

SHAPE DISTRIBUTION FEATURES FOR POINT CLOUD ANALYSIS - A GEOMETRIC HISTOGRAM APPROACH ON MULTIPLE SCALES

R. Blomley^{a,*}, M. Weinmann^a, J. Leitloff^a, B. Jutzi^a

^a Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT),
Englerstr. 7, 76131 Karlsruhe, Germany - {rosmarie.blomley, martin.weinmann, jens.leitloff, boris.jutzi}@kit.edu

Commission III, WG III/4

KEY WORDS: LIDAR, Point Cloud, Features, Geometric Feature Design, Multiscale, Probability Histogram

ABSTRACT:

Due to ever more efficient and accurate laser scanning technologies, the analysis of 3D point clouds has become an important task in modern photogrammetry and remote sensing. To exploit the full potential of such data for structural analysis and object detection, reliable geometric features are of crucial importance. Since multiscale approaches have proved very successful for image-based applications, efforts are currently made to apply similar approaches on 3D point clouds. In this paper we analyse common geometric *covariance* features, pinpointing some severe limitations regarding their performance on varying scales. Instead, we propose a different feature type based on *shape distributions* known from object recognition. These novel features show a very reliable performance on a wide scale range and their results in classification outnumber *covariance* features in all tested cases.

1 INTRODUCTION

Contemporary laser scanning systems provide 3D point clouds with increasing accuracy and point density that may contribute significantly to the huge potential of remote sensing in environmental applications. Therefore there is a growing need for efficient 3D geometric characterisation, structural analysis and interpretation of such data.

Depending on the application, supervised and unsupervised classification approaches may be pursued, yet all of them rely on descriptive features. Following techniques well-known from image processing, the analysis of invariant moments has been applied to study the geometric properties of 3D point cloud data (Maas and Vosselman, 1999). In particular, second-order moments, represented by the covariance matrix or structure tensor (West et al., 2004) are increasingly popular in geometric feature extraction (Jutzi and Gross, 2009; Toshev et al., 2010; Niemeyer et al., 2012). A good overview of currently used features for 3D point cloud analysis and a comprehensive study of their classification relevance is given by Chehata et al. (2009) and Mallet et al. (2011). These features can be grouped into height-based features, geometric features derived from covariance or local plane estimations and sensor specific features such as full-waveform or echo-based features. In those studies, height-based features are generally ranked very important. Geometric features such as covariance and local plane features have to be calculated from a certain local neighbourhood around each point in question. In fact, according to scale selection studies (Demantké et al., 2011; Gressin et al., 2012), those features perform best when calculated from a particularly homogeneous neighbourhood determined by optimisation of the local dimensionality-based entropy.

Meanwhile, landscape classification tasks usually involve some recognition of complex structures beyond the reach of small homogeneous neighbourhoods. This has led to a change in image-based remote sensing towards object-based and multiscale methods (Blaschke and Hay, 2001; Hay et al., 2005), which is not yet

common in point cloud analysis. Object-based point cloud studies include shape parameterisations similar to 3D Hough transformations (Vosselman et al., 2004) and grouping of points to segments and entities (Reitberger et al., 2009; Xu et al., 2012). Multiscale approaches on point cloud data are often very time-consuming due to their iterative schemes, when the appropriate neighbourhood is sought locally (Pauly et al., 2003; Mitra and Nguyen, 2003; Demantké et al., 2011). Especially in natural environments, an evaluation of multiple scales can prove beneficial as it accounts for the characteristic scales of different structures (Brodu and Lague, 2012).

In remote sensing applications, point cloud analysis is sometimes limited by the number of returns per area, due to which important details may not be resolved. However, it has been pointed out very early that both in cases with object sizes larger than the given resolution and object sizes smaller than the resolution, some characteristic spatial autocorrelation can be expected (Strahler et al., 1986). Thus probabilistic distributions of geometric properties in 3D point clouds may hold more information than locally calculated parameters. Reaching beyond locally homogeneous neighbourhoods, histogram distributions have already been successfully used in computer vision sciences (Tombari et al., 2010). In image-based keypoint description, the SIFT algorithm is a prominent example for robustness and effectiveness achieved by a set of local histograms (Lowe, 2004). In 3D point clouds, existing histogram approaches are limited to surface keypoint description, as they rely on surface normal vectors (Tombari et al., 2013; Rusu et al., 2009).

The aim of this research is to introduce novel reliable geometric features for volumetric point cloud classification, which perform well at multiple scales. We therefore adopt a proposal from object recognition, using histograms of randomly sampled geometric measures, called *shape distributions* (Osada et al., 2002), as features within a local neighbourhood. Taking the measures proposed in the original approach as a rotation-invariant metric, the histograms may display correlations represented in this metric as characteristic signatures. Dissenting from the original proposal, the histogram binning is chosen by an adaptive and automated procedure to avoid arbitrary or data specific thresholds.

*Corresponding author.

In Section 2, the main conceptual components of this approach are set out. Even though height features are of paramount importance for real applications (Chehata et al., 2009; Mallet et al., 2011), they will not be included in the following investigations. For the sake of clarity, only rotation-invariant geometric features (like those in the original *shape distribution* proposal) will be compared. For practical purposes, the improvement achievable by height-based features should be independent of the features discussed here and would be beyond the scope of this paper. Section 3 briefly describes the conducted experiments based on commonly available urban area benchmark data. Any findings obtained here are discussed in Section 4. First of all, the results of scale investigations using the *shape distribution* features are evaluated. Secondly, a comparison with presently well-established *covariance* features is conducted. This comparison covers Support Vector Machine classifications at multiple scales, an estimation of the entropy-based optimum neighbourhood size for these features (Demantké et al., 2011) and a feature relevance assessment (Weinmann et al., 2013). Finally, the novel features' performance is evaluated and compared against the *covariance* features' result.

2 METHODOLOGY

The first part of the proposed methodology (Section 2.1) is intended to illustrate the need for a novel geometric feature type suitable for multiscale investigations. Section 2.2 therefore covers both a current geometric feature set (Section 2.2.1) and the concept of this novel approach (Sections 2.2.2 and 2.2.3). Since our investigations focus on feature design, classification methods are only used to evaluate the features practical performance and are therefore discussed in the experimental part.

2.1 Scale Investigations

In geometric point cloud analysis, any single observation (element) can only be interpreted by its relationship to other elements and its probability of belonging to a certain object class. Yet any object class displays characteristic coherence at different spatial scales (Hay et al., 2005). Therefore it is crucial to establish features that can represent structural characteristics on multiple scales. Features calculated on some scales may reproduce special autocorrelation characteristics that are not observed at other scales. A multiscale approach should therefore lead to enhanced classification results. For classification, one or more scales have to be determined, at which the sampling pattern of the sensor can resolve unique properties of the object class in question. Some approaches utilize a scale space representation for this purpose. Since this is not trivial for 3D cases (Tombari et al., 2013), we clarify that we use scale as a neighbourhood size parameter only.

Moreover, the shape of the considered neighbourhood has to be chosen carefully (Filin and Pfeiffer, 2005). For the purpose of aerial laser scanning (ALS), a cylindrical neighbourhood is found to be preferable, as it allows the features to capture the height distribution of the surrounding point cloud (Shapovalov et al., 2010). The varying cylinder radii to be investigated are chosen as $2^{n/2}$ m with $n \in \mathbb{N}$ and $-4 \leq n \leq 11$. Thereby all possibly resolved structural scales should be covered, as the radius ranges between 0.25 m (within the lateral placement accuracy of the laser scanner) and 45 m (above most object sizes). However, the features that will be presented in Section 2.2.2 are easily adapted to different laser scanning applications, as they are insensitive to the overall neighbourhood shape.

2.2 Feature Design

Feature design is of central importance to every knowledge representation and classification. In the following section a widely used geometric feature type for point cloud analysis will be introduced to point out some of its important properties (Section 2.2.1). Subsequently a novel geometric feature type will be introduced, first describing its origin in object recognition (Section 2.2.2) and afterwards a further contribution for its adaption as feature for classification applications (Section 2.2.3).

2.2.1 Covariance Features. Most present approaches using 3D geometric features employ features derived from the local covariance matrix representing second-order invariant moments within the point positions. The covariance matrix is calculated from N observations $A_{1,2,3}$ as follows:

$$[Cov]_{ij} = \frac{\sum_{i=1}^N (A_i - \bar{A}_i) \cdot (A_j - \bar{A}_j)}{N}, \quad (1)$$

where $i, j \in [1, 2, 3]$ and \bar{A}_i holds the mean of all observations in the respective dimension. Subsequent principal component analysis is used to determine linearly uncorrelated second-order moments in an orthogonal eigenvector space. The corresponding eigenvalues $\lambda_{1,2,3}$ then hold a great potential to calculate local features including dimensionality (linearity, planarity and sphericity) and other measures such as omnivariance, anisotropy and eigenentropy. The eigenvalues, sorted as $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ and the measures listed in Equation 2, will be referred to as *covariance* features.

$$\begin{aligned} \text{Linearity: } L_\lambda &= \frac{\lambda_1 - \lambda_2}{\lambda_1} \\ \text{Planarity: } P_\lambda &= \frac{\lambda_2 - \lambda_3}{\lambda_1} \\ \text{Sphericity: } S_\lambda &= \frac{\lambda_3}{\lambda_1} \\ \text{Omnivariance: } O_\lambda &= \sqrt[3]{\lambda_1 \lambda_2 \lambda_3} \\ \text{Anisotropy: } A_\lambda &= \frac{\lambda_1 - \lambda_3}{\lambda_1} \\ \text{Eigenentropy: } E_\lambda &= - \sum_{i=1}^3 \lambda_i \ln(\lambda_i) \\ \text{Sum of } \lambda\text{s: } \Sigma_\lambda &= \lambda_1 + \lambda_2 + \lambda_3 \\ \text{Change of curvature: } C_\lambda &= \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \end{aligned} \quad (2)$$

Yet it is especially important for these features to be derived from a suitably chosen neighbourhood size. It is the nature of second-order moments that the distance of one element from the mean contributes quadratically (c.f. Equation 1) and therefore elements in the vicinity are far less important than those further away. Since the principal component analysis is an orthogonal and thereby unitary transformation, the resulting eigenvalues are sensitive to the original scaling. Demantké et al. (2011) and Gressin et al. (2012) show evidence that a suitable spherical neighbourhood size can be found by minimization of the Shannon entropy based on dimensionality-features. Yet it remains to be shown if this optimum neighbourhood size for *covariance* features corresponds to the characteristic scale of any structure. To advance further research in the field, other geometrical features more suited to multiple scales are indispensable.

2.2.2 Shape Distributions. The characteristic scale of complex and partially random structures may not always be identical to the optimum neighbourhood size of *covariance* features. Yet to reveal such patterns, a statistical distribution of randomly sampled values may be more suitable than single values such as *covariance* measures. Therefore we adapt a concept proposed and developed by Osada et al. (2002). The key idea is to use random sampling of simple geometric measures to obtain a signature of the neighbourhood around each point to classify as a histogrammetric *shape distribution*. Adapting the aforementioned reference, we investigate the following geometric measures:

- **D1**: distance between any random point and the centroid of all considered points,
- **D2**: distance between two random points,
- **D3**: square root of the area of a triangle between any three random points,
- **D4**: cubic root of the volume of a tetrahedron between any four random points,
- **A3**: angle between any three random points.

In the practical implementation presented in Sections 3.3 and 3.4, the resulting values will be called *shape values*. The resulting histogram therefore represents the probability distribution of the taken geometric measures within the observed sample and should reveal repeating structures by a more frequent occurrence of some values. By this approach, feature extraction is reduced to a simple random sampling procedure. Such features are fast to calculate, mirror- and rotation-invariant and robust regarding outliers, noise and varying point density due to application-specific scanning or flight patterns. Notably, classification steps thereafter resemble a comparison of the probability distributions of these simple geometric measures.

2.2.3 Adaptive Histogram Binning. Dissenting from the original *shape distribution* proposal, we use an adaptive histogram binning approach to achieve maximum variance of significant observations from the gross of the total dataset. Above all, this step ensures a scale-independent performance. For this purpose, a simple histogram equalization procedure, known from image processing applications (Gonzales and Wood, 2002), is adapted. For all measured values at a linear binning scope m_k , with $k = 0, \dots, L - 1$ and L the number of bins, a transformation function $T(m_k)$ to a non-linear binning scope is found in such a way that a histogram of any large number of random samples is equally distributed. The transformation function is defined as

$$T(m_k) = \sum_{j=0}^k p_m(m_j), \quad (3)$$

where p_m is defined as the probability of occurrence of a value within m_k from a large number of samples n :

$$p_m(m_k) = \frac{n_k}{n}, \quad (4)$$

where n_k is the number of occurrences within m_k .

3 EXPERIMENTS

To test the proposed features on commonly accessible data, a benchmark dataset is evaluated. Due to the fundamental nature of our investigations and the exclusive use of geometric features, results will not be compatible to others already published in the benchmark context. We aim to submit classification results including other features for evaluation in future work and therefore focus on internal evaluation only. The overall experimental procedure, featuring used data (Sections 3.1 and 3.2), feature calculation (Sections 3.3 and 3.4), as well as classification (Section 3.5), will be presented here, before results are discussed in Section 4.

3.1 Benchmark Dataset

To test the performance of this genuine feature type, an ALS dataset contained in the ISPRS benchmark on urban classification at the Vaihingen test site (Cramer, 2010) is used. This data was

acquired in August 2008 with a Leica ALS50 system at a mean flying height of 500 m with a 45° viewing angle, 30 % overlap of flight strips and a mean point density of 4 pts./m^2 .

3.2 Reference Data and Point Cloud Access

In order to test the features' performance by supervised classification, some ground truth reference is mandatory. Voxel-based reference data was obtained from Gerke and Xiao (2014), and reference points were selected as those data elements closest to the centre of each voxel. The four classes and their respective number of elements are:

$16 \cdot 10^4$ points	for building roofs
$7 \cdot 10^4$ points	for trees
$6 \cdot 10^4$ points	for vegetated ground
$12 \cdot 10^4$ points	for sealed ground

For easy access to all point cloud elements contained in a cylindrically shaped neighbourhood around each element at different radii, a k -dimensional tree structure is used.

3.3 Shape Value Calculation

As we pursue random sampling, some limitations have to be set. Thus we limit ourselves to 255 pulls of geometric *shape values* per histogram. This limits computational efforts and minimizes the possible impact of varying point densities within the dataset. For a cylindrical neighbourhood in the given data 50% of all points have this number of neighbours at a cylinder radius of ~ 2.5 m, suggesting that this number should allow for a representative description. For big neighbourhoods, a relatively small percentage of all points in the volume element are thereby considered. This prevents false high recognition values due to overlap between the neighbourhood of adjacent points and overfitting in reference areas smaller than the considered neighbourhood.

Moreover the number of bins per shape histogram has to be specified. A large number of bins will allow for sophisticated neighbourhood descriptions, but for a small amount of points the signature may then not be descriptive. Therefore we decide to use 10 bins per shape histogram in this proposal, yet an optimization of both bin and sampling number may still be sought in future investigations.

3.4 Histogram Binning Thresholds

To avoid arbitrary threshold selection and ensure a balanced performance of *shape distributions* at multiple scales, an automated choice of histogram binning thresholds (Section 2.2.3) is implemented. Random samples are taken at each radius, and the corresponding *shape values* collected for adaptive binning calculation. After 500 pulls, no significant change of the adapted binning thresholds is observed. Those thresholds are used thereafter to form *shape distributions* from *shape values* sampled within each neighbourhood, as described in Section 2.2.2 and Figure 1.

3.5 Classification

Since our investigations focus on feature design, we aim for a simple, commonly used and universal classification procedure, knowing that better results may be achieved by other methods in the field. Furthermore, the chosen classifier should ideally be tolerant to irrelevant and redundant attributes, which may be significant as a large number of feature values is used in the present histogram approach. Considering those requirements, *Support Vector Machines (SVMs)* are a suitable choice. They are easy to

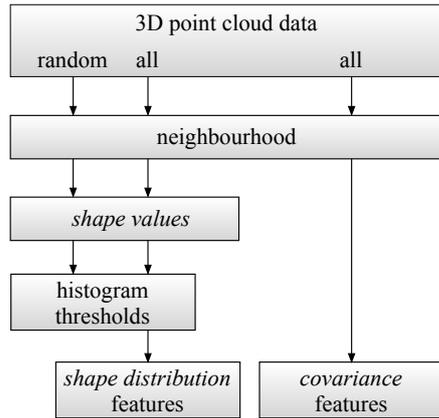


Figure 1: Feature calculation workflow. To calculate *shape distribution* features, histogram binning thresholds have to be acquired first as described in Section 3.4.

use, generally produce very accurate results and cope well with possible correlation between features (Kotsiantis, 2007).

To investigate whether some classes are particularly well described by features from certain scales, separate one-against-all distinctions are better suited than one multiclass classification. The identification of each separate class is performed using a one-against-all binary SVM classifier provided by the LIBSVM package (Chang and Lin, 2011). The classifier uses a radial basis function kernel and depends on two parameters, namely γ , representing the width of the Gaussian kernel function and C , a soft margin parameter allowing for some misclassifications. To ensure a smooth classification procedure, all feature data are scaled to a range between zero and one. A grid search for optimal values of γ and C is completed by evaluation of the cross-validation accuracy on a threefold partition of the training data. The grid search and subsequent training of a classifier with the best respective parameters is performed on a subset containing 1,000 data points of each class to avoid a bias by unbalanced reference data distribution. Afterwards, the performance of any selected classifier is tested on all labelled training data ($4.1 \cdot 10^5$ points).

4 EVALUATION

A central enquiry within our work, namely an analysis of *shape distribution* features at different scales, will be presented in Section 4.1. Afterwards the same analysis is conducted based upon commonly used *covariance* features in Section 4.2.1. This comparison is supported by an estimation of the dimensionality-based optimum neighbourhood for *covariance* features (Section 4.2.2) and classifier-independent feature relevance assessment for both feature types (Section 4.2.3). Eventually, the best classification results of both geometric feature types are evaluated on point cloud level in Section 4.3. *Shape distribution* results are visualised at different settings to analyse the overall performance.

4.1 Scale Analysis

The classification procedure described in Section 3.5 is performed for all four classes contained in the reference data. Features are calculated from different sizes of the cylindrical neighbourhood around each point, as described in Section 3.3 and 3.4.

To ensure consistency and compatibility with other leading publications, each classification result is evaluated using the metrics

described by Rutzinger et al. (2009): *Completeness*, *correctness* and *quality* are calculated from the number of correctly identified (TP), correctly rejected (TN), falsely identified (FP) and falsely rejected (FN) elements.

$$Comp. = \frac{TP}{TP + FN} \quad (5)$$

$$Corr. = \frac{TP}{TP + FP} \quad (6)$$

$$Quality = \frac{1}{1/Comp. + 1/Corr. - 1} \quad (7)$$

The application of such metrics allows for an in-depth analysis of the classification results and a precise comparison of the features' performance on different classes. Classification results for all investigated neighbourhood sizes are presented in Figure 2.

As the resulting graphs are smooth and generally display an even peak-like distribution, it is presumed that *shape distribution* features are a suitable choice to evaluate geometrical properties of point clouds over a wide spatial neighbourhood scale. No prominent peaks occur to suggest a strong pattern or scale preference for individual classes in this data. Optimal results are achieved at the following neighbourhood radii:

- 2.0 m for roofs
- 1.4 m for trees
- 2.8 m for vegetated ground
- 2.8 m for sealed ground

Interestingly all those maxima fall within a similar range. This generally descriptive size for *shape distribution* features is significantly higher than the neighbourhood size used for *covariance* features in ALS. The values of 1.0m used by Jutzi and Gross (2009) and 0.75 m used by Niemeyer et al. (2012) agree well with findings presented later in Section 4.2.2. Yet a suitable scale may well depend on the feature type chosen. When pursuing a *covariance* approach, homogeneous surroundings at a smaller scale will cause more distinctive features, whereas *shape distributions* may rather show repeating or significant patterns at a larger scale.

Comparing different classes, it is notable that trees are detected with the highest *completeness* but slightly lower *correctness* and perform best in regard to *quality*. For all other classes, *correctness* is higher than *completeness*. This indicates, that relatively few points are misassigned, but several points are missed. Whereas for roofs and sealed ground a reasonable amount is still detectable, vegetated ground performs significantly lower due to both low *completeness* and *correctness*.

4.2 Comparison with Covariance Features

Linearly independent second-order invariant moments $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ and features calculated from them as described in Section 2.2.1, are derived for the whole range of neighbourhood sizes around each point. Prior to classification, all features are rescaled so that all but 1 % of outliers lie between zero and one.

4.2.1 Comparison of Scale Analysis. Using exclusively *covariance* features, it is not possible to conduct a cohesive analysis spanning the same scale range as shown for *shape distribution* features on this dataset. Classification results are displayed in Figure 3. For small neighbourhoods the *covariance* features are not separable by the SVM classifier. As at these radii more than 25 % of all points have four or less neighbours, this is not surprising, since with less than four elements no three invariant moments may be calculated. This finding of generally better classification performance at higher numbers of neighbouring elements agrees

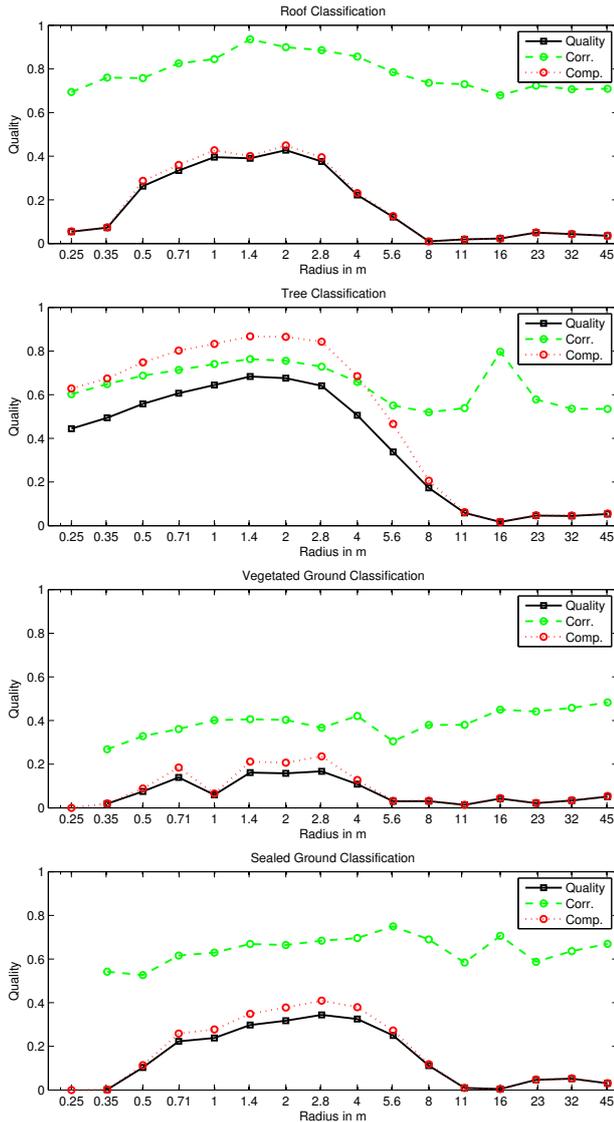


Figure 2: Evaluation of classification results exclusively employing *shape distribution* features for different neighbourhood sizes. Quality measures are calculated according to Section 4.1 and plotted against the radius of a cylindrical neighbourhood.

well with findings published in (Niemeyer et al., 2011), where *covariance* features are used among others.

A very intriguing observation is to be seen at large neighbourhood sizes, where quality increases dramatically. Yet this increase cannot be said to result from a generally better classification performance. Both C and γ parameters of the SVM are very high for these results, indicating an overfitting (Kotsiantis, 2007). This can be explained by a high overlap between the neighbourhood of points within the same reference area, as the reference areas are much smaller than the neighbourhood for feature calculation. As explained in Section 2.2.1, *covariance* features are highly influenced by elements far away from the mean of all observations. For a homogeneous point distribution in the area, the extra number of points in a bigger circle increases more than quadratically by the increase of radius, and the distance of all those extra points in turn contributes quadratically to the covariance matrix. Therefore the neighbourhood of relatively close points has a huge overlap and the resulting features are nearly identical, causing the observed overfitting. Further studies, taking only a reduced random

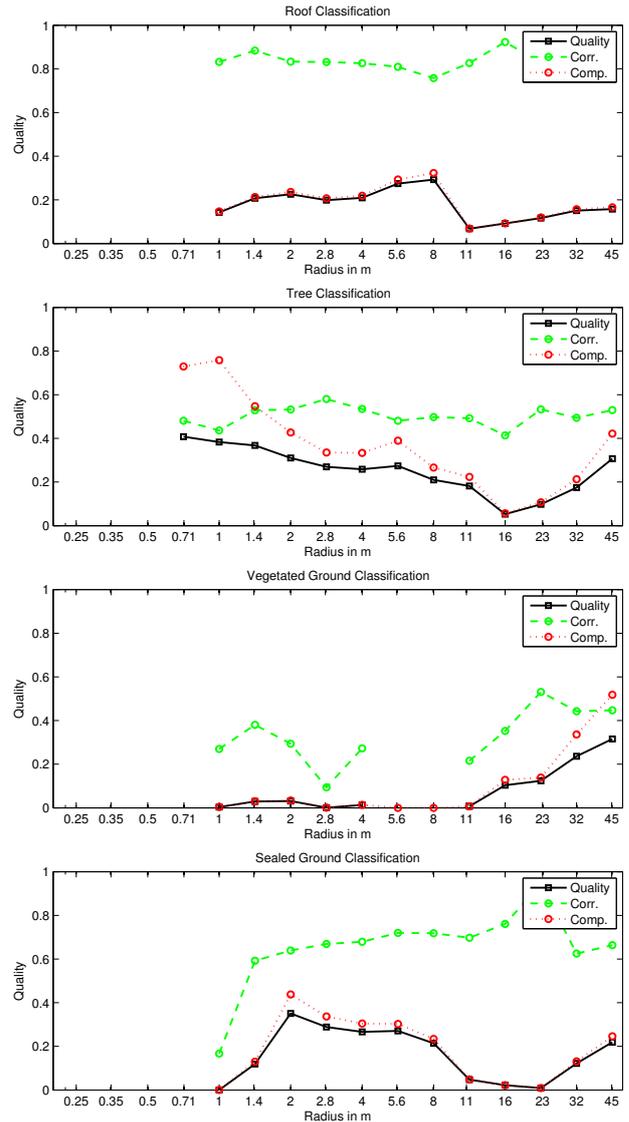


Figure 3: Evaluation of classification results exclusively employing *covariance* features for different neighbourhood sizes. As explained in Section 4.2.1 the quality increase above 16 m radius is not reliable outside of reference areas.

subset per neighbourhood for *covariance* calculation, did not display an increase in *completeness* and *quality* for high neighbourhood radii but were otherwise identical. Therefore increased classification quality for radii above 16 m is regarded as erroneous.

Generally the classification results based on *covariance* features are less smooth than those determined from *shape distribution* features. Rooftops could perform best at 8 m, but not as good as the *shape distribution* result. Sealed ground shows a spike at 2 m, reaching a similar result to the *shape distributions*. Trees are best detected at lower radii, but not as well as when using *shape distributions*, and vegetated ground is virtually undetectable.

4.2.2 Dimensionality-Based Scale Selection. Following an argument stated in (Demantké et al., 2011), the optimum local neighbourhood size may be found by minimizing the absolute value of the Shannon entropy based on dimensionality *covariance* features:

$$E_{Shannon} = -L_{\lambda} \cdot \ln(L_{\lambda}) - P_{\lambda} \cdot \ln(P_{\lambda}) - S_{\lambda} \cdot \ln(S_{\lambda}) \quad (8)$$

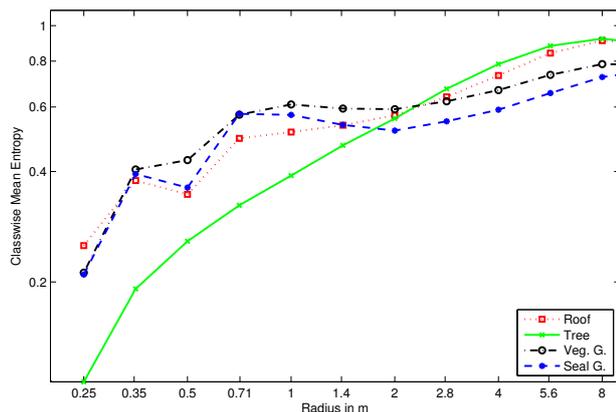


Figure 4: Classwise mean Shannon entropy calculated according to Equation 8 for different neighbourhood sizes.

For radii below 16 m the classwise mean Shannon entropy, ignoring ill-defined values due to $\lambda_i = 0$, is shown in Figure 4. For trees, there is no minimum to be found, whereas for other classes a slight minimum occurs at a neighbourhood radius of 0.5 m. Due to the limited point density of the data set however, no feature separation can be achieved on this scale.

4.2.3 Comparison of Feature Relevance. To further evaluate our features’ performance independent of the used classification scheme, a filter-based feature relevance assessment is performed. The procedure follows (Weinmann et al., 2013). Seven filter-based feature relevance measures are evaluated, each resulting in a relevance rating for all elements of the feature vector. In this case 61 feature vector elements have to be compared (5 *shape distribution* types with 10 feature values each and 11 *covariance* features). The applied score functions evaluate the relation between the values of a feature vector element for all observations and the respective class labels. Tested measures are c_χ from a χ^2 independence test, the Fisher score c_{Fisher} describing the ratio of interclass and intraclass variance, the Gini Index c_{Gini} as a statistical dispersion measure, the Information Gain measure c_{IG} revealing the dependence in terms of mutual information, the Pearson correlation coefficient c_{Pearson} derived from the degree of correlation between a feature and the class labels, the ReliefF measure c_{ReliefF} revealing the contribution of a certain feature to the separability of different classes, and the c_t measure derived from a t -test for checking how effective a feature is for separating different classes.

Since all relevance measures follow different metrics, the value for relative importance was deduced from the ranking order among all feature vector elements. Afterwards, the mean of all importance values was taken for every feature vector element, resulting in a mean importance. A value of one would be achieved if a feature vector element was rated the most important feature by all relevance measures, and zero if it was always rated least important. The mean of all mean importance values belonging to one feature type group is plotted in Figure 5. To avoid any bias by unbalanced reference data, a subset containing 1,000 points per class is investigated.

Comparing the five different *shape distribution* types, all printed as slashed lines, it is clearly seen that the angle between any three random points A3 is only weakly descriptive at those neighbourhood radii that showed the best classification performance in Section 4.1. The volume between any four random points D4 is of great importance here. Obviously the different classes in this test could be best separated by distinctive probability distributions of

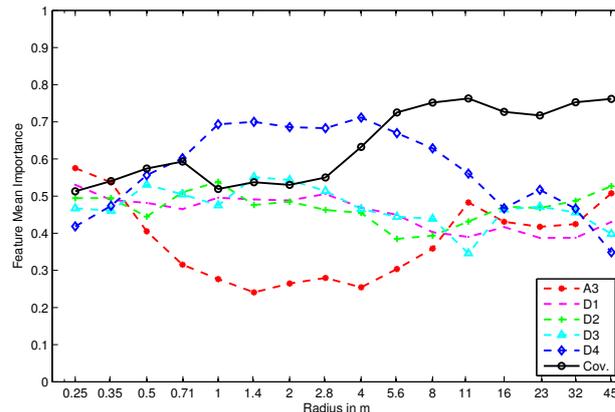


Figure 5: Mean rank of mean feature relevance per feature group for different neighbourhood sizes. A3, D1, D2, D3 and D4 are the *shape distribution* features described in Section 2.2.2, and Cov. the *covariance* features as described in Section 4.2.

random volumetric measures. However, at very small scales, angular and lower dimensional measures like D1, D2 and D3 are of more importance.

As for *covariance* features, a different behaviour can be observed. Below ~ 3 m the importance is roughly the same as for D1, D2 and D3 *shape distribution* features. The slight peak in importance at 0.71 m corresponds well to the optimum neighbourhood size derived from entropy measures (c.f. Figure 4). For higher radii a steep increase followed by high constant importance is measured. This corresponds directly to the scale at which the performance of *shape distribution* features decreases in the SVM classifications (c.f. Figure 2).

4.3 Internal Evaluation on Point Cloud

As the approach with varying scales implies that distinctive neighbourhood sizes may be class-dependent, four separate classification results from different best neighbourhood sizes (c.f. Sections 4.1 and 4.2.1) have to be considered. Combining these separate results necessitates the choice between complete labelling and higher label accuracy. Since the chosen subset of classes may not be complete, we choose only to regard elements with a label probability higher than 50 % as labelled. Therefore some elements may not belong to any class. Also, some points may be found belonging to two or more classes. In this case, the label probability is weighted by the quality of the respective binary classifier before choosing the maximum result.

4.3.1 Quantitative Evaluation. An evaluation was conducted for both *shape distribution* and *covariance* features. Rejection rate (percentage of unidentified elements) and overall accuracy of both results each combined from four binary classifications are shown in Table 1. In this regard the *shape distribution* features obviously outperform the *covariance* features, since the rejection rate is significantly lower whilst the overall accuracy is increased. Not only can more elements be identified, but also more of these elements are identified correctly.

	rejection rate	overall accuracy
<i>shape distribution</i> features	38.4 %	75.6 %
<i>covariance</i> features	56.6 %	63.1 %

Table 1: Rejection rates and overall accuracies for classifications based on *shape distribution* and *covariance* features.

known \ pred.	Roof	Tree	Veg. G.	Seal. G.	Comp.
Roof	77530	10094	4036	6749	78.8 %
Tree	1052	61474	80	47	98.1 %
Veg. G.	4983	4271	8576	11493	29.2 %
Seal. G.	6994	3771	8634	45032	69.9 %
<i>Corr.</i>	85.6 %	77.2 %	40.2 %	71.1 %	
<i>Quality</i>	69.6 %	76.1 %	20.4 %	54.4 %	

known \ pred.	Roof	Tree	Veg. G.	Seal. G.	Comp.
Roof	31544	37225	0	10993	39.5 %
Tree	1040	46941	0	78	97.7 %
Veg. G.	2867	7073	0	13101	0 %
Seal. G.	4699	10306	0	41953	73.7 %
<i>Corr.</i>	78.6 %	46.2 %	-	63.4 %	
<i>Quality</i>	35.7 %	45.7 %	-	51.7 %	

Table 2: Confusion matrices of combined classification results using *shape distribution* (top) and *covariance* (bottom) features, ignoring all cases in which an element could not be detected in any of the classes.

For a more extensive analysis of the different classes' performance, the resulting confusion matrices as well as *completeness*, *correctness* and *quality* are shown in Table 2. The most significant increase is observed for the detection of building roofs, where *quality* almost doubles, mainly due to an increase in *completeness*. The significant *quality* increase for trees is mainly due to increased *correctness*, whereas sealed ground is detected with comparable *quality*. Vegetated ground lacks a comparison, as it can not be detected at all by *covariance* features.

4.3.2 Qualitative Analysis. Figures 6 and 7 depict *shape distribution* classification results as coloured point clouds for qualitative analysis in particular areas. Both the gable roof and tree visible in Figure 6 are generally well classified. Only minor errors occur at the ridge of the roof, where some points are mistaken for vegetated ground, and at the rim of the roof, where some points are misclassified as tree. This is in good agreement with the findings of Figure 2, indicating that trees are generally covered by *great completeness*, but *correctness* is lacking. The high-rise buildings depicted in Figure 7 show a misclassification of flat roofs as sealed and vegetated ground. This is not surprising, as *shape distributions* do not incorporate knowledge about a predominant direction. Therefore, a flat roof and its edge to the ground have the very same characteristics as flat ground and the adjoining edge of a house. Except for the existing confusion between sealed and vegetated ground, those examples explain all main off-diagonal contributions to the confusion matrix.

These misclassification examples demonstrate, that classification results are expected to increase significantly by the use of other, non-geometric features and especially by height distribution features. The latter should enhance results dramatically, since it introduces a predominant direction and avoids the confusion between building roofs and ground. Therefore an evaluation in the benchmark context is intended in a future extended approach.

5 DISCUSSION

At first glance, a main shortcoming of the proposed approach seems to be the high amount of points that cannot be assigned to any class in the evaluated benchmark case. Yet this is inherent in the exclusive examination of their rotation invariant geometrical context without prior assumptions about a predominant direction or a knowledge-based interpretation of man-made objects.

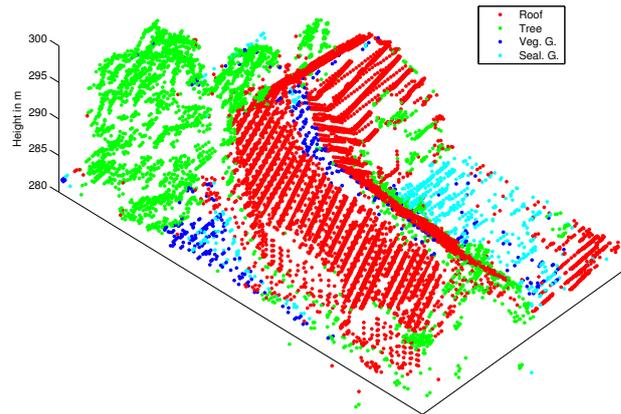


Figure 6: Gable roof and adjacent trees, generally well labelled solely according to geometric *shape distribution* features.

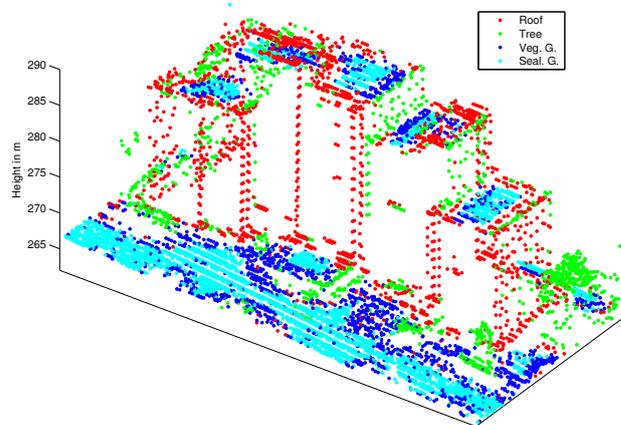


Figure 7: High-rise buildings. Without the use of further features, flat roofs are mistaken as ground.

The class-dependent scale investigations in this urban area context did not yield novel insights about structural characteristics. On the one hand this may be different for other applications, for example in structural vegetation analysis. On the other hand, a combination of different scale features in a unified approach may also enhance classification results for point clouds. For example, the findings of the conducted feature relevance assessment suggest a combination of angular A3 features at very small neighbourhood sizes with volumetric D4 features at ranges between 0.5 m and 16 m around each point in question for this dataset. Broad multiscale studies might be an alternative to extensive approaches optimising the size of every local neighbourhood.

Otherwise *shape distributions* prove to be a robust feature, providing classification results that outperform *covariance* features in all tested cases. They are less sensitive to low point numbers at small scales and are less dependent on an optimally chosen neighbourhood size, thus facilitating studies over wide spatial scales. At huge neighbourhood sizes they are less prone to overfitting within the reference areas due to a high overlap between the neighbourhood of different points. For this urban classification test, *shape distribution* features provide an increase of $\sim 35\%$ in absolute *quality* for roofs, $\sim 30\%$ for trees and $\sim 20\%$ for vegetated ground.

6 CONCLUSION AND OUTLOOK

In this paper *shape distributions* are introduced as a new feature type for geometric point cloud analysis, that outperforms compatible existing features. As they resemble a histogram approach, they require more feature vector elements, yet they hold the potential to describe structures beyond the relatively small reach of homogeneous neighbourhoods well described by *covariance* features. This is especially important when the point density is limited, as in the case of aerial laser scanning data.

In future work we plan to optimise the feature design for remote sensing applications. We aim to combine *shape distribution* features from multiple scales as well as other features types such as height distributions and full-waveform features. Especially vegetation analysis represents a promising field of application for multiscale approaches. Effective classification schemes such as random forests, boosting algorithms or Markov random fields may be used to further enhance classification results.

ACKNOWLEDGEMENTS

The Vaihingen data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer, 2010): <http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html>.

The authors would like to thank Markus Gerke for providing labelled reference data for training and internal testing.

References

- Blaschke, T. and Hay, G. J., 2001. Object-oriented image analysis and scale-space: Theory and methods for modeling and evaluating multiscale landscape structure. *International Archives of Photogrammetry and Remote Sensing XXXIV*, Part 7/W5, pp. 22–29.
- Brodu, N. and Lague, D., 2012. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing* 68, pp. 121–134.
- Chang, C.-C. and Lin, C.-J., 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2(3), pp. 27:1–27:27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Chehata, N., Guo, L. and Mallet, C., 2009. Airborne lidar feature selection for urban classification using random forests. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII*, Part 3/W8, pp. 207–212.
- Cramer, M., 2010. The DGPF test on digital aerial camera evaluation - overview and test design. *Photogrammetrie - Fernerkundung - Geoinformation* 2, pp. 99–115.
- Demantké, J., Mallet, C., David, N. and Vallet, B., 2011. Dimensionality based scale selection in 3D lidar point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII*, Part 5/W12, pp. 97–102.
- Filin, S. and Pfeiffer, N., 2005. Neighborhood systems for airborne laser data. *Photogrammetric Engineering and Remote Sensing* 71(6), pp. 743–755.
- Gerke, M. and Xiao, J., 2014. Fusion of airborne laserscanning point clouds and images for supervised and unsupervised scene classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 87, pp. 78–92.
- Gonzales, R. C. and Wood, R. E., 2002. *Digital Image Processing*. 2. edn, Prentice-Hall, Upper Saddle River, NJ, chapter Histogram Equalization, pp. 91–95.
- Gressin, A., Mallet, C. and David, N., 2012. Improving 3D lidar point cloud registration using optimal neighborhood knowledge. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I*, Part 3, pp. 111–116.
- Hay, G. J., Castilla, G., Wulder, M. A. and Ruiz, J. R., 2005. An automated object-based approach for the multiscale image segmentation of forest scenes. *International Journal of Applied Earth Observation and Geoinformation* 7, pp. 339–359.
- Jutzi, B. and Gross, H., 2009. Nearest neighbor classification on laser point clouds to gain object structures from buildings. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII*, Part 1-4-7/W5, pp. 207–213.
- Kotsiantis, S. B., 2007. Supervised machine learning: A review of classification techniques. *Informatica* 31, pp. 249–268.
- Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, pp. 91–110.
- Maas, H.-G. and Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. *ISPRS Journal of Photogrammetry and Remote Sensing* 54, pp. 153–163.
- Mallet, C., Bretar, F., Roux, M., Soergel, U. and Heipke, C., 2011. Relevance assessment of full-waveform lidar data for urban area classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, pp. 71–84.
- Mitra, N. J. and Nguyen, A., 2003. Estimating surface normals in noisy point cloud data. *Proceedings of the Annual Symposium on Computational Geometry*, pp. 322–328.
- Niemeyer, J., Rottensteiner, F. and Soergel, U., 2012. Conditional random fields for lidar point cloud classification in complex urban areas. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I*, Part 3, pp. 263–268.
- Niemeyer, J., Wegner, J., Mallet, C., Rottensteiner, F. and Soergel, U., 2011. Conditional random fields for urban scene classification with full waveform lidar data. In: Stilla, U., Rottensteiner, F., Mayer, H., Jutzi, B., Butenuth, M. (Eds.), *Photogrammetric Image Analysis*, ISPRS Conference - Proceedings. Lecture Notes in Computer Science, Vol. 6952, Springer, Heidelberg, Germany, pp. 233–244.
- Osada, R., Funkhouser, T., Chazelle, B. and Dobkin, D., 2002. Shape distributions. *ACM Transactions on Graphics* 21(4), pp. 807–832.
- Pauly, M., Keiser, R. and Gross, M., 2003. Multi-scale feature extraction on point-sampled surfaces. *Computer Graphics Forum* 22(3), pp. 81–89.
- Reitberger, J., Schnörr, C., Krzystek, P. and Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing* 64, pp. 561–574.
- Rusu, R. B., Blodow, N. and Beetz, M., 2009. Fast point feature histograms (FPFH) for 3D registration. *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1848–1853.
- Rutzinger, M., Rottensteiner, F. and Pfeifer, N., 2009. A comparison of evaluation techniques for building extraction from airborne laser scanning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 2(1), pp. 11–20.
- Shapovalov, R., Velizhev, A. and Barinova, O., 2010. Non-associative Markov networks for 3D point cloud classification. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII*, Part A, pp. 103–108.
- Strahler, A. H., Woodcock, C. E. and Smith, J. A., 1986. On the nature of models in remote sensing. *Remote Sensing of Environment* 20, pp. 121–139.
- Tombari, F., Salti, S. and Di Stefano, L., 2010. Unique signatures of histograms for local surface description. In: Daniilidis, K., Maragos, P., Paragios, N. (Eds.), *ECCV 2010*, Part III. Lecture Notes in Computer Science, Vol. 6313, Springer, Heidelberg, Germany, pp. 356–369.
- Tombari, F., Salti, S. and Di Stefano, L., 2013. Performance evaluation of 3D keypoint detectors. *International Journal of Computer Vision* 102, pp. 198–220.
- Toshev, A., Mordohai, P. and Taskar, P., 2010. Detecting and parsing architecture at city scale from range data. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 398–405.
- Vosselman, G., Gorte, B. G. H., Sithole, G. and Rabbani, T., 2004. Recognising Structure in Laser Scanner Point Clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI*, Part 8/W2, pp. 33–38.
- Weinmann, M., Jutzi, B. and Mallet, C., 2013. Feature relevance assessment for the semantic interpretation of 3D point cloud data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences II*, Part 5/W2, pp. 313–318.
- West, K. F., Lersch, J. R., Pothier, S., Triscari, J. M. and Iverson, A. E., 2004. Context-driven automated target detection in 3-d data. *Proceedings of SPIE* 5426, pp. 133–143.
- Xu, S., Oude Elberink, S. and Vosselman, G., 2012. Entities and features for classification of airborne laser scanning data in urban area. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I*, Part 4, pp. 257–262.