, D. (2012): 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. In: ISPRS Journal of Photogrammetry and Remote Sensing, 68(2012), 121 – 134.

Chehata, N.; Guo, L.; Mallet, C. (2009): Airborne lidar feature selection for urban classification using random forests. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-3/W8, 207 – 212.

Demantké, J.; Mallet, C.; David, N.; Vallet, B. (2011): Dimensionality based scale selection in 3D lidar point clouds. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-5/W12, 97 – 102.

Guo, B.; Huang, X.; Zhang, F.; Sohn, G. (2014): Classification of airborne laser scanning data using JointBoost. In: ISPRS Journal of Photogrammetry and Remote Sensing, 92(2014), 124 – 136.

Guyon, I.; Elisseeff, A. (2003): An introduction to variable and feature selection. In: Journal of Machine Learning Research, 3(2003), 1157 – 1182.

Hebert, M.; Bagnell, J. A.; Bajracharya, M.; Daniilidis, K.; Matthies, L. H.; Mianzo, L.; Navarro-Serment, L.; Shi, J.; Wellfare, M. (2012): Semantic perception for ground robotics. In: Proceedings of SPIE 8387, 1 – 12.

Jutzi, B.; Gross, H. (2009): Nearest neighbour classification on laser point clouds to gain object structures from buildings. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-1-4-7/W5 (CD-ROM).

Khoshelham, K.; Oude Elberink, S. J. (2012): Role of dimensionality reduction in segment-based classification of damaged building roofs in airborne laser scanning data. In: Proceedings of the International Conference on Geographic Object Based Image Analysis, 372 – 377.

Linsen, L.; Prautzsch, H. (2001): Local versus global triangulations. In: Proceedings of Eurographics, 257 – 263.

Mallet, C.; Bretar, F.; Roux, M.; Soergel, U.; Heipke, C. (2011): Relevance assessment of full-waveform lidar data for urban area classification. In: ISPRS Journal of Photogrammetry and Remote Sensing, 66(2011)6, S71 – S84.

Munoz, D.; Bagnell, J. A.; Vandapel, N.; Hebert, M. (2009): Contextual classification with functional max-margin Markov networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 975 – 982.

Niemeyer, J.; Rottensteiner, F.; Soergel, U. (2012): Conditional random fields for lidar point cloud classification in complex urban areas. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I-3, 263 – 268.

Pu, S.; Rutzinger, M.; Vosselman, G.; Oude Elberink, S. (2011): Recognizing basic structures from mobile laser scanning data for road inventory studies. In: ISPRS Journal of Photogrammetry and Remote Sensing, 66(2011)6, 28 – 39.

Serna, A.; Marcotegui, B. (2013): Urban accessibility diagnosis from mobile laser scanning data. In: ISPRS Journal of Photogrammetry and Remote Sensing, 84(2013), 23 – 32.

Shapovalov, R.; Velizhev, A.; Barinova, O. (2010): Non-associative Markov networks for 3D point cloud classification. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-3A, 103 – 108.

Velizhev, A.; Shapovalov, R.; Schindler, K. (2012): Implicit shape models for object detection in 3D point clouds. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I-3, 179 – 184.

Weinmann, M.; Jutzi, B.; Mallet, C. (2013): Feature relevance assessment for the semantic interpretation of 3D point cloud data. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences II-5/W2, 313 – 318.

Weinmann, M.; Jutzi, B.; Mallet, C. (2014): Semantic 3D scene interpretation: A framework combining optimal neighborhood size selection with relevant features. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences II-3, 181 – 188.

West, K. F.; Webb, B. N.; Lersch, J. R.; Pothier, S.; Triscari, J. M.; Iverson, A. E. (2004): Context-driven automated target detection in 3-D data. In: Proceedings of SPIE 5426, 133 – 143.

Xiong, X.; Munoz, D.; Bagnell, J. A.; Hebert, M. (2011): 3-D scene analysis via sequenced predictions over points and regions. In: Proceedings of the IEEE International Conference on Robotics and Automation, 2609 – 2616.

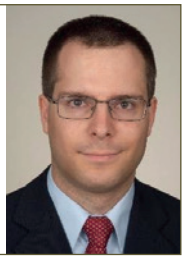
Xu, S.; Oude Elberink, S. J.; Vosselman, G. (2012): Entities and features for classification of airborne laser scanning data in urban area. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I-4, 257 – 262.

Zhou, L.; Vosselman, G. (2012): Mapping curbstones in airborne and mobile laser scanning data. In: International Journal of Applied Earth Observation and Geoinformation, 18(2012)1, 293 – 304.

Martin Weinmann

KARLSRUHE INSTITUTE OF TECHNOLOGY (KIT)
INSTITUTE OF PHOTOGRAMMETRY AND REMOTE SENSING

Englerstraße 7 | 76131 Karlsruhe
martin.weinmann@kit.edu



Clément Mallet

UNIVERSITÉ PARIS-EST, IGN, SRIG, MATIS

73 avenue de Paris | 94160 Saint-Mandé, France
clement.mallet@ign.fr

Stefan Hinz

KARLSRUHE INSTITUTE OF TECHNOLOGY (KIT)
INSTITUTE OF PHOTOGRAMMETRY AND REMOTE SENSING

Englerstraße 7 | 76131 Karlsruhe
stefan.hinz@kit.edu

Boris Jutzi

KARLSRUHE INSTITUTE OF TECHNOLOGY (KIT)
INSTITUTE OF PHOTOGRAMMETRY AND REMOTE SENSING

Englerstraße 7 | 76131 Karlsruhe
boris.jutzi@kit.edu

Manuskript eingereicht: 30.01.15 | Im Peer-Review-Verfahren begutachtet