Unsupervised Classification of Raw Full-Waveform Airborne Lidar Data by Self Organizing Maps

Eleonora Maset(✉), Roberto Carniel, and Fabio Crosilla

Dipartimento di Ingegneria Civile e Architettura (DICA),
Università degli Studi di Udine, Via delle Scienze, 206, 33100 Udine, Italy
eleonora.maset.1@gmail.com,
{roberto.carniel,fabio.crosilla}@uniud.it

Abstract. The paper proposes a procedure based on Kohonen's Self Organizing Maps (SOMs) to perform the unsupervised classification of raw full-waveform airborne LIDAR (Light Detection and Ranging) data, without the need of extracting features from them, that is without any preprocessing. The proposed algorithm allows the classification of points into three classes ("grass", "trees" and "road") in two subsequent stages. During the first one, all the raw data are given as input to a SOM and points belonging to the category "trees" are extracted on the basis of the number of peaks that characterize the waveforms. In the second stage, data not previously classified as "trees" are used to create a new SOM that, together with a hierarchical clustering algorithm, allows to distinguish between the classes "road" and "grass". Experiments carried out show that raw full-waveform LIDAR data were classified with an overall accuracy of 93.9%, 92.5% and 92.9%, respectively.

Keywords: Airborne LIDAR data · Full-waveform · Unsupervised classification · Self organizing maps

1 Introduction

The aerial laser scanner is an instrument used to survey the ground level morphology and the size and shape of natural and man-made objects, that exploits the time of flight of a reflected very short laser pulse (4 ns for the sensor used in this study), usually of wavelength between 0.8 and 1.55 μm. During its path, the laser ray can be reflected by more than one surface, placed at different heights. The earliest laser scanners could register just one return echo for each emitted one, later instruments allowed the use of 5/6 reflections for each emitted pulse. Since 2004, a new category of instruments are available on the market, the so called full-waveform airborne laser scanners, that are finally able to record the entire waveform of the reflected signal. The shape and size of the received waveform is related to the reflectance characteristics of the surface. Recording the complete waveform of the incoming pulse means that it is possible to obtain more information about geometrical and

V. Murino and E. Puppo (Eds.): ICIAP 2015, Part I, LNCS 9279, pp. 62–72, 2015.
DOI: 10.1007/978-3-319-23231-7_6

physical characteristics of the target hit by the laser ray, that can be useful for the classification of the 3D sampled points.

Over the last years, several classification methods have been proposed in the literature using full-waveform data and the features extracted from them [6]. Among these, we mention decision trees [5] or simple thresholds both set up manually [18] and automatically [1]. These methods exploit features extracted from the waveforms, such as amplitude, pulse width and number of pulses, and they have the advantage of not requiring assumptions regarding the distribution of input data. Other methods are based on statistical learning classifiers like Support Vector Machines (SVM, [11]), which belong to non-parametric methods and perform non-linear classification. This algorithm is well suited for high dimensional problems with limited training set. Höfle et al. [8] use instead an artificial neural network classifier consisting of a single hidden layer of neurons and trained by back propagation.

The present paper proposes the application of Kohonen's Self Organizing Maps (SOMs), a kind of neural networks introduced in the 80s of the last century by Kohonen [10] as a method for clustering and visualization of high dimension datasets. The basic principle of SOMs is that a higher level knowledge organization can be carried out by learning using algorithms that perform a self-organization in a spatial order. A SOM consists of a bi-dimensional grid of a predetermined number of equally spaced nodes, that can vary from few units to few thousand, according to the nature of the data set. Each node, also known with the term "neuron", is represented by a vector (the so called "code vector") with the same size of the vectors that constitute the data of the sample to be analyzed and is connected with the other neurons to form the network. During the training process the SOM evolves by changing the neurons' vector values as the data vectors sequentially enter into the process [2].

In the past, Self Organizing Maps found application in multispectral imagery classification and, more recently, in the Earth sciences ([9], [4]). However, methods based on the SOM concept have not been yet extensively exploited for the classification of LIDAR point clouds. Salah et al. [16] apply Self Organizing Maps for building detection from LIDAR data and multispectral aerial images. Zaletnyk et al. [19] exploit the SOM algorithm to investigate in particular the correlation between the shape of the LIDAR waveforms, using various statistical parameters (amplitude, standard deviation, skewness and kurtosis), and the properties of the reflecting surface. Toth et al. [17] and Molnar et al. [12] apply a 2x2 neural grid SOM to classify four different features types (trees, grass, roof and pavement).

In all these works, some features have been preliminary extracted from the original full-waveform data and submitted to the SOM procedure. On the contrary, in this paper the raw full wave data vectors have been directly analyzed and classified. Although Molnar et al. [12] stated in their paper the impracticability to directly use the original waveforms as input for a classification procedure, this paper shows a successful result thanks to the implementation of a SOM with a number of neurons much greater than the number of the required classes.

In [12] a network with a number of nodes equal to the number of classes is used. Tests carried out in this work have shown that presenting the raw waveform data to a SOM of small dimensions leads to a high error in the classification. Large networks, instead, allow to use the original signal for the classification procedure, despite the waveforms can be different for each reflection.

On the other hand, considering a number of nodes much greater than the number of classes, implies the necessity to successively apply a proper algorithm for clustering the nodes.

2 Methodology

As mentioned before, the classification method proposed in this paper is based on the SOM algorithm. SOMs are a particular kind of unsupervised artificial neural networks that promote self-organization of data vectors in a spatial order by suitable learning algorithms. The first stage is the so called global initialization, in which a map of predefined size, not organized at all, is considered. At the beginning, a first set of random values - usually extracted from the experimental data - is assigned to each node vector. Successively, three iterative processes are repeated for all the data vectors until a global convergence is reached ([2], [3]).

The first one is the *competitive process*, that aimed at finding the neuron whose code vector is nearest to the input vector in the n-dimensional space, where n is the dimensionality of input data. The winning neuron (Best Matching Unit, BMU) determines the spatial location of a topological neighbourhood of excited neurons on the map. To carry out the *competitive process*, it is necessary to introduce a discriminant function able to measure the similarity between the input vector $x_i(t)$ at the t-th iteration and each neuron of the map. The winning neuron at the t-th iteration is the one having the greatest similarity with the input vector $x_i(t)$. It has been demonstrated [7] that various similarity criteria of two functions, $f(y)$ and $g(y)$, including the sum of squared differences and the correlation coefficient, are related to the cross-correlation function $R_{f,g}(\tau)$ at $\tau = 0$:

$$R_{f,g}(\tau) = \int f(y)\, g(y+\tau)\, dy \tag{1}$$

Thus, they cannot provide any information about patterns that are shifted relatively to each other. In this work, a generalized expression for similarity [7], $S_{f,g}(\tau)$, is used, which is based on a weighted cross-correlation function, a weighting function $z(\tau)$ normalized with the product of the two weighted autocorrelation functions, that is:

$$S_{f,g}(\tau) = \frac{\int z(\tau)\, R_{f,g}(\tau)\, d\tau}{\sqrt{\int z(\tau)\, R_{f,f}(\tau)\, d\tau \int z(\tau)\, R_{g,g}(\tau)\, d\tau}} \tag{2}$$

where $z(\tau)$ is a triangular weighting function of width defined as $z(\tau) = 1 - |\tau|/h$ if $|\tau| < h$ and $z(\tau) = 0$ if $|\tau| \geq h$. The BMU is then the neuron $w_c(t)$ that

maximizes the value of the function $S_{f,g}(\tau)$. Once the *competitive process* has selected the BMU, the next step is the so called *cooperative process*. During this stage, a neighbouring function, $h_{c(x),j}(t)$, determines how strongly the various neurons are connected to the winner at a certain iteration t. A typical choice of $h_{c(x),j}(t)$ is the Gaussian function [13].

The last process is the *adaptive* one. During the adaptive process an adjustment of the neuron vector values is carried out in order to minimize the distance of each data input from the corresponding neuron of the map and to slowly allow the map to be partitioned into relevant clusters at the end of the process. Usually the model applied is the following [13]:

$$w_j(t+1) = w_j(t) + \alpha(t)\, h_{c(x),j}(t)\, [x_i(t) - w_j(t)] \tag{3}$$

where $w_j(t+1)$ is the j-th updated neuron vector at iteration t+1, $\alpha(t) = \alpha_0 \cdot \exp(-t/T)$ is the learning-rate factor parameter and τ is a new time constant.

The three steps process just presented is applied at each iteration to the entire data set and the entire learning process is stopped when no more substantial changes to the code vectors are observed.

As pointed out before, the purpose of the proposed method here is to classify LIDAR data into three categories: "trees", "grass" and "road". Each input data is the entire waveform of the reflected signal, i.e. a vector whose components are the amplitude values registered and stored by the instrument at a certain sampling interval. No preprocessing procedure is applied to the data. More details about the structure of the data to be classified will be given in Sec. 3.

First of all, the extraction of the waveforms reflected from trees is performed, exploiting the fact that waveforms belonging to this class are generally characterized by the presence of two or more echoes, unlike those of the road and grass, which include only one echo. Nevertheless, as highlighted in [11], waveforms recorded by the receiver are affected by the "ringing effect", i.e. after the peak corresponding to a reflecting surface, a small secondary maximum due to the effects of the hardware waveform processing chain can be seen. So it is not possible to distinguish between trees and the other two classes only on the basis of the number of peaks of each raw signal, without first dealing with the "ringing" problem. As suggested in [15], the "ringing effect" can be recognized and removed if its amplitude is smaller than a certain fraction of the amplitude of the first peak and if the second peak is closer than a certain distance to the first one. The experiments carried out during this work have shown that, due to the variability of the waveforms reflected, it is difficult to identify a single threshold value to be applied to the raw data, which allows to eliminate only false peaks. For this reason, the SOM algorithm, described above, is used a first time, performing the training of the map with all the data set. A threshold value is then applied to the code vectors of the map, rather than directly to the raw data, in order to remove the "ringing effect". Since the code vectors of the SOM are a sort of "abstraction" of the data, the great variability of the waveforms is smoothed, and a more reliable threshold value can be identified.

At the end of this stage, the data associated with the code vectors that still have more than one peak after the application of the threshold, are labeled as trees. The remaining, i.e. those characterized by a single peak, are not classified. Tests have shown that using the SOM algorithm for this purpose can significantly reduce the error in the classification of the points belonging to the class "road" and increase the overall accuracy of the classification. In fact, if the "ringing effect" is removed applying the threshold value directly to the raw waveforms, even taking into account the fact that the second false peak occurs between 10 and 12 ns after the first one, the overall accuracy is reduced by 10%.

Once the points belonging to the category "trees" have been separated, it is possible to classify waveforms reflected by grass and roads. A new network of the same size of the previous one is created and only the waveforms that have not been labeled in the previous stage are used as input. As suggested in [16], the chosen size of the map is high (e.g. 15x15). In fact, small networks result in some unrepresented classes, while large networks lead to an improvement in the overall classification accuracy. After each piece of data has been associated with one neuron on the new map, a hierarchical clustering procedure, performed by the *agnes* function (R environment [14], *cluster* package), is applied. Each code vector is initially considered a small one-node cluster [3], that is then progressively merged with similar clusters, until only a unique cluster, containing all the code vectors, is formed. At each stage the two most similar clusters are combined. The result is a graph called "dendrogram", where the clustering process described above can be seen proceeding from bottom to top (Fig. 1). The clustering procedure can be interrupted at any vertical level by establishing an appropriate threshold. In this work, the threshold is chosen in order to determine two clusters on the map: one is associated with the class "grass" and the other with the class "road". Finally, the two obtained clusters need to be "labelled"; an automatic way to perform the interpretation of the clusters can be based on the maximum value of the average vector that represents the cluster. Surfaces covered with grass have higher reflectivity than the road; for this reason, the cluster represented by the average vector that has the highest value of amplitude is labeled as grass. It is important to underline that the information related to the amplitude is exploited only to label the two classes obtained at the end of the clustering procedure; the entire process that leads to the separation between waveforms belonging to the category "road" and waveforms belonging to the category "grass" is totally independent from this feature. Figure 2 shows the identification of the two clusters on the map.

3 Experiments and Results

The proposed method has been validated for three test areas in Horn, Austria. The data set was acquired by the company RIEGL Laser Measurement System GmbH with the sensor RIEGL LMS-Q680i during a flight at the altitude of 800 m. The emitted laser pulse has a length of 4 ns and each return waveform is composed of 60 samples, with 1 ns sampling interval. The data format used as input in the proposed procedure is amplitude vs. time series samples.

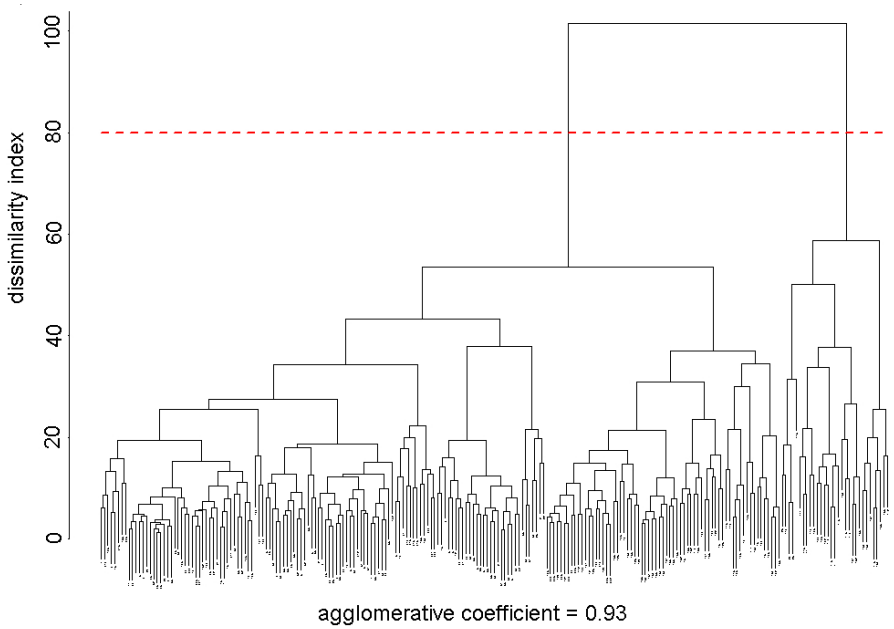

Fig. 1. Dendrogram summarizing the clustering of nodes of the SOM performed during the second stage of the classification algorithm.

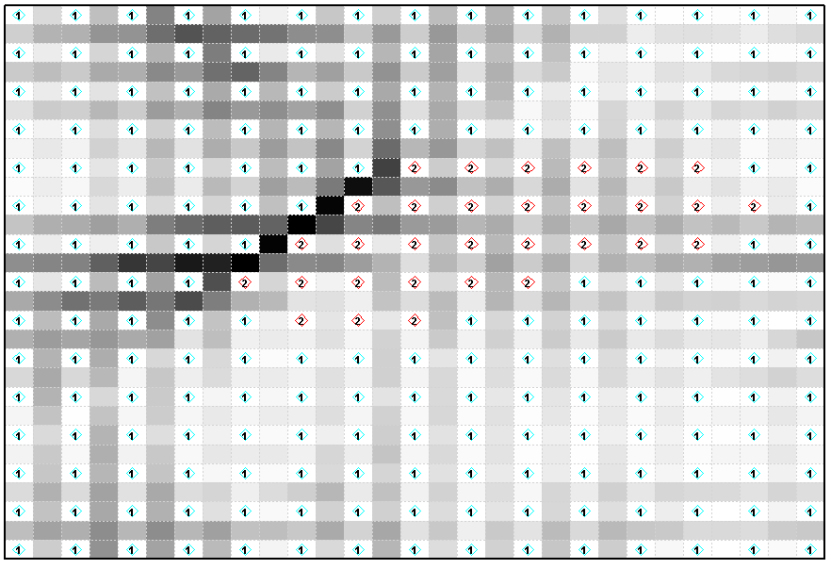

Fig. 2. SOM showing the nodes assigned to each of the clusters. Cells which do not contain a number use a gray scale to show the similarity of the nearby neurons code vectors: light gray indicates a strong similarity and vice versa.

The classification procedure preliminarily requires the operator to choose the size of the Self Organizing Map. Maps of different dimensions were tested, from 2x2 to 25x25. Results showed that maps of size 15x15 are a good trade-off between classification accuracy and execution time. In fact, using larger maps does not lead to a significant improvement of the results, while it increases considerably the computational costs. For these reasons, the dimension of the maps chosen for all the experiments is 15x15. In order to increase the precision of the classification method, the training process of the SOM networks is repeated for each data set.

The first study area is represented in Fig. 3. 15000 waveforms were extracted from the complete data set and used for the validation of the algorithm, and among these 2926 (19.5%) were randomly picked and manually classified for evaluation purposes. The waveforms recorded by the instrument were given as input to a first SOM network of size 15x15. At the end of the training process, the code vectors associated with each neuron were normalized to the maximum value, and the threshold 0.075 was applied to them. In this way, waveforms with more than one peak were classified as trees. At the end of this stage, 12850 out of 15000 waveforms had not yet been classified. These were then used as input data for a second Self Organizing Map of size 15x15. As described in Sec. 2, once the data have been projected on the new map, a hierarchical clustering algorithm was applied and the dendrogram was produced. Finally, the two clusters obtained through the dendrogram were interpreted on the basis of the maximum amplitude value of the average code vector related to each cluster and the classes “road” and “grass” were identified.

Fig. 3. First study area.

In order to verify the accuracy of the classification method, the results obtained with the proposed algorithm were compared with the manually performed classification. As shown in Table 1, the overall classification accuracy for this study area is 93.9%. The error matrix highlights that the proposed method allows to classify the data belonging to the “road” and “grass” categories with very high accuracy, respectively 97.6% and 98.6%. However, the algorithm shows more difficulty in discriminating waveforms relating to trees. This may be due to

the fact that some waveforms reflected from the trees can present only one peak, or secondary peaks have low amplitude and they are not distinguished from the noise. In these cases, the method is not able to correctly establish the class, and these data are erroneously assigned to the categories "road" or "grass".

Table 1. Error matrix for the first study area, computed with 2926 manually classified points.

	Grass	Trees	Road	Total	User's accuracy
Grass	1451	86	7	1544	94.0%
Trees	13	595	3	611	97.4%
Road	22	48	701	771	90.9%
Total	1486	729	711	2926	
Producer's accuracy	97.6%	81.6%	98.6%		93.9%

Figure 4 represents the second test area. In this case 10000 waveforms were extracted from the complete data set and among these 1086 (10.9%) were randomly picked and manually classified. A new SOM of size 15x15 was trained, using this second data set. At the end of the first stage, carried out as described in the previous experiment, 8870 out of 10000 waveforms had not yet been classified. These were then used as input data for the second stage of the algorithm, which led to the results reported in Table 2.

Fig. 4. Second study area.

In this case, the overall accuracy is equal to 92.5%, while the producers accuracy related to each category is 93.9% for the class "grass", 91.9% for the class trees and 89.4% for the class "road". Unlike the previous example, the results of this experiment show that also the points belonging to the class trees are identified with high accuracy, more than 90%.

The third test area is shown in Fig. 5. 12000 waveforms were used in this experiment and among these 1588 (13.2%) were randomly picked and manually classified. The classification procedure was carried out as described above, creating a first SOM network of size 15x15 that led to the identification of 1688 points belonging to the class "trees", and a second map of dimension 15x15 for

Table 2. Error matrix for the second study area, computed with 1086 manually classified points.

	Grass	Trees	Road	Total	User's accuracy
Grass	559	15	24	598	93.5%
Trees	2	226	2	230	98.3%
Road	34	5	220	259	84.9%
Total	595	246	246	1086	
Producer's accuracy	93.9%	91.9%	89.4%		92.5%

the distinction between the class "grass" and the class "road". Table 3 shows the confusion matrix, computed from the comparison between the results obtained with the proposed algorithm and the manually performed classification.

Fig. 5. Third study area.

Table 3. Error matrix for the third study area, computed with 1588 manually classified points.

	Grass	Trees	Road	Total	User's accuracy
Grass	801	3	42	846	94.7%
Trees	27	344	8	379	90.8%
Road	27	5	331	363	91.2%
Total	855	352	381	1588	
Producer's accuracy	93.7%	97.7%	86.9%		92.9%

The overall accuracy achieved is 92.9%, which is a value similar to the previous cases. The accuracy with which points belonging to the classes grass and road were identified, 93.7% and 86.9% respectively, is similar to the second experiment, while the accuracy related to the class trees is very high, equal to 97.7%.

4 Discussion and Conclusion

The goal of this work was to implement a method for the classification of data acquired with a full-waveform airborne laser scanner, based on Kohonen's Self

Organizing Map. The algorithm proposed in this paper uses the entire waveform recorded by the instrument, without any pre-processing, in order to distinguish the data into three categories: "trees", "grass" and "road". Although Molnar et al. [12] stated in their paper the impracticability to directly use the original raw data as input for a classification procedure, since the waveforms can be really different for each reflection, this paper shows a successful result thanks to the implementation of a SOM with a number of neurons much greater than the number of the required classes. In fact, large networks manage to overcome the problem of the variability of the waveforms and the classification is then made possible by the subsequent application of the hierarchical clustering algorithm to the nodes of the map. One may argue about the necessity to pass through a SOM procedure and not directly to the clustering of the original data. In this regard, applying a clustering algorithm to the nodes of the grid and not directly to the original data vectors, is surely much more reliable since neuron's values are in some way "averaged" values.

The experiments carried out to validate this method showed that, on average, the overall accuracy provided by the proposed algorithm is 93.1%, and the three categories "grass", "trees" and "road" can be distinguished with an average producer's accuracy of 95.1%, 90.4% and 91.6%, respectively, and an average user's accuracy of 94.1%, 95.5% and 89%, respectively. It is important to emphasize that the method is almost fully automatic, and the user needs only to choose the threshold for the elimination of the "ringing effect". Unlike other methods ([5], [17], [19]), it is not required to extract features from the waveforms of the data set.

Moreover, since the SOM is a tool that realizes an unsupervised classification, any manual pre-classification of a subset of data is not necessary, as instead required by methods that perform a supervised classification.

If the user is interested only in a particular category of points, this method can be very useful because it allows to extract preliminarily some classes of data and then to process only the waveforms of interest and not the entire data set.

At the time the method is applicable to the classification of points belonging to areas without buildings. Further research is being carried out in order to extend the applicability of the algorithm.

References

1. Alexander, C., Tansey, K., Kaduk, J., Holland, D., Tate, N.J.: Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing **65**(5), 423–432 (2010)
2. Carniel, R., Barbui, L., Jolly, A.D.: Detecting dynamical regimes by Self-Organizing Map (SOM) analysis: an example from the March 2006 phreatic eruption at Raoul Island, New Zealand Kermadic Arc. Boll. Geof. Teor. Appl. **54**(1), 39–52 (2013)
3. Carniel, R., Jolly, A.D., Barbui, L.: Analysis of phreatic events at Ruapehu volcano, New Zealand using a new SOM approach. J. Volc. Geoth. Res. **254**, 69–79 (2013)

4. Carniel, R.: Characterization of volcanic regimes and identification of significant transitions using geophysical data: a review. Bull. Volc. (2014) (in press)
5. Ducic, V., Hollaus, M., Ullrich, A., Wagner, W., Melzer, T.: 3D vegetation mapping and classification using full-waveform laser scanning. In: Workshop on 3D Remote Sensing in Forestry, Vienna, Austria, February 14–15, pp. 222–228 (2006)
6. Fieber, K.D., Davenport, I.J., Ferryman, J.M., Gurney, R.J., Walker, J.P., Hacker, J.M.: Analysis of full-waveform LiDAR data for classification of an orange orchard scene. ISPRS Journal of Photogrammetry and Remote Sensing **82**, 63–82 (2013)
7. de Gelder, R., Wehrens, R., Hageman, J.A.: A generalized expression for the similarity spectra: application to powder diffraction pattern classification. Journal of Computational Chemistry **22**(3), 273–289 (2001)
8. Höfle, B., Hollaus, M., Hagenauer, J.: Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data. ISPRS Journal of Photogrammetry and Remote Sensing **67**, 134–147 (2012)
9. Klose, C.D.: Self organizing maps for geoscientific data analysis: geological interpretation of multidimensional geophysical data. Computational Geosciences **10**(3), 265–277 (2006)
10. Kohonen, T.: Self-organised formation of topologically correct feature map. Biological Cybernetics **43**, 56–69 (1982)
11. Mallet, C., Bretar, F., Roux, M., Soergel, U., Heipke, C.: Relevance assessment of full-waveform lidar data for urban area classification. ISPRS Journal of Photogrammetry and Remote Sensing **66**(6), S71–S84 (2011)
12. Molnar, B., Laky, S., Toth, C.: Using full waveform data in urban areas. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **38**(Part3/W22), 203–208 (2011)
13. Ortiz, A., Górriz, J.M., Ramírez, J., Salas-González, D., Llamas-Elvira, J.M.: Two fully-unsupervised methods for MR brain image segmentation using SOM-based strategies. Applied Soft Computing **13**(5), 2668–2682 (2013)
14. R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation, Austria 3–900051-07-0 (2014). www.R-project.org
15. Reitberger, J., Krzystek, P., Heurich, M.: Full waveform analysis of small footprint airborne laser scanning data in the bavarian forest national park for three species classification. In: Workshop on 3D Remote Sensing in Forestry, Vienna, Austria, February 14–15, pp. 229–238 (2006)
16. Salah, M., Trinder, J., Shaker, A.: Evaluation of the Self-Organizing Map classifier for building detection from Lidar data and multispectral aerial images. Journal of Spatial Science **54**(2), 15–34 (2009)
17. Toth, C., Laky, S., Zaletnyik, P., Grejner-Brzezinska, D.: Compressing and classifying LIDAR waveform data. International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences XXXVIII (1), Paper 130 (2010)
18. Wagner, W., Hollaus, M., Briese, C., Ducic, V.: 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. International Journal of Remote Sensing **29**(5), 1433–1452 (2008)
19. Zaletnyik, P., Laky, S., Toth, C.: LiDAR waveform classification using self-organizing map. In: ASPRS, San Diego, CA, April 26–30 (2010)