Building Height Retrieval From VHR SAR Imagery Based on an Iterative Simulation and Matching Technique

Dominik Brunner, *Student Member, IEEE*, Guido Lemoine, *Senior Member, IEEE*, Lorenzo Bruzzone, *Senior Member, IEEE*, and Harm Greidanus

Abstract—Experimental airborne synthetic aperture radar (SAR) systems achieve spatial resolutions of approximately 10 cm, whereas the new spaceborne very high spatial resolution (VHR) SAR sensors onboard the TerraSAR-X and COSMO-SkyMed satellites achieve spatial resolutions down to 1 m. In VHR SAR data, features from individual urban structures (i.e., buildings) can be identified by their characteristic settings in urban settlement patterns. In this paper, we present a novel concept for the height estimation of generic man-made structures from single detected SAR data. The proposed approach is based on the definition of a hypothesis on the height of the building and on the simulation of a SAR image for testing that hypothesis. A matching procedure is applied between the estimated and the actual SAR image in order to test the height hypothesis. The process is iterated for different height assumptions until the matching function is optimized, and thus, the building height is estimated. The efficiency of the proposed method is demonstrated on a set of 40 flat- and gable-roof buildings using two submeter VHR airborne and two 1-m resolution TerraSAR-X SAR scenes all acquired from the same residential area in Dorsten, Germany. The results show that, in the absence of string disturbing effects, the method is able to estimate the height of flat- and gable-roof buildings in the submeter data to the order of a meter, while the accuracy for the meter resolution spaceborne data is lower but still sufficient to estimate the number of floors of a building.

Index Terms—Building detection, damage assessment, height extraction, remote sensing, synthetic aperture radar (SAR), urban areas, very high geometrical resolution images, very high spatial resolution (VHR) SAR, 3-D reconstruction.

I. INTRODUCTION

RAPID situation assessment after natural disasters (e.g., earthquakes and tsunamis) and violent conflict events (e.g., war-related destruction) is crucial for initiating effective emergency response actions. Remote sensing satellites equipped with optical and synthetic aperture radar (SAR) imaging sensors can provide important information due to their abil-

ity to map affected areas of interest quickly and in a censorship-free manner.

Current spaceborne optical sensors, such as Ikonos, Quickbird, and WorldView, have meter and submeter spatial resolutions. These sensors fall into the passive optical system category and depend on sun illumination and cloud-free weather conditions to acquire useful imagery. In contrast, active SAR sensors can acquire imagery independently of illumination conditions and with a relative insensitivity to weather conditions. Until recently, spaceborne commercial SAR sensors were only capable of imaging the Earth surface with a spatial resolution no better than 9 m. This changed after the new very high spatial resolution (VHR) SAR sensors onboard the TerraSAR-X [1] and COSMO-SkyMed [2] satellites were launched in 2007, providing SAR imagery with spatial resolutions down to 1 m. In such imagery, features from individual urban structures, such as buildings, can be identified in their characteristic settings in urban settlement patterns (e.g., residential areas, city centers, and industrial parks). Current experimental airborne SAR systems even reach spatial resolutions of about 0.1 m [3].

Urban building detection provides an indirect measure for population density, which is an essential parameter in impact assessment that drives emergency response actions. Both the spatial extent of urbanized areas and the spatial characterization of building volume are crucial parameters to estimate affected population, estimate infrastructural damage, and enumerate economic losses resulting from the emergency event. Building volume is the product of spatial extent of a built-up structure and its height. Furthermore, the height of a building is a structural indicator about the status of a building after the event, e.g., whether it is still structurally intact. Hence, height determination of buildings is a key issue in postemergency event information retrieval in urban areas. Successful height characterization of buildings in VHR SAR data, therefore, will add substantial value to operational remote sensing applications in emergency response.

Several building height retrieval techniques have already been proposed for VHR SAR imagery in the literature. Semi-automatic methods for the height estimation in detected VHR SAR imagery by means of shadow or layover analysis are proposed in [4]–[6], while methods in [7]–[9] make use of interferometric SAR (InSAR). The use of stereoscopic SAR (radargrammetry) is proposed in [10] and [11]. Recently, methods based on multiaspect SAR data, in which the same area

is imaged from different flight paths, have been proposed in [12] and [13]. A method based on multiaspect InSAR data is presented in [14]. In [15], the use of multiaspect polarimetric InSAR data is investigated. First results with circular SAR are shown in [16]. The presented methods have in common that the achieved accuracy improves with the use of multidimensional data. However, the performance of a proposed methodology is typically presented for a small set of test data, usually comprising only single or few buildings, leaving a general applicability of the method in doubt.

The height extraction by radiometric analysis of the typical double-bounce reflection of a building (see Section II for details) using an electromagnetic scattering model [17] based on the geometrical optics (GO) and the Kirchhoff physical optics (PO) approximations [18, pp. 922–1033] for a simplified rectangular flat-roofed building is demonstrated in [19]. This method has the potential to extract the height of the building accurately from a single image but needs extensive *a priori* knowledge of the material and surface roughness properties (i.e., dielectric constant, rms height, and correlation length) of the building and its surrounding, which may not always be available.

SAR simulators [20]–[22] are suitable not only for the analysis of scattering phenomena but also as part of information extraction methodologies for actual SAR imagery. In [23], for instance, the polarimetric GrecoSAR simulator is deployed to detect vessel scattering hotspots, which are then used to classify ships in actual SAR imagery. As an extension, GrecoSAR was tested in [24] for the simulation of urban structures. In the case of building reconstruction from multiaspect InSAR data, Soergel *et al.* [25] propose an iterative procedure based on the predictions of height maps compared to the actual digital elevation model (DEM).

In this paper, which generalizes and extends the work presented in [26] and [27], we propose a novel automatic 3-D reconstruction concept for the extraction of the height of man-made structures from single detected SAR (power) imagery under the following assumptions: 1) A map with the location of the building is available and 2) the width, length, and roof type of the investigated building are known. The approach is based on a "hypothesis generation–rendering–matching" procedure. A series of hypotheses is generated and rendered by a SAR imaging simulator, taking into account the acquisition parameters of the actual VHR SAR data. The simulations are compared to the actual VHR SAR data; the estimated height corresponds to the hypothesis for which the simulated image best matches with the actual scene. The novelty of the presented concept consists in the use of single detected VHR SAR images instead of multidimensional data (e.g., interferometric, polarimetric, and multiaspect). It is worth noting that the use of a single detected VHR SAR image for height estimation can support a wide range of current applications, including the use of new spaceborne SAR sensors such as TerraSAR-X and COSMO-SkyMed. Furthermore, the potential use of single detected SAR data can provide significant economic efficiencies in emergency response (e.g., speed and cost).

The proposed height estimation process is applicable to different building shapes under the full range of aspect angles,

i.e., the angle between the front wall of the object and the SAR sensor azimuth direction. One of the key characteristics of the proposed procedure is the simultaneous consideration of the major scattering characteristics of the man-made structure in SAR (i.e., layover and shadow areas and multibounce contributions) for estimating the height. We demonstrate the performance and the properties of our approach by analyzing a set of 40 flat- and gable-roof buildings in submeter VHR airborne SAR images and in 1-m resolution TerraSAR-X images for an urban area in Dorsten, Germany.

The remainder of this paper is structured as follows. In Section II, we review the fundamentals of microwave backscattering of an idealized building. In Section III, we describe the proposed approach in detail. We introduce the test data in Section IV. Section V discusses the results of the method before we finish with some conclusions and an outline of future work in Section VI.

II. PROPERTIES OF IDEALIZED BUILDINGS IN SAR IMAGES

In SAR imagery, typical urban structures are affected by layover, double-bounce, and shadowing effects, which relate to the ranging geometry of radar sensors. To highlight these effects, Fig. 1 shows examples of the backscattering range profiles of a simple flat-roof building model, which is a rectangular box with uniformed surfaces and flat surroundings, with common width w and different heights h viewed by a SAR sensor with incidence angle θ: a shows the return from the ground, b highlights the double bounce caused by the dihedral corner reflector that arises from the intersection of the building vertical wall and the surrounding ground, c indicates single backscattering from the front wall, d depicts the returns from the building roof, and e represents the shadow area from which there is no return from the building or the ground. The symbols l $[l = h \cdot \cot(\theta)]$ and s $[s = h \cdot \tan(\theta)]$ denote the lengths of the areas affected by layover and shadow in the ground-projected image space, respectively. For the backscattering of flat-roof buildings, three different situations can be observed according to the boundary condition $h < w \cdot \tan(\theta)$ [4], [25]. If this condition is fulfilled [Fig. 1(a)], part of the roof scattering d is superimposed on the scattering from the ground a and the front wall c in the region $a + c + d$, while there is a region d which is only characterized by returns from the roof. In the case of $h = w \cdot \tan(\theta)$ [Fig. 1(b)], all of the roof contribution d is sensed before the double-bounce area in such a way that there is a homogeneous layover area $a + c + d$, which has contributions from the ground, the building front wall, and the roof. If $h > w \cdot \tan(\theta)$ [Fig. 1(c)], all roof contributions are sensed before the double-bounce area again, with the difference that the layover area is split in an area $a + c + d$, which has contributions from the ground, the front wall, and the roof, and an area $a + c$, which only has backscatter from the ground and the front wall of the building.

The scattering effects of a gable-roof building are different from what is observed for a flat-roof building [4], [28]. Fig. 2 shows three examples of backscattering profiles from a gable-roof building with roof inclination angle α for different incidence angles. The major difference with respect to flat-roof
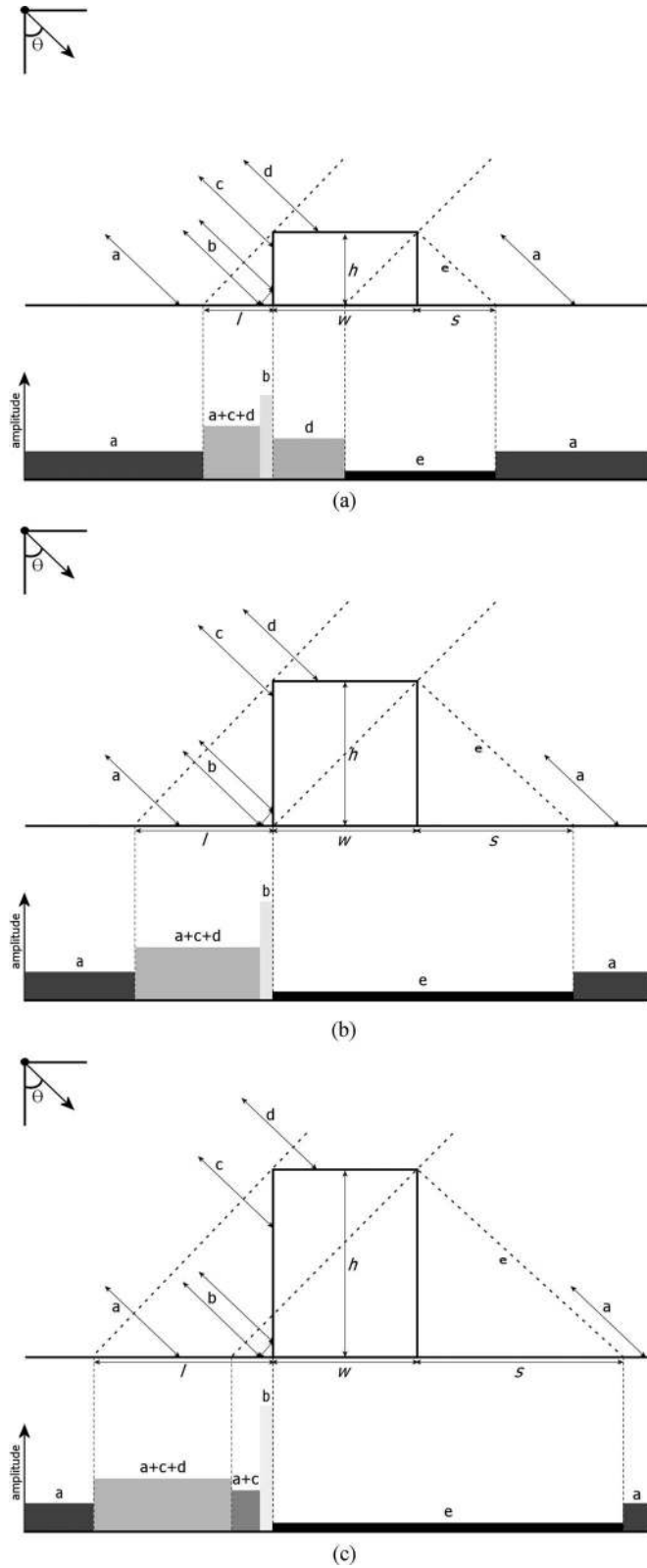
Fig. 1. Scattering from a simple flat-roof building model with width w and different heights h: Ground scattering a, double bounce b, scattering from vertical wall c, backscattering from roof d, shadow area e, length of layover area in ground-projected image space l, and length of shadow area in ground-projected image space s. The gray values in the backscattering profiles correspond to the relative amplitudes. (a) $h < w \cdot \tan(\theta)$. (b) $h = w \cdot \tan(\theta)$. (c) $h > w \cdot \tan(\theta)$.
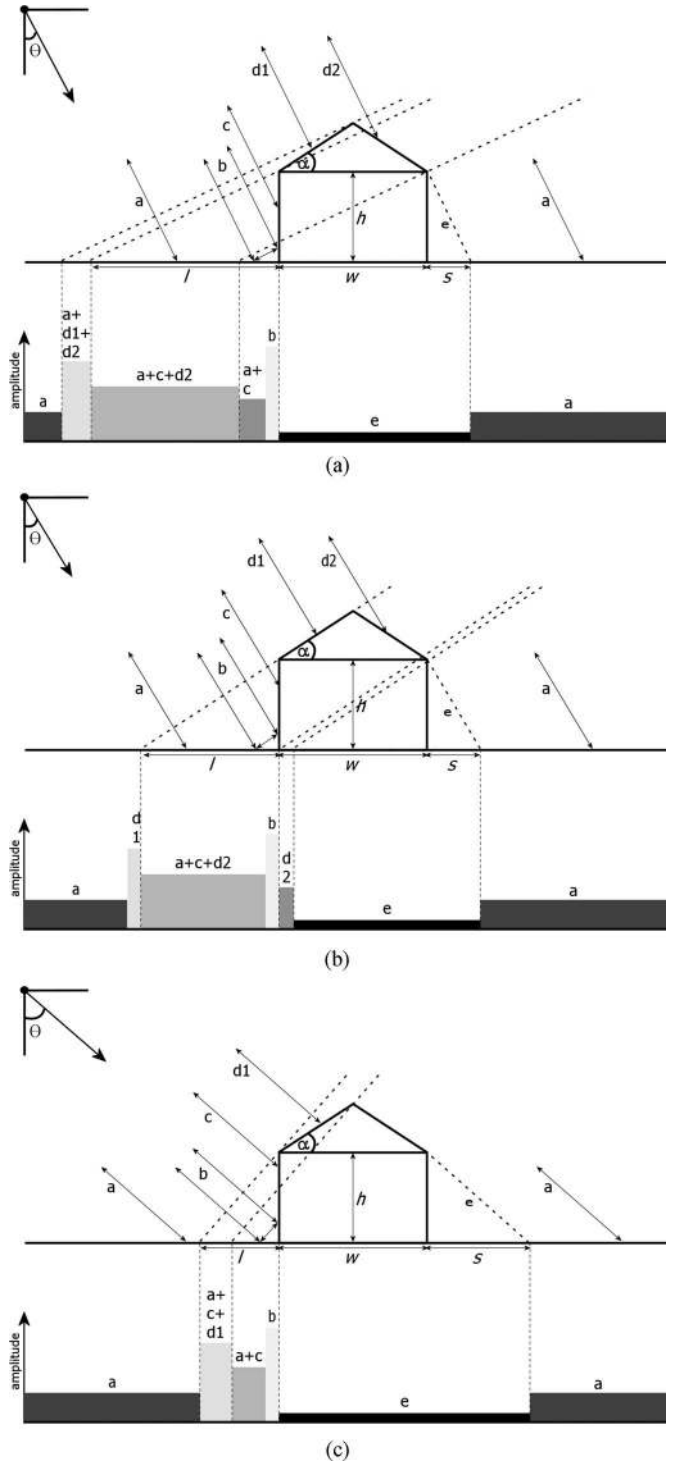

Fig. 2. Examples of backscattering profiles from a gable-roof building with roof inclination angle α at various incidence angles. The legend is similar to the one for flat-roof buildings in Fig. 1. $d1$ denotes the scattering from the side of the roof which is oriented toward the sensor, while $d2$ represents the scattering from the part of the roof which faces away from the sensor. The gray values in the backscattering profiles correspond to the relative amplitudes. (a) $\theta < \alpha$. (b) $\theta = \alpha$. (c) $\theta > \alpha$.

buildings is the presence of a second bright scattering feature, which is closer to the sensor than the double bounce, resulting from direct backscattering $d1$ from the part of the roof which is oriented toward the sensor. For incidence angles which are not equal to the inclination angle of the roof [Fig. 2(a) and (c)],
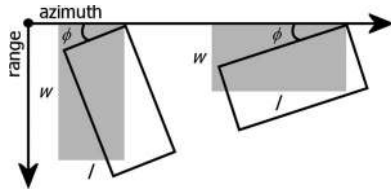
Fig. 3. Definition of width w, length l, and aspect angle ϕ of a building. The buildings in gray are oriented parallel to the azimuth direction with $\phi = 0°$, while the buildings with the solid black lines were rotated counterclockwise by ϕ with respect to the azimuth direction.

this feature is extended to an area, while in the case of $\theta = \alpha$ [Fig. 2(b)], these contributions return to the sensor at the same instance of time and are therefore integrated into a bright line that is similar to the double bounce of a building. In the case that $\theta > \alpha$, there is no backscattering $d2$ from the part of the roof which faces away from the sensor, since this part is occluded by the front side of the building.

The viewing configuration of a sensor with respect to the building is defined not only by the incidence angle of the sensor but also by the orientation of the building with reference to the azimuth direction, known as aspect angle. In Fig. 3, which also highlights the definition of the dimensions of a building, we describe this orientation in terms of the planar dimensions (width and length) of the building and the aspect angle ϕ. Assuming a building having the wall oriented toward the sensor that is parallel to the azimuth direction ($\phi = 0°$), the length l of a building is given by the dimension of the building in the azimuth direction, while its width w is the corresponding dimension in the range direction. The angle ϕ ($0° \leq \phi < 90°$) is defined as the angle between the wall with length l and the azimuth direction, considering a counterclockwise rotation of the building. Hence, the planar geometry of a building is given with respect to the azimuth direction of the sensor by the triplet (w, l, ϕ). For spaceborne acquisitions, the possible aspect angles of a building are fixed by descending and ascending passes, while for airborne measurements, the aspect angle is defined by the flight track of the airplane and the squint angle of the antenna. Hence, airborne acquisitions permit more flexibility for varying ϕ than spaceborne measurements, which can be an advantage for missions where buildings need to be investigated from a predefined viewing configuration.

As shown in Figs. 1 and 2, the double-bounce effect is a significant characteristic of buildings in VHR SAR signals [17]. It indicates the presence of a building and appears in correspondence with its front wall, so that it can be used as a feature for the automatic detection and reconstruction of buildings from SAR data [19]. However, the strength of the double-bounce effect depends on both the height of the building (i.e., the higher the building, the stronger the double bounce) and the aspect angle. Theoretical models for the double bounce of a building show a quadratic dependence of its radar cross section on the building height [17], [29]. An empirical study on the relationship between the strength of the double bounce and the aspect angle is presented in [30], highlighting that the double-bounce contribution drops off significantly if the aspect angle is increased in the lower aspect angle range (up to $10°$), while it decays moderately for higher angles.
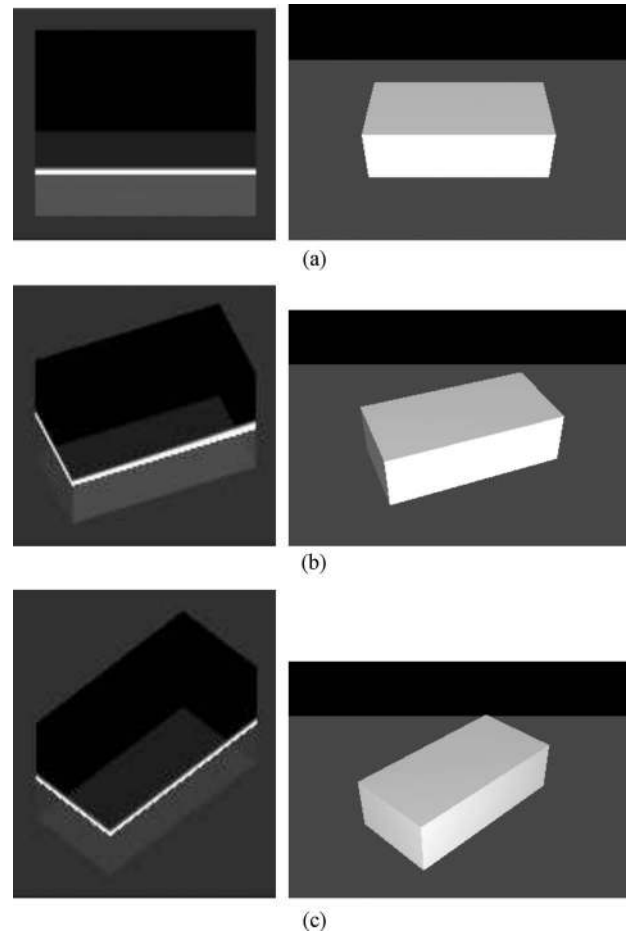

Fig. 4. Simulations (without speckle) of a flat-roof building model [corresponding to Fig. 1(a)] with dimensions $w = 50$ m, $l = 100$ m, and $h = 30$ m with $\theta = 50°$ and 1.0-m azimuth and slant-range resolutions. The images in the left column show the simulations with viewing direction from the bottom, while the images in the right column show the corresponding 3-D models as they would appear visually. (a) $\phi = 0°$. (b) $\phi = 22.5°$. (c) $\phi = 45°$.

The aspect angle mainly influences the appearance of the layover and shadow areas. In Fig. 4, we present the results of simulations (for more details on the simulator, see Section III-B) of a flat-roof building model with dimensions $w = 50$ m, $l = 100$ m, and $h = 30$ m at (a) $\phi = 0°$, (b) $\phi = 22.5°$, and (c) $\phi = 45°$, which reflects the situation shown in the right part of Fig. 3. The simulations were performed with $\theta = 50°$ and 1.0-m azimuth resolution δ_a and slant-range resolution δ_r corresponding to a configuration supported by current spaceborne sensors like TerraSAR-X or COSMO-SkyMed. The images on the left side show the simulation results with viewing direction from the bottom, while the images in the right column display the corresponding 3-D views of the building, as it would appear visually. Since we want to highlight the major scattering effects of buildings in SAR, we suppressed the calculation of speckle. In the situation of $\phi = 0°$, the shadow and layover areas have a rectangular shape, which changes with increasing aspect angle to L-shape. The area at which there is only backscattering from the roof also has a rectangular shape at $\phi = 0°$, but it changes for $\phi > 0°$ to a parallelogram. Note that, in these simulations, the relative strength of the double bounce may be overestimated for the cases of larger aspect angles ($\phi = 22.5°$ and $\phi = 45.0°$) [30].
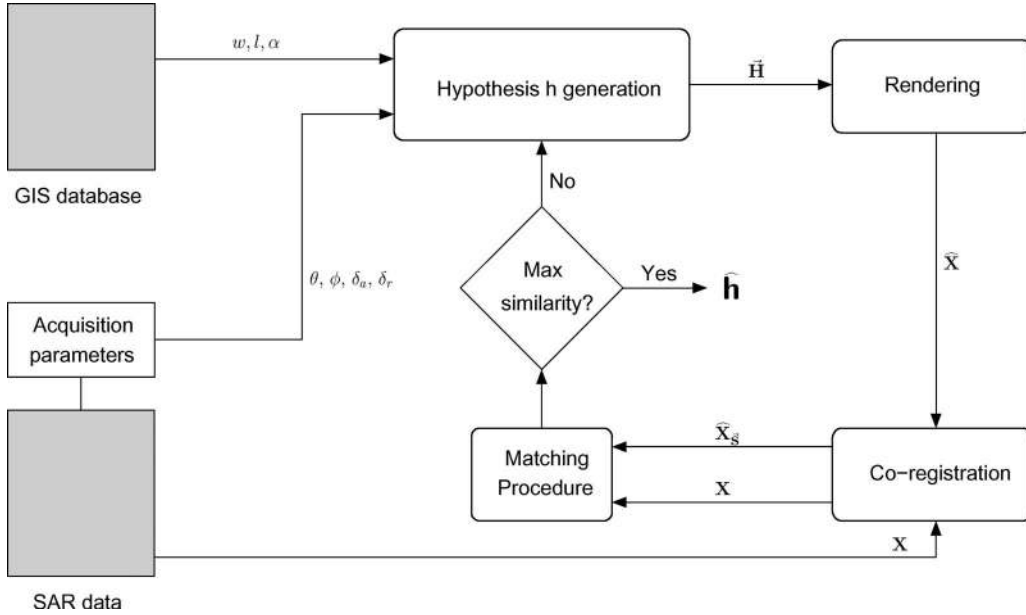
Fig. 5.   Block scheme of the proposed methodology for building height estimation from single detected VHR SAR data.

III. PROPOSED METHODOLOGY FOR BUILDING HEIGHT ESTIMATION FROM VHR SAR IMAGES

Let h and h_{true} be an assumption of the height and the true height of the analyzed building, respectively. Let $\mathbf{X}$ denote the true SAR image, and let $\widehat{\mathbf{X}}$ be a simulated SAR image at building height h. In order to find the best estimate $\widehat{h}$ for the height of the building, we define a "hypothesis generation–rendering–matching" approach, which is shown in Fig. 5. A building is simulated at different heights and compared to the actual scene, under the assumption that the simulated and actual scenes are coregistered. The simulation requires the knowledge of both a set of parameters related to the acquisition of the actual SAR image and a set of parameters related to the shape and size of the building. The last set also includes the hypothesis on the height of the building h. Thus, a simulation is defined by $\vec{H} = \{w, l, h, \alpha, \theta, \phi, \delta_a, \delta_r\}^{\text{T}}$, where w is the width and l the length of the building, α the pitch of the roof, θ is the incidence angle, ϕ is the aspect angle, δ_a is the azimuth resolution, and δ_r is the slant-range resolution. The final estimated height of the object corresponds to the hypothesis which matches best with the actual scene and is given by

$$\widehat{h} = \arg\max_{h} \left\{ \mathbf{M}\left[\widehat{\mathbf{X}}(\vec{H}), \mathbf{X}\right] \right\} \qquad (1)$$

with $\mathbf{M}$ as the matching function. The highest value of $\mathbf{M}$ corresponds to the best match between the hypothesis and the actual scene. To calculate the match between the simulation and the actual scene, both images need to be coregistered. In practice, coregistration and height estimation are similar tasks which can be executed at the same time in the matching procedure. The value of the measure $\mathbf{M}$ for which the best coregistration between a simulation and the actual scene is achieved is also the final match value for this pair, expressing in a quantitative way how well the simulation fitted with the actual scene. Since the viewing configuration at which the object under investigation was sensed in the actual scene is

modeled by the SAR simulator, only translations are considered as transformation. Hence, the 1-D optimization problem in (1) becomes a 3-D problem

$$\widehat{h} = \arg\max_{h, \vec{s}} \left\{ \mathbf{M}\left[\widehat{\mathbf{X}}_{\vec{s}}(\vec{H}), \mathbf{X}\right] \right\} \qquad (2)$$

where $\widehat{\mathbf{X}}_{\vec{s}}$ denotes the translation of the image $\widehat{\mathbf{X}}$ by the 2-D vector $\vec{s} = (dx, dy)^{\text{T}}$ associated with the coregistration process.

To solve the maximization problem of (2), we use the multidimensional Nelder–Mead [31] (or downhill simplex) function optimization method. Alternatively, the use of simulated annealing [32] could also be considered to limit the effects of local maxima, although this would increase computational costs. In order to avoid instability in the similarity measure, an averaging over the similarity values in a predefined height interval, for instance $[h - 0.4\text{ m}; h + 0.4\text{ m}]$ sampled in 0.2 m steps, can be performed. The three methodological steps, namely, hypothesis generation, rendering, and matching, are described next.

A. Hypothesis Generation

For the height estimation process, only the height parameter is variable in $\vec{H}$, while the other parameters are constant throughout the estimation procedure. Planar dimensions w and l are derived from a GIS database (e.g., cadastral maps, digitized maps from independent ancillary data, and optical remote sensing images), which contains the footprint of the building. For flat-roof buildings $\alpha = 0°$, while for gable-roof buildings, α is chosen according to the characteristic roof inclination angle for the investigated area, i.e., $35°$ for the test region considered in this paper. The incidence angle and the SAR sensor parameters are defined by the acquisition conditions of the actual SAR scene from which the height of the building is extracted. ϕ is obtained by combining the information from the GIS database with the information on the flight track of the airplane in the case of an airborne acquisition or with the information

about the orbit of the satellite in the case that the actual scene was acquired by a spaceborne sensor. A number of hypotheses are generated for the same building during the maximization of (2). This can be achieved implicitly by the function optimization method which jointly maximizes for h and $\vec{s}$. For a better performance, an explicit hypothesis generation can be performed by iterating h in a predefined range of expected building heights with a given step size. Thus, the function optimizer has only the task to coregister a rendered hypothesis with the actual scene by varying $\vec{s}$.

B. Rendering

For evaluating which hypothesis matches best with the actual scene, a SAR simulator is employed, which renders the hypothesis into the geometry of a SAR image. First, a 3-D model is generated from the information in $\vec{H}$, taking into account the building parameters w, l, h, and α. Second, the 3-D model is triangulated so that, in the third step, its backscattering can be simulated considering the parameters θ, ϕ, δ_a, and δ_r specified in $\vec{H}$.

Our application scenario aims at extracting building information from SAR scenes over areas where surface roughness parameters and the dielectric properties of the materials in the scene are generally unknown *a priori*. Thus, electromagnetic models such as GO, PO, integral equation method [33], or the finite-difference time-domain method [34] cannot be adopted to calculate the backscattering. Hence, an adjustable mixture model of Lambertian and specular reflection is used to calculate the backscattering from the surface and building model. Rather than calculating the absolute radiometric effects related to material properties and surface roughness parameters, this simplified scattering model approximates the relative differences in backscatter while retaining the dominant geometrical effects of surface and dihedral scattering. Note that any simulator, which can calculate the effects related to the SAR geometry, can be employed, irrespective of its exact radiometric model.

Our SAR simulator uses ray tracing in order to determine which surfaces of a generic object are visible. It can handle any complex object composed of spheres, planes, and triangles or any arbitrary combination of these objects. The simulator optionally includes multiple bounce scattering and can therefore distinguish between single- and dual-bounce reflections. Speckling effects are neglected in our approach.

The output of the simulator is a 2-D rectangular image, whose dimension is determined such that it includes the scattering effects of the simulated object (i.e., single-bounce contributions, shadow, layover, and double bounce) plus a border area, which contains backscattering from the ground. Note that the ratio r between the number of pixels belonging to the scattering effects from the object (foreground pixels) and the number of pixels belonging to the ground scattering of the surrounding (background pixels) varies for different buildings or for the same building measured with different viewing configurations. This can be observed, for instance, in Fig. 4 for three simulations of the same building, which differ only in ϕ: r of Fig. 4(a) is larger than r of Fig. 4(b), which is larger than r of Fig. 4(c). We will highlight the consequence of this effect and propose a solution to compensate for it in the next section.

C. Matching

In order to estimate the height of a building according to (2), we need to optimize the match between the simulated image and the actual scene with respect to h and $\vec{s}$. Image matching and registration are two operations which are closely linked to each other. A slave image, which must be coregistered to a master image, is translated so that the match between these two images reaches a maximum similarity based on a chosen similarity measure. Hence, the matching between two images is an integral part of a coregistration method, so that we can jointly optimize for h and $\vec{s}$.

For image matching, two types of methods exist: area- and feature-based methods [35]. Area-based methods calculate directly the correlation between all (or a subset of) samples in the two corresponding images. For instance, Barat *et al.* [36] propose a method for pattern matching based on a profiling approach using morphological transforms. Feature-based methods, instead, first extract structural information such as lines and edges from the images to be compared and then, in the second step, match them in the feature space. Depending on the underlying data, various features are in use, such as tie points [37], gradients computed from gray-scale intensity images [38], fractal features based on fractal theory [39], and higher level features such as the shape of objects derived from their edge information [40]. The use of the scale-invariant feature transform method, which extracts features that are invariant to image scale and rotation and are robust with respect to affine distortions, change in 3-D viewpoint, addition of noise, and change in illumination, is proposed in [41]. This method shows good performance for optical images, while it has a decreased accuracy for SAR images, depending on the content of the SAR scene [42].

Our matching task is faced with two challenges: 1) comparing the actual SAR data with speckle to synthetic images without speckle, i.e., the geometries of the images are similar, but the local statistics in the comparison are different, and 2) the radiometry of the simulated image does not match with that of the actual scene.

We proposed a feature-based method in [43], which is based on the extraction of shadow areas and edges. As match criterion, we used the normalized cross-correlation coefficient [44]. The drawback of feature-based methods is the dependence of the effectiveness and stability of the feature extraction procedures on parameter settings, which is particularly critical for SAR images. Therefore, we propose in this paper an area-based method based on mutual information (MI) for $\mathbf{M}$ in (2). MI is a measure derived from information theory, which is suitable for multimodality image matching/registration tasks, which was independently proposed in [45] and [46] for the registration of multimodality medical images and studied by Xie *et al.* [47] for its application in the SAR domain. The MI $MI(\widehat{\mathbf{X}}, \mathbf{X})$ between $\widehat{\mathbf{X}}$ and $\mathbf{X}$ is given by

$$MI(\widehat{\mathbf{X}}, \mathbf{X}) = H(\widehat{\mathbf{X}}) + H(\mathbf{X}) - H(\widehat{\mathbf{X}}, \mathbf{X}) \qquad (3)$$

where $H(\widehat{\mathbf{X}})$ and $H(\mathbf{X})$ are the entropies of $\widehat{\mathbf{X}}$ and $\mathbf{X}$, respectively, and $H(\widehat{\mathbf{X}}, \mathbf{X})$ is their joint entropy. By using x and $\widehat{x}$ to

denote the pixel values in the measured and simulated images, respectively, the entropies can be computed by

$$H(\widehat{\mathbf{X}}) = -\sum_{\widehat{x}} p_{\widehat{\mathbf{X}}}(\widehat{x}) \log p_{\widehat{\mathbf{X}}}(\widehat{x}) \qquad (4)$$

$$H(\mathbf{X}) = -\sum_{x} p_{\mathbf{X}}(x) \log p_{\mathbf{X}}(x) \qquad (5)$$

$$H(\widehat{\mathbf{X}}, \mathbf{X}) = -\sum_{\widehat{x}, x} p_{\widehat{\mathbf{X}}, \mathbf{X}}(\widehat{x}, x) \log p_{\widehat{\mathbf{X}}, \mathbf{X}}(\widehat{x}, x) \qquad (6)$$

where $p_{\widehat{\mathbf{X}}}(\widehat{x})$ and $p_{\mathbf{X}}(x)$ are the marginal probability mass functions and $p_{\widehat{\mathbf{X}}, \mathbf{X}}(\widehat{x}, x)$ is the joint probability mass function. They can be calculated by

$$p_{\widehat{\mathbf{X}}, \mathbf{X}}(\widehat{x}, x) = hist(\widehat{x}, x) \bigg/ \sum_{\widehat{x}, x} hist(\widehat{x}, x) \qquad (7)$$

$$p_{\widehat{\mathbf{X}}}(\widehat{x}) = \sum_{x} p_{\widehat{\mathbf{X}}, \mathbf{X}}(\widehat{x}, x) \qquad (8)$$

$$p_{\mathbf{X}}(x) = \sum_{\widehat{x}} p_{\widehat{\mathbf{X}}, \mathbf{X}}(\widehat{x}, x) \qquad (9)$$

where $hist$ denotes the joint histogram of the two images. The reason for the independence of this similarity measure to the absolute intensity values of the two images is that the MI is only sensitive to the occurrence of the same pairs of intensity values in $\widehat{\mathbf{X}}$ and $\mathbf{X}$.

Depending on the speckle filtering of the SAR images, the number of bins for the joint histogram is a noncritical parameter. It should be chosen so that the joint histogram has, on average, at least one entry per bin [47]. Since the lowering of the number of bins has a comparable effect to a low-pass filter, the number of bins should decrease the more the data are affected by speckle. The simulations are without speckle, so that we choose 256 bins for $\widehat{\mathbf{X}}$. Since we apply a speckle filter in the preprocessing step to the actual SAR data (Section IV), we only decrease the number of bins for $\mathbf{X}$ to 128. A test with 64 bins did not yield an increased matching accuracy.

For the coregistration of $\widehat{\mathbf{X}}_{\mathbf{s}}$ and $\mathbf{X}$, we allow subpixel accuracy, which means that we allow shifts in the x- and y-directions, which do not match the grid spacing of the image. Therefore, it is necessary to interpolate the values for the pixels that do not coincide with a grid point of the original raster. With respect to the coregistration of two images with MI, the partial volume (PV) interpolation method was proposed in [48], where, instead of interpolating new intensity values, the joint histogram is updated directly. A series of empirical tests showed that PV outperforms in accuracy other methods such as bilinear interpolation.

In Section III-B, we highlighted that the ratio r between foreground and background pixels is not constant for different buildings and viewing configurations. In the case of a building with $\phi = 0°$ [Fig. 4(a)], r has a relative high value, which means that the matching (and, hence, the height estimation) is dominated by the scattering of the object itself. Instead, for
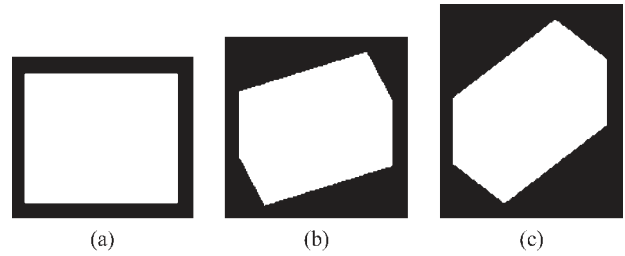

Fig. 6. Object masks of the corresponding simulations in Fig. 4. (a) $\phi = 0°$. (b) $\phi = 22.5°$. (c) $\phi = 45°$.
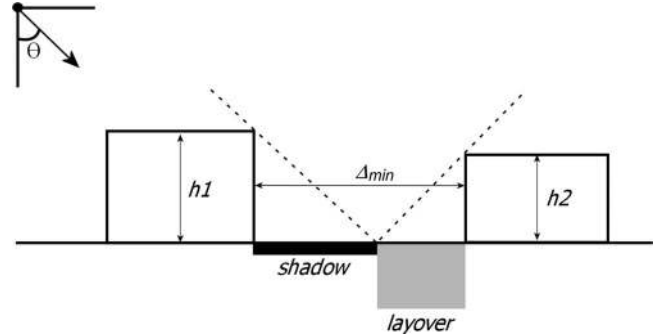

Fig. 7. Minimum distance $\Delta_{\min}$ required between two buildings in order that their scattering effects do not interfere.

$\phi = 45°$ [Fig. 4(c)], the value of r is relatively low, implying that the estimation procedure is influenced more significantly by the background than by the foreground pixels. Since an optimal height estimation accuracy is only achieved for a certain tradeoff between fore- and background pixels, we have to ensure that the matching procedure always uses the same r. To fix r, we defined a binary object mask, where all foreground pixels have the value one and all background pixels have the value zero. This object mask is generated by the simulator as a secondary result of the simulation run. Fig. 6 shows the corresponding object masks for the simulations of Fig. 4. By expanding the object masks using the morphological dilation operator [49] and by only considering, for the calculation of the MI value, those pixels which have the value one in the expanded object mask, r can be fixed for the matching procedure for different buildings and viewing configurations. As a structuring element for the dilation operator, we use a disk whose size is determined separately for each simulation to fulfill the desired r value. A series of tests showed that the best accuracy for the height estimation is achieved for $r = 1$, which means that the number of foreground pixels is equal to the number of the background pixels.

D. Constraint for Obtaining Reliable Estimations

The proposed method estimates the height of an individual building by simulating the expected SAR signature of a simplified building model and evaluating the match with the actual scene. However, the simulation process does not consider the effects of other man-made or natural structures in the surrounding of a building under investigation, which might have an impact on its actual backscattering. The minimum distance $\Delta_{\min}$
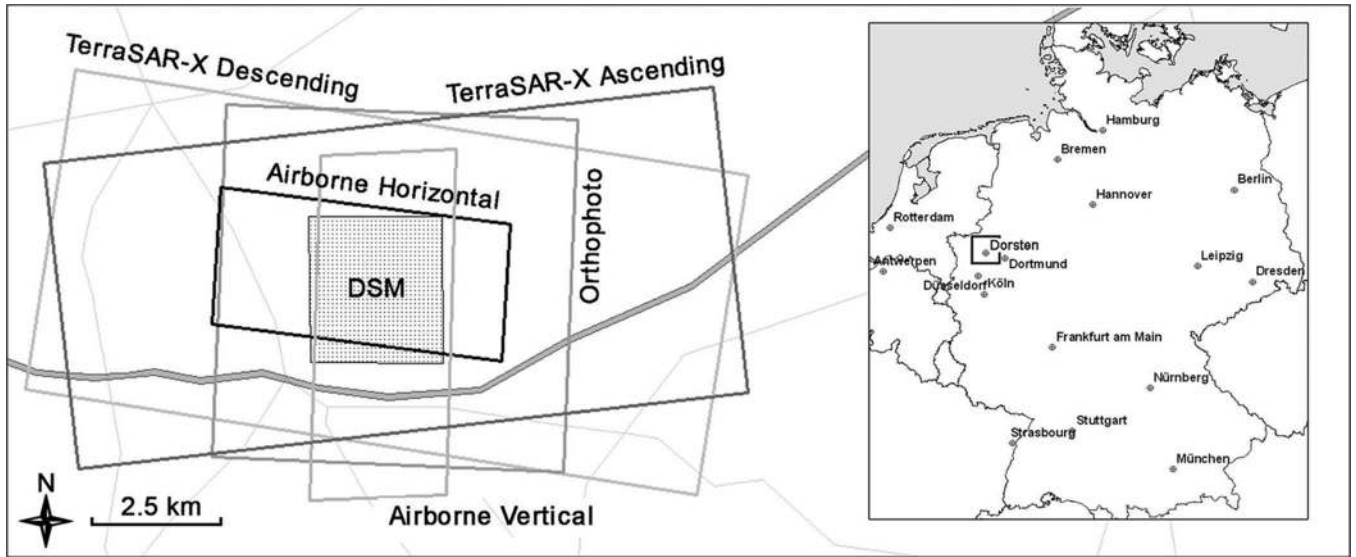
Fig. 8. Overview of data set from Dorsten. Note that, for the spaceborne data, the azimuth dimension of the images is smaller than the range dimension, while for the airborne data, the azimuth dimension is larger compared to the range dimension.

required between two buildings in order that their scattering effects do not interfere with each other is given by [25]

$$\Delta_{\min} = h_1 \cdot \tan(\theta) + h_2 \cdot \cot(\theta) \qquad (10)$$

where h_1 is the height of the building at the sensor close side and h_2 denotes the height of the building, which is behind the first building (see Fig. 7). Hence, optimal accuracy for the height estimation process for a building can only be achieved if the condition

$$\Delta_{\text{act}} > \Delta_{\min} \qquad (11)$$

is fulfilled, where Δ_{act} denotes the actual distance between the buildings.

IV. DATA SET DESCRIPTION

The test area chosen was a subset of the city of Dorsten $(51°40'18'' \text{ N}, 6°59'34'' \text{ E})$, Germany, for which we considered both dual-aspect airborne and a pair of ascending and descending spaceborne VHR SAR data. Ancillary data, which were used to retrieve the building footprint parameters for initializing the simulator, were provided by an orthophoto acquired on June 9, 2006 with a 30-cm resolution. Furthermore, ground data were manually collected in combination with a Lidar digital surface model (DSM) with approximately 10-cm vertical resolution. An overview of the composed data set is shown in Fig. 8.

We consider two types of building structures for which we estimate the heights: flat- and gable-roof buildings. All buildings are assumed to be individual buildings with rectangular footprints. To evaluate the performance of the method under a variety of conditions, we choose 40 individual industrial and apartment buildings with different shapes at various aspect and incidence angles, which we categorized in three groups. Category A contains flat-roof buildings, and category B contains gable-roof buildings. Category C contains the buildings which do not fit the structural assumptions of our building models completely. This category includes buildings which

have a nonrectangular footprint (such as buildings with a tower attached to it) or buildings which have nonuniform heights. The majority of the selected buildings are gable-roof buildings, which is the prevailing type of structure for residential houses in this area. Only few flat-roof buildings could be identified, some of which are apartment buildings (flats) and some are industrial or commercial structures (e.g., factories and stores). Six buildings belong to category A, 27 buildings belong to B, and seven buildings were classified as C. To distinguish between the different buildings in the various categories, we use the naming scheme ⟨Category⟩-⟨Number of building in category⟩, e.g., A-3 denotes the third building in the category for flat-roof buildings.

Fig. 9 shows a subset of the orthophoto and the corresponding DSM with some example buildings for two of the three categories: buildings denoted by A-1, A-2, A-3, and A-4 are flat-roof buildings, while buildings B-1–B-5 and B-18–B-23 belong to the class of gable-roof buildings. An example of a building belonging to category C is shown in Fig. 14 in Section V-A3.

The two airborne SAR scenes taken by the AeS-1 sensor from Intermap Technologies [50], for which the corresponding subset of Fig. 9 is shown in Fig. 10 (in slant-range geometry), were acquired in X-band on March 13, 2003 with 16-cm azimuth and 38-cm slant-range resolutions in HH polarization. The incidence angle ranges over the swath from $28°$ (near range) to $52°$ (far range). The dual-aspect data were measured in almost perpendicular flight paths with a right-pointing antenna so that the flight path for the "horizontal" scene was approximately from west to east, while the "vertical" scene was measured from north to south. The overlapping area in the two scenes, where we focus on in this paper, is about $2.3 \times 2.3 \text{ km}^2$ and includes a medium-dense residential urban area and several smaller industrial zones.

In order to use the MI as similarity measure for SAR image registration/matching, speckle reduction is essential [47]. Hence, we preprocessed the airborne data by multilooking the

(a)

(b)

Fig. 9. (a) Subset of the orthophoto (top corresponds to north), showing examples of buildings for two categories: Buildings A-1, A-2, A-3, and A-4 belong to the category of flat-roof buildings, while buildings B-1–B-5 and B-18–B-23 are classified as gable-roof buildings. (b) Corresponding subset of the DSM. In the lower left corner, it can be noted that, due to some manual editing, the height information of some buildings was removed (e.g., buildings B-18–B-23). For these buildings, manual height measurements were carried out using a laser device. (Orthophoto: © Geobasisdaten: Landesvermessungsamt NRW, Bonn, 2007; DSM: © Fugro NPA, 2003).

image by four samples in azimuth and two samples in range direction, which resulted in an equivalent number of looks of 2.59 and an approximately square pixel spacing (64 cm in azimuth and 76 cm in range). Furthermore, we speckle filtered the image with the Gamma MAP filter [51] and the mean shift filter proposed in [52], which acts mainly on shadow areas.

The TerraSAR-X spaceborne data, for which the corresponding subset of Fig. 9 is shown in Fig. 11, were acquired in a high-resolution spotlight mode with an azimuth and a slant-range resolution of 1.1 and 1.2 m, respectively. The data were processed so that the azimuth and the slant-range spacing is 0.9 m. The descending scene was acquired on December 13, 2007 with θ varying from 53.4° to 54.1° over the swath, while
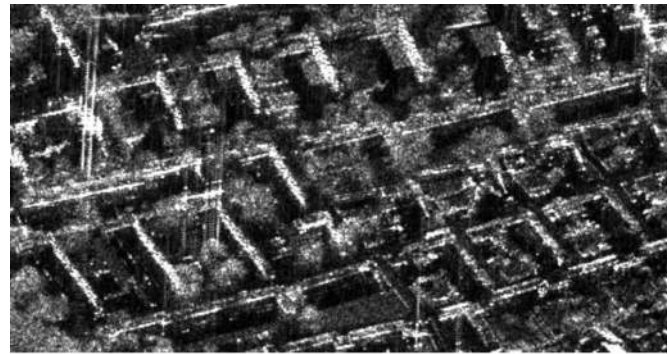
(a)

(b)

Fig. 10. Subset of the airborne SAR scenes in slant-range geometry. (a) Horizontal scene (acquisition from west to east with right-looking sensor) with viewing direction from the top. (b) Vertical scene (acquisition from north to south with a right-looking sensor) with viewing direction from the right side. (© Intermap Technologies GmbH, 2003).

the ascending scene was taken on January 22, 2008 with θ in the range of 50.3°–51.0°. Due to the lower resolution of the spaceborne data, we did not multilook the data before speckle reduction. Hence, the preprocessing of the data was limited to the application of the Gamma MAP and the mean shift filter.

V. RESULTS

The results of the height estimation process for our test data set are presented in this section. Section V-A lists and

Fig. 11.   Subset of TerraSAR-X scenes in slant-range geometry, which corresponds to the subsets shown in Fig. 9. (a) Ascending (viewing direction from left). (b) Descending (viewing direction from right). (TerraSAR-X image: © Infoterra GmbH/DLR, 2007–2008).
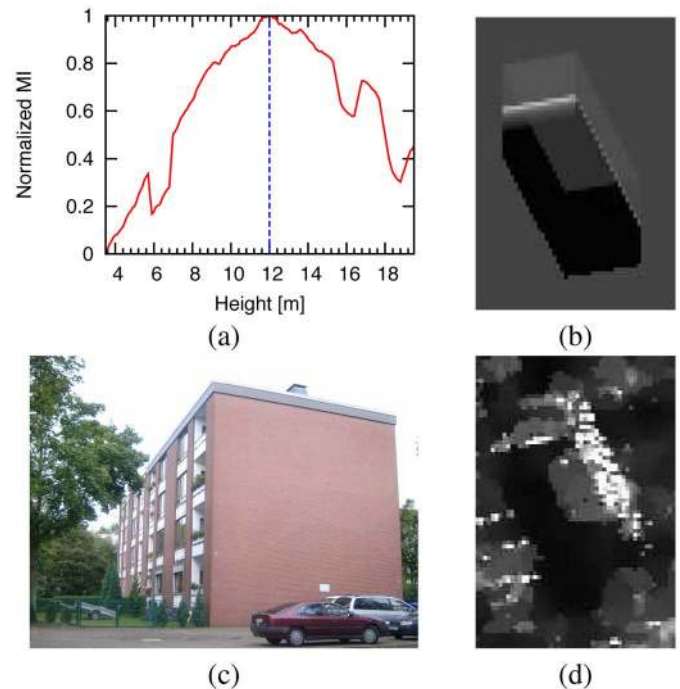

Fig. 12.   Results for building A-2. (a) Plot of the MI values normalized between zero and one, with a maximum at 12.0 m. (b) Simulation of the flat-roof building model with a 12.0-m height. (c) Photograph from the outside of the building. (d) Subset of airborne VHR SAR scene showing building A-2, which is coregistered to the simulation shown in (b). (SAR image: © Intermap Technologies GmbH, 2003).

in Section V-C the results for the ascending and descending TerraSAR-X scenes.

A. Horizontal Airborne Scene

1) Flat-Roof Buildings (Category A): To highlight the results of our method in detail, we show in Fig. 12 the output generated by the method for building A-2, which is also shown in Fig. 9. A photograph of the building is shown in Fig. 12(c), and further details are listed in Table I. The plot in Fig. 12(a) shows MI values normalized between zero and one,[1] for hypotheses with a height range from 3 to 20 m using a 0.1-m step size. The graph shows a good match around the true height of the building (12.5 m), while it drops off with increasing difference between the simulated and actual heights. The global maximum is at 12.0 m, which is 0.5 m lower than the true height. Fig. 12(b) shows the simulation of the building at the estimated height in comparison to the actual SAR scene [Fig. 12(d)], which is coregistered with the simulation. It is obvious that the Lambertian-specular mixture model used in the simulation does not reproduce the correct radiometry of the actual SAR scene. However, qualitative differences in scattering effects, and their characteristic image areas in the SAR geometry (i.e., double bounce, layover, and shadow), are reflected accurately in the simulation. Since we use MI as similarity measure, the difference in the radiometry between simulation and actual scene does not seem to degrade the accuracy of the height estimate.

discusses the results of the three groups of buildings for the horizontal airborne scene in greater detail. In Section V-B, we highlight the results for the vertical airborne scene, particularly with respect to the differences compared to the results from the horizontal scene. To investigate the impact of the lower resolution of the TerraSAR-X data with respect to the airborne data on the accuracy of the height estimation, we summarize

---

[1]Note that this is not the same normalized MI measure proposed by Studholme *et al.* [53].

TABLE I
RESULTS FOR FLAT-ROOF BUILDINGS (CATEGORY A) FOR HORIZONTAL AIRBORNE SCENE

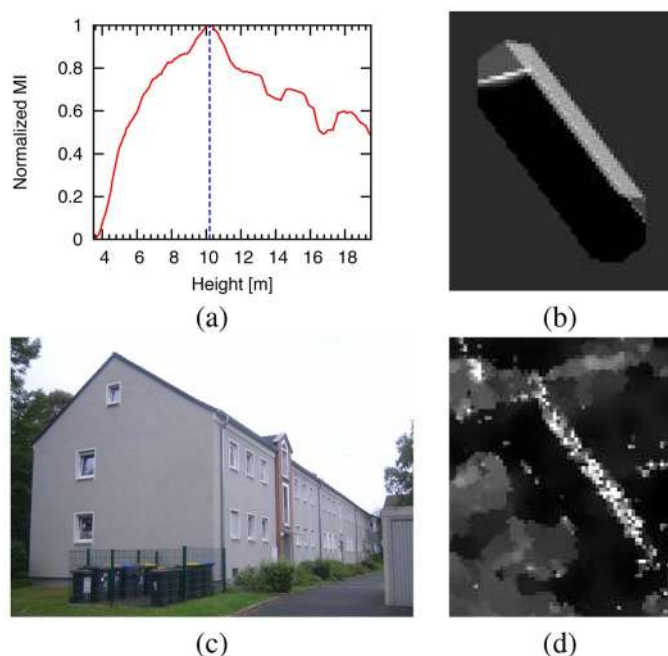| Building	Width	Length	ϕ	Local θ	Height	Estimate	Difference	Relative Difference
A-1	12.1 m	35.8 m	16.6°	45.4°	12.5 m	12.4 m	-0.1 m	-0.8%
A-2	12.4 m	35.9 m	16.3°	45.4°	12.5 m	12.0 m	-0.5 m	-4.0%
A-3	12.3 m	36.3 m	16.6°	45.6°	10.0 m	9.5 m	-0.5 m	-5.0%
A-4	12.4 m	36.0 m	15.5°	45.7°	10.0 m	14.4 m	4.4 m	44.0%
A-5	15.8 m	45.7 m	14.0°	43.0°	7.0 m	5.4 m	-1.6 m	-22.9%
A-6	37.8 m	13.1 m	30.4°	43.0°	5.1 m	5.1 m	0.0 m	0.0%
Mean and standard deviation							0.3 ± 2.1 m	1.9 ± 22.3%

Fig. 13. Results for building B-18. (a) Plot of the normalized MI values with a maximum at 10.2 m. (b) Simulation of the flat-roof building model with a 10.2-m height. (c) Photograph from the outside of the building. (d) Subset of airborne VHR SAR scene showing building B-18, which is coregistered to the simulation shown in (b). (SAR image: © Intermap Technologies GmbH, 2003).

A summary of the results of the proposed height estimation procedure for the buildings in category A is given in Table I, together with the details (dimensions and the corresponding viewing configuration of the sensor) of the buildings.

2) Gable-Roof Buildings (Category B): The category of gable-roof buildings contains the largest distinct set of buildings in the test data set. Fig. 13 shows, in detail, the results for building B-18, which is also shown in Fig. 9. A photograph of the building is shown in Fig. 13(c), and further characteristics are listed in Table II. The plot in Fig. 13(a) shows the match value for the same height hypotheses as in Section V-A1. Similar to the plot in Fig. 12(a), the maximum value at 10.2 m is close to the true height of the actual building (9.5 m). Comparing the simulation in Fig. 13(b) to the actual building shown in Fig. 13(d), it can be noticed again that the SAR image geometries of the two images match well, while there is a significant difference in the radiometries of the two images, which justifies the use of MI as a measure for the matching procedure.

The results of the height estimation for the 27 gable-roof buildings are summarized in Table II. The mean difference is

0.9 m, which indicates that the method has a small tendency to overestimate heights. The standard deviation is 1.5 m, which is slightly better than what was achieved for the flat-roof buildings.

The method provides consistent results when matching buildings at different aspect angles. Let us consider, for example, buildings B-1 and B-4 in Fig. 9, which have approximately the same size and structure. Local incidence angles are more or less the same, whereas B-1 is turned about 96° clockwise with respect to B-4. The results for the height estimation, with an estimation difference of 1.5 m for B-1 and 1.0 m for B-4, are good matches for both buildings, demonstrating the robustness of the method with respect to varying aspect angle.

3) Buildings for Which the Structural Type Does Not Fit the Models (Category C): In this paper, we assume that buildings have a rectangular footprint and have either a flat or a gable roof. However, these simplifying assumptions do not match all actual buildings. To investigate the performance of the proposed height estimation procedure for buildings that have a different structure than our assumptions, but are approximated in the simulation step by rectangular flat- or gable-roof buildings, we summarize in Table III the results for seven buildings.

Building C-1 is a gable-roof building, but with a very low roof inclination angle. It is approximated by a flat-roof building. In this case, the height is underestimated by 1.3 m, which is in the range of the standard deviations for flat- or gable-roof buildings.

Building C-2 has a rectangular footprint, but with a tower attached to it, which is oriented toward the SAR sensor. Furthermore, the roof structure is not a classical gable-roof structure, where two sides of the roof are inclined, but a hipped roof where all four sides of the roof are inclined. We approximate this structure with a gable-roof building, neglecting the tower, which results in an underestimation of 1.6 m.

Building C-3 [Fig. 14(a) and (b)] is a flat-roof building, which has three different heights. The major part (middle part) of the building is 7.0 m high, while the left and right parts are approximately 3.2 and 0.5 m lower than the main part. We approximate this building with a standard flat-roof building, which is, given the complex signature in the actual scene, a significant simplification. This is reflected in the estimation result, which overestimates the height by 2.9 m.

The main part of C-4 is an industrial rectangular flat-roof building, which has several lower building parts attached to it. For the estimation of the height of the main part, we neglected the structures in the surrounding and assumed the building to be a standalone rectangular flat-roof building, which resulted in an underestimation of 2.1 m.

TABLE II
RESULTS FOR GABLE-ROOF BUILDINGS (CATEGORY B) FOR THE HORIZONTAL AIRBORNE SCENE

Building	Width	Length	ϕ	Local θ	Height	Estimate	Difference	Relative Difference
B-1	9.7 m	39.7 m	19.5°	46.1°	9.5 m	11.0 m	1.5 m	15.8%
B-2	30.3 m	10.5 m	21.2°	45.9°	9.5 m	13.7 m	4.2 m	44.2%
B-3	45.2 m	11.3 m	20.5°	45.8°	9.5 m	10.6 m	1.1 m	11.6%
B-4	45.2 m	10.8 m	24.8°	45.6°	9.5 m	10.5 m	1.0 m	10.5%
B-5	50.0 m	10.4 m	17.3°	46.2°	9.5 m	12.5 m	3.0 m	31.6%
B-6	9.8 m	16.9 m	45.5°	40.9°	9.5 m	10.3 m	0.8 m	8.4%
B-7	10.2 m	17.5 m	45.5°	41.1°	9.5 m	10.2 m	0.7 m	7.4%
B-8	9.9 m	17.5 m	45.5°	41.2°	9.5 m	10.6 m	1.1 m	11.6%
B-9	32.0 m	10.3 m	52.0°	40.2°	9.5 m	9.6 m	0.1 m	1.1%
B-10	24.6 m	12.7 m	37.9°	39.8°	10.5 m	10.5 m	0.0 m	0.0%
B-11	32.0 m	10.4 m	52.5°	40.5°	9.5 m	9.0 m	-0.5 m	-5.3%
B-12	32.0 m	10.3 m	52.4°	40.7°	9.5 m	9.6 m	0.1 m	1.1%
B-13	32.6 m	10.2 m	51.9°	40.9°	9.5 m	8.7 m	-0.8 m	-8.4%
B-14	11.3 m	27.6 m	40.1°	41.2°	9.5 m	9.5 m	0.0 m	0.0%
B-15	10.6 m	27.6 m	48.2°	41.8°	9.5 m	8.1 m	-1.4 m	-14.7%
B-16	27.1 m	15.9 m	35.5°	42.2°	12.0 m	12.3 m	0.3 m	2.5%
B-17	27.3 m	9.8 m	41.1°	42.0°	9.5 m	11.2 m	1.7 m	17.9%
B-18	10.0 m	48.1 m	23.8°	46.7°	9.5 m	10.2 m	0.7 m	7.4%
B-19	9.9 m	52.1 m	22.8°	46.6°	9.8 m	11.0 m	1.2 m	12.2%
B-20	10.0 m	48.3 m	23.6°	46.5°	9.8 m	8.5 m	-1.3 m	-13.3%
B-21	32.1 m	9.9 m	27.9°	46.7°	9.8 m	15.5 m	5.7 m	58.2%
B-22	46.9 m	10.6 m	23.5°	46.9°	9.5 m	10.1 m	0.6 m	6.3%
B-23	9.7 m	22.8 m	22.0°	46.5°	9.5 m	9.5 m	0.0 m	0.0%
B-24	10.9 m	38.4 m	11.1°	50.4°	9.5 m	8.8 m	-0.7 m	-7.4%
B-25	11.2 m	35.8 m	13.3°	50.5°	9.5 m	11.5 m	2.0 m	21.1%
B-26	47.6 m	10.0 m	8.1°	50.2°	9.5 m	11.0 m	1.5 m	15.8%
B-27	31.6 m	9.1 m	10.0°	50.5°	9.5 m	10.8 m	1.3 m	13.7%
Mean and standard deviation							0.9 ± 1.5 m	9.2 ± 16.2%

TABLE III
RESULTS FOR BUILDINGS WHERE THE STRUCTURAL SHAPE OF THE ACTUAL BUILDING DOES NOT MATCH OUR
RECTANGULAR FLAT- OR GABLE-ROOF MODELS (CATEGORY C) FOR THE HORIZONTAL AIRBORNE SCENE

Building	Width	Length	ϕ	Local θ	Height	Estimate	Difference	Relative Difference
C-1	25.5 m	10.8 m	21.1°	44.0°	12.5 m	11.2 m	-1.3 m	-10.4%
C-2	38.8 m	12.0 m	31.4°	43.2°	14.5 m	12.9 m	-1.6 m	-11.0%
C-3	42.7 m	21.5 m	33.4°	45.7°	7.0 m	9.9 m	2.9 m	41.4%
C-4	36.1 m	13.1 m	38.4°	45.0°	12.6 m	10.5 m	-2.1 m	16.7%
C-5	12.0 m	60.9 m	27.4°	43.4°	12.5 m	11.1 m	-1.4 m	-11.2%
C-6	11.5 m	57.6 m	40.9°	41.3°	12.0 m	12.7 m	0.7 m	5.8%
C-7	11.1 m	56.5 m	27.3°	45.4°	12.9 m	8.2 m	-4.7 m	-36.4%
Mean and standard deviation							-1.1 ± 2.4 m	-5.5 ± 24.1%

Each of the building groups C-5, C-6, and C-7 consists of three row houses with similar dimensions, which are not arranged in a perfect row, but are slightly staggered. In these cases, we did not estimate the height for each building in the group separately but considered a group as one individual flat- or gable-roof building. This implies that we do not model the correct footprint of the building group for the simulations but approximate it by one rectangle. In Fig. 14(c) and (d), we show the building group C-6 consisting of three gable-roof buildings in the orthophoto and SAR image, respectively. In this situation, the estimated height is 0.7 m higher than the true height. The gable-roof building group C-7 has the buildings positioned significantly staggered to each other. Hence, the approximation as single gable-roof building is quite rough, which can be seen in the significant underestimation by 4.7 m.

The overall mean and standard deviation for this category is -1.0 ± 2.4 m. This is somewhat less accurate than that in the other two categories but still demonstrates that the method is relatively robust with respect to the structural assumption of the buildings.

4) Quality of Height Estimation: In order to detect outliers and to derive a representative overall assessment of the accuracy of the height estimation procedure, we carry out a statistical analysis of the results for the 33 buildings of categories A and B. Since category C only contains buildings which do not fit the considered models, we do not take them into account for the assessment of the overall accuracy.

Fig. 15(a) shows a normal quantile–quantile-plot (Q-Q-Plot) for the estimation differences for the 33 buildings of categories A and B. It highlights that the differences are normal distributed, with some outliers above 4 m. By using the Chauvenet's criterion [pp. 166–168, 54] to detect statistical outliers from the set of observations, we identified the estimations for the buildings A-4, B-2, and B-21 as outliers.

Buildings A-4 and B-2 were overestimated by 4.4 and 4.2 m, respectively, since they are largely surrounded by tall trees. This
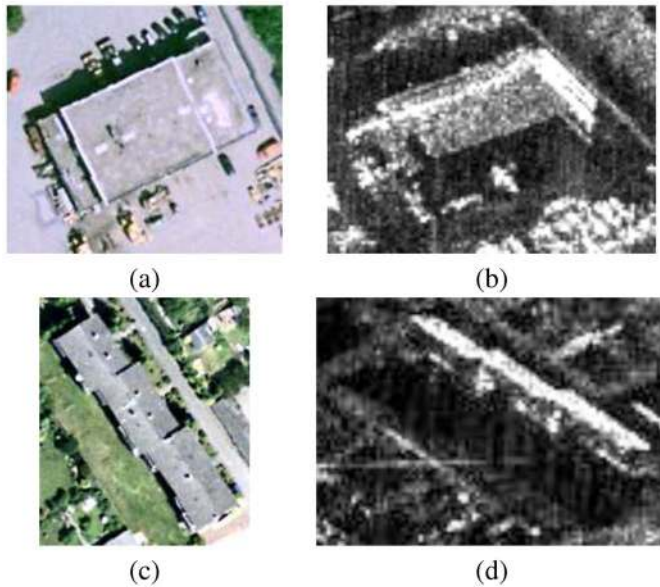
Fig. 14. (a) Building C-3 in orthophoto. (b) Building C-3 in SAR image (viewing direction from top). (c) Building C-6 in orthophoto. (d) Building C-6 in SAR image (viewing direction from top). (Orthophoto: © Geobasisdaten: Landesvermessungsamt NRW, Bonn, 2007; SAR image: © Intermap Technologies GmbH, 2003).
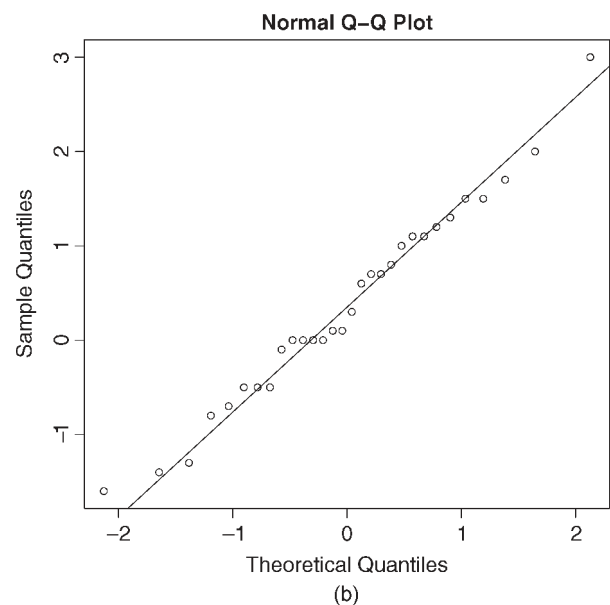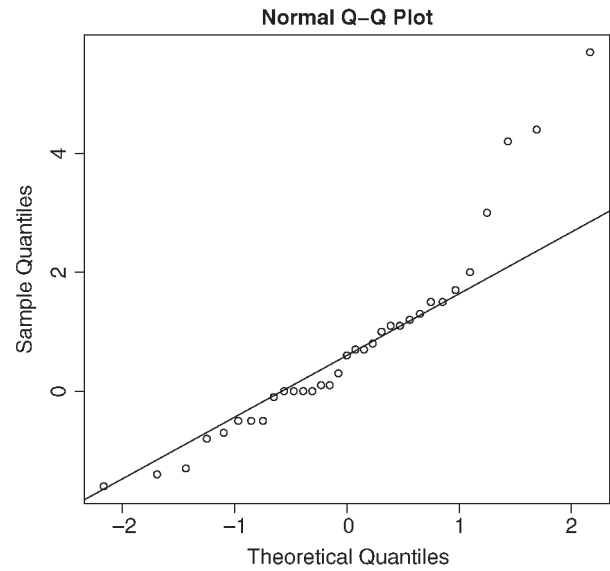

Fig. 15. (a) Q-Q-Plot for the 33 buildings of categories A and B. (b) Q-Q-Plot for the buildings of categories A and B, where three outliers were removed.

leads to a SAR signature which is different from the signature of a building not affected by objects in the surrounding (which can be observed in Fig. 10(a) by comparing the signature of B-2, for example, to the signature of building B-4). Since we do not model trees in the simulation procedure, they have an impact on the accuracy of the height estimation.

Building B-21 was overestimated by 5.7 m since it is surrounded by a relatively smooth surface, giving it a low backscatter, similar to the shadow. Hence, the matching function does not capture well the edge of the shadow region of the building signature in the actual SAR image, leading to the overestimate in height. Hence, the shadow region of a building seems to be important for the method to estimate the correct height.

Fig. 15(b) shows the Q-Q-Plot of the estimation differences where the three outliers were removed, which shows a good correspondence of the set with the normal distribution. The mean of the reduced set is 0.4 m, which demonstrates that the method has no significant preference for over- or undersegmentation, while the standard deviation of 1.0 m highlights the good estimation performance of the proposed approach.

5) Effects of Trees: Backscattering from trees positioned near a building tends to superimpose on the backscattering signature of the buildings and therefore affect the accuracy of the height estimation. We analyze buildings A-3 and A-4, which are very similar and have identical viewing configurations. The amount and the density of trees are similar for both buildings, whereas the relative locations of the trees are different [see Fig. 9(a) and (b)]. Since the sensor images the buildings from the top of the image, the majority of trees, which are in the immediate surrounding of building A-3, are located in its layover area. For building A-4, instead, the majority of the trees which are close to the building are located behind

the building, affecting its shadow area. The estimation errors of −0.5 m (A-3) and 4.4 m (A-4) confirm that the shadow area of the building plays an important role in the height estimation.

The density of trees surrounding the building weighs on the accuracy of the height estimation as well. Consider the buildings B-2, B-4, and B-5, which are three gable-roof buildings with equal heights, located close to each other and with the same orientation toward the sensor. Building B-4 has only some trees in the front, while building B-5 has a higher density of trees in the front and some additional trees in the back, which are not as close and dense as for building B-2, which is completely surrounded by trees. The height estimation for B-4 shows a difference to the actual height of 1.0 m, a difference of 3.0 m for building B-5, and 4.2 m for building B-2,

TABLE IV
ACTUAL AND MINIMUM DISTANCES BETWEEN BUILDINGS
A-1 AND A-4 FOR THE VERTICAL AIRBORNE SCENE

Buildings	Δ_{act}	Δ_{min}
A-1 $\leftrightarrow$ A-2	22.7 m	26.6 m
A-2 $\leftrightarrow$ A-3	19.8 m	23.0 m
A-3 $\leftrightarrow$ A-4	19.7 m	21.2 m

TABLE V
ACTUAL AND MINIMUM DISTANCES BETWEEN BUILDINGS A-1 AND A-4
FOR THE ASCENDING AND DESCENDING TERRASAR-X SCENES

Buildings	Δ_{act}	Δ_{min} ascending	Δ_{min} descending
A-1 $\leftrightarrow$ A-2	22.7 m	25.5 m	26.2 m
A-2 $\leftrightarrow$ A-3	19.8 m	22.4 m	24.3 m
A-3 $\leftrightarrow$ A-4	19.7 m	20.3 m	20.9 m

demonstrating the limitation of the height estimation method in situations that do not conform with the model assumptions.

B. Vertical Airborne Scene

The vertical scene of the dual-aspect data set was acquired such that the angle between the flight paths of the horizontal and vertical scenes is 84.5°. Hence, if the aspect angle of a building was defined in the horizontal scene by the angle between the azimuth direction and the short wall of the building, it will be given in the vertical scene by the angle between the azimuth direction and the long wall of the building (see also Fig. 3). The local incidence angle of a building might change significantly between the horizontal and vertical scenes depending on the location of the building. In the extreme case, the same building might be measured in one scene with 28° incidence angle, while it is measured in the other scene with 52°.

For category A, the height estimation procedure resulted in a mean error of 0.8 m with a large standard deviation of 5.4 m. The reason for the low accuracy is buildings A-1, A-2, A-3, and A-4 which have estimation errors ranging from −7.5 to 7.5 m. This is caused by the fact that these buildings are too close to each other, so that condition (11) is not fulfilled. A part of the shadow area of building A-1 is superimposed by the layover area of building A-2. The actual and minimum distances between the buildings A-1 and A-4 are reported in Table IV.

The height estimation for the 27 buildings in category B resulted in a mean error of −0.4 m with a standard deviation of 1.89 m. These values are in the same order of magnitude as the ones from the horizontal scene.

For category C, the mean value of the error is 0.0 m with a standard deviation of 3.2 m. The reason for the higher standard deviations for categories B and C with respect to the horizontal scene is the incidence angle, which is, on average, 44.5° for the horizontal scene and 37.8° for the vertical scene (considering buildings in categories B and C). The lower mean incidence angle for the vertical scene causes that the shadow areas are smaller with respect to the ones in the horizontal scene, confirming again the relative importance of the shadow feature for the height estimation.

C. Results for TerraSAR-X Scenes

The goal of the analysis of the TerraSAR-X scenes is mainly to investigate the effect of using a lower resolution spaceborne image compared to an airborne image. The ascending and descending scenes were acquired approximately from a north–south and a south–north orbit, respectively, which are, from an orientation point of view, quite similar to the flight path from which the vertical airborne scene was measured

(see Fig. 8). In fact, the vertical airborne and the descending TerraSAR-X scenes were both acquired with a right-looking sensor from a similar north-to-south path. The ascending TerraSAR-X scene was also acquired with a right-looking sensor but from a south–north track, so that the buildings were measured from nearly the opposite side with respect to the descending TerraSAR-X and the vertical airborne scene. This implies that a building was measured with similar aspect angles throughout the ascending and descending TerraSAR-X and the vertical airborne scenes. The incidence angle varies within each of the TerraSAR-X scenes by only 1° over the swath, while there is a difference of 3° between the ascending and descending TerraSAR-X scenes (see Section IV).

In the ascending TerraSAR-X scene, the mean and the standard deviation of the difference in height estimation for category A are −3.3 and 5.8 m, respectively. The results for the descending TerraSAR-X scene show a similar mean of −3.4 m with a standard deviation of 4.3 m. The low accuracy for this category has the same reason as for the buildings in category A in the vertical airborne scene. The buildings A-1–A-4 are located too close to each other so that the scattering effects of different buildings overlap with each other (see Section V-B). In Table V, we summarize the minimum distances required according to (11) and compared these to the actual values. Even though Δ_{min} between buildings A-3 and A-4 is quite similar to Δ_{act}, the height estimate of A-4 is imprecise due to a high density of trees in the immediate surrounding of the two buildings.

The mean values for category B for the ascending and descending scenes are 1.9 and −0.5 m, respectively. The standard deviations are 3.1 and 3.4 m, respectively. The fact that they are very similar in both scenes points out a constant height estimation accuracy.

For the ascending scene, the estimation procedure for category C resulted in a mean value of −2.2 m and a standard deviation of 4.6 m. Those figures are −0.8 and 1.3 m, respectively, for the descending scene. Since they are in the same order of magnitude as for categories A and B, the structural differences from the basic building assumptions are maybe less critical in lower resolution VHR spaceborne data.

The results for the TerraSAR-X data show that meter resolution VHR SAR data are not sufficient to get an accurate height estimate for the building dimensions that were investigated in this paper. Nevertheless, if the height of a single floor of a building is approximately known, the method permits the estimation of the number of floors of the building. This information can be of use, for instance, to distinguish between different types of buildings, such as residential housing, apartment buildings, industrial buildings, or skyscrapers.

TABLE VI
OVERVIEW OF HEIGHT ESTIMATION ERRORS FOR THE DIFFERENT BUILDING CATEGORIES AND ANALYZED SCENES.
A* AND B* SHOW THE VALUES FOR FLAT- AND GABLE-ROOF BUILDINGS, RESPECTIVELY, WHERE BUILDINGS
WHICH WERE IDENTIFIED AS OUTLIERS AND WHICH DO NOT FIT THE MODEL ASSUMPTIONS ARE REMOVED

Category	Width range	Length range	Height range	Airborne		TerraSAR-X	
				Horizontal	Vertical	Ascending	Descending
A (6) *A**	12.1 m - 15.8 m	35.8 m - 45.7 m	5.1 m - 12.5 m	0.3 ± 2.1 m -0.5 ± 0.6 m	0.8 ± 5.4 m -	-3.3 ± 5.8 m -	-3.4 ± 4.3 m -
B (27) *B**	9.7 m - 15.9 m	16.9 m - 52.1 m	9.5 m - 12.0 m	0.9 ± 1.5 m 0.6 ± 1.1 m	-0.4 ± 1.9 m -0.2 ± 1.7 m	1.9 ± 3.1 m 1.3 ± 2.8 m	-0.5 ± 3.4 m -0.8 ± 2.4 m
C (7)	10.8 m - 21.5 m	25.5 m - 60.9 m	7.0m - 14.5 m	-1.0 ± 2.4 m	0.0 ± 3.2m	-2.2 ± 4.6 m	-0.8 ± 1.3 m

D. Computational Aspects of Method

The computational complexity of the proposed approach is still significant. A typical performance of the full height estimation is illustrated with the run-time results for building A-2 estimate using the airborne horizontal scene. The platform used is a PC with two dual-core 2.8-GHz Intel Xeon CPUs and 3-GB RAM running the 32-bit Linux operating system (Ubuntu 8.04). Considering an explicit hypothesis generation from 3.0 to 20.0 m in 0.1-m steps (171 hypotheses), the simulation process takes 67 min, while the matching procedure takes 16 min.

In order to avoid simulating a building at a certain viewing configuration more than once, we store all simulation results in a relational database management system (RDBMS). If the combination of a building and viewing configuration is retrievable from the RDBMS, then the simulation is simply loaded from the database. Otherwise, a new simulation is triggered and added to the database. In this way, a library of SAR building signatures is generated over time, decreasing, with increasing database size, the number of simulations needed per estimation cycle. An alternative solution to speed up the simulation process, which could be seamlessly combined with the RDBMS, may use a very fast simulator based on a graphical processing unit, as proposed, for instance, in [55], achieving simulation run times on the order of milliseconds.

The matching procedure for a single building is a linear process (executed on a single CPU), in which the position $(\vec{s})$ for which the best coregistration for one hypothesis is found can be used as initialization for coregistering the subsequent hypothesis. However, additional building matches can be run in parallel. We use a clustered computing environment with 32 CPU cores. This leads to a gain in performance which is roughly proportional to the number of CPU cores available, apart from some minor overhead due to the task distribution in the cluster.

VI. DISCUSSION AND CONCLUSION

In this paper, we have proposed a novel concept for building height estimation from single VHR SAR detected images and tested it on a representative set of residential area urban structures in Dorsten. The approach is based on a hypothesis generation–rendering–matching procedure, in which a series of building hypotheses with varying heights is rendered by a SAR imaging simulator and the results are matched with the actual scene. The estimated height is given by the hy-

pothesis whose simulation matches best with the actual scene. The simulator is designed to calculate effects related to the SAR geometry without modeling the exact radiometry, since the use of detailed electromagnetic scattering models would imply the need for extensive *a priori* knowledge about the roughness parameters and dielectric constants of the surfaces in the scene. Such detailed parameters are generally not available in real-world operational scenarios. The MI approach for matching model and observation is well suited in this context because it is sensitive to the spatial arrangement of features rather than to the absolute radiometry of the scattering effects.

We have demonstrated the efficiency and generic nature of the proposed concept using dual-aspect airborne and ascending and descending TerraSAR-X VHR SAR scenes, all covering the same test area. A test data set made up of 40 buildings, containing flat- and gable-roof buildings at different viewing configurations (i.e., various aspect and incidence angles), was used. To evaluate the robustness of the method with respect to the simplified assumptions on building structure, we also included in the test data set buildings that only partially met our assumed rectangular flat- or gable-roof building models.

In Table VI, we list a summary of the accuracies for the three categories of buildings achieved in the different scenes. Considering the results for categories A* and B* (buildings which were categorized as flat- or gable-roof buildings, excluding the buildings whose results were identified as outliers or which do not match the model assumptions) for the two submeter resolution airborne data, the standard deviation of the height estimation is 1.4 m, which means that the method has a good overall estimation quality. The corresponding mean difference between estimation and actual height is 0.1 m, indicating that the proposed method has no tendency to over- or underestimate the height. The overall standard deviation of the buildings in category B* in the two TerraSAR-X scenes is 2.8 m. This shows that, for meter resolution VHR SAR data, the method can only provide rough height estimates, which can be used, for instance, to estimate the number of floors of the buildings. Such information is still of interest to characterize urban landscapes.

The detailed analysis of the categories A and B highlighted that the method can handle buildings with these two structural types. The results for the buildings which do not meet our assumptions (category C) show that the method is able to tolerate some degree of deviation from the assumptions, with

the estimation results getting less accurate for higher degrees of simplification. This means that, with a reliable information on the footprint and the type of building, an accurate estimate of the height can be achieved, while in the case of a limited availability of this information, the method can still provide a rough estimate.

The approach is shown to be insensitive to the aspect angle of a building. This is an important characteristic because buildings in urban settings are typically not oriented in a systematic way. Airborne SAR has a higher flexibility to acquire imagery for multiple aspects, although this is usually costly. For satellite SAR, aspect angles in imagery are limited by the ascending and descending orbits of the satellite.

The analysis showed that the approach favors larger incidence angles. This is explained by the shadow areas, which become larger with increasing (shallower) incidence angles, suggesting an important role of the shadow area for an accurate estimate of the height. Furthermore, the results demonstrated that large trees, whose backscattering interferes with the SAR shadow area of a building, decrease the accuracy of the method, which confirms the relatively important role of the shadow feature. One of the reasons for this is the approximation of the layover area by a homogeneous area. However, in reality, it is rather heterogeneous due to windows, balconies, and other structures present at the front wall of a building having different materials and composing several smaller corner reflectors. Hence, the similarity between the simulated and the actual layover area is lower compared to the similarity between a simulated and an actual shadow area. The importance of the shadow areas could be relaxed by modeling the building facade more accurately using facade grammar approaches [56]. The drawback of this would be the increased complexity of the building model and the need for additional *a priori* information on the building facades. Regarding the presence of trees or other disturbing objects in the neighborhood of a building, it would be relatively straightforward to log them at the footprint-capturing stage, using VHR optical orthocorrected imagery. That information would then be useful to filter height estimation results.

The method assumes that buildings are isolated. The rendering procedure does not consider interferences between different buildings, which arise if they are positioned close enough so that, for example, their shadow and layover areas are not separated any more but superimposed in a mixed area. This imposes the constraint that a building needs to have a minimum distance to a neighboring building (only if they are in the same azimuth position) so that no backscattering interferences from the different buildings occur. If this constraint is not fulfilled, the height estimation for the building is not accurate. The minimum distance between two buildings depends on the height of the neighboring buildings and on the local incidence angle: 1) The higher the buildings, the larger the minimum distance, and 2) the more shallow the incidence angle, the larger the minimum distance and vice versa. Hence, we suggest to acquire the VHR SAR imagery for rural and medium dense urban areas with a shallow incidence angle. In dense urban areas with low buildings instead, the data should be measured with a steep incidence angle, which relaxes the min-

imum distance constraint at the cost of a decreased estimation accuracy. The approach is not suitable for dense urban areas with high buildings.

In order to generate the hypotheses, we require information (i.e., footprint and type of building) derived from ancillary data. This information can either be provided as cadastral maps or can be directly extracted from the SAR image or other VHR optical data. With the growing global availability of VHR data from urban areas, these requirements, while demanding, appear to be realistic.

We want to stress that we addressed in this paper an automatic information extraction scenario that is capable of dealing with different types of buildings at various viewing configurations. The proposed method was designed with a minimum number of constraints and minimal requirements on the data. Taking into account the ambitious objective of this paper and the fact that no *a priori* information on the height of buildings is used, we believe that the achieved quality of the estimation results is reasonable. Moreover, this paper has yielded first quantitative evidence of what can be expected from the new meter resolution spaceborne SAR sensors in terms of automatic information extraction in urban settings.

We are currently extending the use of the method to merge the different height estimates from the different aspects in the multiaspect data set. This aims at improving the accuracies for buildings that are affected by neighboring objects in one, but not in another aspect. Furthermore, we plan to apply this method in damage assessment scenarios to detect buildings which are structurally damaged or completely destroyed.

References

[1] M. Eineder, H. Breit, T. Fritz, B. Schättler, and A. Roth, "TerraSAR-X SAR products and processing algorithms," in *Proc. IEEE IGARSS*, Seoul, Korea, 2005, vol. 7, pp. 4870–4873.

[2] G. F. De Luca, G. Marano, M. Piemontese, B. Versini, F. Caltagirone, G. Casonato, A. Coletta, and M. De Carlo, "Interoperability, expandability and multi mission-sensor COSMO-SkyMed capabilities," in *Proc. IEEE IGARSS*, Barcelona, Spain, Jul. 23–27, 2007.

[3] A. R. Brenner and L. Roessing, "Radar imaging of urban areas by means of very high-resolution SAR and interferometric SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2971–2982, Oct. 2008.

[4] A. J. Bennett and D. Blacknell, "Infrastructure analysis from high resolution SAR and InSAR imagery," in *Proc. 2nd GRSS/ISPRS Joint Workshop Remote Sens. Data Fusion Over Urban Areas*, Berlin, Germany, May 22–23, 2003, pp. 230–235.

[5] T. Balz and N. Haala, "SAR-based 3D reconstruction of complex urban environments," *Int. Arch. Photogramm. Remote Sens.*, vol. 34, pp. 181–185, 2003. [Online]. Available: http://www.ifp.uni-stuttgart.de/publications/2003/balz alsdd2003.pdf

[6] F. Tupin, "Extraction of 3D information using overlay detection on SAR images," in *Proc. 2nd GRSS/ISPRS Joint Workshop Remote Sens. Data Fusion Over Urban Areas*, Berlin, Germany, May 22–23, 2003, pp. 72–76.

[7] P. Gamba, B. Houshmand, and M. Saccani, "Detection and extraction of buildings from interferometric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 1, pp. 611–618, Jul. 2000.

[8] P. Gamba and B. Houshmand, "Digital surface models and building extraction: A comparison of IFSAR and LIDAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 4, pp. 1959–1968, Jul. 2000.

[9] C. Tison, F. Tupin, and H. Maître, "A fusion scheme for joint retrieval of urban height map classification from high-resolution interferometric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 2, pp. 496–505, Feb. 2007.

[10] E. Simonetto, H. Oriot, and R. Garello, "Rectangular building extraction from stereoscopic airborne radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2386–2395, Oct. 2005.

[11] U. Soergel, E. Michaelsen, A. Thiele, E. Cadario, and U. Thoennessen, "Stereo analysis of high-resolution SAR images for building height estimation in cases of orthogonal aspect directions," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 5, pp. 490–500, Sep. 2009.

[12] R. Hill, C. Moate, and D. Blacknell, "Estimating building dimensions from synthetic aperture radar image sequences," *IET Radar, Sonar Navig.*, vol. 2, no. 3, pp. 189–199, Jun. 2008.

[13] F. Xu and Y.-Q. Jin, "Automatic reconstruction of building objects from multiaspect meter-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pp. 2336–2353, Jul. 2007.

[14] A. Thiele, E. Cadario, K. Schulz, U. Thoennessen, and U. Soergel, "Building recognition from multi-aspect high-resolution InSAR data in urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3583–3593, Nov. 2007.

[15] S. Sauer, L. Ferro-Famil, A. Reigber, and E. Pottier, "Multi-aspect POLInSAR 3D urban scene reconstruction at L-band," in *Proc. 7th EUSAR*, Friedrichshafen, Germany, Jun. 2008, vol. 3.

[16] H. Oriot and H. Cantalloube, "Circular SAR imagery for urban remote sensing," in *Proc. 7th EUSAR*, Friedrichshafen, Germany, Jun. 2008, vol. 3.

[17] G. Franceschetti, A. Iodice, and D. Riccio, "A canonical problem in electromagnetic backscattering from buildings," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1787–1801, Aug. 2002.

[18] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, vol. 2, *Radar Remote Sensing and Surface Scattering and Emission Theory*. Norwood, MA: Artech House, 1982.

[19] G. Franceschetti, R. Guida, A. Iodice, D. Riccio, and G. Ruello, "Building feature extraction via a deterministic approach: Application to real high resolution SAR images," in *Proc. IEEE IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, pp. 2681–2684.

[20] H. Hammer, T. Balz, E. Cadario, U. Soergel, U. Thoennessen, and U. Stilla, "Comparison of SAR simulation concepts for the analysis of high resolution SAR data," in *Proc. 7th Eur. Conf. Synthetic Aperture Radar*, 2008, pp. 213–216.

[21] H.-J. Mametsa, F. Rouas, A. Berges, and J. Latger, "Imaging radar simulation in realistic environment using shooting and bouncing rays technique," in *Proc. SPIE Conf. SAR Image Anal. Model. Tech. IV*, F. Posa, Ed., Feb. 2002, vol. 4543, pp. 34–40.

[22] G. Franceschetti, A. Iodice, D. Riccio, and G. Ruello, "SAR raw signal simulation for urban structures," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1986–1995, Sep. 2003.

[23] G. Margarit, J.-J. Mallorqui, J. M. Rius, and J. Sanz-Marcos, "On the usage of GRECOSAR, an orbital polarimetric SAR simulator of complex targets, to vessel classification studies," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 12, pp. 3517–3526, Dec. 2006.

[24] G. Margarit, J. J. Mallorqui, and C. Lopez-Martinez, "GRECOSAR, a SAR simulator for complex targets: Application to urban environments," in *Proc. IEEE IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, pp. 4160–4163.

[25] U. Soergel, U. Thoennessen, and U. Stilla, "Iterative building reconstruction from multi-aspect InSAR data," in *Proc. ISPRS Working Group III/3 Workshop—3-D Reconstruction From Airborne Laserscanner and InSAR Data*, Dresden, Germany, Oct. 8–10, 2003, vol. XXXIV.

[26] D. Brunner, G. Lemoine, and L. Bruzzone, "Extraction of building heights from VHR SAR imagery using an iterative simulation and match procedure," in *Proc. IEEE IGARSS*, Boston, MA, Jul. 7–11, 2008, pp. IV-141–IV-144.

[27] "Building height retrieval from airborne VHR SAR imagery based on an iterative simulation and matching procedure," in *Proc. SPIE Conf. Remote Sens. Environ. Monit. GIS Appl. Geology VIII*, Cardiff, U.K., Sep. 15–18, 2008, p. 711 00F.

[28] A. Thiele, E. Cadario, K. Schulz, and U. Thoennessen, "Feature extraction of gable-roofed buildings from multi-aspect high-resolution InSAR data," in *Proc. IEEE IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, pp. 262–265.

[29] Y. Dong, B. Forster, and C. Ticehurst, "Radar backscatter analysis for urban environments," *Int. J. Remote Sens.*, vol. 18, no. 6, pp. 1351–1364, Apr. 1997.

[30] D. Brunner, L. Bruzzone, A. Ferro, J. Fortuny, and G. Lemoine, "Analysis of the double bounce scattering mechanism of buildings in VHR SAR data," in *Proc. SPIE Conf. Image Signal Process. Remote Sens. XIV*, Cardiff, U.K., Sep. 15–18, 2008, vol. 7109, pp. 710 90Q-1–710 90Q-12.

[31] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, Jan. 1965.

[32] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, May 13, 1983.

[33] A. K. Fung, *Microwave Scattering and Emission Models and Their Applications*. Boston, MA: Artech House, 1994.

[34] K. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propag.*, vol. 14, no. 3, pp. 302–307, May 1966.

[35] G. Oller, P. Marthon, and L. Rognant, "Correlation and similarity measures for SAR image matching," in *Proc. 10th Int. Symp. Remote Sens.*, Barcelona, Spain, Sep. 8–12, 2003.

[36] C. Barat, C. Ducottet, and M. Jourlin, "Pattern matching using morphological probing," in *Proc. 10th ICIP*, Barcelona, Spain, 2003, pp. 369–372.

[37] J. S. De Bonet and A. Chao, "Structure-driven SAR image registration," in *Proc. SPIE Conf. Algorithms Synthetic Aperture Radar Imagery V*, E. Zelnio, Ed., Sep. 1998, vol. 3370, pp. 109–119.

[38] W. G. Eppler, S. M. Louie Marcus, J. Petersen, and D. W. Paglieroni, "Fast normalized cross-correlation of complex gradients for autoregistration of multi-source imagery," in *Proc. Amer. Soc. Photogramm. Remote Sensing Annu. Meeting*, Washington, DC, May 22–26, 2000.

[39] J. Zheng, X. Bai, and X. Zhao, "SAR image matching based on fractal theory," in *Proc. 7th ICSP*, Beijing, China, 2004, pp. 1965–1968.

[40] F. Chen, H. Zhang, and C. Wang, "A novel feature matching method in airborne SAR image registration," in *Proc. IEEE IGARSS*, Seoul, Korea, Jul. 25–29, 2005, pp. 4722–4724.

[41] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[42] B. Wessel, M. Huber, and A. Roth, "Registration of near real-time SAR images by image-to-image matching," in *Proc. Photogramm. Image Anal.*, U. Stilla, Ed., Munich, Germany, Sep. 19–21, 2007, p. 179.

[43] D. Brunner, G. Lemoine, and L. Bruzzone, "Height estimation of man made structures using hybrid VHR optical and SAR imagery," in *Proc. EARSeL Joint Workshop, Remote Sens.—New Challenges of High Resolution*, C. Juergens, Ed., 2008, pp. 186–193.

[44] J. Lewis, "Fast normalized cross-correlation," in *Proc. Vision Interface*, 1995, pp. 120–123.

[45] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," in *Proc. Inf. Process. Med. Imag.*, 1995, pp. 263–274.

[46] P. Viola and W. Wells, "Alignment by maximization of mutual information," in *Proc. 5th ICCV*, 1995, pp. 16–23.

[47] H. Xie, L. E. Pierce, and F. T. Ulaby, "Mutual information based registration of SAR images," in *Proc. IEEE IGARSS*, Toulouse, France, Jul. 21–25, 2003, pp. 4028–4031.

[48] F. Maes, A. Collignon, D. Vandeermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.

[49] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Berlin, Germany: Springer-Verlag, 2003.

[50] M. Schwaebisch and J. Moreira, "The high resolution airborne interferometric SAR AeS-1," in *Proc. 4th Int. Airborne Remote Sens. Conf. Exhib.*, Ottawa, ON, Canada, Jun. 21–24, 1999, p. 8.

[51] A. Lopes, E. Nezry, R. Touzi, and H. Laur, "Structure detection and statistical adaptive speckle filtering in SAR images," *Int. J. Remote Sens.*, vol. 14, no. 9, pp. 1735–1758, Jun. 1993.

[52] F. Cellier, H. Oriot, and J.-M. Nicolas, "Introduction of the mean shift algorithm in SAR imagery: Application to shadow extraction for building reconstruction," in *Proc. Earsel Workshop 3D Remote Sens.*, Porto, Portugal, Jun. 10–11, 2005.

[53] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognit.*, vol. 32, no. 1, pp. 71–86, Jan. 1999.

[54] J. R. Taylor, *An Introduction to Error Analysis*, 2nd ed. Sausalito, CA: Univ. Sci. Books, 1997.

[55] T. Balz and U. Stilla, "Hybrid GPU based single- and double-bounce SAR simulation," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 10, pp. 3519–3529, Oct. 2009.
[56] N. Ripperda and C. Brenner, "Data driven rule proposal for grammar based facade reconstruction," in *Proc. PIA*, Munich, Germany, Sep. 19–21, 2007, vol. 36.

Dominik Brunner (S'07) received the Diploma in technical computer science from the University of Applied Sciences Hof, Hof, Germany, in 2004.

From 2004 to 2006, he was a Software Engineer with SAP AG, Walldorf, Germany. Since 2006, he has been with the European Commission Joint Research Centre, Ispra, Italy, and also with the Remote Sensing Laboratory, Department of Information Engineering and Computer Science, University of Trento, Trento, Italy. His research interests include the field of remote sensing, pattern recognition, and image processing in support to damage assessment and emergency response, in particular the analysis of very high spatial resolution SAR data of urban areas.

Guido Lemoine (S'92–M'95–SM'08) is an Agricultural Engineer by training (1987, Soil Science, Wageningen University, Wageningen, The Netherlands). His main expertise is in applied remote sensing, first developed as a research topic (microwave backscattering of soils) and later as a commercial activity in a cofounded remote sensing and GIS consultancy. Since he joined the European Commission Joint Research Centre, Ispra, Italy, in 1997, he has further developed his remote sensing and informatics expertise in agricultural statistics and subsidy control applications, in fishery monitoring, and, since 2006, in civil security applications. His current work focuses on integration of very high resolution optical and SAR data in crisis response and collaborative geospatial analysis and fast computing methods for near real time mapping of crisis event impact.

Lorenzo Bruzzone (S'95–M'98–SM'03) received the M.S. degree (*summa cum laude*) in electronic engineering and the Ph.D. degree in telecommunications from the University of Genoa, Genoa, Italy, in 1993 and 1998, respectively.

From 1998 to 2000, he was a Postdoctoral Researcher with the University of Genoa. In 2000, he joined the University of Trento, Trento, Italy, where he is currently a Full Professor of telecommunications. He teaches remote sensing, pattern recognition, and electrical communications. He is the Head of the Remote Sensing Laboratory, Department of Information Engineering and Computer Science, University of Trento. His current research interests are in the area of remote-sensing image processing and recognition (analysis of multitemporal data, feature extraction and selection, classification, regression and estimation, data fusion, and machine learning). He conducts and supervises research on these topics within the frameworks of several national and international projects. He is an Evaluator of project proposals for many different governments and scientific organizations. He is the author (or coauthor) of 73 scientific publications in referred international journals, more than 130 papers in conference proceedings, and 7 book chapters. He is a Referee for many international journals and has served on the Scientific Committees of several international conferences.

Dr. Bruzzone is a member of the Managing Committee of the Italian Inter-University Consortium on Telecommunications and a member of the Scientific Committee of the India–Italy Center for Advanced Research. Since 2009, he has been a member of the Administration Committee of the IEEE Geoscience and Remote Sensing Society. He is also a member of the International Association for Pattern Recognition and of the Italian Association for Remote Sensing (AIT). He was a Guest Editor of a Special Issue of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING on the subject of the analysis of multitemporal remote-sensing images (November 2003). He was the General Chair and Cochair of the First and Second IEEE International Workshop on the Analysis of Multi-temporal Remote-Sensing Images (MultiTemp) and is currently a member of the Permanent Steering Committee of this series of workshops. Since 2003, he has been the Chair of the SPIE Conference on Image and Signal Processing for Remote Sensing. From 2004 to 2006, he served as an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. Currently, he is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He ranked first place in the Student Prize Paper Competition of the 1998 IEEE International Geoscience and Remote Sensing Symposium (Seattle, July 1998). He was a recipient of the Recognition of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Best Reviewers in 1999.

Harm Greidanus received the Ph.D. degree in astrophysics from Leiden University, Leiden, The Netherlands, in 1989, on the subject of kinematics and multispectral observations of supernova remnants.

After that, he worked in remote sensing research at the TNO Physics and Electronics Laboratory, The Hague, The Netherlands, until 2003 and joined the European Commission Joint Research Centre, Ispra, Italy. He has worked mostly with imaging radar (ground based, airborne, and satellite) but also with optical and infrared, mostly for maritime applications (vessel detection, bathymetry, surface waves, and currents) but also for land mapping.