Remote Sensing of Environment 165 (2015) 175–185

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Reconstruction of individual trees from multi-aspect TomoSAR data

Michael Schmitt *, Muhammad Shahzad, Xiao Xiang Zhu *

Helmholtz Young Investigators Group "SiPEO", Technische Universitaet Muenchen (TUM), Arcisstr. 21, 80333 Munich, Germany

ARTICLE INFO

Article history:
Received 9 January 2015
Received in revised form 8 May 2015
Accepted 16 May 2015
Available online 25 May 2015

Keywords:
Synthetic aperture radar (SAR)
Multi-aspect
SAR tomography
Trees
3D reconstruction
Forested areas
Point cloud segmentation

ABSTRACT

The localization and reconstruction of individual trees as well as the extraction of their geometrical parameters is an important field of research in both forestry and remote sensing. While the current state-of-the-art mostly focuses on the exploitation of optical imagery and airborne LiDAR data, modern SAR sensors have not yet met the interest of the research community in that regard. This paper presents a prototypical processing chain for the reconstruction of individual deciduous trees: First, single-pass multi-baseline InSAR data acquired from multiple aspect angles are used for the generation of a layover- and shadow-free 3D point cloud by tomographic SAR processing. The resulting point cloud is then segmented by unsupervised mean shift clustering, before ellipsoid models are fitted to the points of each cluster. From these 3D ellipsoids the relevant geometrical tree parameters are extracted. Evaluation with respect to a manually derived reference dataset prove that almost 74% of all trees are successfully segmented and reconstructed, thus providing a promising perspective for further research toward individual tree recognition from SAR data.

1. Introduction

The analysis of individual trees in remote sensing data has been a widely studied field of research for quite some years now. This is mainly caused by the fact that it is an important topic for a variety of application fields, among which the most relevant certainly is the task of sustainable forest management: In many countries, single-tree related parameters are used as a basis for forest inventory, e.g. tree species, mean tree height or timber volume. Still, most of these variables are collected manually by measuring sample plots in staff-, time-, and thus cost-intensive field surveys (Shiver & Borders, 1996), although remote sensing-based methods have been investigated for some decades now (Fagan & DeFries, 2009). Another exemplary application is the extraction of individual urban trees in order to generate city tree cadastres or to provide additional objects for geoinformation systems (GIS) and 3D city models (Straub & Heipke, 2001).

Since the mid-1990s an abundance of literature has been published on the detection and localization of individual trees (Chang, Eo, Kim, & Kim, 2013; Chen, Baldocchi, Gong, & Kelly, 2006; Leckie et al., 2005; Pollock, 1996; Wulder, Niemann, & Goodenough, 2000) as well as the delineation of their tree crowns (Culvenor, 2002; Erikson, 2003; Jing, Hu, Li, & Noland, 2012; Koch, Heyder, & Weinacker, 2006; Pouliot, King, Bell, & Pitt, 2002) both from aerial images and LiDAR-derived canopy height models. In contrast to that, the analysis of forested areas on the single-tree level by means of synthetic aperture radar (SAR) remote sensing has not yet met the interest of the community, although modern sensors have reached sub-meter resolutions down to the decimeter-range in recent years, enabling a detailed mapping of natural and urban scenes (Brenner, 2012; Mittermayer, Wollstadt, Prats-Iraola, & Scheiber, 2014). Until now, most of the research exploiting SAR for forest mapping purposes has focused on large-scale forest classification (Perko, Raggam, Deutscher, Gutjahr, & Schardt, 2011) or canopy height model (CHM) reconstruction using X-band sensors (Izzawati, Wallington, & Woodhouse, 2006). In addition, another research direction based on longer-wavelength, canopy-penetrating SAR (e.g. L-band or P-band) is aiming at volume and biomass retrieval using allometric equations (Le Toan, Beaudoin, Riom, & Guyon, 1992), regression models (Englhart, Keuck, & Siegert, 2011), polarimetric SAR interferometry (Neumann, Ferro-Famil, & Reigber, 2010) or SAR tomography (TomoSAR) (Frey, Morsdorf, & Meier, 2008). Only recently, the very first investigations toward single-tree recognition in airborne SAR data have been introduced (Schmitt, Brück, Schönberger, & Stilla, 2013).

Considering the advantages of radar remote sensing, i.e. the active sensor principle, the independence of sunlight and weather conditions, and the potential to map wide swaths in short time, this article intends to further ignite the discussion about the SAR-based mapping of forests on the individual-tree level by proposing a method for the reconstruction of trees from tomographic single-pass SAR data. Exploiting an airborne millimeterwave system equipped with four simultaneously receiving, spatially separated antennas (and thus, multiple independent baselines), first a three-dimensional point cloud is generated by TomoSAR processing. In analogy to the approaches based on 3D LiDAR point clouds, in the

* Corresponding authors.
E-mail addresses: m.schmitt@tum.de (M. Schmitt), muhammad.shahzad@tum.de (M. Shahzad), xiao.zhu@dlr.de (X.X. Zhu).
URL: http://www.sipeo.bgu.tum.de (X.X. Zhu).

http://dx.doi.org/10.1016/j.rse.2015.05.012

second step a clustering algorithm is applied to the TomoSAR point cloud in order to segment the individual trees in 3D space. While the studies of Morsdorf et al. (2004) or Gupta, Weinacker, and Koch (2010) are based on *k*-means clustering, which needs the number of expected clusters and an initialization of their centers as a priori knowledge, in the presented work the unsupervised mean-shift clustering algorithm is used, which enables a fully automatic procedure (Comaniciu & Meer, 2002) and has already successfully been applied to the reconstruction of building facades (Shahzad & Zhu, 2015) and roofs (Shahzad & Zhu, 2014) in TomoSAR point clouds. Finally, rotational ellipsoids are used to model the segmented clusters in order to approximate the tree crown shapes. From these ellipsoids the tree positions, heights and crown diameters can be extracted. This tree reconstruction strategy is evaluated using airborne millimeterwave InSAR data acquired from multiple aspect angles. The purpose of this study is to demonstrate the potential of millimeterwave SAR remote sensing for the reconstruction of individual trees including a first estimate of the achievable accuracy level.

2. Generation of point clouds from airborne tomographic SAR data

2.1. The principle of SAR tomography

The first step of the methodology presented in this paper is the generation of a point cloud of the forested area under investigation from single-pass multi-baseline InSAR data. In single-pass InSAR configurations, two (or more) coherent SAR images are acquired simultaneously by sensors separated by a certain spatial baseline. This simultaneity is necessary because of temporal decorrelation effects, which especially occur over vegetated areas due to wind or growth effects or other, potentially more rapid changes of the scene (Ahmed, Siqueira, Hensley, Chapman, & Bergen, 2011). Since only recently satellite missions offering spaceborne single-pass InSAR data have been realized (Moreira et al., 2004), airborne multi-antenna systems have been the only possibility to acquire such data so far.

As discussed by Hoekman and Varekamp (2001) already, unfortunately the side-looking SAR imaging geometry leads to severe height, displacement and occlusion errors for individual trees due to the well-known layover and shadowing effects (cf. Fig. 1). While the shadowed scene parts can be filled with information by multi-aspect data fusion, the layover-affected resolution cells need to be treated by tomographic SAR inversion: Basically, SAR tomography aims at creating a second synthetic aperture along elevation direction. However, in opposition to the high number of densely and regularly spaced azimuth samples, for the elevation aperture comparably few observations per resolution cell are available in the form of a stack of *N* coregistered single look complex (SLC) SAR images acquired from slightly different antenna positions (Gini & Lombardini, 2005; Reigber & Moreira, 2000; Zhu & Bamler, 2010). The complex measurement stored in a pixel of the *n*th acquisition with a baseline B_n between the respective slave antenna *n* and the master antenna is the integral of the reflected signal weighted by a linear phase term:

$$z_n = \int_{s_{min}}^{s_{max}} x(s) exp(j \cdot \phi_n(s)) ds, \tag{1}$$

where $x(s)$ is the reflectivity function along elevation, $[s_{min}; s_{max}]$ defines the relevant part of the elevation profile, and j is the imaginary unit. $\phi(s)$ is the phase related to system parameters and two-way distance between the scattering center at elevation s and the receiving antenna.

Using SAR sensors operating with rather long wavelengths (e.g. L- or P-band), this enables a fully three-dimensional, canopy-penetrating imaging of vegetation volumes. If, however, shorter wavelengths are used, where canopy penetration is less likely, Eq. (1) can also be approximated by discretizing and sparsifying the reflectivity profile. The measurement vector is then formulated as

$$\mathbf{z} = \sum_{k=1}^{K} x_k \mathbf{a}(s) + \mathbf{n} = \mathbf{A}(\mathbf{s})\mathbf{x} + \mathbf{n}. \tag{2}$$

In this notation, $\mathbf{x} = [x_1, \ldots, x_K]^T$ is the source signal vector containing the complex reflectivities of the *K* discrete scattering contributions, and $\mathbf{n}$ represents complex circularly symmetric Gaussian noise. $\mathbf{A}(\mathbf{s}) = [\mathbf{a}(s_1), \ldots, \mathbf{a}(s_K)]$ is the $N \times K$ system matrix of *K* concatenated steering vectors each of which corresponding to one backscattering source. The task of SAR tomography basically is to invert this imaging model in order to reconstruct either a continuous reflectivity profile in the case of volumetric scenes or a discrete reflectivity profile in case of scenes containing mostly discrete scatterers. Since the tree reconstruction approach presented in this paper is based on point cloud segmentation, here a discrete TomoSAR model is chosen aiming at sparse reflectivity profiles. The utilized approach was first presented in Schmitt & Stilla (2014b) and tested for urban scenes, but has been shown to work well for decimeter-resolution millimeterwave SAR data of forested areas as well (Schmitt & Stilla, 2014a). Its core algorithm is quickly recapitulated in the following section. Due to the low level of canopy penetration provided by millimeterwave SAR, of course the situation would be a little different for very dense forests: In this case it would be unlikely that tree crowns overlap with each other or with any ground contribution in the SAR viewing direction, unless they differ strongly in height, or, e.g., if forest clearing appears. For dense forests it could therefore be advisable to only use standard SAR interferometry for height reconstruction followed by the multi-aspect data fusion step described in Section 2.4 of this paper in order to avoid unnecessary computational costs and methodical complexity.

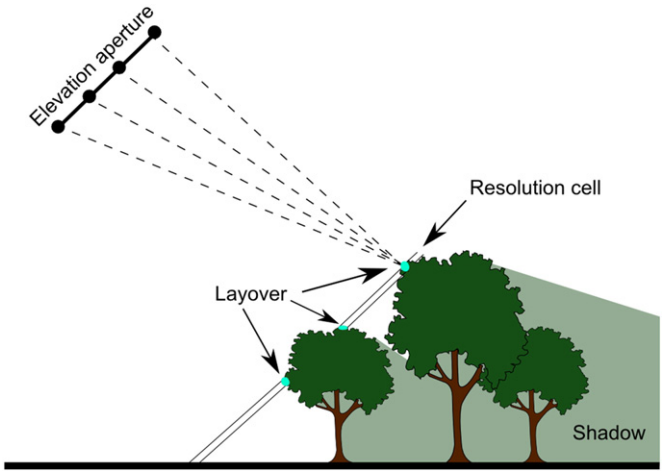

Fig. 1. Side-looking SAR viewing geometry for multi-baseline system illuminating a forest scene. While several scattering components are overlayed in one resolution cell, some trees are fully or partially shadowed by larger trees in front of them.

2.2. Layover separation and height reconstruction

A flowchart of the employed TomoSAR algorithm is shown in Fig. 2, whereas for the scope of this paper only the part enclosed by the rectangle is described. For the remaining pre- and postprocessing steps, the reader is referred to Schmitt and Stilla (2014b). Therefore, here the starting situation is the sample coherence matrix for every pixel in the stack of coregistered SAR images, as well as an estimate $\hat{K}$ of the model order *K*, i.e. of the expected number of scattering contributions in the sparse reflectivity profile. The sample covariance matrix is the

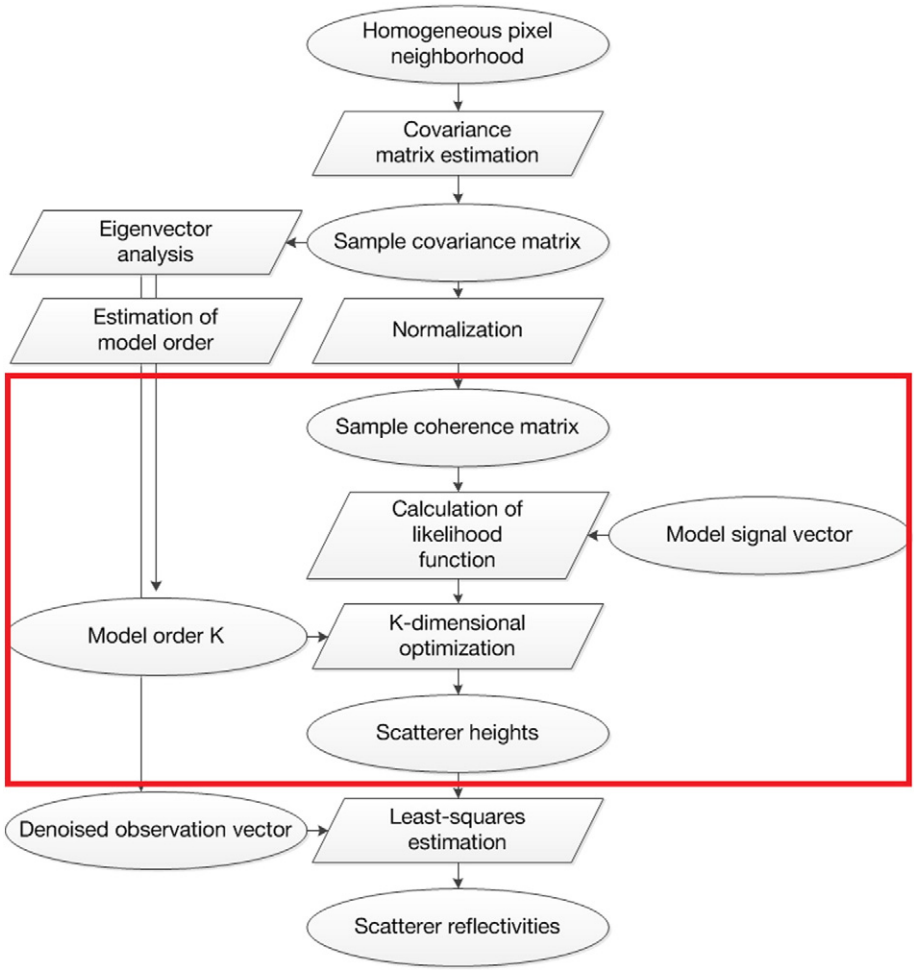

Fig. 2. Flowchart of the TomoSAR algorithm used in the presented work.

result of normalizing the observed sample covariance matrix and is defined as

$$\hat{\boldsymbol{\Gamma}} = \begin{bmatrix} 1 & |\hat{\gamma}_{12}|exp\left(j\hat{\phi}_{12}\right) & \dots & |\hat{\gamma}_{1N}|exp\left(j\hat{\phi}_{1N}\right) \\ |\hat{\gamma}_{21}|exp\left(j\hat{\phi}_{21}\right) & 1 & \dots & |\hat{\gamma}_{2N}|exp\left(j\hat{\phi}_{2N}\right) \\ \vdots & \vdots & \ddots & \vdots \\ |\hat{\gamma}_{N1}|exp\left(j\hat{\phi}_{N1}\right) & |\hat{\gamma}_{N2}|exp\left(j\hat{\phi}_{N2}\right) & \dots & 1 \end{bmatrix}, \quad (3)$$

which means it contains all the interferometric, i.e. phase-related information of the respective pixel: $|\hat{\gamma}|_{ij}$ is the magnitude of coherence between acquisitions i and j, while $\hat{\phi}_{ij}$ is the corresponding interferometric phase. Based on the assumption of circularly-symmetric zero-mean Gaussian SAR pixels, this sample coherence matrix can be used as input to the likelihood function

$$\mathcal{L}\left(\hat{\boldsymbol{\Gamma}}; \tilde{\mathbf{z}}(\mathbf{s})\right) = \frac{1}{\pi^N det\left(\hat{\boldsymbol{\Gamma}}\right)} exp\left(-\tilde{\mathbf{z}}^H(\mathbf{s})\hat{\boldsymbol{\Gamma}}^{-1}\tilde{\mathbf{z}}(\mathbf{s})\right), \quad (4)$$

where $\tilde{\mathbf{z}}(\mathbf{s})$ resembles the steering vector described in Section 2.1, but contains a hypothetical mixture of $\hat{K}$ scattering contributions:

$$\tilde{\mathbf{z}}(\mathbf{s}) = [\tilde{z}_1(\mathbf{s}), \dots, \tilde{z}_N(\mathbf{s})]^T, \quad (5)$$

where

$$\tilde{z}_n(\mathbf{s}) = \sum_{k=1}^{\hat{K}} exp(j\tilde{\varphi}_n(s_k)) \quad (6)$$

and $\tilde{\varphi}_n(s_k)$ is the absolute signal phase caused by scatterer k at elevation s_k minus any "flat earth" phase component potentially measured for the reference surface at elevation $s = 0$. Using the elevation-to-height conversion $h = s \cdot sin(\theta)$, where θ is the off-nadir angle of the radar line-of-sight, Eq. (4) can be maximized with respect to the unknown scatterer heights $\hat{\mathbf{h}} = \left[\hat{h}_1, \dots, \hat{h}_{\hat{K}}\right]$. Since the denominator of Eq. (4) stays constant during the optimization, the simplified estimator finally becomes:

$$\hat{\mathbf{h}} = \underset{\mathbf{h}\in[h_{min};h_{max}]}{\arg\max}\, exp\left(-\tilde{\mathbf{z}}(\mathbf{h})^H\hat{\boldsymbol{\Gamma}}^{-1}\tilde{\mathbf{z}}(\mathbf{h})\right). \quad (7)$$

This maximization can be solved by a simple $\hat{K}$-dimensional grid search: For this, the objective function in Eq. (7) is calculated for $\hat{K}$ hypothetical scattering heights taken from a the pre-defined discretized search interval $[h_{min}; h_{max}]$. Then the maximum is found in the $\hat{K}$-dimensional search space aggregated from the $\hat{K}$ search intervals.

2.3. Geocoding

The result of the TomoSAR inversion described in Section 2.2 is a three-dimensional point cloud in radar geometry, i.e. the data are still

stored in the coordinate system of the SAR imagery, although now for every resolution cell $\hat{K}$ scattering heights are available. This point cloud must then be transformed from the so-called slant range plane to a geodetic reference frame. For this task, the non-linear equation system

$$R - \|\mathbf{p} - \mathbf{s}_m(t)\| = 0 \tag{8}$$

$$\mathbf{v}_M(\mathbf{p} - \mathbf{s}_M(t)) = 0 \tag{9}$$

$$\frac{p_X^2 + p_Y^2}{(a+h)^2} + \frac{p_Z^2}{(b+h)^2} - 1 = 0 \tag{10}$$

has to be solved for the unknown 3D-coordinates $\mathbf{p} = [p_X, p_Y, p_Z]^T$ in world geometry (Schwäbisch, 1998). In this context, R is the slant range distance between $\mathbf{p}$ and the master antenna position $\mathbf{s}_M(t) = [s_{MX}(t), s_{MY}(t), s_{MZ}(t)]^T$ at time t. It has to be mentioned that the sensor velocity $\mathbf{v}_M = [v_{MX}, v_{MY}, v_{MZ}]^T$ in Eq. (9) is modeled as constant and referring to a linearized flight trajectory, while the data are assumed to be zero-Doppler processed. Although this is a simplification not valid for arbitrary mission designs, the extension to non-linear trajectories or non-zero-Doppler processed data is straight-forward. a and b in Eq. (10) denote the semi-major and semi-minor axes of the reference Earth ellipsoid, respectively, while h is the previously reconstructed height of the corresponding scatterer above the ellipsoid's surface.

2.4. Fusion of point clouds from multiple aspects

As described in Section 2.1, the side-looking SAR imaging geometry causes not only layover, but also shadowing effects that lead to scene parts that don't contain any exploitable measurements behind elevated objects. For the analysis of urban areas, multi-aspect data fusion has been demonstrated as a viable solution for this problem (e.g. Schmitt, 2015). Therefore, the same approach is used for fusing point clouds of multiple viewing directions in the context of this work. If several independent point clouds of the forested scene under investigation can be generated by the procedure described in the last sections, they may need to either be registered via their corresponding flight navigation data a priori (Schmitt, Maksymiuk, Magnard, & Stilla, 2013) or using point cloud-based registration methods *a posteriori* (e.g. Makadia, A. P., & Daniilidis, 2006; or Gernhardt, Cong, Eineder, Hinz, & Bamler, 2012). If the data then are aligned in the same common reference system this way, an additional fusion step was proved to be beneficial: By combining the individual point clouds intelligently, not only the accuracy of the 3D data is improved, but also the number of points in the point cloud can be significantly reduced without any information loss. This greatly benefits any further processing due to reduction of computational costs and memory requirements (Schmitt, 2015). This fusion step can be implemented by a voxel-space-based strategy: In order to impose a voxel-space onto the registered, unstructured 3D point clouds, a regular point cloud containing the centers of the desired voxel space has to be created. This is done via a k-d tree of dimension $k = 3$. A k-d tree is a multidimensional binary search tree, which serves as a space-partitioning data structure for storage of information to be retrieved by associative searches in a k -dimensional space (Bentley, 1975): If data are represented as a k-d tree, then each data point is stored as a node in the tree. Every non-leaf node implicitly generates a splitting hyperplane that divides the space into two parts (i.e. half-spaces). Points to the left of this hyperplane are then represented by the left subtree of that node, and points on the right are represented by the right subtree. Since every node in the tree is associated with one of the k dimensions the hyperplane is chosen perpendicular to that dimension's axis. Each split can be denoted using the dimension number and split value, whereas the splits are arranged in order to balance the tree, i.e. its maximum depth is kept as small as possible. If points are queried, the k-d tree search first locates the respective point in its appropriate node and then searches nearby leaves in the tree until it can guarantee that the correct point has been found. Due to their efficient storage structure, k-d trees are particularly useful for multi-dimensional search operations such as range searches or nearest neighbor searches. Besides, an abundance of ready-to-use implementations in different programming languages is available online, many of them open source.

After creating the voxel-space defining point cloud this way, the irregularly sampled 3D TomoSAR point clouds are examined with the established k-d tree by applying a range search based on the Chebychev distance. This distance metric is defined for two points $\mathbf{p} = [p_1, \ldots, p_n]^T$ and $\mathbf{q} = [q_1, \ldots, q_n]^T$ as

$$\lim_{k\to\infty}\left(\sum_{i=1}^{n}|p_i - q_i|^k\right)^{\frac{1}{k}} = \max_{i\in\{1;\ldots;n\}}(|p_i - q_i|), \tag{11}$$

which means it defines the distance between $\mathbf{p}$ and $\mathbf{q}$ as the greatest of their differences along any coordinate dimension. For three-dimensional vectors and a Chebyshev distance of $\frac{d}{2}$, this corresponds to a cube with an edge length of d. Therefore, the distance query results in a list of voxels and the corresponding points from the unstructured point clouds contained in that voxel (see Fig. 3). The final point cloud fusion is then realized by returning the mean value of the coordinates of all points contained in each voxel. The core parameter of this procedure is the edge length d of the voxel cubes, which basically defines both the resolution of the fused point cloud as well as the degree of data and noise reduction. As has been shown in Schmitt (2015), while the combination of multi-aspect data already helps to fill previously shadowed scene parts, this fusion strategy additionally provides a measurable improvement of the 3D accuracy of the reconstructed 3D points, as well as a significant reduction of the number of points.

3. Segmentation and reconstruction of individual trees

3.1. Point cloud segmentation

The 3D points resulting from the TomoSAR processing explained in Section 2 are clustered by the mean shift algorithm as described by

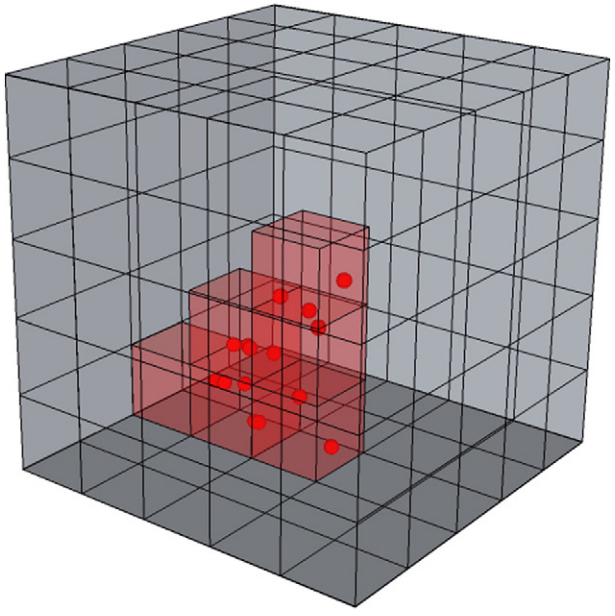

Fig. 3. Discretization of the scene by imposing a voxel space. The red cubes indicate voxels containing at least one 3D TomoSAR point.

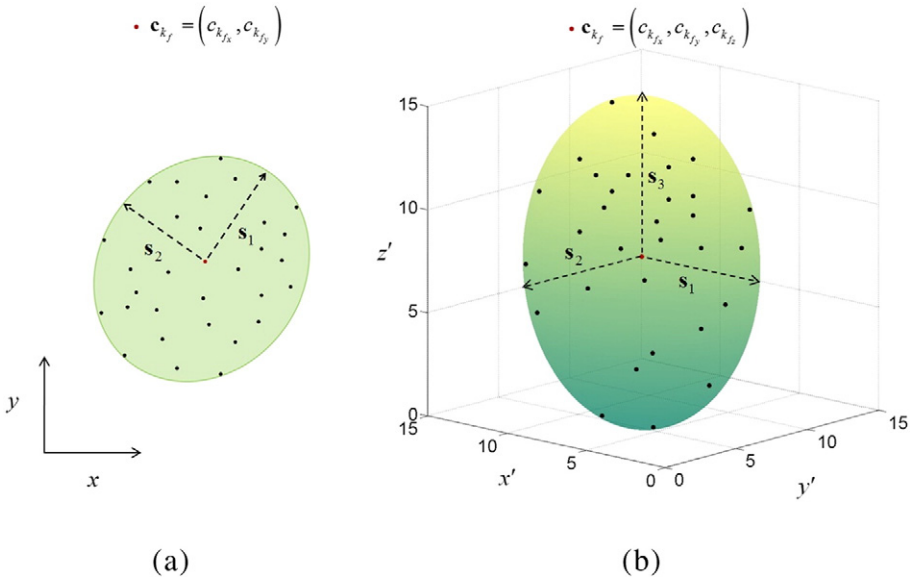

Fig. 4. Illustration of the ellipsoid modeling: (a) MVEE computed using 3D points projected onto the *xy*-plane; $\mathbf{s}_1$ and $\mathbf{s}_2$ are the computed semi-axes of the MVEE. (b) The MVEE of (a) is extruded in *z*-direction both upwards and downwards forming a 3D ellipsoid with a third semi-axis denoted as $\mathbf{s}_3$. *x*′, *y*′ and *z*′ in (b) represent axes of the local coordinate system aligned to the ellipsoid semi-axes. The red points in both (a) and (b) represent the ellipsoid centers $\mathbf{c}_{k_f}$.

Comaniciu and Meer (2002) in order to extract individual trees. Since the tree crowns generally show a comparably high point density when projected into the *xy*-plane, the points are clustered in 2D space, i.e. the feature space consists of *xy*-coordinates only. The kernel density estimate at any point $\mathbf{p}_i$ of the *n* points is given by the expression

$$D_{\mathbf{p}_i} = \frac{c}{nb^3}\sum_{j=1}^{n} g\left(\left\|\frac{\mathbf{p}_i - \mathbf{p}_j}{b}\right\|^2\right), \tag{12}$$

where *b* is the bandwidth parameter and $g(x)$ is a non-negative, non-increasing, piecewise continuous function with definite integral, i.e. $\int_0^\infty g(x)\mathrm{d}\,x < \infty$. Based on the concept of kernels discussed by Cheng (1995) as well as Comaniciu and Meer (2002), the function $g(x)$ is defined as the profile of the radially symmetric kernel $G(x)$ satisfying

$$G(x) = cg\left(\|x\|^2\right), \tag{13}$$

where *c* is a normalization constant ensuring that $G(x)$ integrates to 1. Different kernels, such as the Epanechnikov kernel and the Gaussian kernel can be used to define the density $D_{\mathbf{p}_i}$. Mean shift clustering essentially seeks modes of the kernel density estimates and works iteratively by shifting every data point toward the weighted mean of points within its neighborhood (defined to be cylindrical in the presented case). The shift vector $\mathbf{m}(\mathbf{p}_i)$ always points toward the direction of the maximum increase in the density $D_{\mathbf{p}_i}$ and is computed as

$$\mathbf{m}(\mathbf{p}_i) = \frac{\sum_{j=1}^{n}\mathbf{p}_j exp\left(-\frac{\|\mathbf{p}_i - \mathbf{p}_j\|^2}{b^2}\right)}{\sum_{j=1}^{n} exp\left(-\frac{\|\mathbf{p}_i - \mathbf{p}_j\|^2}{b^2}\right)} - \mathbf{p}_i. \tag{14}$$

The iteration process continues until there is no or only little shift in $\mathbf{m}(\mathbf{p}_i)$ anymore, i.e. the length of the shift vector $\mathbf{m}(\mathbf{p}_i)$ is close to 0. Due to the gradient ascent nature, the mean shift algorithm returns clusters using the concept *attraction of basin*, i.e. those points whose trajectories lead to the same mode form the basin of attraction for that mode and are clustered into one group. The clustering procedure is repeated until all points are assigned to their respective modes.

Clustering via mean shift is a non-parametric procedure in the sense that it does not require the number of clusters a priori, nor does it need any pre-defined model for the shape of the resulting clusters. Nevertheless, it still does require a bandwidth parameter (corresponding to the radius of the kernel), which affects the number of clusters, i.e. the number of modes, that are returned by the algorithm. However, unlike other clustering algorithms such as *k*-means, fuzzy c-means, expectation maximization etc., the bandwidth parameter has some physical meaning and can be set based on prior knowledge such as the expected average diameter of the tree crowns in the scene.

3.2. Ellipsoid modeling

Once the clustering is done, the individual tree clusters are modeled in three dimensions using generalized tri-axial ellipsoids that are aligned to the *z*-axis. For this purpose, parameters of an arbitrarily oriented minimum volume enclosing ellipse (MVEE) are estimated by first projecting points belonging to individual tree clusters onto the

Table 1
MEMPHIS sensor parameters used during the 2013 measurement campaign.

Sensor	MEMPHIS
Carrier frequency	35 GHz (Ka-band)
Wavelength	8.5 mm
Range bandwidth	900 MHz
Nominal depression angle	35°
Resolution	
Azimuth	8.2 cm
Range	16.7 cm
Pixel spacing	
Azimuth	5.3 cm
Range	16.7 cm
Available baselines	5.5 cm, 11 cm, 16.5 cm, 22 cm, 27.5 cm
Approximate ambiguity heights	180 m, 90 m, 60 m, 45 m, 36 m

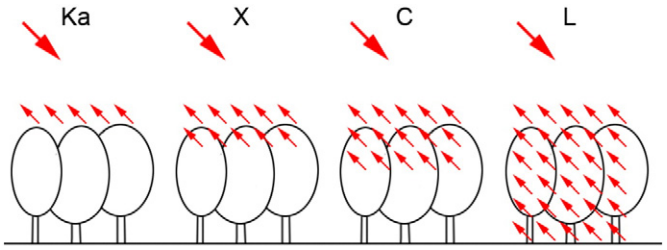

Fig. 5. Expected canopy penetration of common SAR wavelengths.

xy-plane followed by extruding the 2D *xy*-ellipse in *z*-direction to form a 3D ellipsoid. The motivation for expanding the ellipsoid along the *z*-axis is based on geometrical considerations: It is assumed that correct tree models may have an arbitrary orientation in the *xy*-plane, but remain upright or vertical with respect to the ground (cf. Fig. 4). This is based on the assumption that tree trunks are vertical to the ground surface.

3.2.1. Computation of the MVEE

If $\mathbf{K} = \{k_i | i = 1, \ldots, m\}$ denotes m clusters returned by the mean shift algorithm, and $\mathbf{Q} = \{\mathbf{q}_u | u = 1, \ldots, r\}$ denotes the set of r points $\mathbf{q}_u$ belonging to a particular cluster k_f ($f \in i$), then any arbitrarily oriented ellipse ε can be a candidate for the MVEE($\mathbf{Q}$), if and only if all points in $\mathbf{Q}$ lie on or inside its boundary, i.e. if the following condition is satisfied (Kumar & Yildirim, 2005):

$$\left(\mathbf{q}_u - \mathbf{c}_{k_f}\right)^T \mathbf{A} \left(\mathbf{q}_u - \mathbf{c}_{k_f}\right) \leq 1 \quad \text{for } u = 1, \ldots, m. \tag{15}$$

In this equation, $\mathbf{A}$ is a $d \times d$ positive definite matrix, where $d = 2$ in the presented case, and $\mathbf{c}_{k_f}$ is the center of the ellipse surrounding the clustered points $\mathbf{Q}$. The semi-axes $\mathbf{s}_i$ of such an ellipse are given as

$$\mathbf{s}_i = \lambda_i^{-\frac{1}{2}} \mathbf{v}_i, \tag{16}$$

where $\mathbf{v}_i$ denote the eigenvectors of $\mathbf{A}$, which correspond to the directions of the semi-axes. λ_i denotes the eigenvalues of $\mathbf{A}$, which are related to the length of these axes: The length of each axis is equal to $\frac{1}{\sqrt{\lambda_i}}$. The area of an ellipse or volume of an ellipsoid, respectively, is thus directly proportional to $det\left(\frac{1}{\sqrt{\mathbf{A}}}\right)$.

Therefore, in order to obtain an MVEE($\mathbf{Q}$), $det\left(\frac{1}{\sqrt{\mathbf{A}}}\right)$ has to be minimized such that Eq. (15) is satisfied in conjunction with $\mathbf{A}$ being positive definite. In order to solve this minimization, Khachyan's first order algorithm is used, which formulates the problem as optimization using Lagrangian duality (Khachiyan, 1996).

The computed MVEE($\mathbf{Q}$) is extended to the third dimension by extruding it in *z*-axis in order to form a 3D ellipsoid. The *z*-coordinate of the ellipsoid center and its semi-axis length s_3 in *z*-direction are estimated by

$$\mathbf{c}_{k_{f_z}} = \text{median}\left(h_{min,k_f,i}\right) + s_3, \tag{17}$$

where

$$s_3 = \frac{1}{2}\left(\text{median}\left(h_{max,k_f,i}\right) - \text{median}\left(h_{min,k_f,i}\right)\right). \tag{18}$$

In this context, $h_{min,k_f,i}$ and $h_{max,k_f,i}$ ($i = 1, \ldots, N$) are the N lowest heights and the N largest heights of all points in the cluster k_f, respectively.

Once this modeling is complete, the tree parameters *tree height*, *crown diameter*, and *trunk location* can directly be extracted from the ellipsoid model: The tree height is the maximum height of the ellipsoid in *z*-direction, the tree crown radii are calculated by taking the geometric mean of the *x*-and *y*-semi-axes of the ellipsoid, and the *xy*-coordinates of the ellipsoid center point provide the location of the tree trunk. Of course, this is a simplifying model only valid for deciduous trees of approximately ellipsoidal shape, but an extension toward a more general tree model as, e.g., described by Sheng, Gong, and Biging (2001) basically seems possible.

4. Utilized test data

4.1. Airborne millimeterwave sensor

The sensor used for the experimental considerations in this paper is the German MEMPHIS system created by the Fraunhofer Institute for High Frequency Physics and Radar Techniques (Schimpf, Essen, Boehmsdorff, & Brehm, 2002). Of its different modes and configurations, in this work only the basic airborne side-looking configuration with a carrier frequency of 35 GHz (Ka-band) and a bandwidth of 900 MHz is employed. This leads to a slant range resolution of 16.7 cm, whereas the azimuth resolution is 8.2 cm. Since MEMPHIS is equipped with four receiving antennas, it is able to provide multi-baseline InSAR datasets from just a single pass over the scene of interest. Additional parameters of the sensor can be found in Table 1.

For the task of tree reconstruction, this system configuration provides two key advantages: First, it is a single-pass InSAR system, which means the four receiving antennas acquire four images of the same scene simultaneously, thus leading to highly coherent data even for vegetation areas. Second, the system uses radar signals of the millimeterwave domain, which can be expected to provide much less canopy penetration than longer wavelengths (cf. Fig. 5). Therefore, it is possible to use the TomoSAR model based on sparse reflectivity profiles as described in Section 2.

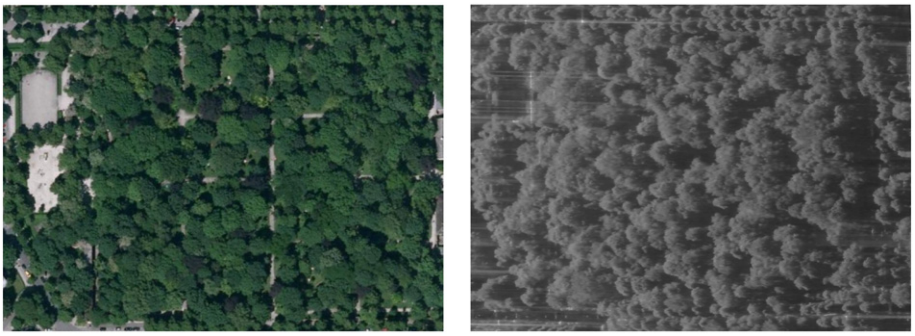

Fig. 6. Test scene "Nordfriedhof" in Munich, Germany: (a) Optical image, (b) MEMPHIS intensity image, range direction from left to right.

Table 2
Parameters of the MEMPHIS test data used in the experiments.

Aspect 1	
Flying altitude	760 m
Heading angle	20°
Number of reconstructed points	1.78 million
Aspect 2	
Flying altitude	760 m
Heading angle	200°
Number of reconstructed points	1.89 million

Table 3
Statistics of the reference trees.

Number of trees	570
Tree crown radii	
Min/max	0.39 m/10.09 m
Mean/median	3.48 m/3.25 m
Tree heights	
Min/max	2.81 m/26.97 m
Mean/median	15.37 m/15.43 m

4.2. Test scene

The available experimental MEMPHIS data were acquired during a campaign over Munich, Germany, in June 2013. The test scene contains the "Alter Nordfriedhof", an abandoned cemetery, which is used as a public park today, with the target coordinates 48°09′13″N, 11°34′13″ E. As can be seen in Fig. 6(a), this approximately 5 ha large areal is characterized by a light planting of deciduous trees (mainly beeches), resembling a grove or little wood. A corresponding SAR intensity image is shown in Fig. 6(b). For the investigations in this paper, multi-baseline InSAR data acquired from two opposing aspects (flight heading angles 20° and 200°, respectively) were processed to a fused 3D point cloud as described in Section 2. The spacing of the height search interval of the TomoSAR optimization step as well as the edge length of the voxel space used for point cloud fusion were both chosen as 0.5 m. A summary of the data parameters can be found in Table 2. In order to emphasize the scope of this study, which is to prove the capability of millimeterwave SAR for individual tree parameter reconstruction rather than tree detection, non-tree points are manually removed from the dataset before any further processing.

5. Experiments and reconstruction results

The results of the experimental assessment of the proposed tree reconstruction procedure are summarized in the following sections. First the results of the segmentation are shown in order to provide an evaluation of the effectiveness of unsupervised mean shift clustering for 3D TomoSAR point clouds. In the second part, the results of the ellipsoid fitting process including the derivation of tree parameters are compiled.

For evaluation purposes, the results have been compared to a manually generated reference dataset that is displayed in Fig. 7. This reference dataset is based on the combination of a helicopter-borne LiDAR point cloud containing approximately 0.16 million points (i.e. 3 points/m^2) and a co-registered orthophoto. From these data, the reference trees were extracted by a human operator as follows:

- Identification of individual trees by simultaneous visual comparison of LiDAR point cloud and orthophoto

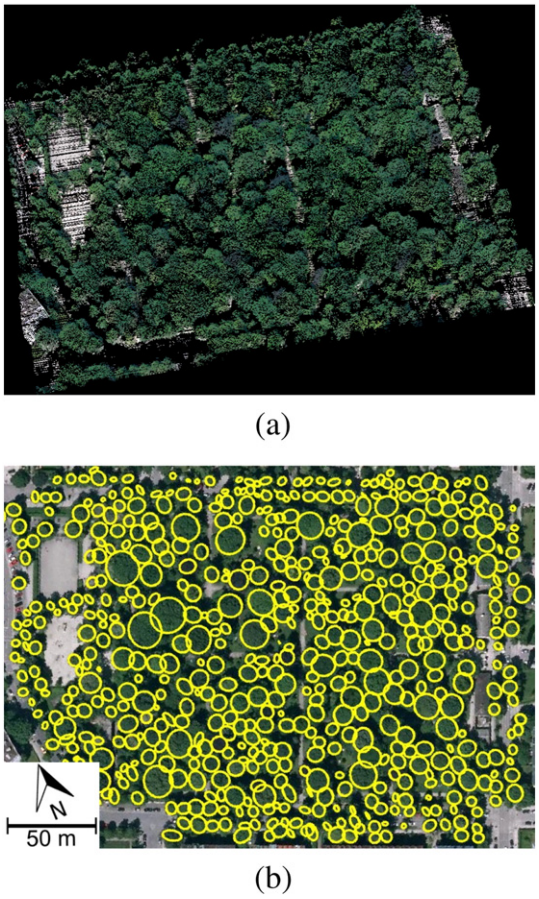

Fig. 7. Reference data of the test scene, created from a LiDAR point cloud and a co-registered orthophoto. (a) Texturized LiDAR point cloud. (b) Every circle indicates one manually extracted reference tree.

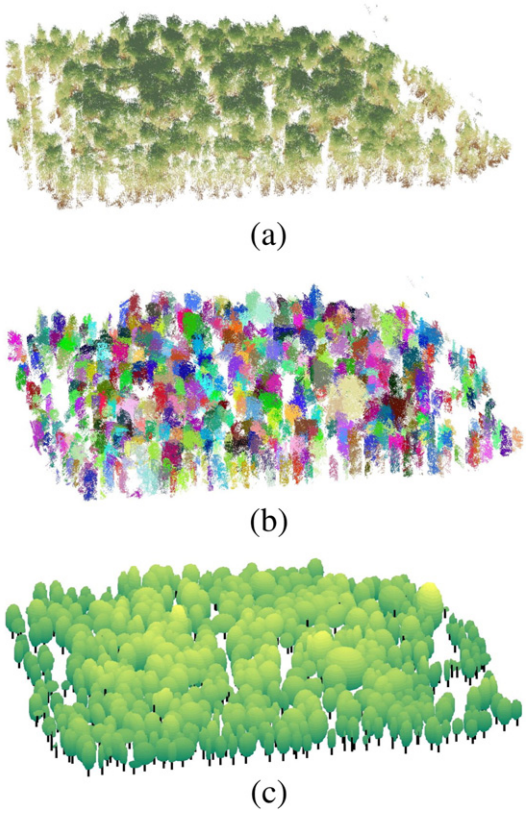

Fig. 8. The scene shown in different processing stages: (a) The 3D point cloud as derived by multi-aspect TomoSAR data fusion; (b) the clustered point cloud; (c) the reconstructed tree models.

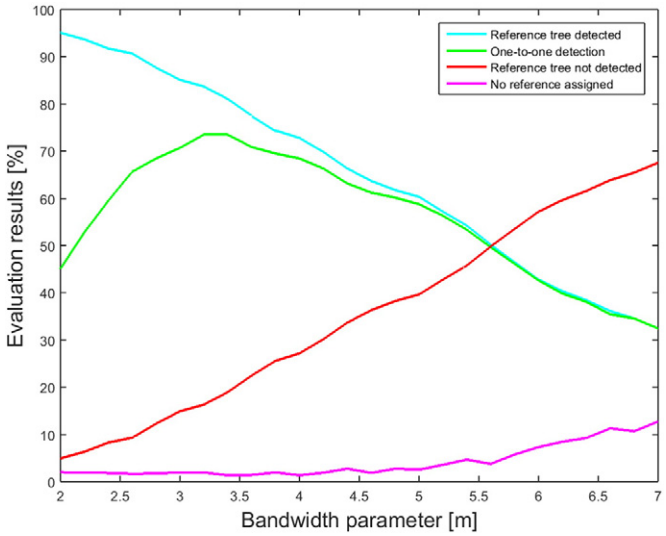

Fig. 9. Determination of the optimal bandwidth parameter by analysis of tree segmentation accuracy for all 570 reference trees. Green indicates perfect matches of one cluster to one reference tree, red missed reference trees, and purple tree clusters that cannot be assigned to any reference tree. In addition, cyan summarizes the overall detected trees, including oversegmented trees.

- Manual measurement of approximate tree crown diameters
- Determination of highest LiDAR point of each tree crown
- Parameterization of each tree by a circle, plus LiDAR-derived tree height
- Modeling of ellipsoids to each tree using the method described in Section 3.2, using only the LiDAR points located within these preliminary tree circles

This way, a reliable set of reference tree models was created, consisting of reference tree locations, heights, and tree crown radii. The statistics of this reference dataset are compiled in Table 3.

5.1. Point cloud generation results

The original TomoSAR point cloud as derived from the test dataset by the method described in Section 2 is displayed in Fig. 8(a). After application of the voxel-space-based fusion procedure (applied with a voxel edge length of 0.5 m), the number of points in the fused point cloud is approximately 1.66 million (corresponding to a point density of about 22 points/m^2). That means that due to redundant observations

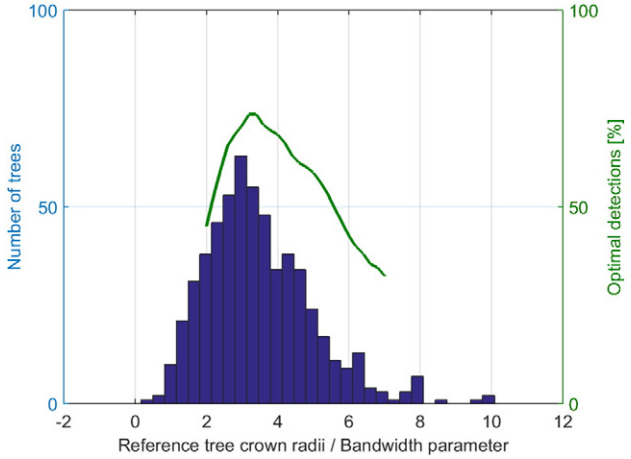

Fig. 10. Correlation of bandwidth parameter and distribution of reference tree radii.

Table 4
Statistics of the tree segmentation for the optimal bandwidth parameter.

Measure	Value
Producer accuracy	73.5%
User accuracy	74.0%
Commission error	1.9%
Omission error	16.3%

and a slight reduction of the spatial resolution, a significant reduction of the data volume (considering more than 3.5 million points existing in the original point clouds) has been achieved. Compared to the LiDAR point cloud mentioned above, the 3D accuracy of the TomoSAR point cloud lies in the sub-meter domain, depending on the evaluation strategy. It has to be mentioned, however, that any evaluation of a TomoSAR point cloud with respect to LiDAR reference data only gives a rough estimate for the 3D accuracy level due to strongly different point densities. Apart from this fact, the accuracy can always further be improved either from the data side, i.e. by providing additional aspects, or from the processing side, i.e. by using a finer spacing during TomoSAR height reconstruction and voxel-space-based fusion.

5.2. Optimal bandwidth parameter selection

Since the only parameter of the mean shift clustering is the bandwidth that is used for generating the clustering kernel, the first set of experiments aimed at determining the optimal bandwidth parameter. Fig. 9 shows a summary of the segmentation results for all 570 reference trees by distinguishing four potential cases: Of course, the desired result of the clustering would be that each reference tree corresponds to one particular cluster, which is referred to as one-to-one detection of this tree. Together with all reference trees, which are oversegmented, i.e. more than one cluster is assigned to them, this adds to the overall rate of detected trees. Then there is the class of missed trees, i.e. reference trees that were not detected as no cluster center falls within their 2D crown outline. Finally, some clusters cannot be assigned to any reference tree at all, which is often referred to as false positives. In the context of this study, it is important to mention that all remaining evaluations are based on one-to-one detections only, whereas oversegmented detections are counted as non-detected trees in order to provide a fair assessment. Analyzing Fig. 9, it becomes obvious that the optimal bandwidth parameter is about 3.2 m, giving an optimal detection rate of 73.5% of the trees, plus oversegmented detections at the rate of 10.2%. Thus, in total 83.7% of all reference trees are discovered, only 16.3% are missed, and 1.9% are false positives.

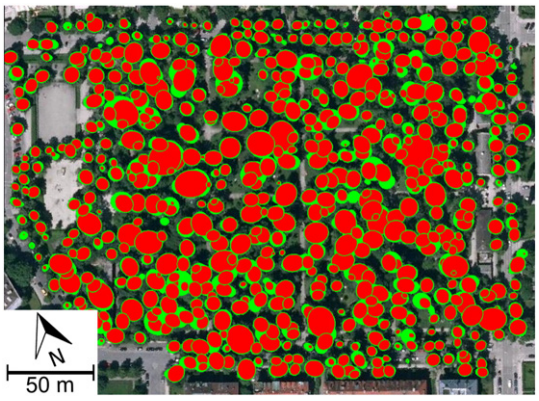

Fig. 11. Ellipsoid models projected onto the 2D reference dataset for one-to-one comparison. The green circles indicate the reference trees, the circles filled in red correspond to the reconstructed trees.

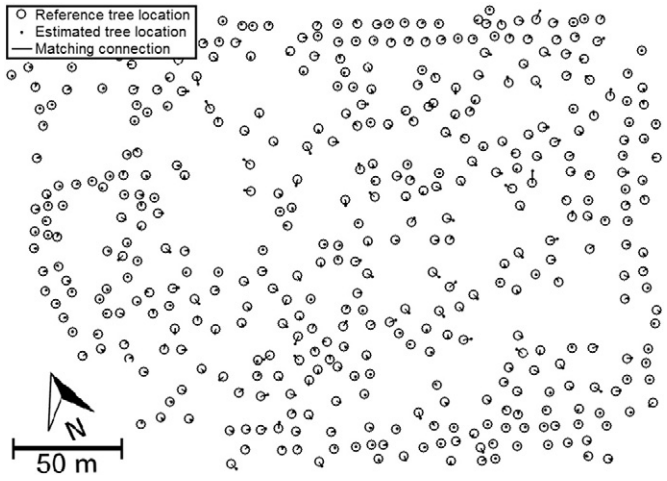

Fig. 12. Localization errors of reconstructed trees and reference trees.

In this context, it is interesting to note that the median tree radius of the reference trees is 3.25 m. The correlation between tree radius distribution and bandwidth parameter is further illustrated in Fig. 10. It can nicely be seen that the optimal bandwidth parameter for the mean shift clustering approximately corresponds to the peak of the tree radius histogram. That means that only light prior knowledge about the expected tree radii of the scene of interest is sufficient to tune the clustering process, while keeping it otherwise fully unsupervised.

5.3. Segmentation results

The result of the mean shift clustering of the point cloud with optimal bandwidth parameter of 3.2 m is displayed in Fig. 8(b). The points have been segmented into 566 clusters, which already resemble individual trees by visual impression. A statistical evaluation of the segmentation based on the one-to-one detections only is summarized in Table 4.

5.4. Ellipsoid modeling results

The final result of the ellipsoid modeling process can be assessed in Fig. 8(c), including tree crowns of different shape and hypothetical stem positions. A projection of the ellipsoids onto the 2D reference data is shown in Fig. 11, while the reference tree positions and the estimated tree locations corresponding to the ellipsoid centers projected into the *xy*-plane are opposed in Fig. 12. A summary of the tree parameter reconstruction errors is given in Table 5. In addition, the error distributions for tree heights and crown radii are shown in Fig. 13. Again it has to be noted that for this quantitative analysis only the optimally segmented trees are used, whereas the oversegmented ones are discarded.

6. Discussion

The experiments presented in this paper show a variety of things:

- It is possible to generate 3D point clouds of forested areas from airborne multi-aspect TomoSAR data if single-pass interferometers and millimeterwaves are employed.
- Depending on the sampling sizes of the SAR pixels, the TomoSAR height reconstruction step and the voxel-space-based fusion, 3D accuracies of the TomoSAR points reach the decimeter domain at very high density.
- Unsupervised clustering of such a point cloud is possible by mean shift clustering if an expectation value of the typical tree crown radii in the investigated scene is available.
- Ellipsoids can be used to model the shapes of deciduous trees, thus providing approximate estimates for core tree parameters such as location, height and diameter.

Table 5
Error statistics of the reconstructed tree parameters.

Errors	Mean	Standard deviation
Height	0.93 m	1.92 m
Radius	0.28 m	0.96 m
x-Localization	0.50 m	1.05 m
y-Localization	−0.29 m	1.20 m

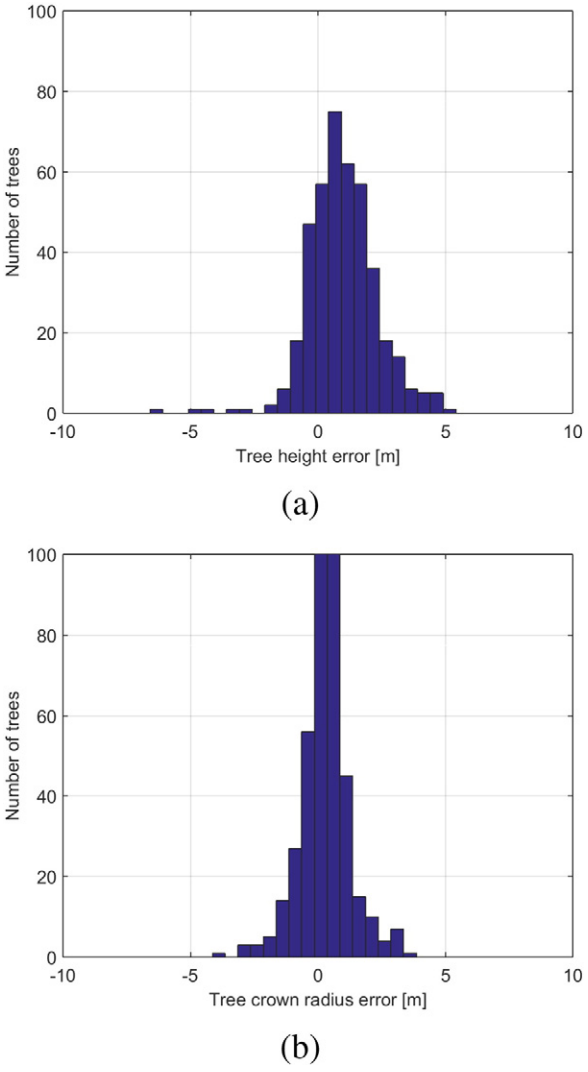

Fig. 13. Distributions of (a) the tree height errors and (b) the crown radii errors. It can be seen that both tree heights and crown radii tend to be slightly overestimated.

Although the results of this study are already very promising, there is still room for further improvement. First of all, it is obvious that the clustering is highly dependent on proper tuning of the bandwidth parameter. Although this can be handled by only light prior knowledge, an adaptive setting of the bandwidth parameter could possibly enhance the segmentation accuracy, in particular concerning oversegmented trees. Secondly, the ellipsoid model of course is only a coarse approximation of real-life tree crowns, and only useful for deciduous trees at that. Here, e.g. a generalized ellipsoid model also accounting for varying crown curvature could help to create a more universal approach and more detailed results. In addition, a more robust estimation of the tree

heights and the crown radii is expected to reduce the over-estimation bias in these parameters significantly.

Concerning the number of missed trees, there is unfortunately always the sensor-inherent limitation: If a small tree is surrounded by large trees on all sides, not even multi-aspect SAR data will help to avoid missing that tree due to the side-looking nature of the SAR imaging process. In such a case, only approaches based on volume tomography might provide a viable solution.

Last but not least, it has to be mentioned that the reference data also provides some potential for erroneous modeling, since no analysis of any kind of data can replace in-situ observations. For example, the smallest tree in the reference data is only about 80 cm wide, i.e. in a real ground truth dataset, it would possibly not have been included at all.

7. Summary and conclusion

In this article, an unsupervised approach for segmentation and reconstruction of individual trees from multi-aspect millimeterwave TomoSAR data has been presented. Starting from at least two stacks of single-pass multi-baseline SAR images, first 3D point clouds are generated by TomoSAR height reconstruction. After geocoding, the point clouds derived from multiple aspects are fused to form a single homogeneous 3D point cloud mostly free from any layover or shadow parts. This point cloud is then segmented by unsupervised mean shift clustering, and for every cluster a three-dimensional ellipsoid is modeled to the contained points. Since these ellipsoids are supposed to serve as satisfying approximations of deciduous tree crowns, three important tree parameters are extracted from each ellipsoid: tree location, tree height and tree crown diameter. Experiments based on an airborne millimeterwave dataset of two opposing aspects acquired over a cemetery in the city of Munich, Germany, have shown that about 74% of all trees are properly segmented and reconstructed by the presented technology. Although the side-looking SAR imaging geometry serves as a system-inherent limitation and leads to the fact that particularly small trees fully surrounded by large trees will always be missed, it is expected that an adaptive selection of the kernel bandwidth during mean shift clustering will further enhance the results. This extension of the mean shift algorithm will be part of future research activities.

In any case, the results presented in this paper are expected to further stimulate the research interest in exploiting SAR imagery for forest remote sensing on the individual tree level.

Acknowledgments

The authors would like to thank Thorsten Brehm and Dr. Stephan Stanko (Fraunhofer Institute for High Frequency Physics and Radar Technology FHR) for providing the MEMPHIS test data. They would also like to acknowledge the effort of Christophe Magnard and Dr. Erich Meier (Remote Sensing Laboratories at the University of Zurich) had focusing the raw SAR data.

This work is supported by the Helmholtz Association under the framework of the Young Investigators Group “SiPEO” (VH-NG-1018, www.sipeo.bgu.tum.de) and the International Graduate School of Science and Engineering, Technische Universität München (Project 6.08: “4D City”).

References

Ahmed, R., Siqueira, P., Hensley, S., Chapman, B., & Bergen, K. (2011). A survey of temporal decorrelation from spaceborne L-band repeat-pass InSAR. *Remote Sensing of Environment*, *115*(11), 2887–2896.

Bentley, J. L. (1975). Multidimensional binary search trees used for associative searching. *Communications of the ACM*, *18*(9), 509–517.

Brenner, A. R. (2012). Ultra-high resolution airborne SAR imaging of vegetation and man-made objects based on 40. *Proceedings of IEEE International Geoscience and Remote Sensing Symposium* (pp. 7397–7400).

Chang, A., Eo, Y., Kim, Y., & Kim, Y. (2013). Identification of individual tree crowns from LiDAR data using a circle fitting algorithm with local maxima and minima filtering. *Remote Sensing Letters*, *4*(1), 30–38.

Chen, Q., Baldocchi, D., Gong, P., & Kelly, M. (2006). Isolating individual trees in a savanna woodland using small footprint LiDAR data. *Photogrammetric Engineering & Remote Sensing*, *72*(8), 923–932.

Cheng, Y. (1995). Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *17*(8), 790–799.

Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(5), 603–619.

Culvenor, D. S. (2002). TIDA: An algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. *Computers & Geosciences*, *28*(1), 33–44.

Englhart, S., Keuck, V., & Siegert, F. (2011). Aboveground biomass retrieval in tropical forests — The potential of combined X- and L-band SAR data use. *Remote Sensing of Environment*, *115*(5), 1260–1271.

Erikson, M. (2003). Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules. *Canadian Journal of Forest Research*, *33*(8), 1557–1563.

Fagan, M., & DeFries, R. (2009). Measurement and monitoring of the world's forests. A review and summary of remote sensing technical capability 2009–2015. *Tech. rep.* Resources for the Future.

Frey, O., Morsdorf, F., & Meier, E. (2008). Tomographic imaging of a forested area by airborne multi-baseline P-band SAR. *Sensors*, *8*(9), 5884–5896.

Gernhardt, S., Cong, X., Eineder, M., Hinz, S., & Bamler, R. (2012). Geometrical fusion of multitrack PS point clouds. *IEEE Geoscience and Remote Sensing Letters*, *9*(1), 38–42.

Gini, F., & Lombardini, F. (2005). Multibaseline cross-track SAR interferometry: A signal processing perspective. *IEEE Aerospace and Electronic Systems Magazine*, *20*(8), 71–93.

Gupta, S., Weinacker, H., & Koch, B. (2010). Comparative analysis of clustering-based approaches for 3-D single tree detection using airborne fullwave LiDAR data. *Remote Sensing*, *2*(4), 968–989.

Hoekman, D. H., & Varekamp, C. (2001). Observation of tropical rain forest trees by airborne high-resolution interferometric radar. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(3), 584–594.

Izzawati, Wallington, E. D., & Woodhouse, I. H. (2006). Forest height retrieval from commercial X-band SAR products. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(4), 863–870.

Jing, L., Hu, B., Li, J., & Noland, T. (2012). Automated delineation of individual tree crowns from LiDAR data by multi-scale analysis and segmentation. *Photogrammetric Engineering & Remote Sensing*, *78*(12), 1275–1284.

Khachiyan, L. (1996). Rounding of polytopes in the real number model of computation. *Mathematics of Operations Research*, *21*(2), 307–320.

Koch, B., Heyder, U., & Weinacker, H. (2006). Detection of individual tree crowns in airborne LiDAR data. *Photogrammetric Engineering & Remote Sensing*, *72*(4), 357–363.

Kumar, P., & Yildirim, E. (2005). Minimum volume enclosing ellipsoids and core sets. *Journal of Optimization Theory and Applications*, *126*(1), 1–21.

Le Toan, T., Beaudoin, A., Riom, J., & Guyon, D. (1992). Relating forest biomass to SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(2), 403–411.

Leckie, D. G., Gougeon, F. A., Tinis, S., Nelson, T., Burnett, C. N., & Paradine, D. (2005). Automated tree recognition in old growth conifer stands with high resolution digital imagery. *Remote Sensing of Environment*, *94*(3), 311–326.

Makadia, A., Patterson, A., & Daniilidis, K. (2006). Fully automatic registration of 3D point clouds. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 1297–1304).

Mittermayer, J., Wollstadt, S., Prats-Iraola, P., & Scheiber, R. (2014). The TerraSAR-X staring spotlight mode concept. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(6), 3695–3706.

Moreira, A., Krieger, G., Hajnsek, I., Hounam, D., Werner, M., Riegger, S., et al. (2004). TanDEM-X: A TerraSAR-X add-on satellite for single-pass SAR interferometry. *Proceedings of IEEE International Geoscience and Remote Sensing Symposium* (pp. 1000–1003).

Morsdorf, F., Meier, E., Kötz, B., Itten, K. I., Dobbertin, M., & Allgöwer, B. (2004). LiDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. *Remote Sensing of Environment*, *92*(3), 353–362.

Neumann, M., Ferro-Famil, L., & Reigber, A. (2010). Estimation of forest structure, ground, and canopy layer characteristics from multi-baseline polarimetric interferometric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(3), 1086–1104.

Perko, R., Raggam, H., Deutscher, J., Gutjahr, K., & Schardt, M. (2011). Forest assessment using high resolution SAR data in X-band. *Remote Sensing*, *3*(4), 792–815.

Pollock, R. J. (1996). *The automatic recognition of individual trees in aerial images of forests based on a synthetic tree crown image model.* Ph.D. thesis The University of British Columbia.

Pouliot, D., King, D., Bell, F., & Pitt, D. (2002). Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. *Remote Sensing of Environment*, *82*(2–3), 322–334.

Reigber, A., & Moreira, A. (2000). First demonstration of airborne SAR tomography using multi-baseline L-band data. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(5), 2142–2152.

Schimpf, H., Essen, H., Boehmsdorff, S., & Brehm, T. (2002). MEMPHIS — A fully polarimetric experimental radar. *Proceedings of IEEE International Geoscience and Remote Sensing Symposium* (pp. 1714–1716).

Schmitt, M. (2015). Three-dimensional reconstruction of urban areas by multi-aspect TomoSAR data fusion. *Proceedings of Joint Urban Remote Sensing Event* (On CD).

Schmitt, M., Brück, A., Schönberger, J., & Stilla, U. (2013). Potential of airborne single-pass millimeterwave InSAR data for individual tree recognition. *Proceedings of 33. Wissenschaftlich-Technische Jahrestagung der DGPF* (pp. 427–436).

Schmitt, M., Maksymiuk, O., Magnard, C., & Stilla, U. (2013). Radargrammetric registration of airborne multi-aspect SAR data of urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing, 86*, 11–20.

Schmitt, M., & Stilla, U. (2014 aa). Generating point clouds of forested areas from airborne millimeterwave InSAR data. *Proceedinigs of IEEE International Geoscience and Remote Sensing Symposium* (pp. 1–4).

Schmitt, M., & Stilla, U. (2014 bb). Maximum-likelihood-based approach for single-pass synthetic aperture radar tomography over urban areas. *IET Radar, Sonar & Navigation, 8*(9), 1145–1153.

Schwäbisch, M. (1998). A fast and efficient technique for SAR interferogram geocoding. *Proceedings of IEEE International Geoscience and Remote Sensing Symposium Proceedings* (pp. 1100–1102).

Shahzad, M., & Zhu, X. (2014). Reconstructing 2-D/3-D building shapes from spaceborne tomographic SAR point clouds. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 15*. (pp. 313–320).

Shahzad, M., & Zhu, X. (2015). Robust reconstruction of building facades for large areas using spaceborne TomoSAR point clouds. *IEEE Transactions on Geoscience and Remote Sensing, 53*(2), 752–769.

Sheng, Y., Gong, P., & Biging, G. S. (2001). Model-based conifer-crown surface reconstruction from high-resolution aerial images. *Photogrammetric Engineering & Remote Sensing, 67*(8), 957–965.

Shiver, B. D., & Borders, B. E. (1996). *Sampling techniques for forest resource inventory*. Hoboken, NJ: Wiley.

Straub, B. -M., & Heipke, C. (2001). Automatic extraction of trees for 3D-city models from images and height data. In E. Baltsavias, A. Gruen, & L. V. Gool (Eds.), *Automatic Extraction of Man-Made Objects from Aerial and Space Images, III*. (pp. 267–277).

Wulder, M., Niemann, K. O., & Goodenough, D. G. (2000). Local maximum filtering. *Remote Sensing of Environment, 73*(1), 103–114.

Zhu, X., & Bamler, R. (2010). Tomographic SAR inversion by L1-norm regularization — The compressive sensing approach. *IEEE Transactions on Geoscience and Remote Sensing, 48*(10), 3839–3846.