

## FREE SHAPE CONTEXT DESCRIPTORS OPTIMIZED WITH GENETIC ALGORITHM FOR THE DETECTION OF DEAD TREE TRUNKS IN ALS POINT CLOUDS

P. Polewski<sup>a, b, \*</sup>, W. Yao<sup>a</sup>, M. Heurich<sup>c</sup>, P. Krzystek<sup>a</sup>, U. Stilla<sup>b</sup>

<sup>a</sup> Dept. of Geoinformatics, Munich University of Applied Sciences, 80333 Munich, Germany - (polewski, yao, krzystek)@hm.edu

<sup>b</sup> Photogrammetry and Remote Sensing, Technische Universität München, 80333 Munich, Germany - stilla@tum.de

<sup>c</sup> Dept. for Research and Documentation, Bavarian Forest National Park, 94481 Grafenau, Germany - marco.heurich@npv-bw.bayern.de

### Commission III, WG III/2

**KEY WORDS:** dead tree detection, laser scanning, precision forestry, feature derivation, evolutionary computation

### ABSTRACT:

In this paper, a new family of shape descriptors called Free Shape Contexts (FSC) is introduced to generalize the existing 3D Shape Contexts. The FSC introduces more degrees of freedom than its predecessor by allowing the level of complexity to vary between its parts. Also, each part of the FSC has an associated activity state which controls whether the part can contribute a feature value. We describe a method of evolving the FSC parameters for the purpose of creating highly discriminative features suitable for detecting specific objects in sparse point clouds. The evolutionary process is built on a genetic algorithm (GA) which optimizes the parameters with respect to cross-validated overall classification accuracy. The GA manipulates both the structure of the FSC and the activity flags, allowing it to perform an implicit feature selection alongside the structure optimization by turning off segments which do not augment the discriminative capabilities. We apply the proposed descriptor to the problem of detecting single standing dead tree trunks from ALS point clouds. The experiment, carried out on a set of 285 objects, reveals that an FSC optimized through a GA with manually tuned recombination parameters is able to attain a classification accuracy of 84.2%, yielding an increase of 4.2 pp compared to features derived from eigenvalues of the 3D covariance matrix. Also, we address the issue of automatically tuning the GA recombination meta-parameters. For this purpose, a fuzzy logic controller (FLC) which dynamically adjusts the magnitude of the recombination effects is co-evolved with the FSC parameters in a two-tier evolution scheme. We find that it is possible to obtain an FLC which retains the classification accuracy of the manually tuned variant, thereby limiting the need for guessing the appropriate meta-parameter values.

### 1. INTRODUCTION

In recent years, using LiDAR point clouds as the basis for forest monitoring and general vegetation mapping tasks has become increasingly popular. Usually, the design of application-specific features is a crucial part of solving the detection or classification problem associated with the task of interest. For point clouds, a number of feature families have been proposed that are based on local geometric properties of surface patches. These features are defined at point level and generally exploit the relationships between surface normal vectors (e.g. spin images (SI) (Johnson and Hebert, 1999), Point Feature Histograms (PFH) (Rusu et al., 2008)) or the eigenvalues of the 3D covariance matrix (Mallet et al., 2011) in the neighborhood around the target point. One potentially problematic aspect of applying these shape descriptors is the requirement that the point density be high enough to retain the fine geometric structures within the captured point cloud. This is less of an issue with high-resolution data obtained by means of TLS or MLS systems, however for ALS point clouds having a resolution below 40 points / m<sup>2</sup>, the neighborhood size which allows a robust estimation of the covariance matrix is of the order of several meters. As a consequence, information about smaller structures may be lost.

In this work, we focus on the problem of detecting elongated objects having a well-defined axis from ALS point clouds. Also, we allow for a poor representation of the object surface within the acquired point set due to sparse sampling. The goal is to define a set of features which provide a comprehensive volumetric descrip-

tion of the target object's point cloud. We assume that point subsets corresponding to single candidate objects are available, e.g. from a prior clustering or connected component calculation step. The task consists of determining for each point subset whether it represents the target object. This requires the definition of an aggregate descriptor for the entire scene, which can be accomplished in several ways. One approach is to use global features based on all points belonging to the scene, e.g. Shape Moments (Saupe and Vranic, 2001) or Shape Distributions (Osada et al., 2002). Building on experiences from 2D object recognition, several authors make use of part-based methods, where local descriptors of parts of the scene are combined into a whole. The Bag-of-Features model (Toldo et al., 2009) implements this paradigm by representing the scene in terms of an unordered set of part descriptors. Unfortunately, part-based methods are not directly applicable in the significantly sparser ALS point clouds, because the local surface descriptors lose most of their discriminative power under the low point density conditions. As an attempt to attain some of the attractive properties of the part-based methods, we propose a new kind of feature family, the Free Shape Context (FSC), which describes the entire point cloud using a sequence of cylindrical neighborhoods arranged around a common axis. We utilize the 3D Shape Context (SC) (Frome et al., 2004) as the building block which characterizes each local neighborhood. This shape descriptor is formed by subdividing a cylindrical volume into bins along 3 dimensions, yielding a histogram reflecting the bin occupancy by points within the scene. The analogy to the local methods is that designing an optimal FSC is based on finding local neighborhoods, corresponding to intervals on the axis, which exhibit the best discriminative capabilities for the object

\*Corresponding author.

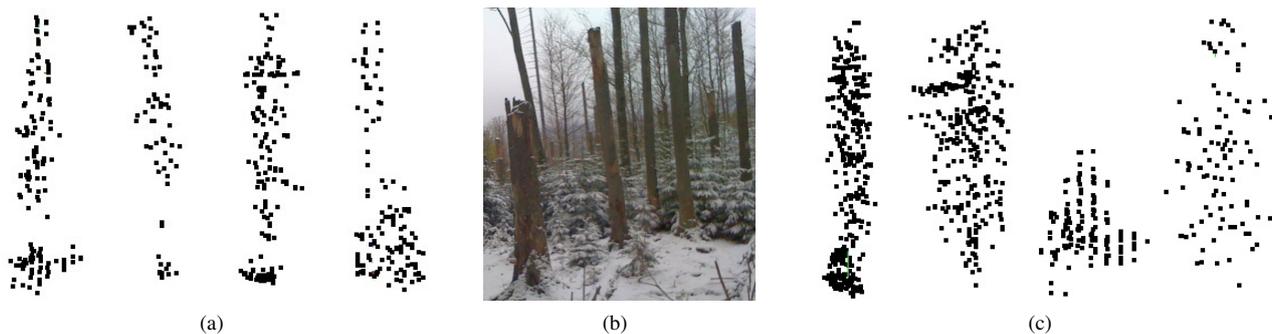


Figure 1: Examples of ALS point clouds of dead trunks (a) and other forest structures (c). The other objects include slender trees, small bushes/ground vegetation, and trees that were poorly captured in the point cloud due to low point density. (b) shows an optical image of several dead trunks surrounded by small trees near the ground.

class of interest. In particular, parts of the axis where no relevant information is usually present can be excluded from contributing anything to the descriptor, preventing the introduction of clutter. Also, the length, radius and number of subdivisions of each local cylindrical neighborhood is independent from all other parts of the descriptor. Due to the required existence of a principal axis, the FSC is primarily applicable to elongated objects such as trees, logs, lamp poles, traffic signs etc. Our choice of the SC over the aforementioned surface descriptors (SI, PFH) stems from the assumption that the objects' surface reflected in the point cloud may be noisy and incomplete, resulting in an inability to reliably estimate the normals. Therefore, the volumetric nature of the SC makes it more suitable for representing the target objects in our scenario. Moreover, the higher number of degrees of freedom (dimensions of the bin divisions) endows the SC with more flexibility to tailor a feature set appropriate for a specific classification problem. Another argument in favor of the SC is its successful application in the related task of detecting fallen tree segments (Polewski et al., 2014, 2015). The high flexibility of the new FSC descriptor comes at a cost of a significantly bigger parameter space, because not only the radii and lengths of the cylindrical neighborhoods, but also their *structure*, i.e. counts, arrangement along the axis, number of embedded cylinders, must be determined through optimization. To tackle this problem, we apply a genetic algorithm (GA) (Goldberg, 1989), a powerful meta-heuristic optimization method which maintains a population of solutions ('individuals') and attempts to *evolve* them towards an optimal solution by promoting high-quality individuals and by recombining existing individuals into new ones through cross-over and mutation, in analogy to an evolutionary process. Among their many successful applications in various fields, GAs have been used to find the optimal neighborhood sizes for features obtained from the 3D covariance matrix (Waldhauser et al., 2014), as well as the optimal parameters of a 3D shape descriptor based on two cylindrical neighborhoods (Wegrzyn and Alexandre, 2013).

The issue of selecting an optimal neighborhood size for feature construction in point clouds has received considerable attention within the remote sensing community. Weinmann et al. (2014) propose an eigenentropy-based scale selection for spherical point neighborhoods. Blomley et al. (2014) analyze the influence of scale on classification accuracy for cylindrical neighborhoods. The authors employ the aforementioned originally *global* Shape Distributions to describe *local* point neighborhoods using an adaptive histogram binning scheme. However, it should be noted that these and other proposed approaches are defined at the *individual point* level, whereas our method aims at describing the entire *object* by concatenating local descriptors of varying complexity.

We test the proposed shape descriptor in the context of the task of detecting standing dead tree trunks. Information about the dis-

tribution of dead wood in forests is of interest in environmental sciences because of its role in biodiversity of plant and animal species (Stokland et al., 2012) and nutrient cycles (Siitonen et al., 2000). Due to their small cross-section area, dead trunks are hardly visible on aerial imagery and hence cannot be reliably detected from this data source. Therefore, full 3D information provided by LiDAR point clouds is advantageous for this scenario. The dead trunks exhibit some variability in appearance due to residual branches and the potential presence of ground vegetation. A further difficulty is their resemblance to slimmer trees (see Figure 1).

The rest of this paper is structured as follows: in Section 2. we describe the overall object detection pipeline, including input data requirements. Section 3. introduces the new Free Shape Context descriptor in detail and outlines the Genetic Algorithm approach used for optimizing the FSC parameters. Section 4. describes the study area, experiments and evaluation strategy. The results are presented and discussed in Section 5. Finally, the conclusions are stated in Section 6.

## 2. OVERALL STRATEGY

The input to our detection pipeline consists of ALS point clouds given in the form of 3D coordinate sets. Both discrete return and full waveform systems may be used, however the point density should be sufficient for distinguishing the objects of interest within the point clouds. Our method is based purely on geometric features and does attempt to utilize radiometric information. The entire processing pipeline is depicted in Figure 2. In the first step, the point cloud is partitioned into subsets which represent single objects. We view the partitioning step as an external procedure and do not focus on it in this study. It can be carried out using any appropriate clustering method. In particular, we use the Normalized-Cut based approach by Reitberger et al. (2009). For the case when the target objects are well separated in the scene, a connected component search such as the pre-processing step described by Velizhev et al. (2012) may also be used. Each point cluster is then processed individually as an object candidate. We first compute the axis of the point cloud using a Sample Consensus method. The object-level features are then calculated and each candidate object is labeled as either target class or non-class. In order to optimize the discriminative power of the features, we also require that a set of objects (outputs from the clustering step) labeled as either class or non-class be provided to serve as training data.

### Calculating object axis

In this step, we find, for each candidate object, a principal axis which best fits the corresponding set of points. Since the cluster-

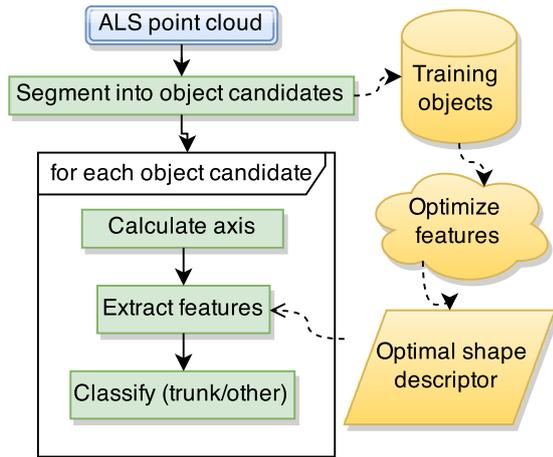


Figure 2: Overview of dead trunk detection pipeline.

ing step is likely to produce imperfect partitions, the presence of clutter and/or parts of other objects is expected within the point cloud. For this reason, we apply an estimation method from the Random Sample Consensus (SAC) family to calculate the axis. As opposed to fitting a model to the entire input data, SAC methods aim at improving robustness to outliers by randomly generating model hypotheses and evaluating how well they fit *subsets* of the data according to a scoring scheme. The score corresponds to the total fit error and hence the lowest-scoring model is returned as the fitting result. The error is usually a function of the number of inlier points within a fixed distance from the model and also of the quality-of-fit to the inliers. Since we wish to maximize both of these possibly contradictory criteria simultaneously, it is necessary to define a tradeoff between them in order to obtain a single criterion function for optimization. Specifically, we make use of the M-estimator SAC variation (Torr and Zisserman, 2000) with a line model, cylindrical neighborhood with radius  $r_{sac}$  and point-to-line perpendicular distance  $d(\cdot, \cdot)$ . The error of line model  $L$  on point set  $P$  is defined as follows:

$$C_P(L) = \sum_{p \in P} \rho(d^2(p, L)), \rho(d^2) = \begin{cases} d^2 & d^2 < r_{sac}^2 \\ r_{sac}^2 & d^2 \geq r_{sac}^2 \end{cases} \quad (1)$$

We additionally introduce a constraint on the maximum angular deviation  $\alpha_{dev}$  of the candidate line models from the world Z coordinate axis to account for the phenomenon of gravitropism in trees. A sample result of the calculated axes for several point clouds is depicted by Figure 3.

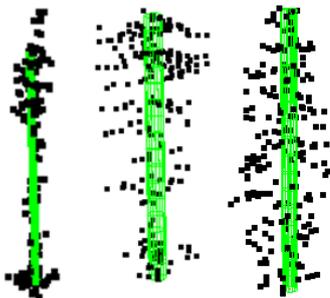


Figure 3: Axes of point clouds determined by Sample Consensus.

### 3. SHAPE DESCRIPTORS

#### 3.1 3D Shape Contexts

Originally introduced by Frome et al. (2004) as a local shape descriptor of surface patches, the 3D Shape Context (SC) attempts to capture the shape distribution inside a volume of a point cloud by subdividing the volume into bins along 3 dimensions and creating a histogram by counting the points in each bin. We replace the original spherical neighborhood centered on a point with a cylindrical neighborhood around an axis. The cylinder is divided along the axial, radial and angular dimensions (see Figure 4). Note that the angular division may cause the SC to lose its rotational invariance, since two rotated copies of the same point cloud with the same axis may in general produce different histogram signatures. To remedy this, a method of defining the common  $0^\circ$  orientation is required. An alternative would be to augment the training set with rotated copies of each labeled example. In this work, for the sake of simplicity we refrain from dividing the cylinder along the angular dimension and thus retain the invariance w.r.t. rotation. The radial subdivisions follow an exponential law of the form:  $r_i = r_0 r_B^i$ , where the base radius  $r_0$  and the exponential base  $r_B$  are user-defined parameters. The remaining parameters consist of the SC's length  $l$ , the number of axial divisions  $n_L$ , and the outer radius  $r_{max}$ .

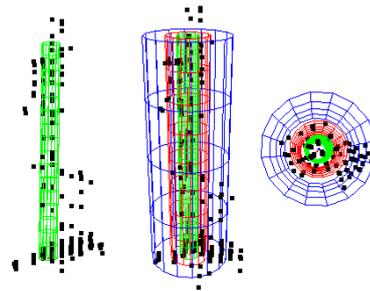


Figure 4: Cylindrical shape context around axis. Center: side view. Right: top view. Only single angular category shown.

#### 3.2 Free Shape Contexts

The standard Shape Context defined in Sec. 3.1 has a uniform structure. However, for a specific detection task, it is unlikely that discriminative information is evenly distributed along the object axis. In particular, some parts may be heavily affected by clutter, which renders them useless as a source of distinctive features. Furthermore, for the parts which are relevant, the optimal neighborhood may significantly vary in size. Based on these two intuitions, we propose a new shape descriptor, the Free Shape Context (FSC). The FSC is essentially a stack of cylindrical volumes arranged around the axis. This can be viewed as a sequence of adjacent standard SCs which are divided only along the radial dimension, i.e. potentially contain embedded cylinders of smaller radii (see Figure 5). The lengths and radii of the constituent SCs are unrelated and may be adjusted independently. Additionally, every cylinder has an activity state which controls whether the associated histogram bin is allowed to emit a feature to the final descriptor. Inactive cylinders can be regarded as buffer elements - they only define the structure of the FSC without affecting feature values. The FSC is parameterized by the ordered sequence of the member SCs' parameters. Each such SC is in turn described by its length and sequence of cylinder radii along with their corresponding active states.

**3.2.1 Parameter space size** The FCS's increased number of degrees of freedom results in a considerably larger parameter

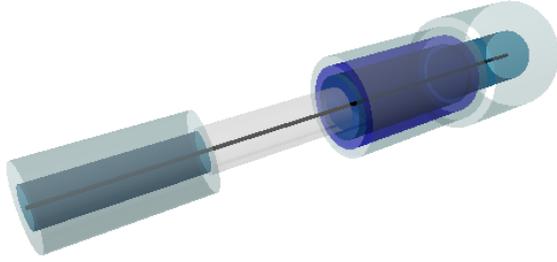


Figure 5: A Free Shape Context. Gray cylinder near middle is inactive - it will not contribute to the shape histogram.

space. To analyze the parameter space, we first introduce some notation. Let  $N_R, N_L$  be the number of quantized radius and length configurations, respectively, that can be examined with a reasonable computational effort. We then have that for an FSC consisting of a single SC with a single cylinder, the number of configurations is  $2N_R N_L$ . The factor 2 arises from the two possible activity states. If we add a new cylinder to the SC, we can set the new radius to any other of the  $N_R$  values, assign it an active or passive state and combine it with any previous configuration, resulting in  $2^2 \binom{N_R}{2} N_L$  options. Following this line of reasoning, the number of configurations for an FSC consisting of one SC with  $i$  cylinders is:  $C_{1,i} = 2^i \binom{N_R}{i} N_L$ . Let  $n_r$  denote the maximal number of radial bins. We can express the total number of FSC configurations consisting of  $k$  SCs using  $C_{1,i}$ :  $C_k = (C_{1,1} + \dots + C_{1,n_r})^k$ . This expression is dominated by the term  $\{2^{n_r} \binom{N_R}{n_r} N_L\}^k$ . It can be easily verified that even for a coarse quantization of the radii and lengths and a small number of SCs, the computational complexity renders any exact or even grid search approach infeasible.

**3.2.2 Approximate optimization** It is possible to alleviate the computational burden by making two simplifying assumptions. The first assumption states that the problem of finding the most discriminative FSC exhibits the optimal substructure property. In practice this means that if the best FSC  $f$  of length  $L$  contains its last shape context  $s$  from positions  $D$  to  $L$  along the axis, then the FSC obtained by removing  $s$  from  $f$  must be the most discriminative FSC of length  $D$ . The second assumption recasts this requirement in the context of cylinder radii within a single SC. The optimal substructure premise gives rise to a dynamic programming strategy for calculating the optimal FSC. Let  $W_{k,j}$  denote the best score of a  $k$ -component FSC of (quantized) length  $j$ , and  $P_{k,j}$  the corresponding FSC parameters. Also, let  $V_{k,j}(P, l)$  be the smallest error of an FSC formed by combining the parameters  $P$  with a new SC consisting of  $k$  cylinders with a total (external) radius  $j$  and length  $l$ . Let  $S_{k,j}^{P,l}$  denote this optimal SC, and  $c_r$  - a cylinder with radius  $r$ . The minimum cost can then be calculated:

$$W_{k,j} = \min\{W_{k,j-1}, \min_{l,u,v} V_{u,v}(P_{k-1,j-l}, l)\} \quad (2)$$

$$V_{k,j}(P, l) = \min\{V_{k,j-1}(P, l), \min_r E(P \oplus [S_{k-1,j-r}^{P,l} \cup c_j])\}$$

In the above,  $\oplus$  denotes appending an SC at the end of an FSC, and  $E(\cdot)$  is the evaluation function (e.g. classification error).

### 3.3 Evolving optimal FSC parameters

Due to the size of the parameter space of the FSC, it is unlikely that any approach which attempts to sample parameters in equal intervals (such as the dynamic algorithm described in Sec. 3.2.2) will attain a near-optimal solution. The parameter space is usually highly heterogeneous, with small isolated regions of high-quality solutions. Also, the optimal substructure assumptions

may not be valid. Therefore, it appears more beneficial to instead apply an optimization strategy which is capable of dedicating more effort to the exploration of promising regions of the solution space at the expense of omitting its weaker parts, thereby making a better use of the allocated resources. One such method is the Genetic Algorithm, a biologically-inspired meta-heuristic which has been successfully applied to the optimization of multi-problems in science and engineering.

**Genetic algorithms** As a representative of evolutionary computation methods, Genetic Algorithms (Goldberg, 1989) are an optimization technique which attempts to mimic an evolutionary process. In the language of the biological metaphor, solutions to the problem of interest are referred to as *individuals* and are encoded using a *genetic representation*. The value of the optimized function  $f$  associated with a solution is called the *fitness function* of the individual. The genetic algorithm acts on a *population* of individuals by iteratively selecting the best elements based on their fitness and subsequently recombining them to create a new generation of individuals. The recombination usually consists of two parts: mutation and crossover. Mutation is a unary operator which randomly alters small parts of the target individual's genome. Crossover is an operator acting on a pair of individuals ('parents') with the purpose of randomly exchanging parts of their genomes. The iteration is repeated until a convergence criterion is met. It is expected that in the course of evolution, the individuals will gradually improve their fitness until a (local or global) extremum of the fitness function is attained.

**3.3.1 FSC genotype representation** We use a variable-length, structured genetic representation  $G(c)$  of the Free Shape Context  $c$ . At the top level, the genome is an ordered sequence of the constituent SCs' sub-genomes:  $G(c) = (G(sc_1), \dots, G(sc_N))$ . Each sub-genome  $G(sc_k) = (l_k, \{[r_{k,i}, a_{k,i}]\}_{i=1..n_k})$  is composed of the length  $l$  and list of the  $n_k$  cylinder radii  $r$  and active states  $a$ . We make the distinction between *structural parameters*, which describe the number of SCs, the number of radii within each SC as well as the active state of each cylinder, and *size parameters*, which encode the cylinder lengths and radii.

**3.3.2 Mutation** For the structural parameters, we employ a simple mutation scheme which randomly adds and removes SCs from the FSC, adds and removes cylinders from a single SC, or flips the active state. All of these events occur with a small probability  $p_{m,1}$ . For the size parameters, a relative mutation strategy is used. A random number  $\kappa$  is drawn from a zero-mean normal distribution with standard deviation  $\sigma_m \in (0, 1)$ , and the target parameter  $p$  is updated according to:  $p \leftarrow p(1 + \kappa)$ . This mutation takes place with a probability  $p_{m,2} > p_{m,1}$ . The mutation rates for the structural and size parameters are distinct because we want the evolutionary process to spend sufficient effort examining the size parameters for a given structure. This is motivated by the fact that the number of structural configurations is small compared to the number of length/radius configurations.

**3.3.3 Crossover** The crossover operator is implemented by a randomized strategy which picks, with uniform probability, one of three procedures. The first two, one-point and two-point crossover, take place only at the top level and involve swapping SCs at corresponding positions in the genome. The crossover points (one or two) are picked at random. Finally, the Fuzzy Recombination (FR) crossover takes place between corresponding SCs at each position in the parents' genomes, with probability  $p_{FR}$ . In this procedure, the lengths and radii of the SCs' cylinders are subjected to the FR operator, which aims at integrating the effects of crossover and mutation. For two values  $p_1, p_2$  and a width parameter  $0 \leq d_{FR} \leq 0.5$ , FR defines a bi-triangular probability

distribution on the output value, with modes located at  $p_1$  and  $p_2$  (Figure 6). Large values of  $d_{FR}$  lead to introducing more diversity into the population, whereas small values benefit the protection of good existing solutions.

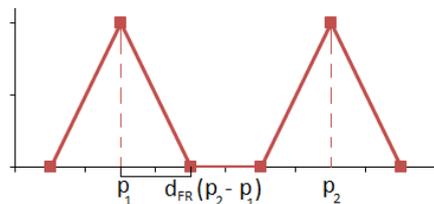


Figure 6: Bi-triangular output distribution induced by Fuzzy Recombination operator.

**3.3.4 Fitness function** The evolution of the Free Shape Contexts is driven by their overall classification performance on a set of labeled objects (point clouds). In particular, we do not partition the labeled pool into training and test sets, but instead apply a cross-validation scheme for calculating the classification error. It is important that the error measure be deterministic in the sense of always obtaining the same error value upon repeated calculation for the same individual. Otherwise, the evolution could be promoting random constellations of training vs. test set partitions which lead to lower errors at the expense of true discriminative performance. For this reason, we apply leave-one-out cross-validation which is free from any randomness.

### 3.4 Adaptive tuning of GA parameters

A key issue in a successful application of a GA to a specific domain is whether the GA can attain a balance between *exploitation* and *exploration*. The former can be understood as the capability to thoroughly search promising regions of solution space, thereby 'exploiting' the neighborhood of high-quality solutions. The latter is associated with a broader strategy of 'exploring' all relevant parts of the solution space as opposed to the evolutionary process getting stuck in a local optimum. The recombination meta-parameters (e.g.  $p_{m,1}, p_{m,2}, d_{FR}$ ) have a decisive influence on the GA's dynamic behavior (Herrera and Lozano, 2001). On the other hand, manually setting these meta-parameters is a non-trivial task usually accomplished through resource-expensive trial-and-error procedures. For these reasons, methods have been developed which are able to adaptively change the meta-parameter settings during the evolution, based on feedback from the evolutionary process. We take an approach inspired by the contributions of Herrera and Lozano (2001) as well as Lee and Takagi (1993). The principal idea is to utilize a fuzzy logic controller (FLC) to calculate the FR width  $d_{FR}$  separately for each application of the FR crossover operator, based on the fitness of the parents and the overall diversity of the entire population. The processing pipeline of the GA augmented with the FLC is shown in Figure 7. The fuzzy rules comprising the optimal FLC are themselves found using a genetic algorithm. Specifically, a two-tier evolution scheme is set up. At the top level, a small population of FLCs is evolved. Each FLC  $c$  is evaluated by executing an instance of the Free Shape Context evolution (bottom level) and letting  $c$  control the FR width. The fitness of  $c$  is the *offline performance* of the bottom-level GA, i.e. the best fitness of any individual found up to the current iteration, averaged over all iterations.

**3.4.1 Fuzzy inference** Fuzzy logic (Klir and Yuan, 1995) is an extension of Boolean logic to multi-valued domains. As opposed to crisp Boolean values of true and false, a fuzzy variable may assume any continuous value from the interval  $[0; 1]$ . This value represents a subjective degree of belief that a proposition is

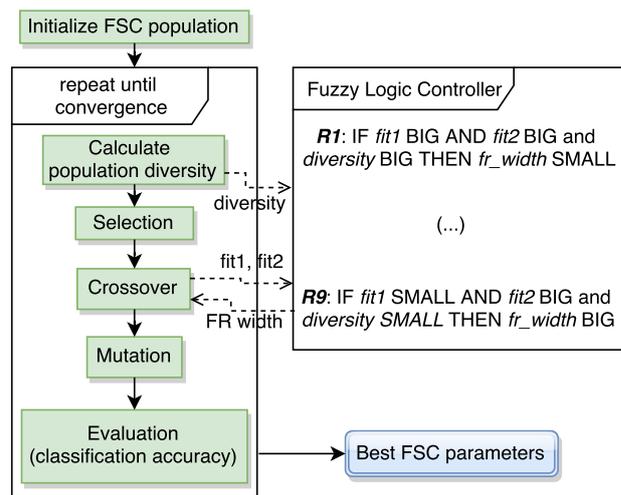


Figure 7: Overview of GA for finding best FSC parameters. FLC adapts the FR width for crossover using 9 fuzzy rules based on the population diversity and the parents' fitness (fit1, fit2).

true, with 0, 1 representing absolute falseness and absolute truth, respectively. This formulation makes fuzzy logic appropriate for modeling uncertain or vague relationships between variables. To perform fuzzy inference, an additional level of abstraction in the form of linguistic variables and fuzzy sets is defined. A linguistic variable represents any quantity that is of interest in the application (e.g. 'temperature'), along with its enumerated possible values (e.g. 'high', 'medium', 'low'). Each variable level is associated with a fuzzy set, which is in turn defined by a fuzzy membership function  $\mu(x) \in [0; 1]$ . Based on linguistic variables  $X, Y, Z$  with levels  $a, b, c$ , fuzzy rules may be defined:

$$IF X IS a AND Y IS b THEN Z IS c$$

The application of fuzzy rules requires defining a fuzzy 'and' operator, implication function, rule output aggregation and defuzzification method for converting the fuzzy rule results into a crisp number.

**3.4.2 Fuzzy logic controller for GA** The FLC responsible for setting the FR width uses three linguistic variables as input. The first two are the ranks (ordinal numbers) of each parent's fitness within the entire population. The two fitness variables assume the values 'high' and 'low'. Their corresponding fuzzy membership functions are fixed (see Figure 8(a)). Following Lee and Takagi (1993), we also introduce the population diversity as an input variable with three levels: 'high', 'medium', and 'low', to enable the FLC to employ different strategies depending on the current stage of the evolution. The fuzzy membership functions of the diversity levels are parameterized by the locations  $d_l$  and  $d_m$  (Figure 8(b)). Finally, the output variable (FR width) also assumes the three levels 'high', 'medium', and 'low', with fuzzy membership functions defined by the parameters  $o_l, o_m$ , and  $o_h$  (Figure 8(c)). For each diversity level, three decision rules are defined which assign an FR width level for every combination of parent fitness ranks (the high-low and low-high cases are symmetric), yielding a total of 9 rules. The output levels of all rules are evolved together with the 5 fuzzy membership location parameters to produce the FLC whose control actions lead the embedded Free Shape Context evolution to attain the best offline performance.

**3.4.3 Measuring population diversity** The population diversity is calculated as the average distance between all pairs of FSCs in the population. The distance between two FSCs  $c_1, c_2$

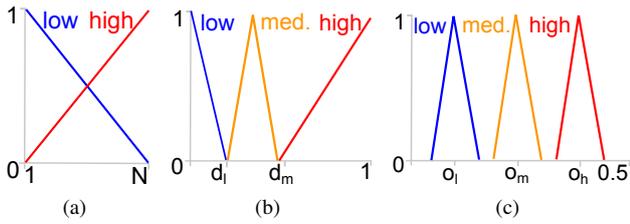


Figure 8: Fuzzy membership functions of (a) fitness ranks ( $N$  - population size) (b) population diversity, (c) output variable.

is normalized on the interval  $[0; 1]$  and is taken as the average distance between the member SCs of  $c_1$  and  $c_2$  at corresponding positions. For two SCs  $s_1, s_2$ , their distance is in turn defined as  $1 - V_i/V$ , where  $V$  is the maximum of  $s_1$  and  $s_2$ 's outer cylinder volumes, and  $V_i$  is the sum of the intersection volumes of the corresponding cylindrical bins in  $s_1$  and  $s_2$ . If the number of SCs within  $c_1$  and  $c_2$  is not equal, each excess SC without a corresponding element is regarded as having the maximum distance of 1 (consistent with an intersection volume of 0).

## 4. EXPERIMENTS

### 4.1 ALS data

For testing the proposed methodology, we used a  $1 \times 1 \text{ km}^2$  plot located in the Bavarian Forest National Park ( $49^\circ 3' 19'' \text{ N}$ ,  $13^\circ 12' 9'' \text{ E}$ ), which is situated in South-Eastern Germany along the border to the Czech Republic. The study was performed in the mountain mixed forests zone consisting mostly of Norway spruce (*Picea abies*) and European beech (*Fagus sylvatica*). The dead wood originated from an outbreak of the spruce bark beetle (*Ips typographus*) in recent years. The airborne full waveform ALS data were acquired using a Riegl LMS-680i scanner in July 2012 with a nominal point density of 30-40 points/ $\text{m}^2$  and pulse rate of 266 kHz. The flying altitude of 650 m led to a footprint size of 32 cm.

### 4.2 Reference data

To create a data set for assessing the performance of each feature family, we first performed the segmentation into individual trees (Reitberger et al., 2009). We then manually selected and labeled a total of 285 objects to serve as the basis for our study, of which 56% were living trees and 44% were dead trunks. The average point counts in the segmented point clouds of the objects were 340 and 208 respectively for living trees and dead trunks. The selection criterion was associated with ensuring a possibly broad coverage of both classes' appearance ranges within the ALS point cloud. The labeling was done based on visual inspection of each object's point subset.

### 4.3 Detailed experiments

We designed two experiments to attain detailed insight into the performance of the proposed Free Shape Context features compared to other competing feature families in the context of our classification problem. In all calculations, we used the same input data, classification model (logistic regression), deterministic evaluation criterion (leave-one-out cross validation), and pre-calculated object axes, therefore the differences between the detection rates can solely be attributed to the features themselves. A value of  $2^\circ$  was used for the angular deviation threshold  $\alpha_{dev}$ .

**4.3.1 Comparison of feature families** In the first experiment, we compare the performances of Free Shape Contexts, standard

(uniform structure) Shape Contexts, and features based on eigenvalues of the covariance matrix (referred to as CE). To understand what part of the potential gain in classification accuracy is associated with the features as opposed to being merely a result of applying a better optimization strategy (genetic algorithm vs. grid search), we tested each feature family in two configurations. The first variation uses parameters optimized by grid search (in case of FSC the dynamic algorithm described in Sec. 3.2.2), while the second configuration is based on GA parameter optimization. In all GA-based experiments, we used the same population size (1000) and generation count (150) as well as selection mechanism (size 5 tournament selection). Because the GA optimization result is non-deterministic, we repeated the experiments 15 times from random initializations. For FSC, the maximum allowed number of member SCs was 12, each with up to 5 radial bins. The meta-parameter settings (Sec. 3.3) were:  $p_{m,1} = 0.03$ ,  $p_{m,2} = 0.15$ ,  $\sigma_m = 0.1$ ,  $p_{FR} = 0.6$ . In this experiment, the FR width was fixed at a value of  $d_{FR} = 0.05$ . This value was chosen empirically from a small set of discrete levels based on the results of preliminary experiments with smaller population sizes and generation counts. For the evolution of standard SCs, we applied a vector of 5 real numbers representing the parameters  $r_0, r_B, l, n_l, r_{max}$  as the SC's genotype, with two-point crossover and Gaussian mutation (such as the size parameter mutation in Sec. 3.3.2). Finally, for the CE features, we defined a simple variation of the Bag-of-Features model (Toldo et al., 2009). For each point within the input point cloud, we calculate the values of the 8 CE features defined in (Weinmann et al., 2014) (Table 1). We then partition the domain of each feature  $i$  into  $b_i$  equidistant bins. Each bin defines a 'visual word' for the associated attribute. The descriptor of the entire scene consists of the decorrelated, concatenated histograms of visual word counts for each feature over all points within the scene. The parameters of this model are the numbers of feature partitions  $b_i$  as well as the common neighborhood radius for calculating the covariance matrix around each point. Note that a value of 0 partitions causes the feature to be omitted from analysis, which enables a simultaneous feature selection. We utilize a similar strategy for genetic representation and recombination as was the case with standard SC evolution. The grid search/dynamic algorithm approximation strategies were carried out with a number of inspected configurations greater than or equal to the number of individuals processed in a single GA execution ( $1000 * 150 = 150000$ ).

**4.3.2 Evaluation of FLC for adaptive control** The goal of the second experiment is to determine whether a Fuzzy Logic Controller can help alleviate the burden of manually tuning the GA meta-parameters. For this purpose, we implemented the two-tier evolution scheme of FLCs controlling the FR width parameter, as described in Sec. 3.4. At the top level, a population of 40 FLCs is maintained. Each FLC is evaluated by executing an internal evolution of FSCs on a reduced population size (100) and generation count (100) to ensure computational feasibility. The initial population of FSCs is the same for every internal evolution to prevent FLCs obtaining an unfair advantage due to a better random initialization instead of better actual performance. We also carried out a baseline experiment to estimate the frequency of obtaining the optimal solution without the FLC using the same computational effort necessary for evolving the FLC. For this purpose, the FSC evolution was repeated on randomly initialized populations with the same GA meta-parameters as during the FLC evolution, except the FR width which was fixed. Finally, we tested whether the FLC evolved optimized on a single starting population would also lead to improving the performance of other, random populations. This was done by generating 25 random populations and for each one performing the evolution with and without the FLC, leading to pairs of related results.

## 5. RESULTS AND DISCUSSION

### 5.1 Comparison of feature families

Table 1 summarizes the overall classification results for all 6 tested combinations of feature families and optimization algorithms. The computational effort is measured in units equal to the number of evaluations of individuals in a single GA run, i.e. 150000, with 1 unit corresponding to about 6 hours of absolute processing time on a quad-core Intel® Xeon® processor with a base frequency of 3.7 GHz. In case of the two SC-based feature types, the genetic algorithm is able to outperform the grid search/dynamic algorithm by up to 4.6 percentage points (pp), using less parameter configuration evaluations. Also, in case of the Free Shape Contexts, the dynamic algorithm is not able to find a solution comparable to one obtained through GA despite a tenfold increase in the used computational resources. This is an expected result due to the discussed high computational complexity of optimizing the FSC parameters (Sec. 3.2.1). For the sake of fairness, it should be noted that because they are stochastic in nature, genetic algorithms do not always converge to the best found solution. In fact, we observed that for all three feature types, out of 15 runs with random initializations, the best solution was only obtained 1-2 times. This emphasizes the difficulty in finding an optimal set of GA meta-parameters which reliably leads to convergence under different initializations. On the other hand, even the mean classification accuracies averaged over all restarts outperform their counterparts from grid search for SC and FSC. Interestingly, the GA seems to fail as an optimizer of CE param-

Descriptor	Effort	Overall accuracy [%]		
		max.	$\mu$	$\sigma$
<b>Shape Context</b>				
Grid Search	1.07	76.1	76.1	0.0
GA	1	80.7	78.7	1.3
<b>Covariance</b>				
<b>Eigenvalue BoF</b>				
Grid Search	1.09	80.0	80.0	0.0
GA	1	77.2	75.6	1.0
<b>Free Shape</b>				
<b>Context</b>				
Dyn. Search	10.44	80.0	80.0	0.0
GA	1	84.2	82.7	0.8

Table 1: Evaluation results for considered features w.r.t. classification accuracy. Workload reflects the number of evaluated parameter sets (1 = number of evaluations in a single GA run).

eters, its best found solution lagging by 2.8 pp behind the grid search result. We hypothesize that this may be due to the CE parameters consisting mostly of integers (numbers of feature divisions) as opposed to the SC and FSC parameters being mostly reals (cylinder lengths and radii). This makes it harder for the evolutionary process to find a gradient of the optimized function w.r.t. the more discretized parameter space. Perhaps a genetic representation or recombination method more suitable for integer variables could remedy this problem. In terms of the best-performing solution, the CE features appear to be inferior to both SC and FSC. This indicates that the point density afforded by ALS point clouds is not sufficient so as to support discriminating between different surface types on a scale finer than several meters. This claim is further supported by the fact that the optimal

CE performance arose from a neighborhood radius of 4.4 m for the covariance matrix calculation (from an available interval of 0.1-5 m), which implies that the no meaningful way to utilize information based on smaller neighborhoods could be found. The proposed Free Shape Context-based features have delivered the best performance. For average quality of the GA-optimized solutions, FSC outperforms SC by 4 pp and CE by 7.1 pp, whereas in case of the best found solution, the respective gains are 3.5 pp and 4.2 pp. Regarding the feature counts associated with each descriptor, the best FSC feature set had a cardinality of 10, equaling the corresponding count for CE and lying significantly below the SC-related count of 30. Combined with the superior performance, this suggests that, compared to its competitors, the FSC descriptor was able to achieve a better balance between finding the relevant, discriminative parts of the target objects and omitting non-informative parts. To investigate the statistical significance of the performance differences, we used the one-tailed binomial test. We regard the outcome of the labeling of  $N$  objects by a classifier with accuracy  $p$  as a random variable with a binomial distribution  $B(N, p)$ . In particular, for the pair FSC-CE, we test if the result of 240 successfully classified objects is likely to have arisen from a distribution  $B(285, 0.8)$ . The obtained  $p$ -value is 0.04, so the difference is significant at a level of 0.05. Similarly, for the pair FSC-SC, the  $p$ -value lies at 0.07. In contrast, the difference between SC and CE is considerably less significant with a  $p$ -value of 0.42.

### 5.2 Evaluation of FLC for adaptive control

The evolution of the Fuzzy Logic Controller for the FR width parameter (Sec. 4.3.2) resulted in an offline performance of 0.82 after 3 generations. The corresponding FLC was able to drive the evolution of the initial reduced population of 100 FSCs over 100 generations to achieve the best-known solution (classification accuracy 84.2%) found by means of the manually tuned GA acting on the normal-sized (1000) population over 150 generations (Table 1). Without the FLC (using a fixed FR width), the same initial population converged to an FSC with classification accuracy 81.4%. This indicates that the FR width meta-parameter had a key role in the evolutionary process and that the FLC was able to successfully control it to produce an increase in solution quality. The total workload needed to evolve the successful FLC was 3 outer generations \* 40 evaluations of 100 FSCs \* 100 generations, which gives a total of 1,200,000 FSC evaluations. To match this in the baseline experiment, we carried out 120 evaluations of 100 FSC individuals over 100 generations, with the same GA meta-parameters except for the fixed FR width. The best solution performance of 84.2% was found in exactly one of the 120 restarts, yielding a frequency of 0.0084. Therefore, both the FLC-based approach and the fixed FR width method led to the same result using the same amount of computational resources. However, it should be pointed out that the meta-parameters used in the baseline experiment, including the FR width, were manually tuned for the feature comparison experiment (Sec. 4.3.1) and therefore can be considered above average in terms of performance. Also, the low success rate of 1/120 suggests that in this instance, attaining the best solution was a lucky occurrence since no attempt is made to adapt the search directions based on intermediate results. In contrast, the FLC evolution actively alters the FR width control strategy based on feedback from the evolutionary process, which may reduce the reliance on pure chance. Finally, we address the FLC's generalization capability to populations it has not been optimized on. We applied the Wilcoxon signed-rank test on the pairs of FLC-controlled and fixed FR width results, obtaining a  $p$ -value of 0.34, which indicates that the differences between the FLC-based and fixed results were not statistically significant. The FLC was not successful in leading the evolutionary process to the

optimal solution in any of the 25 random retries, which means that the FLC's design was tightly coupled with initial population which it was exclusively optimized on. Therefore, its overall generalization capability is poor.

Finally, the implicit assumption about the correctness of the calculated object axis should be addressed. For the dead trunks, the SAC-based axis calculation method usually succeeds since the tree crown is missing and many stem hits are available. However, problematic cases may arise in the presence of dense ground vegetation and an asymmetric residual branch structure, leading to an erroneous axis. For living trees, this is even more of a concern. In this study, we did not investigate the axis quality and its influence on the classification accuracy, but the fact that both axis-based descriptors outperformed the axis-independent covariance features indicates some level of robustness to axis displacement.

## 6. CONCLUSIONS

This paper introduced a new kind of shape descriptor for classification tasks in sparse point clouds, the Free Shape Context (FSC), which is based on a sequence of cylindrical 3D Shape Contexts organized around a common axis. We have shown that owing to its increased number of degrees of freedom, the FSC can exploit characteristic regions of the target objects while disregarding cluttered or non-informative regions, yielding good discriminative power and simultaneously keeping the cardinality of the feature set at low levels. Specifically, the proposed FSC descriptor was able to outperform, in a statistically significant manner, a simple Bag-of-Features model based on covariance matrix eigenvalue features, and a standard 3D Shape Context descriptor. Our results show that the FSC can be a viable alternative to local features based on eigenvalues of the point neighborhood's covariance matrix for inference tasks in ALS point clouds. The higher number of degrees of freedom leads to an increased computational complexity, which prevents using exact or grid search approaches for optimizing the FSC parameters. We have demonstrated that a genetic algorithm (GA) is suitable for this task. Also, the successful application of a Fuzzy Logic Controller for adapting the Fuzzy Recombination width meta-parameter confirms previous results from the field of evolutionary computation showing that the meta-parameter selection may be partially automatized by combining GA with fuzzy logic. In the future, it would be interesting to confirm the FSC's discriminative capabilities on other kinds of remote sensing data. Also, the FSC's expressive power could benefit from introducing subdivisions in the angular dimension, although this would require a strategy for ensuring rotational invariance around the axis. Finally, to ease the computational burden of finding the optimal FSC parameters, instead of directly estimating the classification error induced by an FSC, an approximate quality assessment scheme could be used.

## REFERENCES

- Blomley, R., Weinmann, M., Leitloff, J. and Jutzi, B., 2014. Shape distribution features for point cloud analysis - a geometric histogram approach on multiple scales. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II-3, pp. 9–16.
- Frome, A., Huber, D., Kolluri, R., Bülow, T. and Malik, J., 2004. Recognizing Objects in Range Data Using Regional Point Descriptors. In: *Proceedings of the European Conference on Computer Vision (ECCV)*.
- Goldberg, D. E., 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. 1st edn, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- Herrera, F. and Lozano, M., 2001. Adaptive genetic operators based on coevolution with fuzzy behaviors. *IEEE Trans. Evolut. Comput.* 5(2), pp. 149–165.
- Johnson, A. and Hebert, M., 1999. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Trans. Pattern Anal.* 21(5), pp. 433–449.
- Klir, G. J. and Yuan, B., 1995. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall, Upper Saddle River, USA.
- Lee, M. A. and Takagi, H., 1993. Dynamic control of genetic algorithms using fuzzy logic techniques. In: *Proceedings of the 5th International Conference on Genetic Algorithms*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 76–83.
- Mallet, C., Bretar, F., Roux, M., Soergel, U. and Heipke, C., 2011. Relevance assessment of full-waveform lidar data for urban area classification. *ISPRS J. Photogramm. Remote Sens.* 66(6), pp. S71 – S84.
- Osada, R., Funkhouser, T., Chazelle, B. and Dobkin, D., 2002. Shape distributions. *ACM Trans. Graph.* 21(4), pp. 807–832.
- Polewski, P., Yao, W., Heurich, M., Krzystek, P. and Stilla, U., 2014. Detection of fallen trees in ALS point clouds by learning the Normalized Cut similarity function from simulated samples. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II-3, pp. 111–118.
- Polewski, P., Yao, W., Heurich, M., Krzystek, P. and Stilla, U., 2015. Detection of fallen trees in ALS point clouds using a Normalized Cut approach trained by simulation. *ISPRS J. Photogramm. Remote Sens.* 105, pp. 252 – 271.
- Reitberger, J., Schnörr, C., Krzystek, P. and Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS J. Photogramm. Remote Sens.* 64(6), pp. 561 – 574.
- Rusu, R., Marton, Z., Blodow, N. and Beetz, M., 2008. Learning informative point classes for the acquisition of object model maps. In: *10th International Conference on Control, Automation, Robotics and Vision*, pp. 643–650.
- Saupe, D. and Vranic, D., 2001. 3D model retrieval with spherical harmonics and moments. In: B. Radig and S. Florczyk (eds), *Pattern Recognition, Lecture Notes in Computer Science*, Vol. 2191, Springer Berlin Heidelberg, pp. 392–397.
- Siitonen, J., Martikainen, P., Punttila, P. and Rauh, J., 2000. Coarse woody debris and stand characteristics in mature managed and old-growth boreal mesic forests in southern Finland. *Forest Ecol. Manag.* 128(3), pp. 211 – 225.
- Stokland, J. N., Siitonen, J. and Jonsson, B. G., 2012. *Biodiversity in Dead Wood*. Cambridge University Press. Cambridge Books Online.
- Toldo, R., Castellani, U. and Fusiello, A., 2009. A Bag of Words Approach for 3D Object Categorization. In: A. Galalowicz and W. Philips (eds), *Computer Vision/Computer Graphics Collaboration Techniques, Lecture Notes in Computer Science*, Vol. 5496, Springer Berlin Heidelberg, pp. 116–127.
- Torr, P. and Zisserman, A., 2000. Mlesac: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding* 78(1), pp. 138 – 156.
- Velizhev, A., Shapovalov, R. and Schindler, K., 2012. Implicit shape models for object detection in 3D point clouds. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* I-3, pp. 179–184.
- Waldhauser, C., Hochreiter, R., Otepka, J., Pfeifer, N., Ghuffar, S., Korzeniowska, K. and Wagner, G., 2014. Automated Classification of Airborne Laser Scanning Point Clouds. In: S. Koziel, L. Leifsson and X.-S. Yang (eds), *Solving Computationally Expensive Engineering Problems*, Springer Proceedings in Mathematics & Statistics, Vol. 97, Springer, pp. 269–292.
- Wegrzyn, D. and Alexandre, L., 2013. A Genetic Algorithm-Evolved 3D Point Cloud Descriptor. In: J. Ruiz-Shulcloper and G. Sanniti di Baja (eds), *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Lecture Notes in Computer Science*, Vol. 8258, Springer, pp. 92–99.
- Weinmann, M., Jutzi, B. and Mallet, C., 2014. Semantic 3D scene interpretation: A framework combining optimal neighborhood size selection with relevant features. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II-3, pp. 181–188.