

© 2012 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Title: Automatic Detection and Reconstruction of Building Radar Footprints from Single VHR SAR Images

This paper appears in: IEEE Transactions on Geoscience and Remote Sensing

Date of Publication: 2013

Author(s): Adamo Ferro, Dominik Brunner, Lorenzo Bruzzone

Volume:51, Issue: 2

Page(s): 935-952.

DOI: 10.1109/TGRS.2012.2205156

Automatic Detection and Reconstruction of Building Radar Footprints from Single VHR SAR Images

Adamo Ferro, Dominik Brunner, *Member, IEEE*, Lorenzo Bruzzone, *Fellow, IEEE*

Abstract—The spaceborne synthetic aperture radar (SAR) systems Cosmo-SkyMed, TerraSAR-X and TanDEM-X acquire imagery with very high spatial resolution (VHR), supporting various important application scenarios, such as damage assessment in urban areas after natural disasters. To ensure a reliable, consistent and fast extraction of the information from the complex SAR scenes, automatic information extraction methods are essential. Focusing on the analysis of urban areas, which is of prime interest of VHR SAR, in this paper we present a novel method for the automatic detection and 2D reconstruction of building radar footprints from VHR SAR scenes. Unlike most of the literature methods, the proposed approach can be applied to single images. The method is based on the extraction of a set of low-level features from the images and on their composition to more structured primitives using a production system. Then, the concept of semantic meaning of the primitives is introduced and used for both the generation of building candidates and the radar footprint reconstruction. The semantic meaning represents the probability that a primitive belongs to a certain scattering class (e.g., double bounce, roof, facade) and has been defined in order to compensate for the lack of detectable features in single images. Indeed, it allows the selection of the most reliable primitives and footprint hypotheses on the basis of fuzzy membership grades. The efficiency of the proposed method is demonstrated by processing a 1-m resolution TerraSAR-X spotbeam scene containing flat- and gable-roof buildings at various settings. The results show that the method has a high overall detection rate and that radar footprints are well reconstructed, in particular for medium and large buildings.

Index Terms—Building detection, building reconstruction, production system, very high spatial resolution (VHR), synthetic aperture radar (SAR), urban areas, remote sensing.

I. INTRODUCTION

IN the last decade, very high spatial resolution (VHR) spaceborne remote sensing sensors (e.g., QuickBird, Worldview-2, Cosmo-SkyMed) acquiring data with meter or sub-meter resolutions became widely available. These data have the potential to be employed for various important application scenarios, such as the monitoring of changes in urban areas [1], [2], the characterization of urban areas (e.g., slum mapping) [3], [4], the surveillance of the effects of violent conflicts [5], and the crisis management after natural disasters (e.g., earthquakes) [6], [7]. For the latter application scenario, spaceborne VHR synthetic aperture radar (SAR) sensors, such as Cosmo-SkyMed [8] and TerraSAR-X [9], are of particular

interest, due to their independence on the solar illumination and the relative insensitivity to the weather conditions [10], [11].

One of the main drawbacks of VHR SAR is the complexity of the images, mainly owing to the speckle effect and the side-looking geometry of the SAR sensor, hampering the interpretation of the data by non-SAR experts. This is especially true for urban areas, where the data are mainly characterized by layover, multibounce and shadowing effects of the buildings. Therefore, to support the widespread usage of VHR SAR, robust automatic information extraction methods are essential.

Different techniques for building detection and reconstruction from VHR SAR images have been presented in literature. For instance, Soergel *et al.* [12] proposed an iterative technique for building reconstruction from interferometric SAR (InSAR) data which is based on the detection of the combined occurrence of a strong scattering line and a shadow area in correspondence of an elevated region. Cellier *et al.* [13] presented a building reconstruction technique for InSAR data based on building hypothesis management. The developed method uses a tree of hypotheses, which is simplified according to a set of semantic rules. Thiele *et al.* [14] proposed an approach to building detection which uses orthogonal multi-aspect InSAR images. The approach is based on the detection of edges and their combination to building footprints. A method for the extraction of buildings and the estimation of their height from stereoscopic airborne radar images was presented by Simonetto *et al.* in [15], while in [16] a building extraction method using dual-aspect SAR data was presented. An algorithm for building reconstruction from multi-aspect polarimetric SAR (PolSAR) images was presented by Xu and Jin [17]. The polarimetric information is exploited by employing an edge detector effective on polarimetric images. The retrieved edges are then parameterized by means of the Hough transform to generate the building footprint hypotheses. Wang *et al.* [18] developed a method for the detection of buildings from single-aspect PolSAR data combining edge and area features with Markov random fields. Hill *et al.* [19] presented a semi-supervised method for the estimation of building dimensions in VHR SAR temporal scenes based on the analysis of the shape of building shadows. Another method based on shadow analysis which exploits InSAR data and is suitable for high or isolated buildings was proposed by Tison *et al.* [20]. A building detection method using an orthophoto and an InSAR image based on conditional random fields is presented in [21]. Techniques for the 3D reconstruction of buildings using VHR optical data for the 2D building footprint reconstruction and a single VHR SAR scene for the building height extraction were presented in [22] and [23].

A. Ferro and L. Bruzzone are with the Department of Information Engineering and Computer Science, University of Trento, Via Sommarive 5, I-38123, Trento, Italy, e-mail: adamo.ferro@disi.unitn.it, lorenzo.bruzzone@ing.unitn.it.

At the time of the research, D. Brunner was also with the Department of Information Engineering and Computer Science, University of Trento, Via Sommarive 5, I-38123, Trento, Italy, e-mail: dominik@brunneronline.org

Manuscript received Month day, year; revised Month day, year.

All the above-mentioned works addressed the problem of building detection and reconstruction in VHR SAR images by relying on the availability of ancillary or multi-sensor data (e.g., optical imagery), polarimetric SAR, interferometric SAR, or multi-dimensional airborne data which implies that the area under investigation is imaged more than once with different viewing configurations (changed incidence and/or aspect angle). This represents a limitation for application scenarios with stringent timing restrictions that do not allow the acquisition of multi-dimensional SAR data (e.g., emergency response). For these reasons, research on the detection and extraction of buildings from single VHR SAR data is important. To our knowledge, only very few papers addressed the problem of building detection with one single meter resolution SAR images only. One of the few related works using single VHR SAR images was presented by Quartulli and Datcu [24], and was based on a stochastic geometrical model and a *a posteriori* probability maximization (MAP). Recently, a method for L-shape building footprint extraction from single SAR images was proposed in [25]. This method fails in the detection of buildings if they do not show L-shaped returns. Moreover, it considers only bright lines and discards other relevant features, such as bright areas and shadows.

In this paper, which generalizes and extends the work presented in [26] and [27], we propose a novel method for the detection and reconstruction of building radar footprints from detected VHR SAR images. Unlike most of the literature methods, it can be applied to single images. Moreover, it is suitable to be used with data acquired by currently operational spaceborne SAR sensors. In this context, radar footprint refers to the characteristic scattering signature of buildings in SAR. The method integrates the concepts of basic feature extraction and their composition to more structured primitives using a production system [28], [29]. In order to compensate for the lack of detectable features in single images, the concept of semantic meaning of the primitives is introduced and used to generate building candidates and reconstruct radar footprints. The semantic meaning represents the probability that a primitive belongs to a certain scattering class (e.g., facade, double bounce) and allows the selection of the most reliable primitives and footprint hypotheses on the basis of fuzzy membership grades.

The main novelties and advantages of the proposed method are: i) the capability to accurately detect the radar footprint of individual buildings using only one SAR scene without the need for ancillary data, ii) the possibility to estimate the reliability of the detected features and footprint hypotheses through a set of fuzzy functions, iii) the flexibility to handle gable- and flat-roof buildings at different sizes and at various settings, and iv) the expansibility of the approach, which allows the definition of new scattering classes and rules according to specific image characteristics or user requirements. These characteristics make the approach valuable for supporting different application scenarios, e.g., damage assessment after crisis events and change detection in urban areas. In addition, as shown later in the paper, the method is suited to the implementation on computer clusters, thereby making it possible almost-real-time applications.

Some steps in our proposed method have similarities with existing work. For instance, the method presented in [13] is based on hypothesis management. Since their approach relies on sub-meter resolution InSAR data the hypotheses are based on different information (combination of height and topology) compared to ours (presence and semantic meaning of scattering features). Moreover, we introduce a way to quantitatively evaluate the hypotheses to automatically select the best one, which is missing in [13]. Similar to our approach, the method in [24] uses the layover and double-bounce features for the reconstruction of buildings. However, this method is based on a global MAP estimation using Monte Carlo methods, while the approach proposed in this paper exploits also the shadow information and introduces the concept of semantic meaning and membership grade for each primitive and footprint hypothesis. Moreover, such a work was intended as a tool for the investigation of the limits and merits of information extraction from single images, and was not optimized for building reconstruction purposes.

The radar footprint map extracted with the proposed method can be used to derive different information, such as the build-up presence index. It can also be used as a feature in the classification of the build-up areas (e.g., according to residential and commercial areas). Indeed, radar footprints in single SAR images lack the information about the exact dimensions (length, width, height) and the location of the 2D optical footprint of buildings. In order to derive them, the method could be combined with an iterative simulation and matching scheme as presented for instance in [22] for the building height extraction. In this context, the capability of the proposed method to extract the individual scattering contributions of a building in the SAR image could be used to improve the matching function as the simulator is also able to distinguish between the different contributions.

The remainder of this paper is structured as follows. In Section II we review the characteristics of flat- and gable-roof buildings in VHR SAR imagery. In Section III, we present the proposed methodology in detail, while Section IV discusses the processing of full VHR SAR scenes using a grid computing infrastructure. In Section V we demonstrate the performance and the properties of our approach by processing and analysing a large 1-meter resolution TerraSAR-X spotlight mode scene from Dorsten, Germany, which is characterized by different types of buildings at various settings. Finally, in Section VI, we draw the conclusions and give an outlook for future work.

II. MODELING OF BUILDING RADAR FOOTPRINTS IN SINGLE DETECTED VHR SAR IMAGES

The key characteristics of buildings in SAR are the layover, double-bounce, and shadowing effects which are caused by the side-looking and ranging properties of SAR sensors. To illustrate this, Fig. 1 shows a schematic view of the scattering profile of a simplified flat-roof building model. In this figure, the building in the middle, which is modeled as a rectangular box, is imaged by a sensor with incidence angle θ . The annotations a refers to backscattering from the ground surface surrounding (in this 2D figure before/behind) the building. *acd*

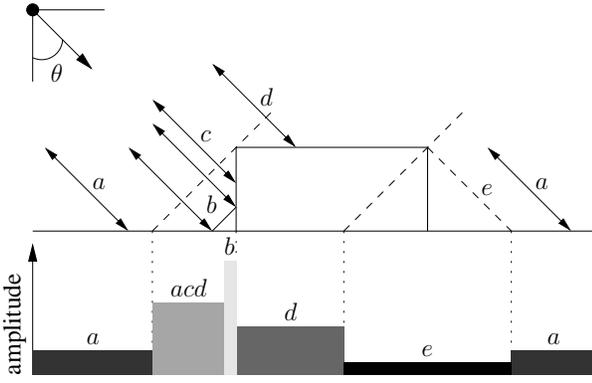


Fig. 1. Scattering model for a flat-roof building with viewing direction from left. The different gray areas at the bottom of the figure symbolize the amplitudes.

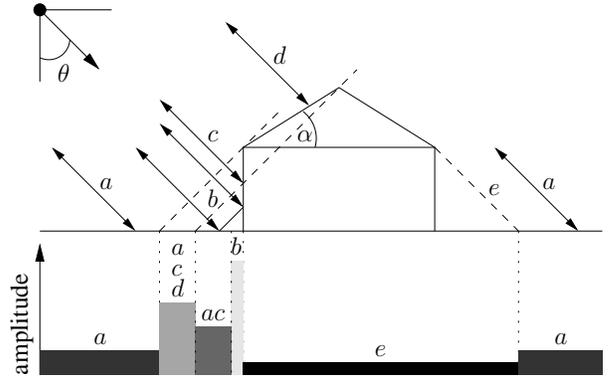


Fig. 3. Scattering model for a gable-roof building with viewing direction from left. Here, the roof inclination angle α is smaller than θ . The different gray areas at the bottom of the figure symbolize the amplitudes.

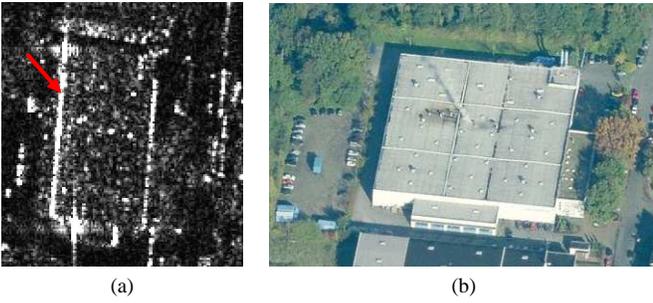


Fig. 2. Example of a flat-roof building. (a) Building in 1-meter resolution TerraSAR-X data with viewing direction from left (Infoterra). The double bounce line is highlighted with a red arrow. (b) The same building in an optical image (Microsoft).

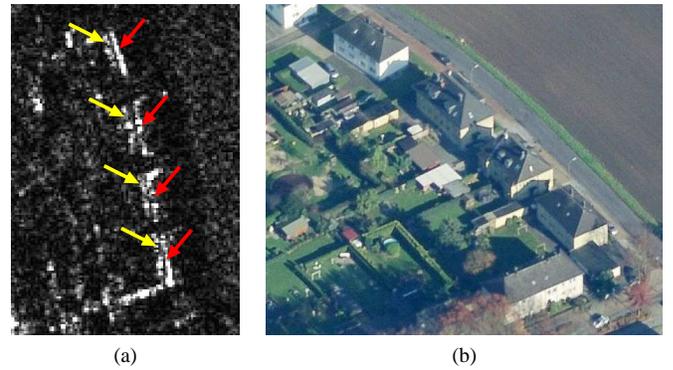


Fig. 4. Examples of gable-roof buildings with small aspect angles. (a) Buildings in 1-meter resolution TerraSAR-X data with viewing direction from left (Infoterra). The double-bounce and roof scattering lines are highlighted with red and yellow arrows, respectively. (b) The same buildings from (a) in an optical image (Microsoft).

denotes the layover area where scattering from the ground, from the vertical building front wall and from parts of the flat roof are superimposed since these parts have the same distance to the sensor. The vertical front wall and the surface area in front of the building compose a corner reflector resulting in the bright double-bounce effect b . The scattering area that is only characterized by scattering from the roof is denoted by d . The elevated building occludes parts of the surface behind the building from the radar beam, resulting in the shadow area e . This backscattering profile is flexible with respect to a number of parameters [22]. For instance, for very high buildings there is typically no area d as the part of the roof is entirely included in the layover area. An example of radar footprint of an industrial flat-roof building is shown in Fig. 2. The main scattering mechanisms are visible (i.e., layover, double bounce, scattering from roof and shadow). However, additional features appear (e.g., bright spots on the roof due to metallic structures). The figure also shows examples of interference due to other targets, in this case tall trees. In fact, both the layover and the shadow areas of the footprint are partially masked by the trees that surround the building.

For gable-roof buildings the theoretic scattering signature is slightly different [30], [31]. As shown in Fig. 3, the signature has a second bright scattering feature acd at the sensor close side resulting from direct backscattering from the roof. The extent and the strength of this feature depends on the

relationship between θ and the roof inclination angle α . For $\alpha = \theta$ the strength of this feature is maximum, whereas its extent is minimum. Moreover, we found that in actual 1-meter resolution TerraSAR-X and Cosmo-SkyMed data this second bright scattering area is also detectable for buildings with a high aspect angle (the angle between the building wall facing the sensor and the azimuth direction). This is illustrated in Fig. 4 and Fig. 5, where we show actual scattering signatures from gable-roof buildings with small and large aspect angles, respectively.

In Fig. 1 and Fig. 3 the double-bounce feature is very pronounced. A detailed analysis of the characteristics of the double bounce of buildings with actual TerraSAR-X data and theoretic electromagnetic scattering models presented in [32] showed that this feature has a significant dependency on the building aspect angle. The double bounce has a strong signature for buildings with low aspect angles (the wall facing the sensor is almost parallel to the azimuth direction). Then it decays significantly in a narrow range of aspect angles, while it drops moderately for larger aspect angles. The method presented in this paper will take into account this non-linear relationship between the strength of the double bounce and the aspect angle.



Fig. 5. Examples of gable-roof buildings with large aspect angles. (a) Buildings in 1-meter resolution TerraSAR-X data with viewing direction from left (Infoterra). The scattering lines due to the roofs are highlighted with yellow arrows. (b) The same buildings from (a) in an optical image (Microsoft).

III. PROPOSED TECHNIQUE FOR THE AUTOMATIC DETECTION AND RECONSTRUCTION OF BUILDING RADAR FOOTPRINTS

The proposed technique for the automatic detection and reconstruction of building radar footprints from single VHR SAR images is suited for meter-resolution data. Buildings are assumed to be approximately regular parallelepipeds, with rectangular base, or compositions of parallelepipeds. The minimum building size which can be handled by the algorithm depends on the specific building characteristics. As a reference, buildings with a base with a main side shorter than 10 m and a height lower than 5 m with no relevant scattering centers are likely to be not detected in meter-resolution images. The radar footprints corresponding to very tall buildings have a high probability to be detected. However, additional features and rules would be necessary (with respect to the algorithm specifications reported in this paper) in order to handle properly those situations. The algorithm does not require the buildings to be isolated. However, it may provide better results on isolated buildings. In fact, such buildings usually show a clear shadow feature, which is exploited by the algorithm to improve the detection performance. Very close buildings may be detected as single structures, as we will show in Section V.

The proposed technique is composed of six main steps: i) preprocessing and feature extraction, ii) generation of primitives, iii) analysis of primitives, iv) generation of building radar footprint hypotheses, v) selection of hypotheses, vi) 2D radar footprint reconstruction. Fig. 6 shows a block scheme representing the proposed processing chain. In the following we describe in detail each step. In this paper we present the algorithm optimized for the application to meter-resolution SAR images. However, the general structure of the algorithm is suitable to handle also higher resolution data. We highlight throughout the paper the modifications which would be necessary to apply the algorithm to sub-meter resolution images.

A. Preprocessing and Feature Extraction

In the preprocessing, the input image is first radiometrically calibrated. Although this step is not strictly necessary, it permits to define the algorithm parameters to be used with SAR images of different datasets and data products acquired by either the same or different sensors. Afterwards, the image is filtered with a Gamma MAP filter [33] in order to reduce the signal variability due to speckle. Both the unfiltered and filtered images are used by the algorithm. The basic features

composing building radar footprints in VHR SAR images are extracted from the calibrated image. According to the aforementioned assumptions on building shapes, these are bright linear features with different thicknesses, and dark areas. The former are usually related to double-bounce scattering or, as the line thickness increases, to layover areas, where the roof or the facade scattering may be dominant depending on the building characteristics. The latter are due to building shadows and low-return areas (e.g., roads, rivers, lakes). These features are sufficient to describe the main parts of a building radar footprint in meter-resolution images. However, as far as resolution increases, other scattering effects due to small structures become visible (e.g., point scatterers due to pipes on walls) and other types of features may be extracted to increase the detection performance of the algorithm. In the following, the techniques used for the extraction of bright linear features and dark areas are described in detail.

1) *Extraction of bright linear features*: The extraction of bright linear features is performed on the unfiltered image by means of the line detector proposed by Tupin *et al.* in [34]. This detector is based on a three-region sliding-window approach and is a well-known algorithm specifically developed for SAR images. In this paper we use as reference for the window size the dimension of the central region, and assume that the lateral regions have the same width and length (see Fig. 7). The length has been set to ten times the resolution of the image and 16 directions have been considered for the window. As we are interested in both thin and thick linear features, the detector is applied T times with different increasing window sizes w_t ($t = 1, \dots, T$). Each filtering is performed independently. The result of each filtering is a detection map, which is then thresholded, obtaining binary linear regions which thickness is related to w_t . Such regions are vectorized using a rectangular approximation. This is performed by approximating the region skeletons with lines and using such lines as the axis of rectangles of width w_t . The region skeletons are extracted according to [35]. In Fig. 8 we show an example of the detection on a meter resolution SAR image of an urban area using $w_t = 5$ m. The intermediate results are also shown. For each rectangle r , the local contrast value C_r is calculated on the filtered image as:

$$C_r = \left[\frac{1}{M_{\text{in}}} \sum_{i \in A_{\text{in}}} x_i \right] \cdot \left[\frac{1}{M_{\text{out}}} \sum_{i \in A_{\text{out}}} (1 - x_i) \right] \quad (1)$$

where A_{in} and A_{out} are the inner region of the rectangle and an outer thick border surrounding it, respectively. The thickness

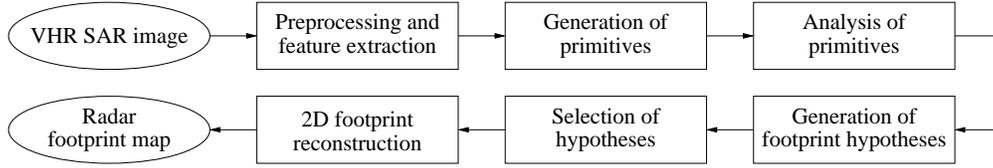


Fig. 6. Block scheme of the processing chain of the proposed technique for the automatic detection and reconstruction of building radar footprints in single VHR SAR images.

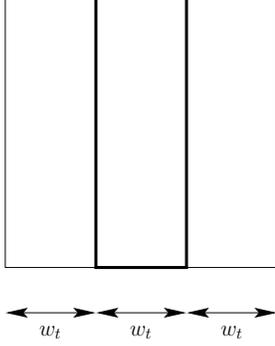


Fig. 7. Definition of the window used by line detector.

of A_{out} is defined as $\frac{w_t}{2}$. M_{in} and M_{out} are the number of pixels contained in the regions A_{in} and A_{out} , respectively; and x_i represents the pixel amplitude value normalized between 0 and 1. For the normalization the image amplitude dynamic range has been thresholded to cover the 99.5th percentile of the original image histogram in order to reduce the effect of very bright point scatterers. C_r is a measure of the contrast between the pixels contained in the rectangles and their surrounding. The higher the difference between the mean amplitude of the two regions, the higher the value of C_r . This measure has been proposed in [36] and has been used in [15] for the case of binary images.

As a result of the T filterings we obtain T vector maps containing rectangles corresponding to bright linear features with different thicknesses. These maps are thus merged in one map. It is possible that the same real bright objects are detected independently for different w_t , resulting in overlapping rectangles in the merged map. In order to reduce the number of rectangles, a downselection step is performed by means of a production net. For each combination of two rectangles (i, j) the net tests the following conditions: i) the width of the two rectangles is similar, and ii) the two rectangles overlap. Condition i) is met when:

$$|w_i - w_j| < \delta w_{\text{max}} \quad (2)$$

where w_i and w_j are the widths of the rectangles, and δw_{max} is an user-defined threshold (see Fig. 9). Condition ii) is fulfilled when:

$$A_{\cap} > A_i \cdot A_t \wedge A_{\cap} > A_j \cdot A_t \quad (3)$$

where A_i and A_j are the areas of the rectangles, $A_{\cap} = A_i \cap A_j$ (see Fig. 9), and A_t is a value belonging to the range $(0, 1)$ set by the user. When conditions (2) and (3) are fulfilled, the net discards the rectangle with the lowest

contrast, which is the rectangle associated to the lowest value of C_r .

For the choice of the values of δw_{max} and A_t , values on the order of 3 m and 0.5 are suggested, respectively. Moreover, in our experiments a number of $T = 7$ filterings using equally spaced w_t between 3 and 15 m has given a good detection of the linear bright features in the test images using a fixed threshold equal to 0.4 for all the considered w_t .

It is worth noting that this downselection step is not strictly necessary for the correct operation of the proposed technique. However, it greatly reduces the number of extracted bright linear features, thus improving the overall performance in terms of execution time and memory requirements of the technique.

2) *Extraction of dark areas*: Dark areas are extracted from the unfiltered image by means of mean shift clustering followed by a threshold operation, according to the approach proposed in [37]. This operation selects only the clusters with amplitude values lower than an user-defined threshold x_S . The extracted clusters are then vectorized and a simplification procedure is applied in order to reduce the number of vertexes describing their shape. Such simplification is not strictly necessary, but it allows the algorithm to work with simpler objects reducing the needed amount of memory. In order to select only the dark regions which are likely to be related to building shadows, the algorithm removes the regions which are not located in the sensor-far side of any bright linear feature (previously extracted). This is done by keeping only the dark areas which overlap with the predicted shadow area of the bright features. The predicted shadow area is determined by taking into account the viewing configuration of the SAR. The maximum range size l_S of the expected shadow area is set by the user. Schematic examples of predicted shadow area and selected dark area are shown in Fig. 10. The parameters of the mean shift clustering and the value of x_S have to be selected by analyzing the amplitude of sample pixels belonging to shadow regions in the SAR image. In our experiments, reasonable values for x_S were in the order of -13 – -11 dB.

B. Generation of Primitives

The goal of this step is to generate the primitives that will be used in the following steps as basis for the composition of building radar footprint hypotheses. Starting from the set of simple extracted bright linear features and dark areas, the algorithm merges adjacent features in order to compose bigger objects. This is done by a production system applied to the vector domain, after a conversion from slant range to ground range, and is aimed at compensating for errors in the feature

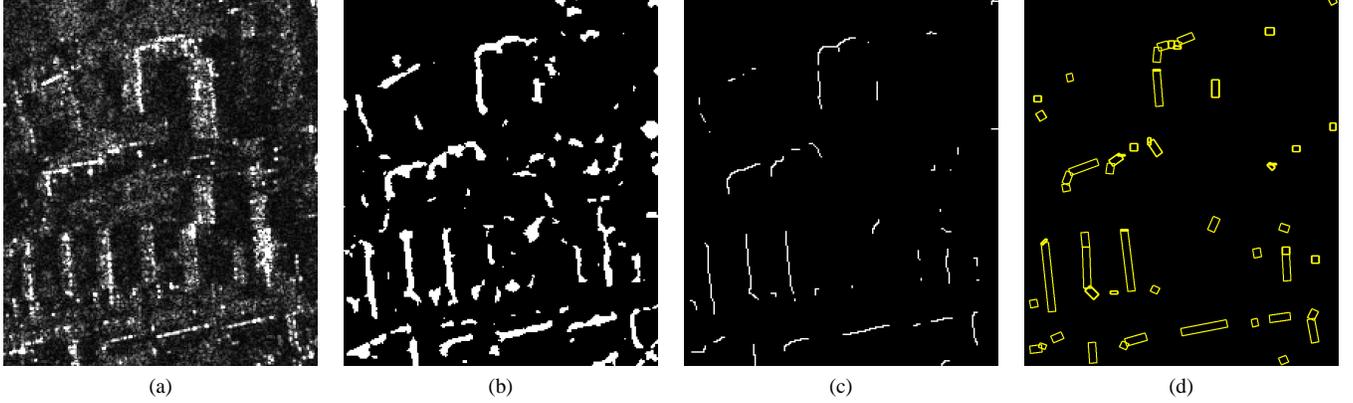


Fig. 8. (a) Meter-resolution TerraSAR-X image of an urban area (Infoterra). (b) Result of the line detection using $w_t = 5$ m. (c) Skeletons of the binary regions shown in (b). (d) Rectangles generated using the skeletons shown in (c).

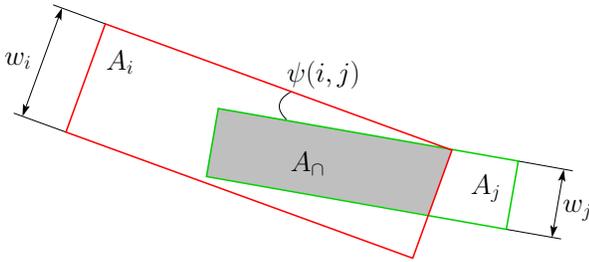


Fig. 9. Measures involved in the rectangle downselection described in the feature extraction step, and in the primitives generation step.

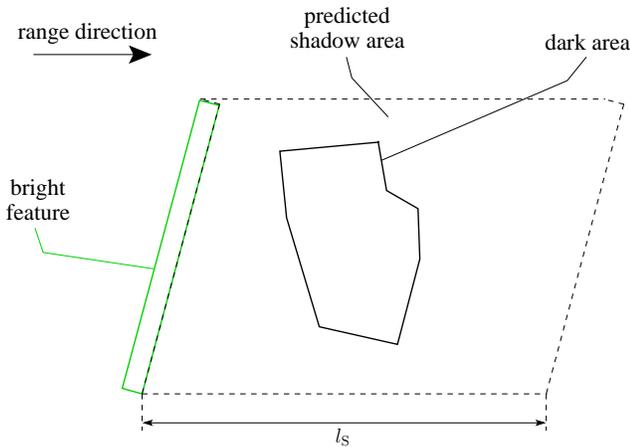


Fig. 10. Schematic examples of predicted shadow area and selected dark area.

extraction step. The conversion from slant to ground range allows us to define the parameters of the method in the ground domain, which is independent on the incidence angle and thus simpler to handle for an end-user. After their generation, composed objects are given as input to the production system. Therefore, multiple compositions with other simple or composed objects are possible. The set of objects and productions involved in the generation of primitives is shown in Fig. 12. The composition of dark areas is described by the production rules P_1 and P_2 . Such rules merge two dark areas when these are adjacent (e.g., their minimum distance is less than 2 m).

In this paper the merging is carried out by calculating the convex hull including the two original features. For the case of bright linear features, merged features are generated as new rectangles that have as principal axis the conjunction of the two extremes of the principal axes of the original features which have the largest relative distance (see Fig. 11). The width of the new rectangles is calculated as the weighted average of the widths of the original features, using as weights their length. The algorithm merges two bright features when the following conditions are fulfilled (P_3 and P_4): i) the features have similar widths, ii) their orientation is approximately the same, iii) the composed object has an orientation that is approximately the same of the original features. Condition i) is equivalent to (2). Condition ii) is fulfilled when $\psi(i, j) < \delta\psi_{\max}$, where $\psi(i, j)$ is the angle between the two linear bright features represented by the rectangles i and j (see Fig. 9), and $\delta\psi_{\max}$ is user-defined and indicates the maximum angle allowed between two features for which they are considered parallel. The value of $\delta\psi_{\max}$ should be on the order of 20° . Condition iii) is satisfied when:

$$\psi(\chi, i) < \delta\psi_{\max} \wedge \psi(\chi, j) < \delta\psi_{\max} \quad (4)$$

where χ is the rectangle corresponding to the composed bright linear feature. It is probable that in this step many bright primitives are generated. In order to reduce their number, a selection procedure as the one described in the previous subsection for bright linear features can be applied.

At the end of this step, for the whole set of simple and composed objects the algorithm stores a set of attributes regarding their size and position, and the amplitude features of the composing pixels (i.e., mean value, coefficient of variation). The set of simple and composed objects (with the related attributes) will be considered as set of primitives for the following steps.

C. Analysis of Primitives

This step aims at evaluating the semantic meaning of the primitives. Here we use the term semantic meaning to describe the membership grade of a certain primitive to belong to a pre-defined scattering class. Different scattering classes are related to different parts of building radar footprints. The choice of

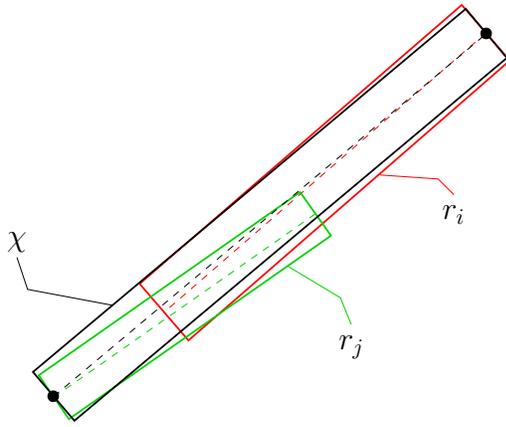


Fig. 11. Schematic example of the merging of two rectangles r_i and r_j into a new rectangle χ .

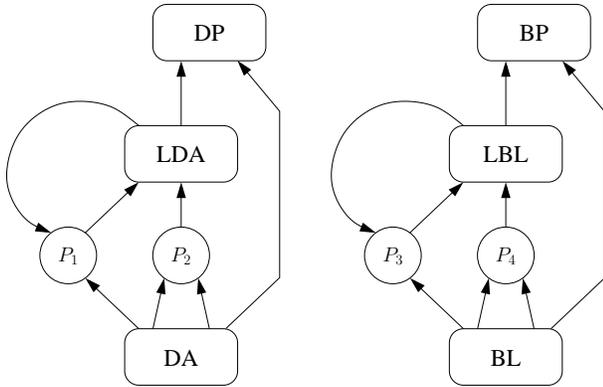


Fig. 12. Production net for the generation of dark primitives (DP) and bright primitives (BP). The inputs of the process are dark areas (DA) and bright linear features (BL), which are composed to large dark areas (LDA) and large bright linear features (LBL), respectively. The whole set of DA, LDA, BL and LBL are selected as primitives.

the set of semantic classes is related to the types of features extracted from the image and, thus, to the image resolution. For the bright primitives (i.e., the primitives obtained from bright linear features) we define four semantic classes: *general line*, *double bounce*, *roof*, and *facade*. For dark primitives (i.e., the primitives obtained from dark areas) only the class *shadow* has been defined. The membership grade of each primitive to belong to a certain semantic class is calculated on the filtered image according to membership functions (MFs) derived empirically for each semantic class. The MFs are functions of the primitive attributes and describe the membership grade of a primitive as a number in the range (0,1). The membership grades to belong to the different semantic classes are calculated independently. Thus, one primitive can have high membership grade for different classes at the same time. The different semantic meanings are managed by the proposed technique in the later stages of the processing chain.

The MFs are defined as a product of sigmoid functions. Each sigmoid factor depends on a specific attribute of the primitives. A generic sigmoid function is defined as follows

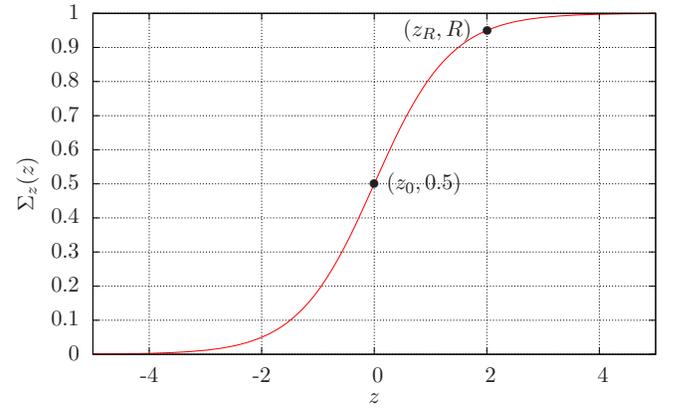


Fig. 13. Example of sigmoid function $\Sigma_z(z)$ defined according to (5). $z_R = 2$, $z_0 = 0$, $R = 0.95$.

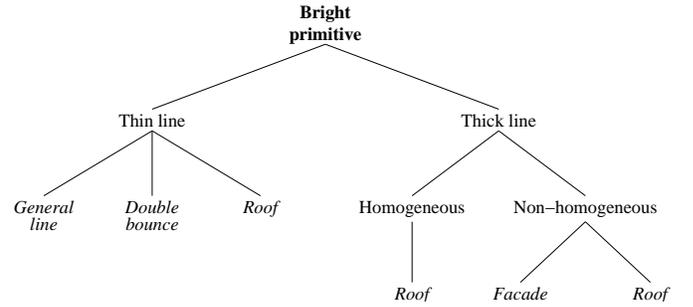


Fig. 14. Tree representing the semantic classes used in this paper for bright primitives.

[38]:

$$\Sigma_z(z) = \frac{1}{1 + e^{-\alpha_z(z-z_0)}} \quad (5)$$

$$\alpha_z = -\frac{\ln(1/R - 1)}{z_R - z_0} \quad (6)$$

where z indicates the attribute which constrains the function (e.g., the coefficient of variation of the amplitude of the pixels contained in the primitive), $\Sigma(z_0) = 0.5$ and $\Sigma(z_R) = R$. The function $\Sigma_z(z)$ gives values in the range (0,1). For each sigmoid function two parameters needs to be specified: the value of z for which the sigmoid returns a high likelihood R (z_R), and the value corresponding to the center of the sigmoid (z_0), implicitly setting the slope of the function. Fig. 13 shows an example of sigmoid function.

The MFs which relate bright primitives to the relative semantic classes are defined according to the tree shown in Fig. 14. The number of sigmoid functions composing the MF for a semantic class is smaller or equal to the number of branches which connect the root to the final leaf. In the following we describe in detail the MFs of each semantic class for both bright and dark primitives, by also suggesting the range of parameters which is most suited for the related scattering class. Unless otherwise stated, such values have been estimated by analyzing the scattering properties of a set of samples of the considered scattering classes manually selected on the meter-resolution TerraSAR-X input images used in this paper. As the images are calibrated, the suggested values related to pixel amplitude can be considered generally valid. In the case of

images acquired at different resolution and/or with a sensor with different characteristics, some of the values should be estimated again. In Section V-E the choice of the parameters is discussed more in detail.

1) *Bright primitives:*

General line: The membership grade of a primitive to the class *general line* depends only on its width. The MF is thus defined as

$$\text{MF}_{\text{GL}} = \Sigma_w^{\text{thin}}(w) \quad (7)$$

where $\Sigma_w^{\text{thin}}(w)$ gives a measure of the membership of the primitive to the high-level class *thin line*, which depends on the primitive width w . According to the definitions of (5) and (6), $\Sigma_w^{\text{thin}}(w)$ is controlled by the parameters w_R^{thin} and w_0^{thin} . The values of these parameters are chosen to give high values when w is small (e.g., $w_R^{\text{thin}} = 5$ m, $w_0^{\text{thin}} = 7$ m for meter-resolution images).

Double bounce: The double-bounce effect appears in VHR SAR images as relatively thin bright lines. It is more evident when the building wall is parallel to the azimuth direction, i.e., its aspect angle is close to zero (see Section II). The MF of the class *double bounce* is thus defined as follows

$$\text{MF}_{\text{DB}} = \Sigma_w^{\text{thin}}(w) \Sigma_\phi^{\text{DB}}(\phi) \quad (8)$$

where the term ϕ is the primitive aspect angle, and $\Sigma_\phi^{\text{DB}}(\phi)$ takes into account the dependence of the double-bounce effect on ϕ . $\Sigma_\phi^{\text{DB}}(\phi)$ has high values when ϕ is close to zero. In such a case, the MFs of the classes *general line* and *double bounce* give very similar values. Proper values for ϕ_R^{DB} and ϕ_0^{DB} are on the order of 10° and 30° , respectively. Such values have been chosen according to our previous studies about the double-bounce effect in VHR SAR images [32].

Roof: The class *roof* is the most specific, as it appears as leaf for every branch combination. This is due to the intrinsic uncertainty given by the fact that we are using only one VHR SAR image and that we are considering meter-resolution images. Indeed, the signature of a building roof could be either a thin line (e.g., in the case of gable-roof buildings with high aspect angle), or a homogeneous rectangular area (e.g., flat roof buildings), or a non-homogeneous rectangular area (e.g., flat roof buildings with metal structures on the roof, which are common for industrial buildings). Therefore, for the class *roof* the final membership grade is calculated as the maximum of the membership grades given by the three MFs corresponding to the three occurrences of the class in the tree. These are defined as:

$$\text{MF}'_{\text{R}} = \Sigma_w^{\text{thin}}(w) \quad (9)$$

$$\text{MF}''_{\text{R}} = \Sigma_w^{\text{thick}}(w) \Sigma_\sigma^{\text{hom}}(\sigma) \quad (10)$$

$$\text{MF}'''_{\text{R}} = \Sigma_w^{\text{thick}}(w) \Sigma_\sigma^{\text{non-hom}}(\sigma). \quad (11)$$

Finally, we obtain:

$$\text{MF}_{\text{R}} = \max \{ \text{MF}'_{\text{R}}, \text{MF}''_{\text{R}}, \text{MF}'''_{\text{R}} \}. \quad (12)$$

The definition of $\Sigma_w^{\text{thick}}(w)$ is complementary to that of $\Sigma_w^{\text{thin}}(w)$. As a requirement, to cover the whole possible range of primitive thicknesses it is necessary that $w_R^{\text{thin}} = w_R^{\text{thick}}$. This assures that any value of w is mapped either in the *thin*

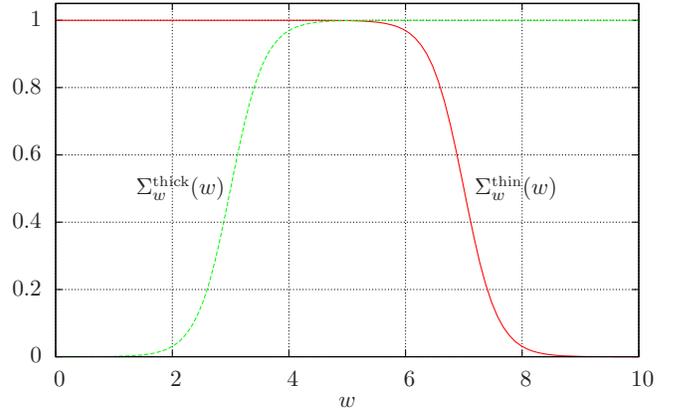


Fig. 15. Complementary MFs $\Sigma_w^{\text{thin}}(w)$ and $\Sigma_w^{\text{thick}}(w)$ used in this paper. $w_R^{\text{thin}} = w_R^{\text{thick}} = 5$, $w_0^{\text{thin}} = 7$, $w_0^{\text{thick}} = 3$, $R = 0.999$.

line or *thick line* classes with high membership grade (greater than R). Fig. 15 shows the behaviors of the complementary MFs $\Sigma_w^{\text{thin}}(w)$ and $\Sigma_w^{\text{thick}}(w)$ that are used in this paper. The same considerations hold for the definition of $\Sigma_\sigma^{\text{hom}}(\sigma)$ and $\Sigma_\sigma^{\text{non-hom}}(\sigma)$, which indicate the degree of membership of a primitive to the classes *homogeneous* and *non-homogeneous*, respectively. These refer to the homogeneity of the pixels contained in the primitive. The homogeneity is measured using as parameter the coefficient of variation σ of the pixels. Reasonable values for w_0^{thick} , $\sigma_R^{\text{hom}} = \sigma_R^{\text{non-hom}}$, σ_0^{hom} and $\sigma_0^{\text{non-hom}}$ are on the order of 2 – 3 m, 0.3 – 0.35, 0.45 – 0.55, and 0.15 – 0.3, respectively. Thanks to these constraints, the tree representing the semantic classes covers all the possible combinations of attributes taken into account in this paper. In the specific case of the class *roof*, (12) shows that the membership grade is always greater or equal to R^2 . This is in line with the aforementioned issue of the uncertainties related to the radar signature of building roofs.

Facade: As reported in the tree of Fig. 14, the semantic class *facade* includes primitives with a relevant width and which pixels have non-homogeneous values. This is the general scattering behavior of building facades, where returns coming from structures like windows or balconies (often made of metal) give a strong textured signature in the radar footprint. As a further constraint, the aspect angle of the building should not be too high (i.e., the building should not be perpendicular to the azimuth direction). Indeed, the facade scattering area in the radar footprint becomes smaller with increasing aspect angles. These factors are taken into account in the definition of the *facade* MF as follows:

$$\text{MF}_{\text{F}} = \Sigma_w^{\text{thick}}(w) \Sigma_\sigma^{\text{non-hom}}(\sigma) \Sigma_\phi^{\text{F}}(\phi) \quad (13)$$

where $\Sigma_\phi^{\text{F}}(\phi)$ models the effect of the building aspect angle ϕ by penalizing primitives with high aspect angles (e.g., $\phi_R^{\text{F}} = 70^\circ$ and $\phi_0^{\text{F}} = 80^\circ$). As mentioned at the beginning of this section, we do not include in our analysis very high buildings, for which the facade scattering area can have different characteristics.

2) *Dark primitives:* For dark primitives only the semantic class *shadow* has been defined. The MF of this class takes into account the mean and the coefficient of variation of the

pixels contained in the primitive. It is defined as:

$$MF_S = \Sigma_m^S(m) \Sigma_\sigma^{\text{hom}}(\sigma) \quad (14)$$

where $\Sigma_m^S(m)$ is the sigmoid functions depending on the pixel mean m . The MF is tuned in order to penalize dark primitives with high mean value and high coefficient of variation. Our experiments pointed out that reasonable values for $\Sigma_m^S(m)$ are $m_R^S \in (-14, -12)$ dB and $m_0^S \in (-9, -8)$ dB.

D. Generation of Building Radar Footprint Hypotheses

In this step the algorithm creates building radar footprint hypotheses starting from the set of primitives. The hypotheses are generated according to a set of rules and the process is performed by means of a production system. Fig. 16 summarizes the generation process. A footprint hypothesis is generated when i) two bright primitives, or ii) two bright primitives and one dark primitive, or iii) one bright primitive and one dark primitive are close each other and have a relative position compatible with the viewing configuration of the SAR sensor (i.e., dark primitives are located in the sensor-far side of bright primitives). The three cases are described by the productions P_5 , P_6 and P_7 of Fig. 16. The generation is thus based only on the vicinity criterion, and many hypotheses are usually created for the same actual building radar footprint. The vicinity is checked by measuring the minimum distance between the primitives. A proper value for the maximum distance allowed between two primitives is the value δd_0 used in the selection of hypotheses (see Section III-E). As it will be shown in Section III-E, on the one hand if the maximum distance allowed between two primitives is greater than δd_0 the probability that many low-score footprint hypotheses are generated is high. On the other hand, a distance threshold shorter than δd_0 would discard hypotheses which may be associated to high scores.

The order in which the bright primitives are aggregated is also taken into account, i.e., at least two hypotheses will be generated for each pair of bright primitives. The choice of using a maximum number of two bright primitives depends on the image resolution and on the types of features used in this paper. In meter-resolution images an average building radar footprint can be usually described effectively by the combinations considered in this paper. In the case more types of features are extracted from the image or decimeter-resolution images are used, more combinations of primitives become relevant.

E. Selection of Hypotheses

As mentioned in the previous subsection, many hypotheses are generated by aggregating the primitives. At this stage the algorithm selects only the most reliable hypotheses, which will be used in the next step as starting point for the 2D radar footprint reconstruction. Therefore, the output of this step is a map containing the detected (but not reconstructed) building radar footprints. This means that the output map is composed by footprint hypotheses which are still not refined.

The reliability of each hypothesis is evaluated on the basis of a score. The score is computed from the membership grades

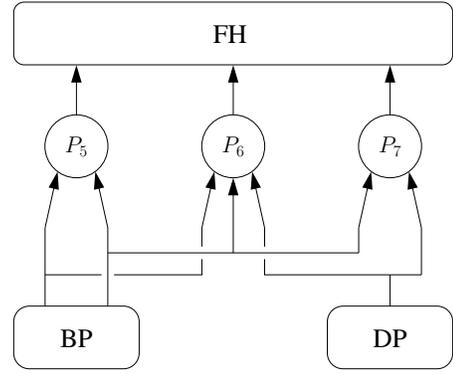


Fig. 16. Production net for the generation of building radar footprint hypotheses (FH) starting from the set of bright primitives (BP) and dark primitives (DP).

of the primitives composing each hypothesis. The general form of the score equation for a building radar footprint hypothesis h is given by:

$$S_h = N_h \max_{p,q} \{X_h(p,q)G_h(p,q)\} W_h \quad (15)$$

where N_h depends on the number of primitives composing the hypothesis, $X_h(p,q)$ and $G_h(p,q)$ are related to the relative position and to the membership grades of the bright primitives, respectively, and W_h depends on the membership grade and position of the dark primitive. (p,q) indicates the combination of the semantic class p of the first bright primitive and the class q of the second bright primitive. All these factors belong to the range $[0, 1]$. The overall value of S_h thus belongs to the same range. In the following we describe in detail each term of the equation:

- N_h : for the case presented in this paper, when the hypothesis h includes three primitives (i.e., the maximum number allowed), then $N_h = 1$. In the case one primitive is missing, it takes the value $N_h = N_h' < 1$, which is set by the user. This term is thus related to the reliability assigned by the user to the candidates composed by a non-complete set of primitives.
- $X_h(p,q)$: this term depends on the relative position of the bright primitives in the radar footprint hypothesis. In this paper, only the classes *general line* and *double bounce* are considered for the first bright primitive, and the classes *roof* and *facade* for the second bright primitive. It is worth noting that the technique also considers the case in which the bright primitives are switched, as in the hypotheses generation step different hypotheses are created taking into account also the order in which the bright primitives are aggregated. If only one bright primitive is present the value of $X_h(p,q)$ is 1. When two bright primitives are included in the hypothesis and the first bright primitive is closer to the SAR flight path than the second primitive, its value is 0. Indeed, for the considered cases, scattering from double bounce or any other linear scattering feature of a building (associated to the first bright primitive) cannot precede in range the scattering from the roof and from the facade (which are associated to the second bright primitive). At most,

the scattering area of double bounce and other lines are contained in that of roof and facade. When this condition is fulfilled, the value of $X_h(p, q)$ is calculated differently depending on the combination (p, q) . In detail, $X_h(p, q)$ is calculated as follows:

$$X_h(p, q) = \begin{cases} \Sigma_{\delta d}^{\text{close}}(\delta d_{fs}) \Sigma_{\delta \psi}^{\text{parallel}}(\delta \psi_{pa}) & p=\text{double bounce}, q=\text{facade} \\ \Sigma_{\delta d}^{\text{close}}(\delta d) \Sigma_{\delta \psi}^{\text{parallel}}(\delta \psi_{\min}) & \text{otherwise} \end{cases} \quad (16)$$

where δd_{fs} is the distance between the first bright primitive and the sensor-far side of the second bright primitive oriented in its principal direction, and $\delta \psi_{pa}$ is the angle between them (see Fig. 17). δd is the distance between the two bright primitives. Distances are measured in terms of minimum distance between the considered objects. The distance to the sensor is calculated considering an infinite line located outside the image with a position and angle compatible with the viewing configuration of the SAR. If one bright primitive overlaps with the other, $\delta d = 0$. $\delta \psi_{\min}$ is defined as:

$$\delta \psi_{\min} = \min \{ \delta \psi_{pa}, \delta \psi_{sa} \} \quad (17)$$

where $\delta \psi_{sa}$ is the angle between the first bright primitive and the secondary axis of the second bright primitive (see Fig. 17). The functions $\Sigma_{\delta d}^{\text{close}}(\delta d)$ and $\Sigma_{\delta \psi}^{\text{parallel}}(\delta \psi)$ give values close to 1 when their argument is small. The definition of $X_h(p, q)$ thus assures that its value is close to 1 when the bright primitives are both close and oriented parallel or perpendicularly to each other. For the combination *double bounce/facade* the condition is more strict and requires that the two primitives have their principal axis oriented in the same direction, and that the supposed double-bounce line is located at the sensor-far side of the facade scattering area (see Section II). Proper values for δd_R and δd_0 are on the order of 3 and 10 m, respectively. Regarding $\delta \psi_R$ and $\delta \psi_0$, values on the order of 10° and 30° are suggested.

- $G_h(p, q)$: this factor depends on the membership grades of the bright primitives composing the footprint hypothesis and on their size. When only one bright primitive is present, it reduces to $G_h(p)$ and its value is equal to the membership grade of the primitive to the class p . If two bright primitives are present, it is calculated in the following way:

$$G_h(p, q) = \frac{G'_h(p, q) + G''_h(p, q)}{2} \quad (18)$$

$$G'_h = \text{MF}_{1,p} \cdot \text{MF}_{2,q} \quad (19)$$

$$G''_h = \frac{A_1 \cdot \text{MF}_{1,p} + A_2 \cdot \text{MF}_{2,q}}{A_1 + A_2} \quad (20)$$

where $\text{MF}_{1,p}$ indicates the membership grade of the first bright primitive to the class p , and $\text{MF}_{2,q}$ is the membership grade of the second bright primitive to the class q . A_1 and A_2 are the areas of the first and of the second bright primitives, respectively. The definition of $G_h(p, q)$ permits to obtain reliable scores also for

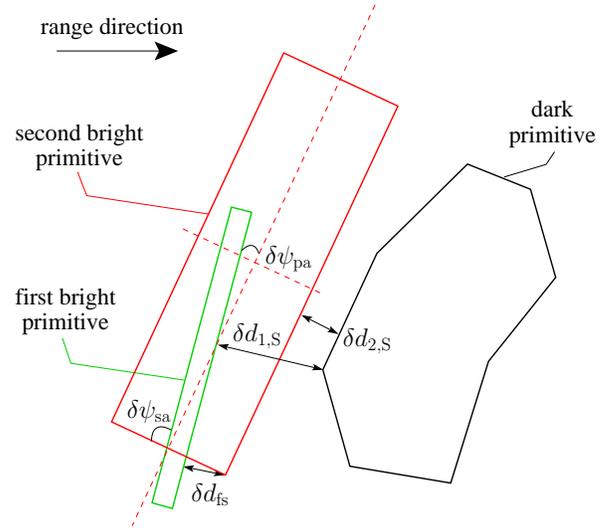


Fig. 17. Measures involved in the calculation of the term $X_h(p, q)$.

particular combinations of bright primitives. For instance, if one of the two bright primitives has a very low membership grade, the term $G'_h(p, q)$ becomes very small and the overall value of $G_h(p, q)$ will be low (in the limit, $G_h(p, q) = \frac{G''_h(p, q)}{2} \leq 0.5$). Instead, the term $G''_h(p, q)$ takes into account the area of the bright primitives. As a result, the value of $G''_h(p, q)$ depends more on the larger bright primitive.

- W_h : this term is function of the membership grade to the class *shadow* of the dark primitive, and on its position in the radar footprint hypothesis with respect to the bright primitives. It is calculated as:

$$W_h = \text{MF}_S \cdot \Sigma_{\delta d}^{\text{close}}(\min \{ \delta d_{1,s}, \delta d_{2,s} \}) \quad (21)$$

where $\delta d_{1,s}$ and $\delta d_{2,s}$ are the distances of the dark primitive from the first and the second bright primitive, respectively (see Fig. 17).

On the basis of the value of S_h , the algorithm deletes all the radar footprint hypotheses for which $S_h < S_{h,\min}$, where $S_{h,\min}$ is an user-defined threshold. After this first selection, many hypotheses with high values of S_h may still overlap in correspondence of actual building radar footprints (e.g., composed by different combinations of primitives). Therefore, the algorithm selects amongst the overlapping hypotheses only the one with the highest value of S_h .

F. 2D Radar Footprint Reconstruction

The 2D radar footprint reconstruction aims at refining the detection of both, the bright part and the dark part (if present) of the footprint hypotheses selected in the previous step. This is performed in order to reduce the effect of imprecisions coming from the feature extraction and primitive generation steps, and to provide reliable outputs which can be used as a starting point to estimate parameters of the buildings, such as their length, width and height (with the limitations imposed by the fact that only a single image is available). The result of this

procedure is thus the final map of the building radar footprints detected and reconstructed from the input VHR SAR image.

As a first step, the algorithm generates for each footprint hypothesis a best-fit rectangle which includes its bright primitives. If only one bright primitive is present, the best-fit rectangle and the bright primitive match. The local contrast C_r of the rectangle is calculated according to (1). Then, the rectangle is translated, rotated, expanded and shrunk with the goal to maximize C_r . The maximization is carried out using a Particle Swarm Optimization approach (PSO) [39], which is a well-known iterative method suited for the optimization of problems without *a priori* assumptions. A similar approach was applied in [15] for binary images and using a different optimization strategy. The rectangles which become smaller than the minimum sizes set in the previous steps of the algorithm are deleted. Moreover, it is possible that some rectangles move and overlap. Therefore, the algorithm deletes overlapping rectangles, and thus the corresponding footprint hypothesis, keeping only the rectangles associated to the hypotheses with the highest scores S_h .

A refinement procedure is carried out also for the dark part of the footprint hypothesis, when it is present. In fact, a good knowledge of the size of the shadow area of a building can be exploited for the retrieval of the building height [22]. The refinement aims at expanding the dark primitive on pixels with amplitude values similar to those of shadows in the sensor-far side area of the reconstructed bright primitive (i.e., in the neighborhood of the reconstructed bright primitive where it is expected the presence of a shadow region). To this end, the center of the dark primitive is used as seed for a region growing algorithm which, starting from an initial circular contour, stretches its border to fit the dark area around the seed. The chosen implementation is a level-set algorithm [40] which moves the contour by including the pixels which have amplitude values in the range $[0, m_R^S]$ (m_R^S has been defined in Section III-C). The resulting regions are cut in the azimuth direction in order to match the extension of the reconstructed bright part of the footprint hypothesis. Indeed, the reconstructed regions are associated to building shadow areas, which cannot be larger than the corresponding buildings in the azimuth direction. The size of the reconstructed dark areas in the range direction depends only on the radiometric measurements in the image. As the proposed technique uses as input only one VHR SAR image and no *a priori* information is available, it is not possible to detect the end of the shadow region by other means. This may lead to shadow areas which are longer than real shadows because of low scattering areas behind the buildings (e.g., roads, parking lots). This problem can be partially mitigated by imposing a maximum shadow range size l_S set by the user as in Section III-A2. Shadows longer than l_S are cut to l_S , and a flag is set to notice the user about the lower reliability of the reconstructed shadow.

IV. ANALYSIS OF LARGE VHR SAR SCENES

The technique proposed in this paper can be used as a preliminary step in many application scenarios, e.g., the detection of changes in urban areas aimed at the quick assessment of

damages after a natural disasters. For these applications it is important to process entire scenes in a fast manner. However, the processing chain described in the previous section is demanding both in terms of computation effort and memory requirements. This reduces the size of the input images that can be analyzed to a small subset of an actual VHR SAR scene, thus limiting the potential application of the method in real scenarios. In particular, the amount of resources required by the proposed technique depends directly on: i) the size of the input image, mainly for the parts of the algorithm based on image filtering and feature extraction (i.e., despeckling and line detection); and ii) the number of primitives and hypotheses generated through the processing chain. The latter is the most relevant factor that defines the complexity of the method. Indeed, the amount of required resources shows a non-linear dependence on the number of objects inserted in the production systems used in the processing chain. Although the number of primitives and hypotheses depends on the size of the input image, it also depends on the type of imaged area. As an example, two images of the same size covering an urban area and a rural area will produce a different number of primitives, with the greater number of primitives from the urban area.

In order to face these problems, we extended the algorithm to operate in a computer cluster infrastructure. In such a framework, the nodes in the cluster process different subsets of the input image in parallel. Each subset contains only few primitives, and thus also a reduced number of footprint hypotheses. This enables us to apply the proposed technique on large scenes in a fast way on state-of-the-art hardware. In Fig. 18 a block scheme of the considered simple architecture is presented. As a first step, the VHR SAR image is split into tiles. Every tile overlaps with its neighbors to assure that buildings located at the tile borders are detected and reconstructed properly at least in one tile. Then, the tiles are distributed across the nodes which independently execute the proposed method. Finally, the results for each tile are merged in order to generate the final radar footprint map for the entire input scene. When footprint hypotheses coming from different tiles overlap on tile borders, the algorithm selects the ones with the highest score S_h .

V. EXPERIMENTAL RESULTS

In this section we show the results obtained by applying the proposed methodology to a real meter-resolution large SAR image. After a brief description of the used dataset, we show and analyze qualitatively the results obtained on the whole image following the grid-computing approach described in Section IV. Then, we focus on two subsets of the image in order to assess quantitatively the accuracy of the method.

A. Dataset Description

The effectiveness of the proposed method has been tested on a TerraSAR-X image of the city of Dorsten, Germany. The image has been acquired in HH polarization in spotlight mode, resulting in a geometrical resolution of approximately 1.1 m \times 1.2 m (azimuth \times slant range). The incidence angle varies

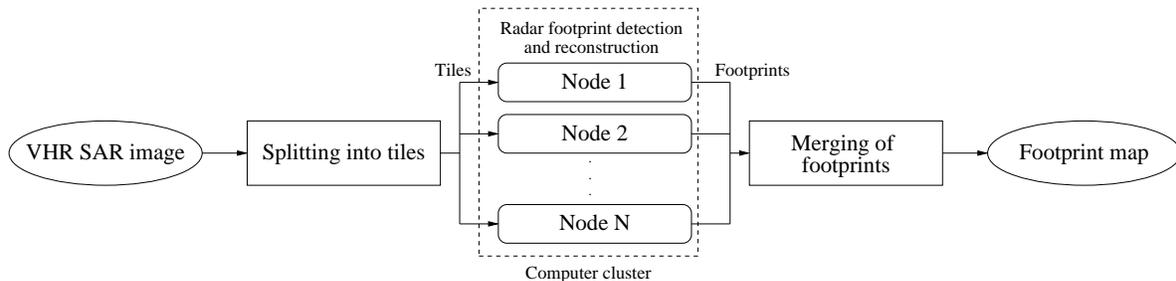


Fig. 18. Proposed computing architecture to perform the building detection and reconstruction method on large VHR SAR scenes.

between 50.3 and 51.0 degrees. The original scene has been cut to a subset of 2800×3712 pixels, covering an area of approximately 10 km^2 . The cut includes both urban and rural areas. Urban areas are characterized by both flat- and gable-roof buildings at various settings. Fig. 19 shows the SAR test image and an optical image corresponding to the same area taken from Google Maps.

B. Results on the Entire Scene

The proposed method has been run using the parameters reported in Table I, Table II and Table III. The values of such parameters have been chosen according to the guidelines given in Section III. The results obtained are shown in Fig. 20. The method shows in overall a high detection rate. False alarms are mostly related to the scattering from objects different from buildings (e.g., trees, garages) that show radar footprints similar to those of buildings. A particular case is represented by bridges, which have been also detected. Such structures can be easily masked, either using *a priori* information about the presence of rivers, or by extracting the rivers directly from the SAR scene [41]. The radar footprints of complex buildings which do not correspond to the rectangular model used in this paper are mostly detected with some reconstruction errors (e.g., the radar footprint has been split in more parts). In general, the proposed method detected and reconstructed quite precisely the radar footprints of medium- and big-size buildings that fulfill the rectangular model. Radar footprints of small adjacent buildings aligned in regular patterns are also detected, but in some cases are considered as belonging to a single building. Small buildings which do not show clear features are not detected by the method. However, considering the use of a single SAR image, the results can be considered qualitatively very satisfactory. Moreover, it is worth noting that if the proposed method is applied in order to derive indexes of the presence of buildings, reconstruction errors (i.e., split and merged buildings) do not represent a critical issue. In order to analyze quantitatively and in greater detail the results achieved by the proposed method, in the following we focus on two subsets of the test image.

C. Results on Subset 1

Fig. 21 shows the area corresponding to the subset 1 in both the SAR and optical images. This area is characterized by both flat- and gable-roof buildings with different sizes and orientations. In particular, the upper part of the image contains

TABLE I
PARAMETERS USED IN THE FEATURE EXTRACTION AND PRIMITIVE GENERATION STEPS IN THE EXPERIMENTS CARRIED OUT WITH THE PROPOSED TECHNIQUE.

Parameter	Value
T	7
w_1, \dots, w_7	3, 5, ..., 15
δw_{\max}	3
A_t	0.5
x_S	-12.2 dB
l_S	30 m
$\delta\psi_{\max}$	20°

TABLE II
PARAMETERS USED IN THE ANALYSIS OF PRIMITIVES STEP IN THE EXPERIMENTS CARRIED OUT WITH THE PROPOSED TECHNIQUE.

Parameter	Value
R	0.999
$w_R^{\text{thin}}, w_0^{\text{thin}}$	5 m, 7 m
$w_R^{\text{thick}}, w_0^{\text{thick}}$	5 m, 3 m
$\phi_R^{\text{DB}}, \phi_0^{\text{DB}}$	$10^\circ, 30^\circ$
$\sigma_R^{\text{hom}}, \sigma_0^{\text{hom}}$	0.3, 0.5
$\sigma_R^{\text{non-hom}}, \sigma_0^{\text{non-hom}}$	0.3, 0.2
$\phi_R^{\text{F}}, \phi_0^{\text{F}}$	$70^\circ, 80^\circ$
$m_R^{\text{S}}, m_0^{\text{S}}$	-13.6 dB, -8.6 dB

mainly medium to large buildings, while the bottom part includes smaller buildings, which are also often joined together and surrounded by gardens with other man-made structures or trees. In order to assess the performance of the proposed technique, we consider the correct/missed and false building detection rates and correlate such results with the size of the buildings. The number of split or merged buildings is also counted. The planar area of the buildings (length \times width) has been estimated using the optical image. The set of buildings present in the investigated area has been divided into three subsets: *small*, *medium* and *large*. Each subset corresponds to a different range of planar areas. Buildings are considered to be small if their planar area is smaller or equal to 200 m^2 , medium if the area is between 200 and 400 m^2 , and large if it is greater than 400 m^2 . Table IV reports the number of buildings for each size class in the subset 1 and the number of buildings correctly detected given by the proposed technique. As it is difficult to measure numerically the correctness of the reconstruction of the building radar footprints, here we only evaluate the detection performance of the algorithm in terms of footprints detected in correspondence of actual buildings. The detections



Fig. 19. (a) TerraSAR-X image used for assessing the effectiveness of the proposed technique (Infoterra). (b) Optical image taken from Google Maps of the investigated area.

TABLE III
PARAMETERS USED IN THE SELECTION OF HYPOTHESES AND 2D RADAR FOOTPRINT RECONSTRUCTION STEPS IN THE EXPERIMENTS CARRIED OUT WITH THE PROPOSED TECHNIQUE.

Parameter	Value
N'_h	0.8
$\delta d_R, \delta d_0$	3 m, 10 m
$\delta \psi_R, \delta \psi_0$	$10^\circ, 30^\circ$
$S_{h,\min}$	0.7

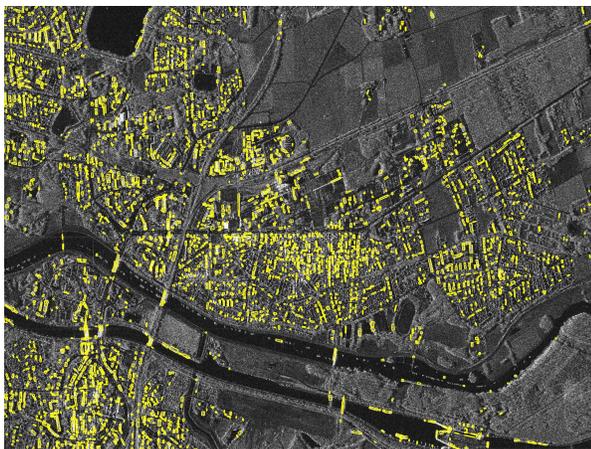


Fig. 20. Building detection and radar footprint reconstruction obtained by the proposed technique on the SAR image of Fig. 19a. Only the bright parts of the reconstructed building radar footprints are shown.

have been checked by comparing the positions of the radar footprints extracted by the algorithm with those manually detected on the SAR image in correspondence of buildings, which positions have been previously estimated from an optical image of the same area. A building is considered detected if the algorithm extracted a footprint which overlaps with the actual building radar footprint. Fig. 21e shows the correct and missed detection on the optical image. As it is not possible to directly match the detected radar footprints to the optical signatures of buildings without information on their height,

this map has been drawn manually using as reference the results obtained by the proposed method for each building in the scene. The results point out that the overall detection rate of radar footprints of the proposed technique is high, especially considering that the method is unsupervised and works on a single meter-resolution VHR SAR image. The performance of the technique is very good for medium and large buildings, while small buildings result in a higher number of missed alarms. This expected result is due to the fact that small buildings in meter-resolution images often do not show the scattering features used by the proposed technique. On the one hand, the number of split buildings is 1, 4 and 2 for the classes *small*, *medium* and *large*, respectively. Therefore, as far as building size increases, the probability that the technique splits the radar footprints in more parts increases. On the other hand, as far as building size decreases, the probability that the radar footprints of adjacent buildings are detected as a single one increases. In fact, the number of merged buildings is 9 for the class *small*, 3 for the class *medium* and 1 for the class *large*. The number of false alarms in building detection is 11. The size of the bright part of the false building radar footprints has been measured on the SAR image and false alarms have been divided in *small*, *medium* and *large* according to the same rules used for building planar sizes. Although the two measurements considered (i.e., area of false building radar footprints and planar area of real buildings) are different, the use for false alarms of the same classes as for real buildings allows us to give an indication on the types of false alarms produced by the proposed method. As shown in Table IV, false alarms are mostly related to small radar footprints. By comparing the SAR image to the optical image it is clear that false alarms usually correspond to other man-made structures (e.g., isolated garages) or trees which show radar signatures that are very similar to those of buildings. Such false alarms are also difficult to be detected by an expert human interpreter without other sources of information (e.g., a reference optical image). The footprints generated by the proposed technique are usually correctly reconstructed for medium and large buildings. As an

example, Fig. 22 shows the refinement of the bright part of a footprint hypothesis after the 2D footprint reconstruction step. For small buildings the radar footprints are often reconstructed with lower precision. Fig. 21c shows also the detected (and reconstructed) building shadows. The proposed technique extracted with good precision most of the shadow areas related to the detected building radar footprints. This result can be used as a starting point for further estimations on the building sizes, e.g., for estimating building heights [22]. However, it is worth noting that in many cases shadow areas are limited by adjacent buildings, thus reducing their usefulness for height extraction purposes.

D. Results on Subset 2

The area corresponding to the subset 2 is shown in Fig. 23. This area is characterized by a large number of trees located along the streets (in Fig. 23d it is possible to see their shadows). Such trees often mask the radar returns also from medium-sized buildings. Moreover, small buildings are usually quite irregular, and show many structures on their walls. This subset is thus a challenging benchmark for the proposed technique. Table IV reports the results obtained for the subset 2, and Fig. 23e shows the correct and missed detections on the optical image. As for the subset 1, the detection rate for the classes *large* and *medium* is very good. For the class *small* performance are less satisfactory. The number of split buildings is 1 for the class *small*, 4 for the class *medium* and 4 for the class *large*; while the number of merged buildings is 8 for the class *small*, 3 for the class *medium* and 0 for the class *large*. The total number of false alarms is 11. As for subset 1, the most of them are related to small false building radar footprints. In overall, considering the issues mentioned at the beginning of this paragraph and the limited amount of information used by the proposed technique, the results can be considered very good. In order to provide a more general view of the results obtained by the proposed method, Table IV also reports the overall results computed by summing the results of the subsets 1 and 2. The total statistic confirms the trend highlighted for the single subsets, i.e., the algorithm has a high detection rate for medium and large buildings, with a limited amount of false alarms, whereas its performance decreases in the case of small buildings, which are associated to most of the total number of false alarms. It is worth noting that it is possible to mitigate this problem by imposing a rule for discarding the footprints smaller than an user-defined minimum footprint size. As a consequence, the number of false alarms would be considerably reduced and the detection of radar footprints of small buildings would not be a target of the method anymore. This is a reasonable strategy to adopt for tuning the proposed technique only on the detection of medium and large buildings.

E. Selection of Algorithm Parameters

The tuning of the parameters has been performed according to the scene investigated. However, some parameters are not strictly related to the image analyzed, and can be set *a priori* following general rules. Moreover, many of the considered

parameters have a clear physical meaning that helps the user to include its prior knowledge on the scene in the detection algorithm. In addition to the guidelines already provided in Section III, in this section we analyze more in detail the role of the parameters of the proposed method.

1) *Feature extraction and primitive generation*: In these steps the main parameters of the proposed technique are related to the detection and generation of bright rectangles, and to the extraction of the shadows. The possible range of values for the window of the line detector w_t should be set between the expected thickness of thin linear features and the maximum size of the buildings which has to be extracted. The sampling of the range of w_t , given by the number of filterings T , should assure that most of the linear features can be effectively modeled with the considered values of w_t . The minimum value for δw_{\max} has to be greater than the width sampling resulting from the definition of the values of w_t . On the one hand, a value smaller than this quantity would not allow the algorithm to downselect effectively the rectangles produced in the feature extraction step. Moreover, the procedure for the generation of primitives would combine only rectangles with approximately the same width. On the other hand, a value much greater than the width sampling would make the algorithm to downselect too many rectangles, and combine features with much different widths. According to our tests, a good choice for the value of δw_{\max} is 1.5 times the width sampling used in the line detection. Regarding the parameters A_t and $\delta\psi_{\max}$, high values for A_t and low angles for $\delta\psi_{\max}$ make conditions (3) and (4) too stringent, respectively. By setting $A_t = 0.5$ and $\delta\psi_{\max} = 20^\circ$ we obtained the best results in our experiments. Note that these settings are general and do not depend on the image under analysis.

As mentioned in Section III-A2, the choice of the value of x_S depends on the characteristics of the shadow regions in the SAR image. The results obtained with different values for x_S showed that the detection and reconstruction of the shadows is not sensitive to slight variations of the parameter. The choice of l_S depends on the maximum expected height of the buildings present in the scene (and thus of their shadows). Thus, this parameter should be set according to the acquisition incidence angle and to prior information on the scene. However, if no *a priori* information is available, a large value can be set. This does not affect significantly the detection of the radar footprints. In fact, footprint hypotheses including dark primitives which are not close to bright primitives (which have been kept in the feature extraction due to a large l_S) are penalized by the term (21) in the selection of hypotheses. Thus, only the reconstruction step is affected by the choice of l_S , as shadows can grow further.

2) *Analysis of primitives*: In this step the main parameters to be set are those related to the membership functions defined for the different scattering classes. The choice of the value of R is not critical, and $R = 0.999$ can be considered as a fixed value. The parameters $w_R^{\text{thin}} = w_R^{\text{thick}}$, w_0^{thin} , and w_0^{thick} used in this paper can also be considered general. Indeed, they are given in meters, so that they do not depend on the resolution of the system. According to our tests, by setting $w_R^{\text{thin}} = w_R^{\text{thick}}$ to a value 2-3 m greater than the expected thickness of the linear

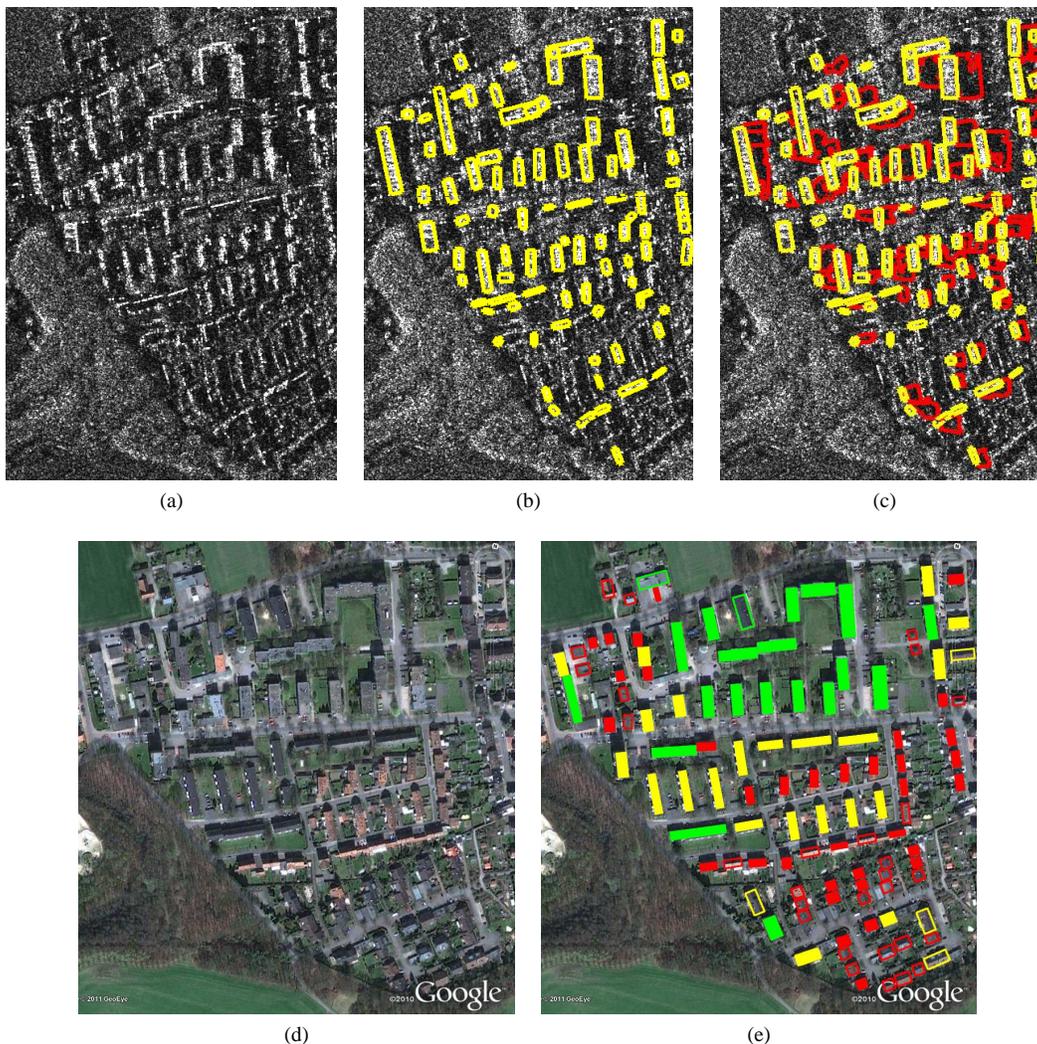


Fig. 21. Subset 1: (a) original TerraSAR-X image of the considered area, viewing direction from left (Infoterra); (b) reconstructed bright parts of the detected building radar footprints on the SAR image; (c) reconstructed building radar footprints on the SAR image: (yellow) bright parts, (red) dark parts; (d) optical image (Google); (e) optical image with detected and missed buildings for each building size class: (green) *large*, (yellow) *medium*, and (red) *small*. Detected and missed buildings are highlighted with filled and empty rectangles, respectively.

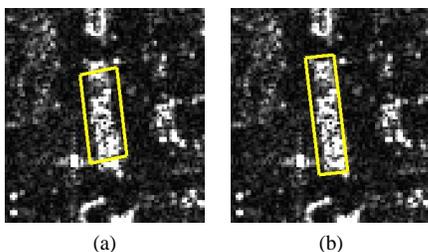


Fig. 22. Example of bright part of a radar footprint hypothesis (a) before, and (b) after the 2D footprint reconstruction step.

signatures due to the double-bounce effect give the best results, as the procedure which creates rectangles from the output of the line detector may overestimate their actual thickness.

The values of the parameters ϕ_R^{DB} , ϕ_0^{DB} , ϕ_R^F , and ϕ_0^F are defined on the basis of our experience in analyzing VHR SAR images. These values are also general, and can be considered valid for most images of urban areas. The main studies carried

out specifically on the relation between the double-bounce effect and the aspect angle of buildings can be found in [32].

The choice of the values $\sigma_R^{\text{hom}} = \sigma_R^{\text{non-hom}}$, σ_0^{hom} , and $\sigma_0^{\text{non-hom}}$ depends on the characteristics of speckle in the considered image. As the membership functions are evaluated on the GMAP filtered image, different parameters apply for different filterings. Similarly, these parameters depend on the image resolution, as speckle develops differently on the same target depending on resolution. For these reasons, the correct choice of these values in terms of capability to model effectively homogeneous and non-homogeneous areas comes after a proper optimization of the GMAP filtering parameters.

The last parameters used in this step are m_R^S and m_0^S . As for x_S , these values depend on the characteristics of shadows in the SAR image. According to our experiments, m_R^S and m_0^S should be set about 1.5 dB lower and 3-4 dB greater than x_S , respectively. This allows one to obtain a quite smooth term $\Sigma_m^S(m)$ in (14). Indeed, the mean amplitude of a dark region corresponding to a shadow may be biased by the interference

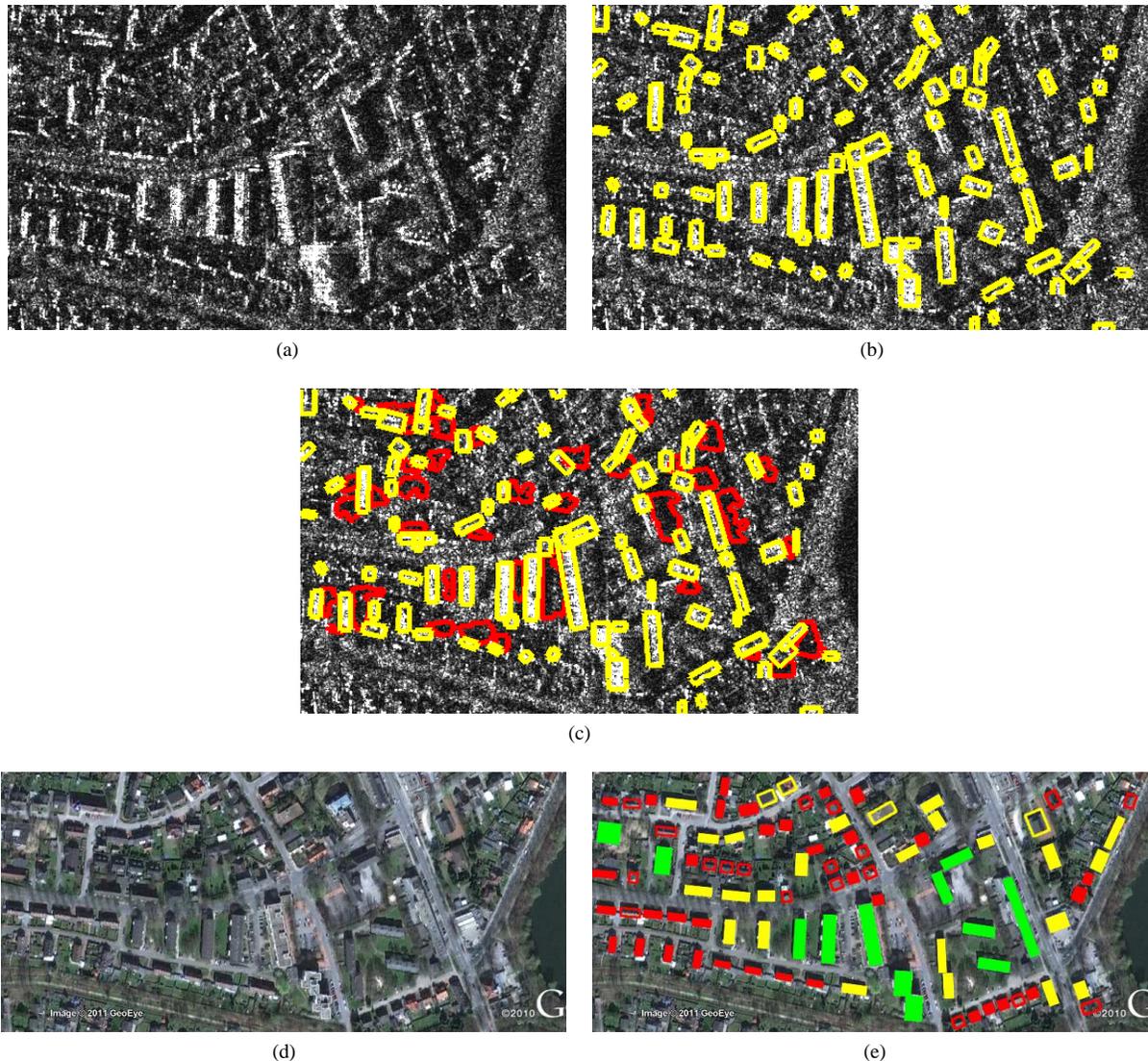


Fig. 23. Subset 2: (a) original TerraSAR-X image of the considered area, viewing direction from left (Infoterra); (b) reconstructed bright parts of the detected building radar footprints on the SAR image; (c) reconstructed building radar footprints on the SAR image: (yellow) bright parts, (red) dark parts; (d) optical image (Google); (e) optical image with detected and missed buildings for each building size class: (green) *large*, (yellow) *medium*, and (red) *small*. Detected and missed buildings are highlighted with filled and empty rectangles, respectively.

of surrounding structures which increases its value. Thus, using a sharp $\Sigma_m^S(m)$ would make the algorithm to discard possible real shadows.

3) *Selection of hypotheses and 2D radar footprint reconstruction*: As mentioned in Section III-E, the parameter N'_h is related to the reliability assigned by the user to the footprint hypotheses composed by only two primitives. In our tests, by setting this parameter to higher values resulted in detection maps with less hypotheses composed by three primitives, as expected. Indeed, increasing the value of N'_h makes three-primitive hypotheses to have higher probability to score lower than those composed by two primitives. Therefore, three-primitive hypotheses have higher probability to be discarded when they overlap with others made up of two primitives. This does not affect significantly the detection rate of the proposed method, but it increases the probability that the extracted footprints are not well-reconstructed (e.g., shadows

are missing even though they were detected). On the contrary, by setting N'_h to low values would increase the number of missed detections. Therefore, the choice of N'_h should be done by the user as a tradeoff between reliability of the reconstruction and detection performance.

The pair of parameters $(\delta d_R, \delta d_0)$ and $(\delta \psi_R, \delta \psi_0)$ are related to the vicinity and relative orientation of the primitives, respectively. The values proposed in this paper can be considered general for the defined scattering classes. Note that, using these values, the sigmoid functions present in (16) and (21) are quite smooth, thus mitigating the effect of possible errors in feature extraction.

The last parameter to be discussed is $S_{h,\min}$. This parameter gives the tradeoff between false and missed detections. According to our tests, the use of high $S_{h,\min}$ results in a greater number of missed detections, as expected. However, the number of false alarms is not reduced significantly. Indeed, these

TABLE IV
ALGORITHM PERFORMANCE FOR SUBSET 1, SUBSET 2 AND SUBSET 1 + SUBSET 2 IN TERMS OF NUMBER OF DETECTED BUILDINGS, FALSE ALARMS, SPLIT AND MERGED BUILDINGS PER BUILDING CLASS.

	Building size	Number of buildings	Detected	False alarms	Split	Merged
Subset 1	Large	21	19	0	2	1
	Medium	26	22	2	4	3
	Small	66	35	9	1	9
Subset 2	Large	12	12	0	4	0
	Medium	27	23	2	4	3
	Small	53	34	9	1	8
Subsets 1+2	Large	33	31	0	6	1
	Medium	53	45	4	8	6
	Small	119	69	18	2	17

are usually related to footprints of other man-made structures, or trees, which actually appear as related to buildings. For this reason, values in the order of 0.6-0.7 are suggested.

F. Computational Load

The test image described in Section V-A has been processed using a cluster composed by 16 AMD® Opteron™ 6172 CPUs, for a total of 192 cores, with 4 GB of RAM per core. The image has been split on tiles of 300×300 pixels with an overlapping offset of 30 pixels with the neighbors. The total number of tiles was thus 154, and each tile was processed by one core. The total processing time was about 45 minutes. With the same infrastructure it is thus possible to process a whole spotlight image of about 6000×10.000 pixels in less than 3 hours. We also tested the proposed technique using a smaller cluster composed by 8 commercial workstations equipped with Intel® Core™ i7-870 quad-core processors and 8 GB of RAM. The total processing time for the test image on this smaller architecture was about 1 hour and 30 minutes, which is a good performance in terms of operational application of the algorithm.

VI. DISCUSSION AND CONCLUSION

In this paper the problem of the detection and reconstruction of building radar footprints in VHR SAR images has been addressed. Unlike many other methods presented in the literature, the proposed technique can be applied to single VHR SAR images. It extends state-of-the-art feature extraction and composition steps to more structured primitives using a production system and by introducing the concept of semantic meaning. This has been done in order to compensate for the lack of information due to the fact that only one VHR SAR image is used as input. The semantic meaning represents the probability that an object belongs to a certain scattering class (e.g., facade, double bounce), and is calculated via fuzzy membership functions. Therefore, it allows the technique to select the most reliable primitives and footprint hypotheses during its processing steps. As a further refinement, the proposed technique also reconstructs the detected radar footprints. The goal of this step is to provide as output a map which can be used as a starting point for further calculations, e.g., the estimation of building widths and lengths. Moreover,

by exploiting the reconstruction of the shadow areas, height retrieval techniques can be also applied to estimate building heights. In order to make it possible to use the proposed technique on large VHR SAR images in near real-time, we also proposed and implemented an infrastructure based on a computer cluster for the processing of large VHR SAR scenes.

The proposed method is suited for meter-resolution SAR images. However, it can be extended and tuned for higher-resolution airborne data by introducing new types of primitives, composed objects and rules. Moreover, new semantic classes for the primitives should be defined, as finer scattering mechanisms become visible in sub-meter data.

The experimental results obtained on a large meter-resolution SAR image confirmed the effectiveness of the proposed technique. In particular, the method shows very high detection rates in the case of medium and large buildings, exhibiting also a good capability to reconstruct their radar footprints. The number of false alarms is limited, and these are mostly related to other man-made structures or trees which show radar signatures similar to those of buildings. For small buildings the proposed technique shows worse detection and reconstruction performance of radar footprints, and an increased number of false alarms. This is mainly due to the low number of features related to small buildings visible in single meter-resolution SAR images. Nonetheless, this is an expected problem, which is mainly related to the need to use sub-meter resolution images for a proper detection of these buildings, rather than to a limitation of the proposed technique. In order to mitigate this problem, it is possible to include a simple rule in the proposed technique for discarding the radar footprints smaller than an user-defined threshold, thereby reducing the number of false alarms and avoiding the detection of small buildings. This is a reasonable strategy to adopt for tuning the proposed method only on the detection of medium and large buildings, on which performances are very accurate.

The proposed approach needs the user to set some parameters which depend on the product under analysis. After this, the method is automatic and can be applied with the same set of parameters to similar products. Guidelines for the selection of the parameters were given throughout the paper. It is worth noting that many relevant parameters have been already selected on calibrated SAR images so that they can be applied to different VHR SAR scenes without the need to

be changed.

The proposed technique is promising for addressing problems in real operative scenarios which exploit the available spaceborne meter-resolution SAR systems (e.g., Cosmo-SkyMed, TerraSAR-X, and TanDEM-X). As an example, it can be used as a preliminary step for a fine estimation of the density of urban areas even from single images or it can be used for the analysis of multitemporal series, e.g., for the detection of changes in urban areas. It is worth noting that the method can work with any viewing configuration of the SAR sensor, as it can handle radar footprints acquired with different incidence and aspect angles. Moreover, it provides results in the vector domain. These factors make it possible to potentially combine the results obtained from SAR acquisitions taken with different viewing angles, or also maps derived from optical images. This would allow a finer detection of buildings and a more precise estimation of building properties. However, the problem of the correct geolocalization of buildings in the different acquisitions should be faced e.g., for the correct merging of the single radar and/or optical footprint maps.

As future developments we plan to extend the proposed technique to both the analysis of multi-aspect acquisitions (e.g., images acquired on ascending and descending orbits) and the integration of interferometric height information in the steps of the processing chain. By this we aim at developing a flexible framework for building detection and radar footprint extraction requiring as minimum only a single SAR scene, but making best use of additional input data if available. We also plan to study the integration of the presented method with state-of-the-art change detection algorithms in order to develop novel reliable approaches to change detection in urban areas using VHR SAR multi-temporal series.

ACKNOWLEDGMENT

The authors would like to thank Franziska Greßler from Infoterra, Germany, for providing the TerraSAR-X data, the Orfeo Toolbox (OTB) [42] and Insight Toolkit (ITK) [43] communities for publishing these libraries as open source products. We also acknowledge the anonymous reviewers for their comments, which helped to improve the paper.

REFERENCES

- [1] S. Marchesi, F. Bovolo, and L. Bruzzone, "A context-sensitive technique robust to registration noise for change detection in VHR multispectral images," *IEEE Transaction on Image Processing*, vol. 19, pp. 1877–1889, Jul. 2010.
- [2] F. Pacifici and W. J. Emery, "Pulse coupled neural networks for automatic urban change detection at very high spatial resolution," in *Proc of the 14th Iberoamerican Conference on Pattern Recognition: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, 2009, pp. 929–942.
- [3] T. Kemper, A. Wania, and M. Pesaresi, "Supporting slum mapping using very high resolution satellite data," in *In Proc. of 33rd International Symposium on Remote Sensing of Environment, Sustaining the Millennium Development Goals*, 2009.
- [4] I. Baud, M. Kuffer, K. Pfeffer, R. Sliuzas, and S. Karuppanan, "Understanding heterogeneity in metropolitan India: The added value of remote sensing data for analyzing sub-standard residential areas," *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, no. 5, pp. 359 – 374, 2010.
- [5] D. Brunner, G. Lemoine, F.-X. Thoorens, and L. Bruzzone, "Distributed geospatial data processing functionality to support collaborative and rapid emergency response," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2, no. 1, pp. 33 – 46, Mar. 2009.
- [6] D. Al-Khudhairy, "Geo-spatial information and technologies in support of EU crisis management," *International Journal of Digital Earth*, vol. 3, no. 1, pp. 16 – 30, Mar. 2010.
- [7] S. Voigt, T. Kemper, T. Riedlinger, R. Kiefl, K. Scholte, and H. Mehl, "Satellite image analysis for disaster and crisis-management support," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1520 –1528, Jun. 2007.
- [8] F. Caltagirone, G. De Luca, F. Covello, G. Marano, G. Angino, and M. Piemontese, "Status, results, potentiality and evolution of COSMO-SkyMed, the Italian Earth Observation constellation for risk management and security," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul. 2010, pp. 4393–4396.
- [9] R. Werninghaus and S. Buckreuss, "The TerraSAR-X mission and system design," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 606–614, 2010.
- [10] D. Brunner, G. Lemoine, and L. Bruzzone, "Earthquake damage assessment of buildings using VHR optical and SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2403–2420, May 2010.
- [11] P. Gamba, F. Dell'Acqua, and G. Trianni, "Rapid damage detection in the Bam area using multitemporal SAR and exploiting ancillary data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1582–1589, Jun. 2007.
- [12] U. Soergel, U. Thoennessen, and U. Stilla, "Reconstruction of buildings from interferometric SAR data of built-up areas," in *ISPRS Conference "Photogrammetric image analysis"*, 17-19 Sept., Munich, Germany, 2003, pp. 59–64.
- [13] F. Cellier, H. Oriot, and J.-M. Nicolas, "Hypothesis management for building reconstruction from high resolution InSAR imagery," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 31 July-4 Aug., Denver, CO, USA, 2006, pp. 3639–3642.
- [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] 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.
- [16] F. Zhang, Y. Shao, and X. Zhang, "Building extraction using dual-aspect high resolution SAR images," in *Proc. 18th International Conference on Geoinformatics*, 2010, pp. 1–5.
- [17] 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.
- [18] Y. Wang, F. Tupin, C. Han, and J.-M. Nicolas, "Building detection from high resolution PolSAR data by combining region and edge information," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 4, 7-11 July, Boston, MA, USA, 2008, pp. IV-153–IV-156.
- [19] 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.
- [20] C. Tison and F. Tupin, "Retrieval of building shapes from shadows in high resolution SAR interferometric images," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 3, 20-24 Sept., Anchorage, Alaska, USA, 2004, pp. 1788–1791.
- [21] J. D. Wegner, R. Hansch, A. Thiele, and U. Soergel, "Building detection from one orthophoto and high-resolution InSAR data using conditional random fields," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 1, pp. 83 –91, march 2011.
- [22] D. Brunner, G. Lemoine, L. Bruzzone, and H. Greidanus, "Building height retrieval from VHR SAR imagery based on an iterative simulation and matching technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1487–1504, Mar. 2010.
- [23] H. Sportouche, F. Tupin, and L. Denise, "Extraction and three-dimensional reconstruction of isolated buildings in urban scenes from high-resolution optical and SAR spaceborne images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3932–3946, Oct. 2011.
- [24] M. Quartulli and M. Datcu, "Stochastic geometrical modeling for built-up area understanding from a single SAR intensity image with meter resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 9, pp. 1996–2003, Sep. 2004.

- [25] F. Zhang, Y. Shao, X. Zhang, and T. Balz, "Building L-shape footprint extraction from high resolution SAR image," in *IEEE Proc. Joint Urban Remote Sensing Event*, 2011, pp. 273–276.
- [26] A. Ferro, D. Brunner, and L. Bruzzone, "An advanced technique for building detection in VHR SAR images," in *SPIE Europe Remote Sensing*, vol. 7477, 31 Aug-3 Sept., Berlin, Germany, 2009, pp. 74 770V–74 770V–12.
- [27] —, "Building detection and radar footprint reconstruction from single VHR SAR images," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 25-30 July, Honolulu, HI, USA, 2010, pp. 292–295.
- [28] E. Michaelsen, U. Soergel, and U. Thoennessen, "Perceptual grouping for automatic detection of man-made structures in high-resolution SAR data," *Pattern Recog. Lett.*, vol. 27, no. 4, pp. 218–225, Mar. 2006.
- [29] C. L. Forgy, "Rete: A fast algorithm for the many pattern/many object pattern match problem," *Artif. Intell.*, vol. 19, no. 1, pp. 17–37, Sep. 1982.
- [30] A. Thiele, E. Cadario, K. Schulz, U. Thoennessen, and U. Soergel, "Feature extraction of gable-roofed buildings from multi-aspect high-resolution InSAR data," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 23-27 July, Barcelona, Spain, 2007, pp. 262–265.
- [31] A. J. Bennett and D. Blacknell, "Infrastructure analysis from high resolution SAR and InSAR imagery," in *Proceedings of the 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Berlin, Germany, May 2003, pp. 230– 235.
- [32] A. Ferro, D. Brunner, L. Bruzzone, and G. Lemoine, "On the relationship between double bounce and the orientation of buildings in VHR SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 612–616, 2011.
- [33] D. Kuan, A. Sawchuk, T. Strand, and P. Chavel, "Adaptive restoration of images with speckle," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 3, pp. 373–383, 1987.
- [34] F. Tupin, H. Maître, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky, "Detection of linear features in SAR images: application to road network extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 2, pp. 434–453, Mar. 1998.
- [35] R. C. Gonzales and R. E. Woods, *Digital Image Processing*. Addison Wesley, 1993.
- [36] E. Simonetto, H. Oriot, and R. Garello, "Extraction of industrial structures and DEM from airborne SAR images," in *Proc. Physics in Signal and Image Processing*, Marseille, France, Jan. 2001, pp. 412–417.
- [37] 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 *Proceedings of EARSel Workshop 3D-Remote Sensing*, Porto, Portugal, Jun. 2005.
- [38] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing*. Prentice Hall, 1997.
- [39] R. C. Eberhart, Y. Shi, and J. Kennedy, *Swarm Intelligence*, ser. The Morgan Kaufmann Series in Evolutionary Computation. Morgan Kaufmann, 2001.
- [40] J. Sethian, *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science, 2nd ed.* Cambridge, UK: Cambridge University Press, 1999.
- [41] M. Silveira and S. Heleno, "Separation between water and land in SAR images using region-based level sets," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 471–475, Jul. 2009.
- [42] J. Inglada and E. Christophe, "The Orfeo Toolbox remote sensing image processing software," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 4, 2009, pp. IV–733–IV–736.
- [43] L. Ibanez, W. Schroeder, L. Ng, J. Cates, Consortium T.I.S., and R. Hamming, *The ITK Software Guide*. Kitware, Inc., 2003.