Wavelet based buildings segmentation in airborne laser scanning data set

Wolfgang Keller[1], Andrzej Borkowski[2]

[1]Institute of Geodesy, Stuttgart University
24D Geschwister-Scholl-St., D-70174 Stuttgart, Germany
e-mail: wolfgang.keller@gis.uni-stuttgart.de
[2]Institute of Geodesy and Geoinformatics
Wroclaw University of Environmental and Life Sciences
Grunwaldzka 53, PL-50-357 Wroclaw, Poland
e-mail: andrzej.borkowski@up.wroc.pl

Abstract: In the recent years three-dimensional buildings modelling based on an raw airborne laser scanning point clouds, became an important issue. A significant step towards 3D modelling is buildings segmentation in laser scanning data. For this purpose an algorithm, based on the multi-resolution analysis in wavelet domain, is proposed in the paper. The proposed method concentrates only on buildings, which have to be segmented. All other objects and terrain surface have to be removed. The algorithm works on gridded data. The wavelet-based segmentation proceeds in the following main steps: wavelet decomposition up to appropriately chosen level, thresholding on the chosen and adjacent levels, removal of all coefficients in the so-called influence pyramid and wavelet reconstruction. If buildings on several scaling spaces have to be segmented, the procedure should be applied iteratively. The wavelet approach makes the procedure very fast. However, the limitation of the proposed procedure is its scale-based distinction between objects to be segmented and the rest.

Keywords: wavelet, airborne laser scanning, segmentation, multi-resolution analysis

1. Introduction

In the recent years three-dimensional modelling of space elements, derived from a raw airborne laser scanning point clouds, becomes an important issue. One of its components is to model buildings based on airborne laser scanning data. 3D models of buildings are increasingly used in many areas, such as promotion of the city, urban planning, shadow modelling, spread of noise in the city etc. A building modelling with its complex topology is still a challenging research topic. In research particular emphasis is placed on both, the efficiency and process automation.

Many studies are dealing with buildings modelling from LiDAR data. A good overview of issues in this area can be found in the book by Vosselman and Maas (2010). Review of methods, algorithms and solutions applicable on the various stages of modelling is given in (Dorninger and Pfeifer, 2008; Haala and Kada, 2010).

Building extraction from airborne LiDAR data is a complex process, which proceeds in several basic steps. In the first step buildings have to be detected in point clouds. Building detection is a classification task which separates them from ground and other objects (Vosselman and Maas, 2010). For this reason different filter techniques can be used. Some of them are linear prediction (Kraus, 2000; Kraus and Pfeifer, 2001; Briese et al., 2002), adaptive TIN models (Axelsson, 2000), mathematical morphology (slope adaptive filtering) (Sithole, 2001; Filin and Pfeifer, 2006), data clustering analysis (Roggero, 2001), or surface energy minimization (Elmqvist, 2002; Borkowski, 2004; Borkowski and Jóźków, 2008). Evaluation and comparison of some filtration methods can be found in the study (Sithole and Vosselman, 2004). Unfortunately, all filtration methods have their limitations and often a manual inspection is needed.

After the points belonging to the buildings have been identified, they have to be extracted from the point cloud and have to be segmented. Segmentation is a process in which the points belonging to the individual buildings are identified. In the following, 2D outlines of building footprints can be derived. In this stage the region growing algorithm and RANSAC algorithm are frequently used (Neidhart and Sester, 2008). Sampath and Shan (2007) use a slope based 1D bi-directional filter for buildings points separation and then region growing approach for the segmentation. Also Forlani et al. (2006) use region growing technique, but they perform a complete segmentation of raw LiDAR data in a multi stage framework at different levels of resolution.

The modelling process and the used algorithms depend also on the type of available data. Scanning data are combined frequently with additional information sources. Oude Elbering and Vosselman (2009) use laser point clouds and topographic maps as an initial information in their workflow from point clouds to 3D models. Awrangjeb et al. (2010) presented an automatic building detection technique using LiDAR data and multispectral imagery. An automatic method for extracting buildings based on the snake model with improved external energy using LiDAR data and aerial images in urban areas is proposed by Kabolizade et al. (2010).

For classification of airborne laser scanning data in urban environment several methods were applied. Frequently no filtering is applied before the classification. For example Melzer (2007) used the unsupervised Mean-Shift classification method, while Mallet et al. (2008) used supervised classification algorithm such as Support Vector Machines. A cascade of binary classifiers based on 3D shape analysis was proposed by Carlberg et al. (2009). Airborne laser scanner data classification can also be performed using geometrical and textural features (Matikainen et al., 2003). Chehata et al. (2009) proposed a method for urban scenes classification using Random Forest algorithm.

The existing algorithms and methods do not work perfectly and they have limitations. Some of them are listed and discussed in (Vosselman and Mass, 2010). Additional data sources, e.g. multispectral data are useful for buildings segmentation. However such a data is not always available. Modern airborne laser scanning systems provide multiple returns and higher point densities. This implicates higher numerical effort, which can be a challenge for some algorithms.

Wavelets are commonly known for their excellent numerical properties. Wavelet algorithms can process very fast huge data sets. Also thresholding is frequently used in several segmentations algorithms. In this context the wavelets analysis seems to be attractive for buildings segmentation in airborne laser scanning data.

Wavelet analysis has been used for airborne laser scanning data filtering (Vu et al., 2002; Wei and Bartels, 2006). Xu et al. (2007) used the wavelet analysis for the automatic extraction of bare earth from point clouds. Wavelets have also been used for compression of full-waveform airborne laser scanning data (Laky et al., 2010).

In this paper, a wavelet segmentation method based on multi-resolution analysis of scanning data sets is proposed. The method concentrates only on buildings, which have to be segmented. All other objects and ground have to be removed. The algorithm works on gridded data and does not need any other pre-processing steps.

2. Brief review of wavelet theory

Before describing the wavelet based algorithm of image segmentation, a brief outline of the necessary parts of wavelet theory will be given. This outline will be only one-dimensional. The generalization to the *n*-dimensional case will be given in section 3.

A function ψ is called a wavelet, if its Fourier transform $\hat{\psi}$ fulfils

$$0 < c_\psi := \int_{-\infty}^{\infty} \frac{\left|\hat{\psi}(\omega)\right|^2}{|\omega|} d\omega < \infty$$

From this definition can be concluded, that a wavelet is an oscillating function, which decays to zero for $x \rightarrow \pm\infty$. Frequently used wavelets are the Haar wavelet or the Daubechies wavelet (Fig. 1).
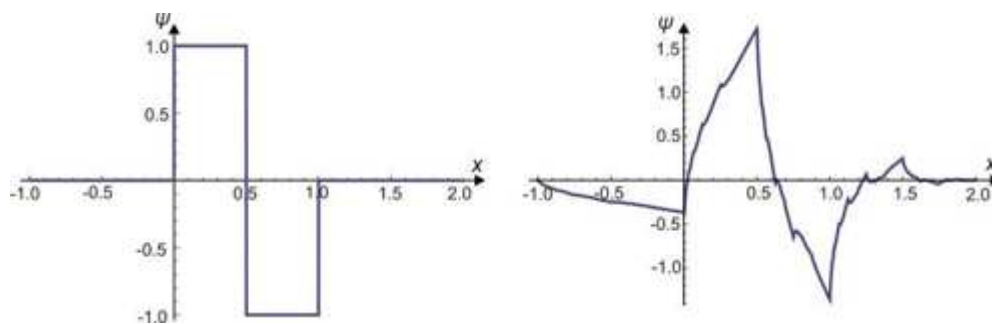

Fig. 1. Haar wavelet (left) and Daubechies wavelet (right)

The continuous wavelet transform of a real function f is defined by

$$W\{f\}(a,b) := \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(x)\psi\left(\frac{x-b}{a}\right)dx$$

This means the wavelet spectrum $W\{f\}(a,b)$ measures the degree of similarity of the function f in the vicinity of b with a scaled version $\frac{\psi(x/a)}{\sqrt{a}}$ of the wavelet ψ. Due to the finite support of the wavelet, the wavelet spectrum does not only inform about the scale of features contained in the function f, but also about the place b, where this feature appears in the signal. As an example we compare the wavelet and the Fourier spectrum (Fig. 3) of the following function (Fig. 2).
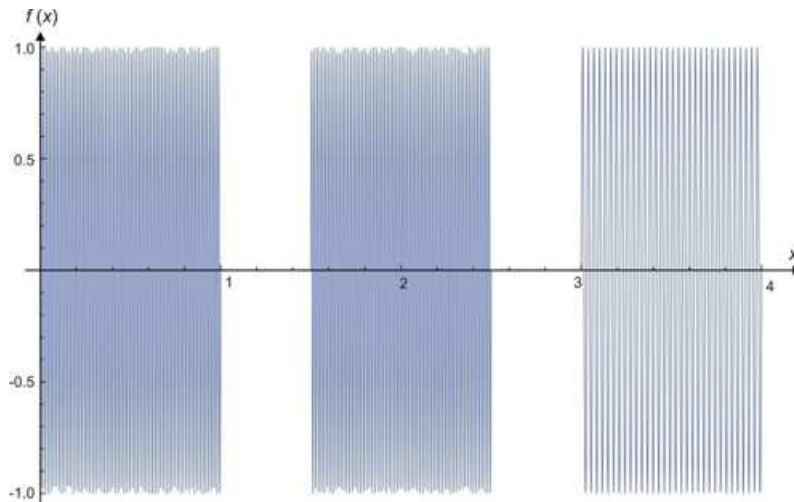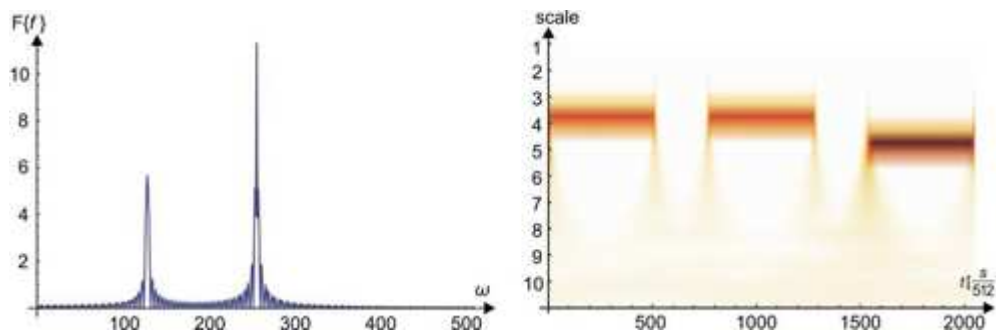

Fig. 2. Signal f with transient patterns

Fig. 3. Fourier spectrum (left) and wavelet spectrum (right) of transient signal f

While the Fourier spectrum informs only about the two wave-lengths contained in the signal, the wavelet spectrum additionally shows at which epochs which wave-lengths appear in the signal.

The continuous wavelet spectrum is well suited for signal interpretation, but expensive to compute and therefore less well suited for signal processing. The break-through of wavelet signal analysis is connected with the discrete wavelet transform, where the wavelet spectrum is computed only on a discrete $\{a, b\}$ grid. This discrete wavelet transform is closely related to conjugate-quadrature filtering. Conjugate quadrature filters try to decompose a signal f into a low-frequency part Hf and a high-frequency part Gf, in such a way that both parts are uncorrelated

$$0 = \int_{-\infty}^{\infty} Hf \cdot Gf dx$$

Of course this is not possible for every wavelet, but only for a certain subclass. Wavelets of this subclass possess a so-called scaling function φ. The scaling function allows to express both, the dilated version of the wavelet as well as the dilated version of itself as linear combination of shifted scaling functions.

$$\varphi(x) = \sqrt{2} \sum_{k=0}^{l} h_k \varphi(2x - k), \quad \psi(x) = \sqrt{2} \sum_{k-0}^{l} g_k \varphi(2x - k)$$

The coefficients $h_k g_k$ of these linear combinations form the weights of the low-pass filter H and the high-pass filter G.

$$(Hf)_k = \sum_{i=0}^{l} h_i f(x_{i+2k}), \quad (Gf)_k = \sum_{i=0}^{l} g_i f(x_{i+2k})$$

The filters H,G decompose the signal f in a low- and a high frequency part. For the low- frequency part the procedure can be repeated.

$$\left(H^2 f\right)_k = \sum_{i=0}^{l} h_i (Hf)_{i+2k}, \quad (HGf)_k = \sum_{i=0}^{l} g_i (Hf)_{i+2k}$$

In the end the signal has been decomposed into uncorrelated frequency bands (see Fig. 4).

If the original signal is a member of a certain set V_0, say the set of all signals not having details smaller than 1, the smoothed signal Hf is an element of V_1, the set of all signals not having details smaller than 2. The difference signal Gf then belongs to the space W_1, the set of all signals containing details only in the range between 1 and 2. In this way, the discrete wavelet transform is related to a so called multi resolution analysis of the space $L_2(R)$

$$0 \subset ... \subset V_0 \subset V_{-1} \subset ... \subset L_2(\mathbb{R})$$

where the scaling spaces V_i are spanned by the shifted and scaled versions of the scaling function φ and the difference spaces W_i are spanned by the shifted and scaled versions of the wavelet ψ.
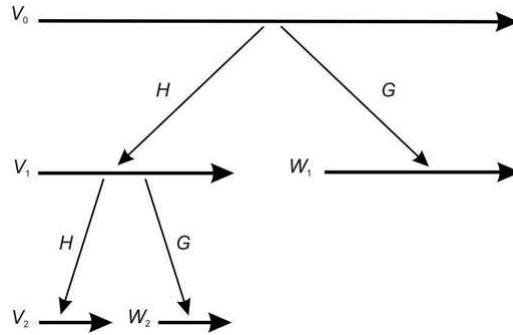

Fig. 4. Multi resolution analysis of the given signal

This multi-resolution analysis is the point of departure for the wavelet segmentation. Since the object to be extracted has an a-priori known range of size it falls into a certain difference space W_j. Therefore the only thing one has to do is to delete all difference spaces $W_i \neq W_j$ and perform the inverse discrete wavelet transform for the modified wavelet spectrum.

This brief and therefore rather imprecise introduction into discrete wavelet analysis will be generalized for n dimensions in section 4.

3. Multi-resolution analysis of $L_2(\mathbb{R}^n)$

Let

$$0 \subset ... \subset V_0 \subset V_{-1} \subset ... \subset L_2(\mathbb{R}) \tag{1}$$

be a multi-scale analysis of $L_2(\mathbb{R})$. The corresponding orthogonal scaling function and the corresponding orthogonal wavelet are denoted by φ and ψ, correspondingly. Their scaling equations are

$$\varphi(x) = \sqrt{2} \sum_{l \in \mathbb{Z}} h_l \varphi(2x - l)$$

and

$$\phi(x) = \sqrt{2} \sum_{l \in \mathbb{Z}} g_l \varphi(2x - l)$$

Based on the multi-scale analysis of $L_2(\mathbb{R})$ we first construct a multi-scale analysis of $L_2(\mathbb{R}^n)$.

Similarly to the one-dimensional case, a multi-resolution analysis of $L_2(\mathbb{R}^n)$ is a nested sequence of subspaces

$$0 \subset ... \subset V_0 \subset V_{-1} \subset ... \subset L_2(\mathbb{R}^n) \tag{2}$$

with

$$\overline{\cup_{m \in \mathbb{Z}} V_m} = L_2 \left(\mathbb{R}^n \right), \quad \cap_{m \in \mathbb{Z}} V_m = \{0\} \tag{3}$$

and

$$f \left(\bullet \right) \in V_m \Leftrightarrow f \left(2 \mathbf{I} \bullet \right) \in V_{m-1} \tag{4}$$

The base space V_0 is again spanned by the integer translates of a scaling function φ

$$\mathrm{V}_0 = \mathrm{span} \left\{ \varphi \left(\bullet - \mathbf{k} \right) \big| \mathbf{k} \in \mathrm{Z}^{\mathrm{n}} \right\} \tag{5}$$

As it is shown in Louis et al. (1998), and Keller (2004), there exist $2^n - 1$ wavelets $\psi^{(1)}, \ldots, \psi^{(2^n-1)}$ which generate the orthogonal complements $W_0^{(j)}$, $j = 1, ..., 2^{n-1}$ of V_0 in V_{-1}

$$V_{-1} = V_0 \otimes \left(\otimes_{j=1}^{2^{n-1}} W_0^{(j)} \right) \tag{6}$$

The scaling function φ and the wavelets $\psi^{(j)}$ of the n-dimensional multi-resolution can be constructed from the known scaling function and wavelet of the one-dimensional multi-resolution analysis in the following way

$$\varphi \left(\mathbf{x} \right) = \varphi \left(x_1 \right) \cdot \varphi \left(x_2 \right) \cdots \varphi \left(x_n \right), \quad \mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n \tag{7}$$

For definition of the wavelets $\psi^{(j)}$ the multi-index

$$\mathbf{d} = (d_1, d_2, \ldots, d_n), \quad d_i \in \{0, 1\} \tag{8}$$

is used and the modulus of $\mathbf{d}$ is defined by

$$|\mathbf{d}| = \sum_{k=1}^{n} d_k 2^{k-1} \tag{9}$$

The tensor wavelets are defined by

$$\psi^{(|\mathbf{d}|)} \left(\mathbf{x} \right) = \left(\varphi \left(x_1 \right) \delta_{0,d_1} + \psi \left(x_1 \right) \delta_{1,d_1} \right) \cdot \left(\varphi \left(x_2 \right) \delta_{0,d_2} + \psi \left(x_2 \right) \delta_{1,d_2} \right) \cdots$$
$$\cdots \left(\varphi \left(x_n \right) \delta_{0,d_n} + \psi \left(x_n \right) \delta_{1,d_n} \right) \tag{10}$$

Their scaled and shifted versions are defined by

$$\varphi_{l,\mathbf{m}} \left(\mathbf{x} \right) = 2^{-nl/2} \varphi \left(2^{-nl} \mathbf{x} - \mathbf{m} \right), \quad \psi_{l,\mathbf{m}}^{(|\mathbf{d}|)} \left(\mathbf{x} \right) = 2^{-nl/2} \psi^{(|\mathbf{d}|)} \left(2^{-nl} \mathbf{x} - \mathbf{m} \right) \tag{11}$$

Now, in $L_2(\mathbb{R}^n)$, there are 2^n scaling equations

$$\varphi \left(\mathbf{x} \right) = 2^{n/2} \sum_{\mathbf{k} \in \mathbb{Z}^n} h_{\mathbf{k}} \varphi \left(2 \mathbf{x} - \mathbf{k} \right) \tag{12}$$

and

$$\psi^{(|\mathbf{d}|)} \left(\mathbf{x} \right) = 2^{n/2} \sum_{\mathbf{k} \in \mathbb{Z}^n} g_{\mathbf{k}}^{(|\mathbf{d}|)} \varphi \left(2 \mathbf{x} - \mathbf{k} \right), |\mathbf{d}| > 0 \tag{13}$$

where the sequences are defined in the following way:

$$h_k = h_{k_1} \cdot h_{k_2} \cdots h_{k_n} \tag{14}$$

and

$$g_{\mathbf{k}}^{(|\mathbf{d}|)} = (h_{k_1}\delta_{0,d_1} + g_{k_1}\delta_{1,d_1}) \cdot (h_{k_2}\delta_{0,d_2} + g_{k_2}\delta_{1,d_2}) \cdots (h_{k_n}\delta_{0,d_n} + g_{k_n}\delta_{1,d_n}) \tag{15}$$

A function $f \in V_{-1}$ has two representations. Firstly a representation with respect to the base of V_{-1}

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} c_{\mathbf{k}}^{(-1)} \varphi_{-1,\mathbf{k}}(\mathbf{x}) \tag{16}$$

and secondly, because of (6) a representation with respect to the bases of V_0 and $W_0^{(j)}$

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} \left(c_{\mathbf{k}}^{(0)} \varphi_{0,\mathbf{k}}(\mathbf{x}) + \sum_{|\mathbf{d}|=1}^{2^n-1} d_{\mathbf{k}}^{(0,|\mathbf{d}|)} \psi_{0,\mathbf{k}}^{(|\mathbf{d}|)}(\mathbf{x}) \right). \tag{17}$$

The terms $\sum_{\mathbf{k}\in\mathbb{Z}^n} c_{\mathbf{k}}^{(0)} \varphi_{0,\mathbf{k}}$ and $\sum_{\mathbf{k}\in\mathbb{Z}^n} d_{\mathbf{k}}^{(0,|\mathbf{d}|)} \psi_{0,\mathbf{k}}^{(|\mathbf{d}|)}$ are the orthogonal projections of f into V_0 and $W_0^{(|\mathbf{d}|)}$, respectively.

The following lemma explains how the coefficients of two different representations of the same function are related to each other.

Lemma 1

$$c_{\mathbf{k}}^{(0)} = \sum_{\mathbf{l}\in\mathbb{Z}^n} h_{\mathbf{l}-2\mathbf{k}} c_{\mathbf{l}}^{(-1)}, \quad d_{\mathbf{k}}^{(0,|\mathbf{d}|)} = \sum_{\mathbf{l}\in\mathbb{Z}^n} g_{\mathbf{l}-2\mathbf{k}}^{(|\mathbf{d}|)} c_{\mathbf{l}}^{(-1)} \tag{18}$$

Proof:

$$c_{\mathbf{k}}^{(0)} = \langle f, \varphi_{0,\mathbf{k}} \rangle = \left\langle f, \sum_{\mathbf{l}\in\mathbb{Z}^n} h_{\mathbf{l}} \varphi_{-1,2\mathbf{k}-\mathbf{l}} \right\rangle = \sum_{\mathbf{l}\in\mathbb{Z}^n} h_{\mathbf{l}-2\mathbf{k}} \langle f, \varphi_{-1,\mathbf{l}} \rangle = \sum_{\mathbf{l}\in\mathbb{Z}^n} h_{\mathbf{l}-2\mathbf{k}} c_{\mathbf{l}}^{(-1)}$$

$$d_{\mathbf{k}}^{(0,|\mathbf{d}|)} = \left\langle f, \psi_{0,\mathbf{k}}^{(|\mathbf{d}|)} \right\rangle = \left\langle f, \sum_{\mathbf{l}\in\mathbb{Z}^n} g_{\mathbf{l}}^{(|\mathbf{d}|)} \varphi_{-1,2\mathbf{k}-\mathbf{l}} \right\rangle = \sum_{\mathbf{l}\in\mathbb{Z}^n} g_{\mathbf{l}-2\mathbf{k}}^{(|\mathbf{d}|)} \langle f, \varphi_{-1,\mathbf{l}} \rangle = \sum_{\mathbf{l}\in\mathbb{Z}^n} g_{\mathbf{l}-2\mathbf{k}}^{(|\mathbf{d}|)} c_{\mathbf{l}}^{(-1)}$$

Then the n-D Mallat-algorithm uses the filters

$$H : \begin{cases} l^2(\mathbb{Z}^n) & \to l^2(\mathbb{Z}^n) \\ \mathbf{c} & \to H\mathbf{c} = \left\{ (H\mathbf{c})_k = \sum_{\mathbf{l}\in\mathbb{Z}^n} h_{\mathbf{l}-2\mathbf{k}} c_{\mathbf{l}} \right\} \end{cases} \tag{19}$$

$$G^{(|\mathbf{d}|)} : \begin{cases} l^2(\mathbb{Z}^n) & \to l^2(\mathbb{Z}^n) \\ \mathbf{c} & \to G^{(|\mathbf{d}|)}\mathbf{c} = \left\{ \left(G^{(|\mathbf{d}|)}\mathbf{c}\right)_{\mathbf{k}} = \sum_{\mathbf{l}\in\mathbb{Z}^n} g_{\mathbf{l}-2\mathbf{k}}^{(|\mathbf{d}|)} c_{\mathbf{l}} \right\} \end{cases} \quad |\mathbf{d}| > 0 \tag{20}$$

The *n*-D-filter applies the one-dimensional *H*-filter successively to each index. The $G^{(1)}$filter applies the one-dimensional *G* filter to the first index and the one-dimensional *H*-filter to all other indicees, and so on. With the help of the Mallat filters the change of the representations can be written in a more compact form

$$\mathbf{c}^{(0)} = H\mathbf{c}^{(-1)}, \mathbf{d}^{(0,|\mathbf{d}|)} = G^{(|\mathbf{d}|)}\mathbf{c}^{(-1)} \tag{21}$$

This compact notation is the key for a schematic description of the change of the representation of a function *f* on different levels of a multi-resolution analysis. This schematic description is given in Figure 5.
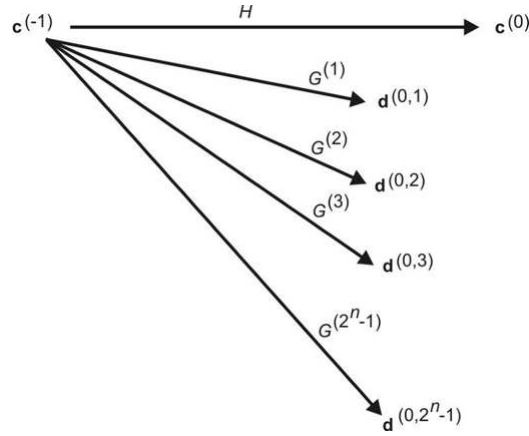
Fig. 5. Transformation of coefficients when changing the base for the representation of a function *f*

We define the *n*-dimensional adjoint Mallat filters by

$$H^* : \begin{cases} l^2\left(\mathbb{Z}^n\right) & \to l^2\left(\mathbb{Z}^n\right) \\ \mathbf{c} & \to H^*\mathbf{c} = \left\{ (H\mathbf{c})_{\mathbf{k}} = \sum_{\mathbf{l} \in \mathbb{Z}^n} h_{\mathbf{k}-2\mathbf{l}}\mathbf{c}_{\mathbf{l}} \right\} \end{cases}$$

$$\left(G^{(|\mathbf{d}|)}\right)^* : \begin{cases} l^2\left(\mathbb{Z}^n\right) & \to l^2\left(\mathbb{Z}^n\right) \\ \mathbf{c} & \to \left(G^{(|\mathbf{d}|)}\right)^*\mathbf{c} = \left\{ \left(G^{(|\mathbf{d}|)}\mathbf{c}\right)_{\mathbf{k}}^* = \sum_{\mathbf{l} \in \mathbb{Z}^n} g_{\mathbf{k}-2\mathbf{l}}^{(|\mathbf{d}|)}\mathbf{c}_{\mathbf{l}} \right\} & |\mathbf{d}| > 0 \end{cases}$$

Theorem 1

$$H^*H + \sum_{|\mathbf{d}|>0} \left(G^{(|\mathbf{d}|)}\right)^* G^{(|\mathbf{d}|)} = I \tag{22}$$

Proof:

$$f = \sum_{\mathbf{k}\in\mathbb{Z}^n} c_{\mathbf{k}}^{(0)} \varphi_{0,\mathbf{k}} + \sum_{|\mathbf{d}|=1}^{2^n-1} \sum_{\mathbf{k}\in\mathbb{Z}^n} \mathbf{d}_{\mathbf{k}}^{(0,|\,\mathbf{d}|)} \psi_{0,\mathbf{k}}^{(|\mathbf{d}|)}$$

$$= \sum_{\mathbf{k}\in\mathbb{Z}^n} \left(H\mathbf{c}^{(-1)}\right)_{\mathbf{k}} \varphi_{0,\mathbf{k}} + \sum_{|\mathbf{d}|=1}^{2^n-1} \sum_{\mathbf{k}\in\mathbb{Z}^n} \left(G^{(|\mathbf{d}|)}\mathbf{c}^{(-1)}\right)_{\mathbf{k}} \psi_{0,\mathbf{k}}^{(|\mathbf{d}|)}$$

$$= \sum_{\mathbf{k}\in\mathbb{Z}^n} \left(H\mathbf{c}^{(-1)}\right)_{\mathbf{k}} \sum_{\mathbf{l}\in\mathbb{Z}^n} h_{\mathbf{l}}\varphi_{-1,2\mathbf{k}+\mathbf{l}} + \sum_{|\mathbf{d}|=1}^{2^n-1} \sum_{\mathbf{k}\in\mathbb{Z}^n} \left(G^{(|\mathbf{d}|)}\mathbf{c}^{(-1)}\right)_{\mathbf{k}} \sum_{\mathbf{l}\in\mathbb{Z}^n} g_{\mathbf{l}}^{(|\mathbf{d}|)}\varphi_{-1,2\mathbf{k}+\mathbf{l}}$$

$$= \sum_{\mathbf{k}\in\mathbb{Z}^n} \left(H\mathbf{c}^{(-1)}\right)_{\mathbf{k}} \sum_{\mathbf{l}\in\mathbb{Z}^n} h_{\mathbf{l}-2\mathbf{k}}\varphi_{-1,\mathbf{l}} + \sum_{|\mathbf{d}|=1}^{2^n-1} \sum_{\mathbf{k}\in\mathbb{Z}^n} \left(G^{(|\mathbf{d}|)}\mathbf{c}^{(-1)}\right)_{\mathbf{k}} \sum_{\mathbf{l}\in\mathbb{Z}^n} g_{\mathbf{l}-2\mathbf{k}}^{(|\mathbf{d}|)}\varphi_{-1,\mathbf{l}}$$

$$= \sum_{\mathbf{l}\in\mathbb{Z}^n} \left(\sum_{\mathbf{k}\in\mathbb{Z}^n} \left(H\mathbf{c}^{(-1)}\right)_{\mathbf{k}} h_{\mathbf{l}-2\mathbf{k}} + \sum_{|\mathbf{d}|=1}^{2^n-1} \sum_{\mathbf{k}\in\mathbb{Z}^n} \left(G^{(|\mathbf{d}|)}\mathbf{c}^{(-1)}\right)_{\mathbf{k}} g_{\mathbf{l}-2\mathbf{k}}^{(|\mathbf{d}|)} \right) \varphi_{-1,\mathbf{l}}$$

$$= \sum_{\mathbf{l}\in\mathbb{Z}^n} \left(\left(H^*H\mathbf{c}^{(-1)}\right)_{\mathbf{l}} + \sum_{|\mathbf{d}|=1}^{2^n-1} \left(\left(G^{(|\mathbf{d}|)}\right)^* G^{(|\mathbf{d}|)}\mathbf{c}^{(-1)}\right)_{\mathbf{l}} \right) \varphi_{-1,\mathbf{l}}$$

$$= \sum_{\mathbf{l}\in\mathbb{Z}^n} \left(\left(H^*H + \sum_{|\mathbf{d}|>0} \left(G^{(|\mathbf{d}|)}\right)^* G^{(|\mathbf{d}|)}\right) \mathbf{c}^{(-1)} \right)_{\mathbf{l}} \varphi_{-1,\mathbf{l}}$$

On the other hand

$$f = \sum_{\mathbf{l}\in\mathbb{Z}^n} c_{\mathbf{l}}^{(-1)} \varphi_{-1,\mathbf{l}}$$

holds, which combines to

$$\left(H^*H + \sum_{|\mathbf{d}|>0} \left(G^{(|\mathbf{d}|)}\right)^* G^{(|\mathbf{d}|)}\right) \mathbf{c}^{(-1)} = \mathbf{c}^{(-1)}$$

and because the coefficient sequence $\mathbf{c}^{(-1)}$ is arbitrary, this proves the theorem.

The meaning of the theorem above is that the adjoint operator comes into the picture, when the representation of a function is changed back to the original level of the multi-resolution analysis. This is schematically shown in Figure 6.
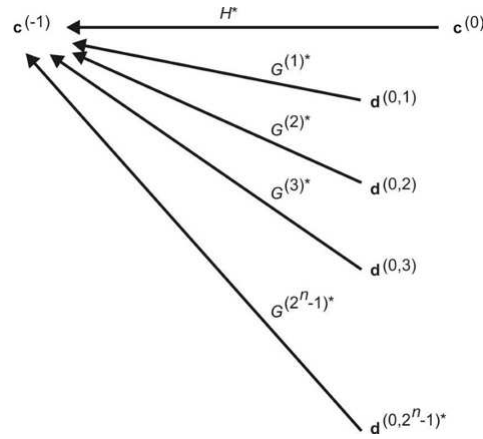
Fig. 6. Transformation of coefficients when back-changing the base for the representation of a function f

4. Wavelet segmentation

4.1. Principle solution

Assume, that

$$L_2\left(\mathbb{R}^2\right) \supset \cdots \supset V_{-1} \supset V_0 \supset V_1 \supset \cdots \supset \{0\} \tag{23}$$

is a multi-resolution analysis (MRA) of $L_2\left(\mathbb{R}^2\right)$. The space V_0 corresponds to the sampling step-width of a given data set f. Supposed, the signal f has only components in a few of the scaling spaces V_k, we are aiming at the detection of those signal parts, which belong to $V_m, m > 0$. The given signal has the representation

$$f(\mathbf{x}) = \sum_{\mathbf{l} \in \mathbb{Z}^2} c_{\mathbf{l}}^{(0)} \varphi(\mathbf{x} - \mathbf{l}) \tag{24}$$

with φ being the scale function of the MRA. The signal part $f_m \in V_m$ has the representation

$$f_m(\mathbf{x}) = \sum_{\mathbf{l} \in \mathbb{Z}^2} c_{\mathbf{l}}^{(m)} \varphi_{m,\mathbf{l}}(\mathbf{x}), \varphi_{k,\mathbf{l}}(x) := 2^{-\frac{m}{2}} \varphi\left(2^{-m}\mathbf{x} - \mathbf{l}\right) \tag{25}$$

The coefficients $\left\{\mathbf{c}^{(m)}\right\}$ can be computed from the given coefficients $\left\{\mathbf{c}^{(0)}\right\}$ by m steps of the Mallat-algorithm

$$\mathbf{c}^{(m)} = H^m \mathbf{c}^{(0)} \tag{26}$$

with

$$H : \begin{cases} l_2\left(\mathbb{R}^2\right) & \to l_2\left(\mathbb{R}^2\right) \\ (H\mathbf{c})_{\mathbf{k}} & = \sum_{\mathbf{l} \in \mathbb{Z}^2} h_{\mathbf{l}-2\mathbf{k}} c_{\mathbf{l}} \end{cases} \tag{27}$$

The signal part f_m does not only contain those signal parts from V_m but also the projection of finer signal parts into V_m. For this reason these projections still have to be subtracted. First, the shadow-projection from $V_j, j < m$ into V_m have to be eliminated. The key to this elimination is the fact, that the magnitude of these shadow projections is significantly smaller than the magnitude of the V_m content of f. Hence, the elimination can be achieved by a thresholding

$$\bar{c}_\mathbf{l}^m = \begin{cases} c_\mathbf{l}^m & , |c_\mathbf{l}^m| > \gamma \\ 0 & , \text{else} \end{cases} \tag{28}$$

The thresholded signal is

$$\bar{f}_m = \sum_{\mathbf{l} \in \mathbb{Z}^2} \bar{c}_\mathbf{l}^m \varphi_{m,\mathbf{l}} \tag{29}$$

Since an element of V_m has no entry in the detail spaces $W_n^{(i)}$

$$f_e = \bar{f}_m - \sum_{n=1}^{m} \sum_{i=1}^{3} \sum_{\mathbf{l} \in \mathbb{Z}^2} \left(\left(G^{(i)} \right)^* \mathbf{d}^{(n,i)} \right)_\mathbf{l} \varphi_{n,\mathbf{l}} \tag{30}$$

● **Example**

In order, to make this procedure more transparent, it will be applied to a very academic example. The data set (Fig. 7) consists of two parts, belonging to V_4 and V_6, respectively.
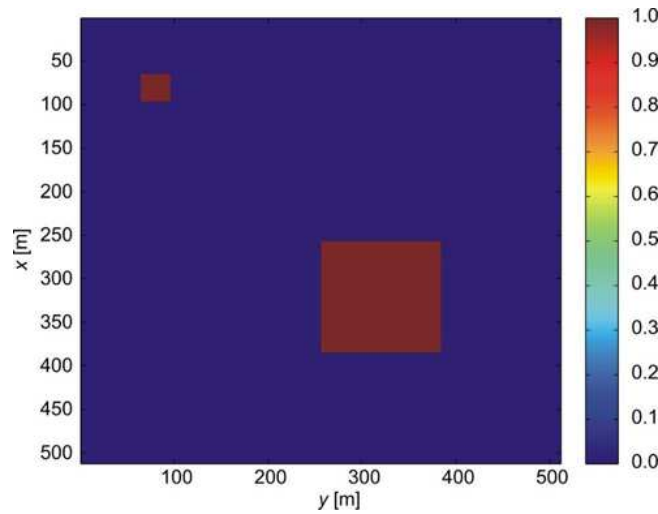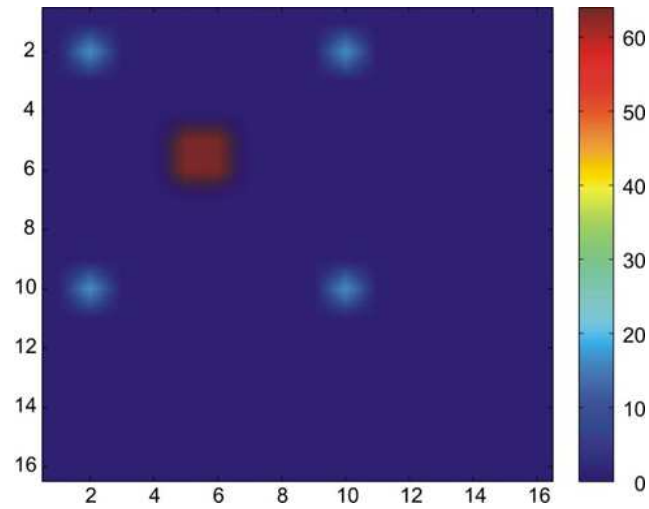

Fig. 7. Data set
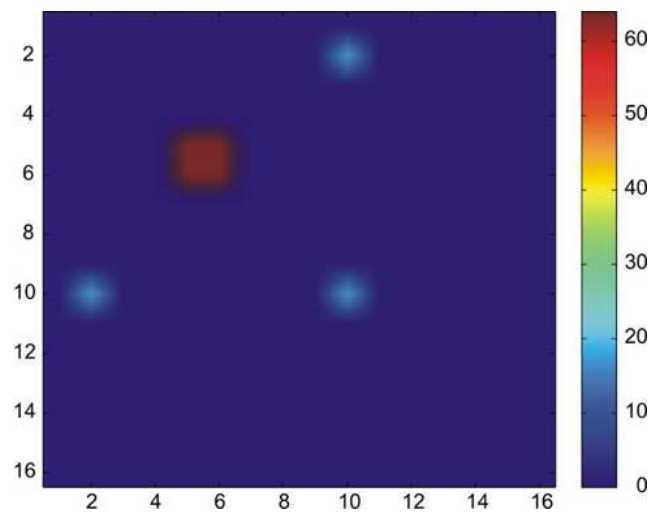
Fig. 8. Contents of smoothing space V_6 (north-west corner) and of the detail spaces $W_6^{(i)}, i = 1, 2, 3$ (remaining corners)

The data consist of a smaller structure in the north-west corner and a larger structure in the south-east corner. After 6 steps of Mallat's algorithm the contents of the space V_6 and of the detail spaces $W_6^{(i)}, i = 1, 2, 3$ are shown in Figure 8.

The dark structure is the V_6 content of the signal, while the light structures are the shadow projections of the smaller structure. In order to remove these shadow projections, the V_6 content has to be thresholded. After thresholding the content of the level 6 spaces are displayed in Figure 9.

Fig. 9. Contents of smoothing space V_6 (north-west corner) and of the detail spaces $W_6^{(i)}, i = 1, 2, 3$ (remaining corners) afterthresholding

Finally, if the details are removed from the threshold dV_6 signal content, the smaller structure is eliminated from the data. As it is shown in Figure 10.
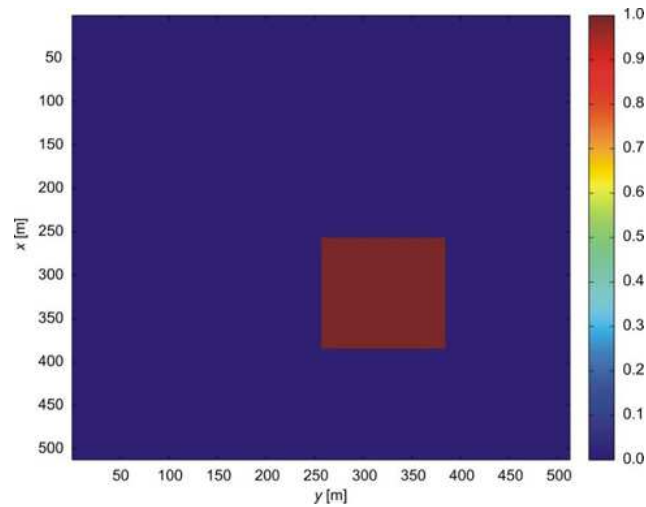

Fig. 10. Segmented data

4.2. Limitation of principal solution and work-arounds

The basic limitation of the MRA is its dyadic structure: All structures are assumed to have sizes, which equal a power of two, times the sampling step size. Because, this limitation is never met in practice, the thresholding has to be carried out not only in one scaling space V_m but also in its adjacent spaces V_{m+1} and V_{m-1}. If some coefficients are removed by threshold the corresponding coefficients in the coarser scaling- and detail spaces also have to be removed. These coefficients can be located by the so-called influence-pyramid. Size and shape of this pyramid is determined by the length of the wavelet filter. Assume, that for the filter coefficients holds

$$h_{\mathbf{l}} = 0, \quad \text{if} \quad l_i < 0 \quad \text{or} \quad l_i > L, \quad \mathbf{l} = (l_1, l_2)$$

In this case we say, the filter has the length $L + 1$.

For an arbitrary coefficient $c_{\mathbf{k}}^{(m)}$ it holds

$$c_{\mathbf{k}}^{(m)} = \sum_{\mathbf{l} \in ([0,L] \times [0,L]) \cap \mathbb{Z}^2} h_{\mathbf{l}} c_{\mathbf{l}-2\mathbf{k}}^{(m-1)} \tag{31}$$

If now by thresholding the coefficient $c_{\mathbf{k}}^{(m)}$ is set to zero, the coefficients

$$c_{\mathbf{l}-2\mathbf{k}}^{(m-1)}, \quad \mathbf{l} \in ([0, L] \times [0, L]) \cap \mathbb{Z}^2 \tag{32}$$

have to be set to zero as well. On the level m the influence-region of the thresholded coefficient $c_{\mathbf{k}}^{(m)}$ is only the coefficient itself

$$I_{\mathbf{l}}^{(m,\mathbf{k})} = \delta_{\mathbf{l},\mathbf{k}} \tag{33}$$

This means, the influence region of the coefficient $c_{\mathbf{k}}^{(m)}$ in the decomposition level m on the coefficients in the decomposition level $m-1$ given by

$$I_{\mathbf{l}}^{(m-1,\mathbf{k})} = I_{\mathbf{l}}^{(m,\mathbf{k})} \otimes \left((\mathbf{1}_L)^T (\mathbf{1}_L)\right) \tag{34}$$

where $\otimes$ is the Konecker product and $\mathbf{1}_L$ is a row-vector with $L+1$ entries, all of them being 1. Each non-zero entry in $I_{\mathbf{l}}^{(m,\mathbf{k})}$ creates an influence region in the level $m-2$ by

$$I_{\mathbf{l}}^{(m-2,\mathbf{k})} = I_{\mathbf{l}}^{(m-1,\mathbf{k})} \otimes \left((\mathbf{1}_L)^T (\mathbf{1}_L)\right) \tag{35}$$

If the levels are stacked vertically, the influence-regions in the different levels form a pyramid, the so-called influence pyramid. Schematically, this is displayed in Figure 11.
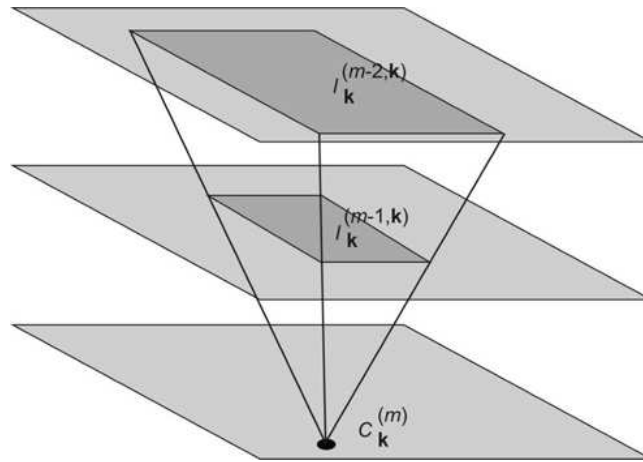

Fig. 11. Influence pyramid

For each coefficient on level m which is set to zero during the thresholding, all coefficients, which fall into its influence pyramid have to be set to zero as well. After this modification of the wavelet spectrum, the wavelet reconstruction yields the segmentation of the given data set into objects of the size corresponding to the chosen level m.

4.3. Wavelet segmentation algorithm

The wavelet segmentation algorithm is described in the flow-chart given in Figure 12. From the given sampling step-width h and the average size of the objects to be detected

a decomposition depth m is chosen so, that $2^m h$ is as close as possible to . As next step a wavelet decomposition up to level $m + 1$ is carried out. In the levels $m − 1$, m, $m + 1$ a thresholding is carried out, aiming at a suppression of the shadow elements in these levels. The indices of the suppressed coefficients are stored. For each of the stored entries the corresponding influence pyramid is constructed. All smoothing coefficients $c_k^{(n)}$ and detail coefficients $d_k^{(i,n)}$, falling into one of these pyramids are set to zero. A wavelet reconstruction of the so modified wavelet spectrum yields the segmented scene, which still can be improved by a morphologic dilatation.
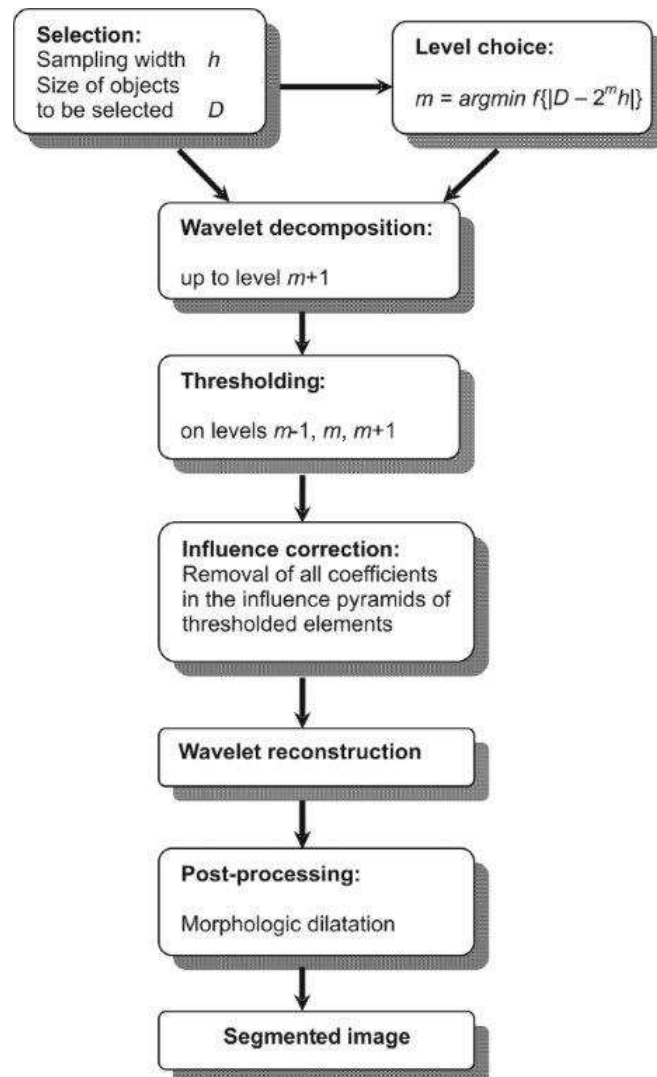
Fig. 12. Flow chart

5. Wavelet segmentation of laser-scanning data

5.1. One-step procedure

In Figure 13 a laser-scanning scene is displayed. The scene is given in a local co-ordinate system and with a pixel size of approximately 0.25×0.25 m^2. The objects to be segmented are buildings with an average size of 10 m. This leads to a level of $m = 5$. In Figure 14 the wavelet spectrum on levels $m = 4$ and $m = 5$ is displayed. In both sub-images the scaling coefficients $c_k^{(n)}$ are arranged in the north-west corner and the detail coefficients $d_k^{(i,n)}$ in the other three corners.
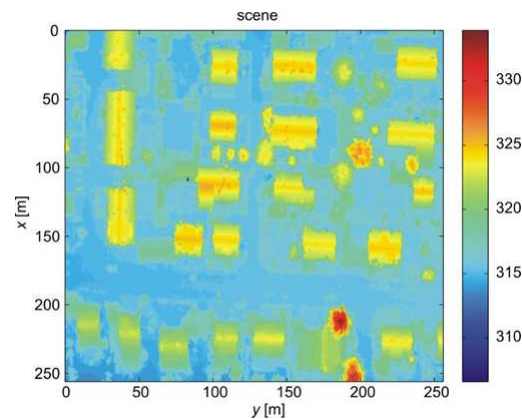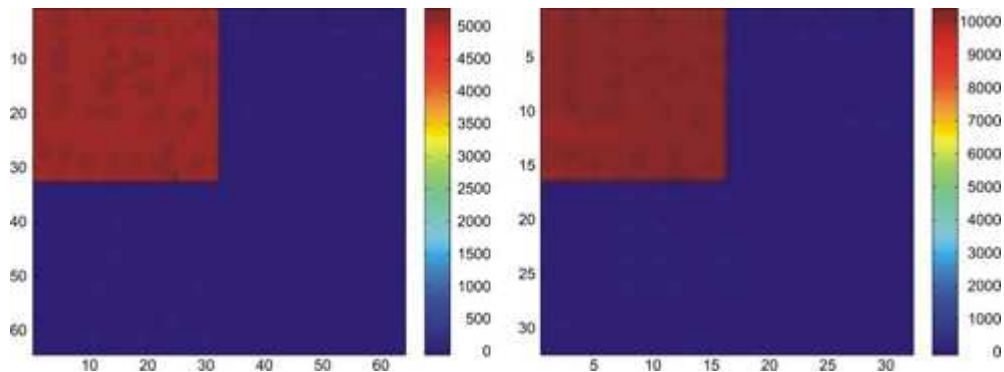

Fig. 13. Laser-scanning scene

Fig. 14. Wavelet spectrum on level $m = 4$ (left) and $m = 5$ (right)

In the scaling coefficients of level $m = 4$ the shapes of the buildings are clearly visible and smaller details have disappeared. But the buildings are not concentrated on level $m = 4$ only, but also on level $m = 5$. The thresholding step is to separate pixels

belonging to buildings from those that do not belong to buildings. In Figure 15 the level 4 coefficient before and after thresholding are compared.
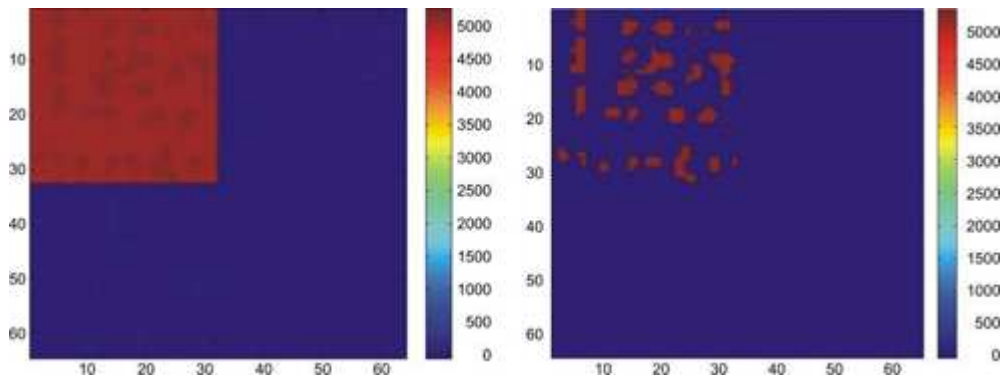

Fig. 15. Wavelet spectrum on level $m = 5$ (left) and thresholded wavelet spectrum on the same level (right)

The thresholding makes the objects to be segmented much clearer visible. But still, there are shadow projections visible in the detail coefficients. If the thresholding is carried out on the levels 4, 5, 6 and the coefficients of the corresponding influence pyramids are set to zero, the wavelet reconstruction yields the segmentation. The result is displayed in Figure 16.
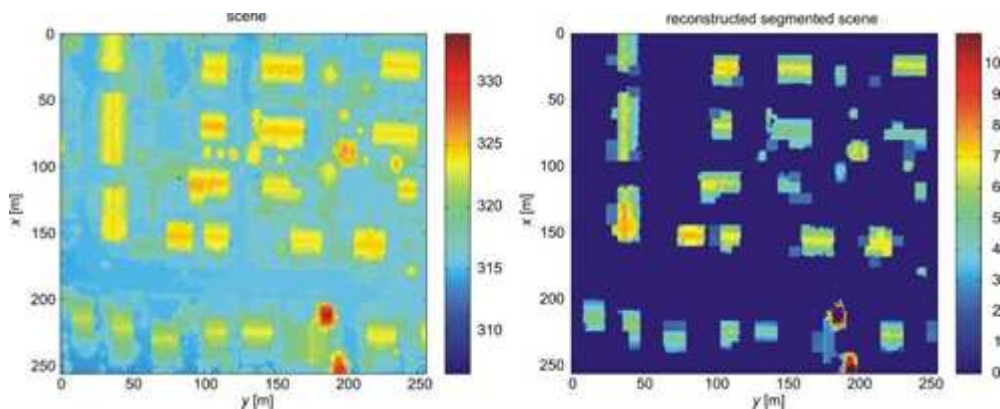

Fig. 16. Original (left) and segmented scene

In the segmented scene the small details, not belonging to buildings, are eliminated. But larger detail, as for instance big trees with crowns having sizes of small buildings, are still contained in the segmented scene. This is unavoidably, because the wavelet segmentation is a size based algorithm. The separation between buildings and treetops can only be achieved by a subsequent texture analysis.

5.2. Two-step procedure

In most scenes not only the elements of a single scaling space V_m but of two V_m, V_n ore more scaling spaces are to be segmented. This can be done by iterative application of the one-step procedure: First the elements of the coarser of the two scaling spaces V_m are segmented. Then the segmented objects are removed from the original scene. And as the following step, the objects of V_n are segmented in the modified scene. In Figure 17 a laser-scanning scene is displayed.
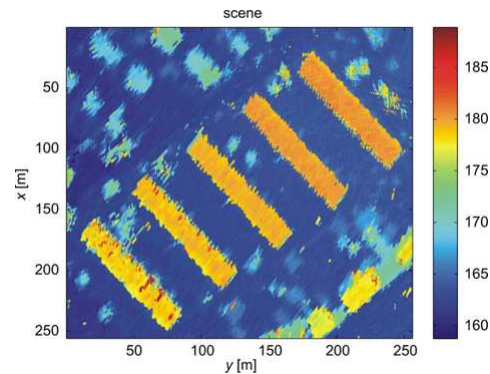

Fig. 17. Laser-scanning scene with objects on two different scales

In the first step the larger objects are to be segmented. For this purpose the scene was decomposed up to level 5. The content of V_5 is shown in Figure 18.
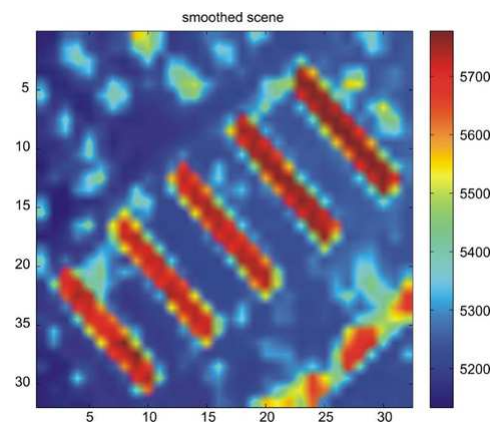

Fig. 18. Smoothed scene

The smoothed scene shows the objects to be segmented but also smaller objects. The two classes of objects can be distinguished by their intensities. The smaller objects appear on a coarser scale only as shadow-projections and have therefore a lower inten-

sity. The shadow-projections can be discriminated by a thresholding. The remaining objects after a morphologic dilatation operation are shown in Figure 19.
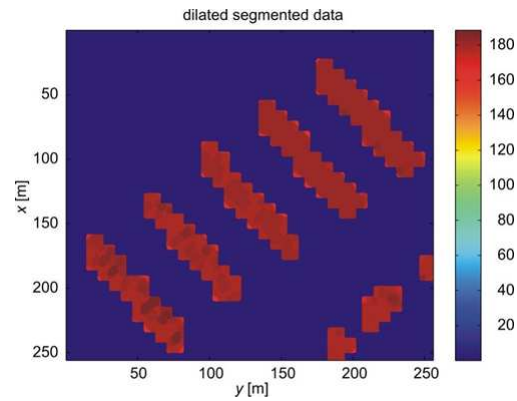

Fig. 19. Dilated thresholded scene

If the values, which fall into the mask of the segmented objects are replaced by the values of the surrounding terrain, the segmented objects are erased from the original scene. The modified scene is shown in Figure 20.
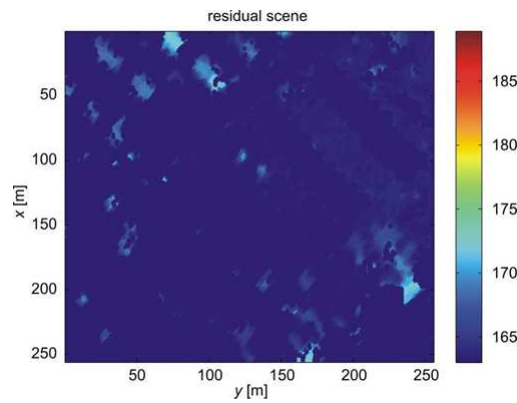

Fig. 20. Residual scene

If the same procedure is repeated for the residual scene, also the smaller objects can be segmented. The result is shown in Figure 21.
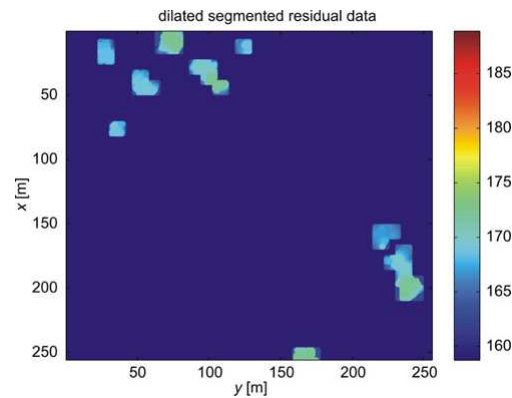
Fig. 21. Segmented residual scene

6. Conclusions and future work

We have developed and tested a wavelet-based procedure to segment buildings from laser-scanning scenes. The wavelet approach makes the procedure very fast and enables the processing of large data sets. The limitation of the procedure is its scale-based distinction between objects to be segmented and the rest. For this reason the algorithm is not able to distinguish between small buildings and trees with large crowns. Here only a texture analysis could be a solution, but at the moment is it not clear, how this can be achieved by wavelet analysis.

Another way to improve the algorithm is to use wavelet-packages instead of wavelets. The key of the numerical efficiency of wavelet decomposition and reconstruction is its dyadic structure. The data is decomposed in data sets of two-times, four-times, 8-times... the sampling width structures. Real objects have never two-, four- or 8-times the size of the sampling width, but also sizes, which lay between these scales. This results in the wide distribution of shadow-projections over different scales. Wavelet packages provide a finer scale division but on the costs of higher numerical effort. Here a suitable balance has to be found.

However, the results in the present form are promising and useful. The result of laser scanning data segmentation using wavelet-based algorithm can be used as initial information for another, more sophisticated method such as region growing method.

References

Awrangjeb M., Ravanbakhsh M., Fraser C.S., (2010): *Automatic detection of residential buildings using LIDAR data and multispectral imagery*, ISPRS Journal of Photogrammetry and Remote Sensing, 65(2010), pp. 457–467.

Axelsson P., (2000): *DEM generation from laser scanner data using adaptive TIN models.* International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII-1/B4, pp. 110–117.

Borkowski A., (2004): *Modellierung von Oberflächenmit Diskontinuitäten*, Deutsche Geodätische Kommission, Reihe C, Heft Nr 575.

Borkowski A., Jóźków G., (2008): *Airborne laser scanning data filtering using flakes*, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII, Part B3b, pp. 179–184.

Briese C., Pfeifer N., Dorninger P., (2002): *Applications of the robust interpolation for DTM determination*, Symposium ISPRS Commision III, Photogrammetric Computer Vision, Graz, 9 – 13 September 2002; International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV/3A, pp. 55–61.

Carlberg M., Gao P., Chen G., Zakhor A., (2009): *Classifying urban landscape in aerial LiDAR using 3D shape analysis*, Proceedings of the IEEE International Conference of Image Processing, IEEE Cairo, Egypt, pp. 1701–1704.

Dorninger P., Pfeifer N., (2008): *A comprehensive automated 3D approach for building extraction, reconstruction and regularization from air borne laser scanning point cloud*, Sensors, Vol. 8, pp. 7323–7343.

Elmqvist M., (2002): *Ground surface estimation from airborne laser scanner data using active shape models*, ISPRS, Commission III, Symposium Photogrammetric Computer Vision, Graz, 9 – 13 September 2002, pp. 114–118.

Filin S., Pfeifer N., (2006): *Segmentation of airborne laser scanning data using a slope adaptive neighbourhood*, ISPRS Journal of Photogrammetry and Remote Sensing, pp. 60, 71–80.

Forlani G., Nardinocchi C., Scaioni M., Zingaretti P., (2006): *Complete classification of row LiDAR data and 3D reconstruction of buildings*, Pattern Analysis and Applications 8, pp. 357–374

Haala N., Kada M., (2010): *An update on automatic 3D building reconstruction*, ISPRS Journal of Photogrammetry and Remote Sensing 65, pp. 570–580

Kabolizade M., Ebadi H., Ahmadi S., (2010): *An improved snake model for automatic extraction of buildings from urban aerial images and LiDAR data*, Computers, Environment and Urban Systems 34, pp. 435–441.

Keller W.,(2004): *Wavelets in Geodesy and Geodynamics*, DeGruyter, Berlin 2004.

Kraus K., (2000): *Photogrammetrie. Band 3. TopographischeInformatonssysteme*, Dümmler, Köln.

Kraus K., Pfeifer N., (2001): *Advanced DTM generating from LIDAR data*, International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV-3/W4, Annopolis, Maryland, 22–24 October, pp. 23–30.

Laky S., Zaletnyik P., Toth C., (2010): *Land classification of wavelet-compressed full-waveform LiDAR data*, International Archives of Photogrammetry and Remote Sensing, Vol. XXXVIII-3A, pp. 115–119.

Louis A.K., Maass P., Rieder A., (1998): *Wavelets*, B.G. Teubner, Stuttgart 1998.

Mallet C., Bretar F., Soergel U., (2008): *Analysis of full- waveform LiDAR data for classification of urban areas*, Photogrammetrie, Fernerkundung, GeoInformation (PFG) (5), pp. 337–349.

Matikainen L., Hyyppa J., Hyyppa H., (2003): *Autimatic detection of buildings from laser scanner data for map updating*, International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences 33(part 3/W13), pp. 218–224.

Melzer T., (2007): *Non-parametric segmentation of ALS point clouds using mean shift*, Journal of Applied Geodesy 1(3), pp. 159–70.

Neidhart H., Sester M., (2008): *Extraction of buildings ground planes from LiDAR data*, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII, Part B2, pp. 405–410.

Oude Elbering S.J., Vosselman G., (2009): *Building reconstruction by terged based graph matching on incomplete laser data: analysis and limitations*, Sensors 9(8), pp. 6101–6118.

Roggero M., (2001): *Airborne laser scanning: Clustering in row data*, International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV-3/W4, Annopolis, Maryland, 22–24 October, pp. 227–232.

Sampath A., Shan J., (2007): *Building boundary tracing and regularization from airborne pointclouds*, Photogrammetric Engineering & Remote Sensing, Vol. 73(7), pp. 805–812.

Sithole G., (2001): *Filtering of laser altimetry data using a slope adaptive filter*, International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV-3/W4, Annopolis, Maryland, 22–24 October, pp. 203–210.

Sithole G., Vosselman G., (2004): *Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds,* ISPRS Journal of Photogrammetry and Remote Sensing, Vol. LIX, pp. 85–101.

Vosselman G., Maas H-G., (2010): *Airborne and terrestrial laser scanning.* Whittles Publishing, Dunbeath.

Vu T.T., Yamazaki F., Tokunaga M., (2002): *Wavelet-based filtering the cloud points derived from airborne laser scanner*, Proc. 23rd Asian Conference on Remote Sensing. Kathmandu, 25–29 November. http://www.gisdevelopment.net/aars/acrs/2002/adp/adp008.asp

Wei H., Bartels M., (2006): *Unsupervised segmentation using Gabor Wavelets and statistical features in LIDAR data analysis*, 18th International Conference on Pattern Recognition, Hong Kong, 20 August.

Xu L., Yang Y., Jiang B., Li J., (2007): *Ground extraction from airborne laser data based on wavelet analysis,* Proceedings of SPIE – the International Society for Optical Engineering, DOI: 10.1117/12.750425

Segmentacja budynków w zbiorze danych lotniczego skaningu laserowego w oparciu o analizę falkową

Wolfgang Keller[1], **Andrzej Borkowski**[2]

[1]Instytut Geodezji, Uniwersytet w Stuttgartcie,
Geschwister-Scholl-Str. 24D, D-70174 Stuttgart
wolfgang.keller@gis.uni-stuttgart.de
[2]Instytut Geodezji i Geoinformatyki,
Uniwersytet Przyrodniczy we Wrocławiu,
Grunwaldzka 53, PL-50-357 Wrocław
andrzej.borkowski@up.wroc.pl

Streszczenie

W ostatnich latach ważnym zagadnieniem staje się trójwymiarowe modelowania budynków w oparciu o dane lotniczego skaningu laserowego. Istotnym krokiem na drodze dochodzenia do trójwymiarowego modelu jest segmentacja budynków w zbiorze danych skaningowych. W tym celu zaproponowano w pracy algorytm bazujący na analizie wielorozdzielczej danych w dziedzinie falkowej. Proponowana metoda koncentruje się wyłącznie na budynkach, które podlegają segmentacji. Wszystkie inne obiekty oraz powierzchnia terenu są usuwane. Algorytm działa na danych opartych o regularna siatkę. Segmentacja oparta o analizę falkową przebiega w następujących krokach głównych: falkowa dekompozycja aż do odpowiednio wybranego poziomu (wybranej skali), progowanie na wybranym i sąsiednich poziomach, usunięcie wszystkich współczynników dekompozycji w obrębie tak zwanej piramidy wpływu i rekonstrukcja falkowa sygnału. Jeśli budynki podlegające segmentacji występują na kilku skalach opisaną procedurę należy zastosować iteracyjnie. Algorytm opary o analizę falkową charakteryzuje się duża szybkością. Jednakże ograniczeniem proponowanej metody jest rozróżnialność obiektów podlegających segmentacji od reszty na tym samym poziomie - tej samej skali.