Building Height Estimation from High Resolution SAR Imagery via Model-Based Geometrical Structure Prediction

Zhuang Wang[1], Libing Jiang[1, *], Lin Lei[1], and Wenxian Yu[2]

Abstract—Height extraction by radar remote sensing is an attractive issue for the building detection and recognition. According to the analysis on the building geometrical properties in the SAR imagery, a novel height estimation algorithm is proposed following a model-based geometrical structure prediction and matching strategy. The range Doppler equation is introduced and simplified for the building 2D geometrical structure prediction in the slant image plane. An evaluation function implementing the ratio of exponentially weighted averages (ROEWA) is also established for the matching between the predicted structure and the observed SAR image. By incorporating the genetic algorithm (GA), the evaluation function is maximized to get the optimal height parameter. The experimental results with the simulated and real airborne and spaceborne SAR images show that the proposed method could efficiently estimate building height from single SAR imagery, and achieve better performance than two popular algorithms with the partial occlusion case.

1. INTRODUCTION

Building detection and reconstruction by remote sensing is a fundamental task in urban monitoring, planning and controlling [1]. Particularly, in the domain of rapid situation assessment after natural disasters (e.g., earthquakes and tsunamis), building height estimation is a key issue in post-emergency event information retrieval, which is crucial for initiating effective emergency response actions.

Although a lot of methods for building height extraction from optical remote sensing images have been presented [2–5], the usefulness of the optical images highly depends on sun illumination and cloud-free weather, which is not always available in some conditions (e.g., Wenchuan earthquake [6]). On the contrary, the synthetic aperture radar (SAR), as an active remote sensing technique, is capable of providing images independent from daytime and weather conditions. With the recent advances in metric and sub-metric resolution SAR sensors (e.g., TerraSAR-X and Cosmo-SkyMed), anthropogenic structures such as individual buildings can be identified from high resolution (HR) SAR images and thus accurate building height retrieval is becoming feasible.

Quite often, there is only one or a few HR SAR images available immediately after the disaster. Consequently, the problem of building height estimation from single HR SAR imagery has to be investigated particularly. The most difficult challenge is to accurately delineate the building boundaries, which are distorted by speckle noise and may be occluded by adjacent structures, such as in the classical monoscopy [7, 8] and radargrammetry [9, 10] methods. Recently, model-based [11–17] approaches have been proposed, which can explicitly incorporates knowledge of the SAR imaging and electromagnetic theory. For example, Guida et al. [11] and Franceschetti et al. [17] calculated the building height analytically through the deterministic electromagnetic equation, in which extensive prior information of the building (e.g., the dielectric constants and the roughness parameters) as well as the surrounding

* Corresponding author: Libing Jiang (readfish0101@163.com).
[1] School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China. [2] School of Electronic, Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China.

terrains are needed explicitly. Brunner [12] iteratively simulated the building images by ray tracing and matched them with the real SAR image. In contrast, Wang [13] introduced the mapping and projection algorithm (MPA) for the building image simulation. Nevertheless, the image simulation procedure is usually of highly computational cost and heavy resource demanding. Zhao [14] instead extracted the layover and shadow from the real SAR imagery and correlated them with the simulated image. Similar to the monoscopy, its usefulness depends on the correct delineation of the building features, which may deteriorate by the speckle noise. Sportouche et al. [15] took advantage of the SAR formation equations as well as the image statistical characteristics and proposed a projection based approach. However, precise backscattering modeling for the urban areas in the HR SAR images is quite difficult, as is discussed by Gao et al. [18].

Furthermore, most of the existing methods have focused on isolated buildings but paid rare attention to partially occluded building cases. In fact, in urban and semi-urban areas, the interference between adjacent buildings usually leads to partial occlusions. The appearance of the occluded buildings in the SAR image is different from the isolated ones, which complicates the height estimation problem.

In this paper, we propose a building height estimation approach by means of model-based building geometrical structure prediction and matching, which requires no assumptions on the HR SAR backscattering distribution or the image simulation. Furthermore, unlike the most existing height estimation methods, the proposed method is capable of extracting the height for both the isolated and partially occluded buildings. The proposed approach is implemented in an iterative manner, which can be decomposed into three phases: 1) height hypothesis generation, 2) building geometrical structure prediction, and 3) geometrical structure matching. Firstly, a series of building height hypotheses are generated randomly. In the following, the building geometrical structures corresponding to the height parameters in the slant image plane are precisely predicted according to the basic range Doppler equation. Finally, the predicted building structure is masked onto the observed SAR image and matched by the evaluation function, which is maximized to find the optimal height hypothesis.

The reminder of this paper is organized as follows. In Section 2, the dominating backscattering mechanisms as well as the building properties in the SAR images are analyzed. In Section 3, a general form for the model-based building height estimation is formulated, and the flowchart of the proposed approach is presented. In Section 4, the proposed height estimation approach is introduced in detail. The experimental results employing the simulated data as well as the real SAR imagery are presented in Section 5, and the conclusions are drawn in Section 6.

2. BUILDING BACKSCATTERING CHARACTERISTICS IN SAR IMAGERY

SAR sensor is characterized by its active side looking illumination and ranging geometry. The projection of the three dimensional building into the two-dimensional slant image plane produces effects such as layover, shadow and foreshortening that distort the appearance of the building in the SAR imagery. Besides, a small incidence angle also leads to the specular reflection. Dense alignment of the buildings along the radar look of sight (RLOS) also causes the multi-path reflection effects [19], as is briefly depicted in Figure 1.

The geometry of the radar collection and the shape of the building influence the appearance of the
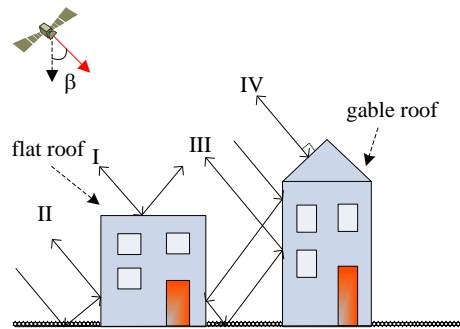
Figure 1. Scattering mechanism on the buildings in the urban area. I. Single bounce reflection. II. Double bounce reflection. III. Multi-path reflection. IV. Specular reflection.

building in the slant image plane. To understand the relationship between building and its shape in the SAR imagery, the range profiles of two common building types are analyzed: the flat roof and the gable roof building. Example of the isolated flat roof building is shown in Figure 2. The layover area a is generated by the single bounce reflections from the front wall, ground and roof, which share the same range resolution cell. The line b is formed by the dihedral corner that arises from the double bounce reflection between the building vertical wall and its front terrain. The area c indicates the single bounce reflections from the front wall, ground or roof. The area d is the shadow where no reflection happens. The relation between the layover and shadow dimension and the building height parameter is as follows

$$h = \begin{cases} s_{\mathrm{r}} \times \cos\beta & \text{if } \beta < \tan^{-1}(w/h) \\ l_{\mathrm{r}} / \cos\beta & \text{if } \beta > \tan^{-1}(w/h) \end{cases} \tag{1}$$

where β is the depression angle, and l_r and s_r are the length of layover and shadow, respectively.

The range profile of the isolated gable roof building is a little different from the flat roof case, as shown in Figure 3. A second bright scattering area arises closer to the sensor than the double bounce line, which is mainly caused by the inclined roof e directly against the radar. In the specific case $\beta = \gamma$, the single bounce reflections from e are assembled in a single range resolution cell, which result in a bright line at the near range. Besides, the alignment of the layover and shadow also varies with different geometrical attributes, as shown in Figures 3(a)–(d). The relationship between the dimension of the layover, shadow and the height parameter is intricate compared to Eq. (1), depending not only on the depression angle β and $\tan^{-1} w/h$ but also on the roof inclined angle γ, as shown in Eq. (2).

$$h = \begin{cases} l_{\mathrm{r}} / \cos\beta & \text{if } \beta > \gamma \\ s_{\mathrm{r}} \times \sin\beta & \text{if } \beta < \gamma \end{cases} \quad \text{and} \quad \beta < \tan^{-1}\mathrm{w/h} \tag{2}$$
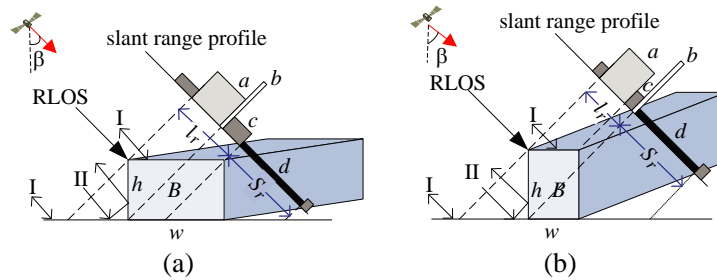

Figure 2. Slant range profile of isolated flat roof building. (a) $\beta < \tan^{-1}(w/h)$. (b) $\beta > \tan^{-1}(w/h)$.
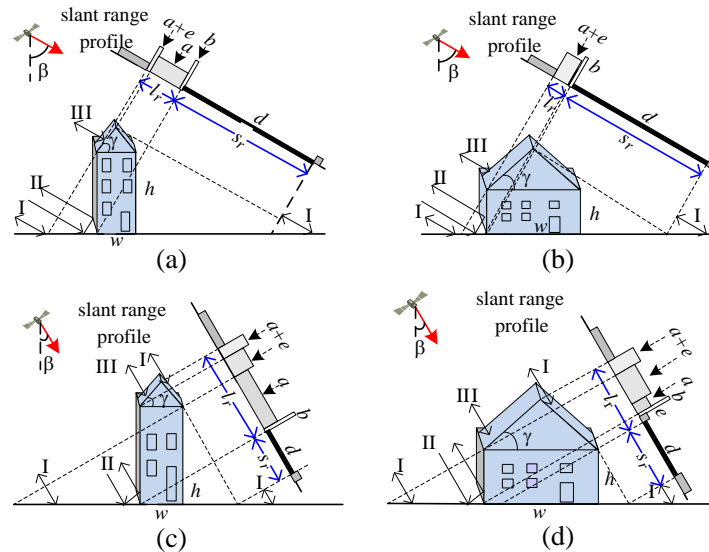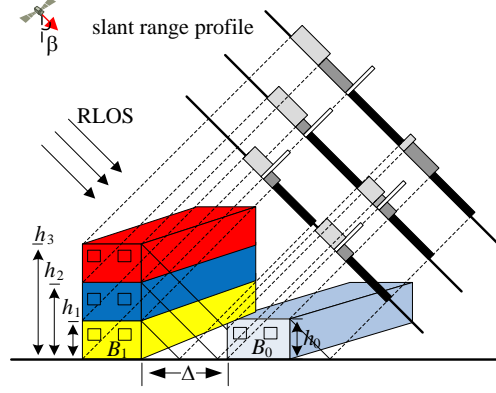

Figure 3. Slant range profile of isolated gable roof building. (a) $\beta > \tan^{-1}(w/h)$ and $\beta > \gamma$. (b) $\beta < \tan^{-1}(w/h)$ and $\beta > \gamma$. (c) $\beta > \tan^{-1}(w/h)$ and $\beta < \gamma$. (d) $\beta < \tan^{-1}(w/h)$ and $\beta < \gamma$.

Figure 4. Slant range profile of buildings with occlusion.

Especially, for the case $\beta > \tan^{-1} w/h$ and $\beta < \gamma$ in Figure 3(c), there is no direct relationship between h and l_r since contributions originated from different building parts (e.g., the front wall, ground, and the inclined roof) are mixed into the layover, which have comparable influence and are hard to distinguish.

Furthermore, in urban or sub-urban areas, the buildings are usually densely aligned to each other, which may cause occlusion and makes their structures quite different from the isolated cases. An example of two neighbored buildings is shown in Figure 4. For the building B_0 and B_1 with height h_0 and h_1 respectively, since the distance between B_0 and B_1 is large enough, their range profile are the same as isolated cases. As the height of B_1 gradually increase to h_2, partial occlusion happens on B_0, which results in a subdivided layover and a weakened double bounce line for B_0, as well as a shortened shadow for B_1. This effect is intensified as B_1 becomes higher. As a result, the double bounce line of B_0 vanishes when the height of B_1 grows to h_3 because its "front wall-ground" double bounce reflection disappears. In this case, the Eq. (1) and Eq. (2) no longer works.

The minimum distance [12] $\Delta_{\min}$ required between buildings without occlusion is

$$\Delta_{\min} = h_1 \tan(\beta) + h_0 \cot(\beta) \tag{3}$$

where h_1 and h_0 is the height of the building at the near and far range of the SAR sensor respectively (see Figure 4).

The examples from Figures 2–4 indicate that the building geometrical structure in the SAR image contains information about the building parameters, which is the fundament of the building height estimation. Figure 4 further illustrates that the geometrical structure in the slant image plane is determined not only by the building's own geometry, but is also interfered by its surrounding buildings. Thus different building cases should be distinguished in advance and treated differently. For the isolated building case, its building structure is solely determined by its own geometry, thus its height parameter can be estimated separately one by one. For the partial occluded building case, its height should be estimated by concerning its own geometry as well as the neighbor buildings simultaneously.

3. PROBLEM STATEMENT OF THE BUILDING HEIGHT ESTIMATION

Canonical building can be explicitly designated by its geometric parameters. Hence we denote a building B as $B = (l, w, h, \alpha, \gamma, C)$, where l is the length, w is the width, h is the height, α is the aspect angle, γ is the roof inclination angle and $C = (x, y, z)$ is the center of the building footprint. We also denote H and β as the height and the depression angle of the SAR sensor. As is discussed in Refs. [11–13, 15–17], the building footprint and the SAR sensor parameters are usually regarded as known *prior* in the problem of building height retrieval. For the simplicity, these prior parameters are denoted in the vector form as $\theta = (l, w, \alpha, \gamma, C)$ and $\varphi = (H, \beta)$.

From a model-based point of view, the problem of image interpretation can be regarded as searching for the parameter that best matches with the observed SAR image. Thus, the general form of the

building height estimation problem can be formulated as the following maximization function

$$\hat{\mathbf{h}} = \arg \max_{h, \vec{S}} L\{I'(B(h, \theta)|\varphi), I\} \tag{4}$$

where I is the observed SAR imagery and $I'(B(h, \theta)|\varphi)$ the simulated image or features corresponding to the hypothesis $B(h, \theta)$ under sensor parameter φ. $L\{\cdot\}$ is the evaluation function which measures the consistency between $I'(B(h, \theta)|\varphi)$ and I. $\mathbf{S} = (dm, dn)^{\mathrm{T}}$ is the translation between $I'(B(h, \theta)|\varphi)$ and I.

Considering the fact that the heights of the adjacent buildings should be estimated simultaneously for the partially occluded building case, the formula in Eq. (4) is rewritten as

$$\hat{\mathbf{h}} = \arg \max_{\mathbf{h}, \vec{S}} L\{I'(B(\mathbf{h}, \theta)|\varphi), I\} \tag{5}$$

where $\mathbf{h} = (h_1, h_2, \ldots, h_{n_{ner}})^T$ is the height vector of all the neighbor buildings $\{B(h_i, q), i = 1, \ldots, n_{ner}\}$. For $n_{ner} = 1$, it is the case of the isolated building, and for $n_{ner} > 1$ it is the case of the occluded building.

As illustrated in Section 2, it is not the entire building image but the shapes of building regions that are straightforwardly related to its geometric parameters. So in this paper, we do not simulate the SAR images. Instead, the geometrical structure of the building in the slant image plane (e.g., the layover, shadow, single reflection area and double bounce line), are predicted according to the building model. Besides, considering the influence of speckle, we neither extract the building contours, but directly mask the predicted structure onto the observed SAR image. Hence, the matching between the predicted building geometrical structure $I'(\mathbf{h})$ and the observed SAR image I in Eq. (5) can be matched by the evaluation function.

The proposed height estimation algorithm can be summarized as follows: Starting from the building model composed of the prior 2D footprints and the height hypothesis, firstly, the isolated and the partially occluded building cases are discriminated according to Eq. (3). For the isolated building case, its geometrical structure is directly calculated through the mapping function. For the partially occluded building case, its visible facets are identified in advance through the "sensor-building" illumination geometry and then the corresponding geometrical structure is also calculated by the mapping function. Subsequently, this predicted structure is projected onto the observed SAR image. Finally, an evaluation function is constructed and maximized by the genetic algorithm (GA), which leads to the solution for the height parameter. The flowchart of the proposed approach is shown in Figure 5.
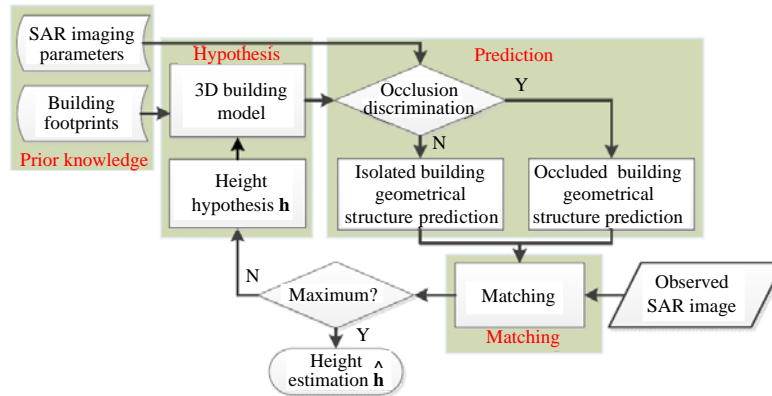

Figure 5. Flowchart of the proposed method.

4. PROPOSED HEIGHT ESTIMATION APPROACH

4.1. Height Hypothesis Generation

During the height estimation process described by Figure 5, only the height parameter $\mathbf{h}$ and translation parameter $\mathbf{S}$ are variable. The building prior parameters $\theta = (l, w, \alpha, \gamma, C)$ can be derived from

cadastral maps, optical remote sensing images or single high-resolution SAR imageries. The sensor prior parameters $\varphi = (H, \beta)$ are defined by the flight height of the airplane as well as the depression angle of the SAR imaging system. During the maximization for Eq. (5), the height parameter hypotheses are uniformly drawn from a predefined height interval $[\mathbf{h}_{min}, \mathbf{h}_{max}]$. Initial height parameter hypothesis can also be provided by the monoscopy method. The translation parameter $\mathbf{S}$ is also generated randomly from a predefined translation range.

4.2. Building Geometrical Structure Prediction

In order to understand the building specific phenomena in the SAR images, it is highly necessary to discriminate the building backscattering types and how they influence the image formation. As is illustrated in Figure 1, there mainly exist three basic reflection types in the urban areas: the single bounce reflection, the double bounce reflection and the multi-path bounce reflection. In most cases [19], the contribution of the multi-path bounce reflection decreases sharply as the number of bounce increase, which is quiet weak with respect to the contribution from the single and double bounce reflections. Hence, the single and the double bounce reflections are the two dominant reflections for the SAR building formation, as shown in Figure 6. This understanding aids the SAR image interpretation in terms of a building's geometric attributes.

For the single bounce reflection, the mapping of building from 3D scene coordinates to the 2D slant image plane is depicted by the basic range Doppler equation, as shown in Figure 7. In detail, for the point $\mathbf{p}_t = (x_t, y_t, z_t)$ in the scene, through the radar imaging by single bounce, its corresponding range coordinates m_t in the slant image plane can be deduced by the following equations,

$$\begin{cases} r_t = \sqrt{(x_t - x_s)^2 + (y_t - y_s)^2 + (z_t - z_s)^2} = r_0 + m_t \cdot \Delta m \\ f_D = -\dfrac{2}{\lambda} \dfrac{v_s(x_s - x_t) + v_s(y_s - y_t) + v_s(z_s - z_t)}{r_t} \end{cases} \quad (6)$$

where r_0 is the near range of the scene; λ is the radar wavelength; Δn and Δm are the range and azimuth resolutions; (x_s, y_s, z_s) and (v_x, v_y, v_z) are the position and velocity of the radar; f_D is the Doppler frequency.

When the radar flight trajectory is perpendicular to its illumination direction, we have $v_x = v_z = 0$ and $f_D = 0$, substituting into Eq. (6) and we can obtain $y_t = y_s$. So the azimuth coordinate n_t of $\mathbf{p}_t$ is as follows,

$$y = y_0 + n_t \cdot \Delta n \quad (7)$$

where y_0 is the initial azimuth coordinate of the scene. From Eqs. (6)–(7), the slant image coordinate
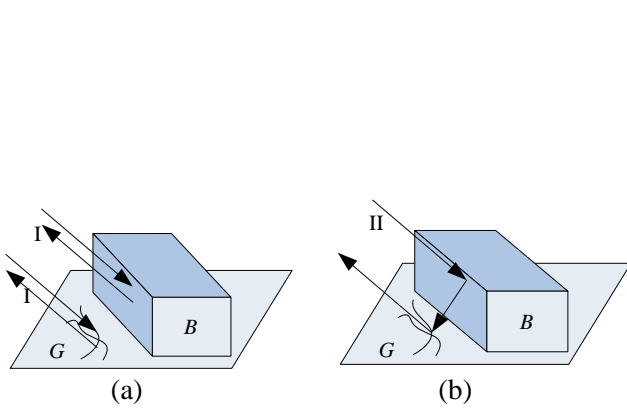
Figure 6. Dominating backscattering mechanisms for canonical building. (a) Single bounce reflection. (b) Double bounce reflection.
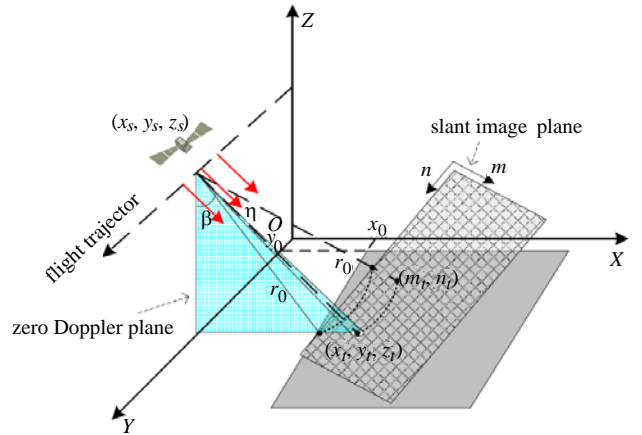
Figure 7. Mapping from the scene coordinates to the slant image plane.

(m_t, n_t) of the point $\mathbf{p}_t = (x_t, y_t, z_t)$ can be calculated,

$$
\begin{cases}
m_t = \dfrac{\sqrt{(x_t - x_s)^2 + (y_t - y_s)^2 + (z_t - z_s)^2} - r_0}{\Delta m} \\
n_t = \dfrac{y - y_0}{\Delta n}
\end{cases}
\tag{8}
$$

The double bounce reflection originates from radar signals following the path "radar-building wall-ground-radar" or vice versa, as shown in Figure 6(b), whose contribution is a highlight line superposed on the near range footprint of the building. Since this double bounce line shares the same resolution cell with the single bounce reflection happened on the building footprint, its coordinate could also be calculated by mapping the near range building footprint to the slant image plane through Eq. (8). So based on the above analysis, the mapping of the building from 3D scene to the 2D slant image coordinates is given by Eq. (8).

Based on the analysis above, the building regions (such as layover, shadow, double bounce line, etc.) in the slant image plane can be calculated by the Eq. (8). Moreover, combining the Eq. (8) with the specific backscattering model such as Lambertian or specular reflection model, the corresponding simulated SAR image can be obtained, as is illustrated in Refs. [12, 13, 16]. However, as is discussed in Section 3, it is not the entire SAR image data but the shape of the building that is straightforwardly related to its geometric parameters. Therefore, in this paper we concentrate on the prediction of the building geometrical structure in the slant image plane from 3D scene coordinates.

An example of mapping is shown in Figure 8. The rectangles of the building facets such as roof and walls are distorted and superimposed, which become parallelograms with corners that are not right angles. The visible facets form the layover and the double bounce line, while the invisible facets determine the shadow, as shown in Figure 8(b). Furthermore, all these points labeled in the slant image plane of Figure 8(c) can be explicitly calculated by the building parameters and the radar illumination geometry. Therefore, the first task for the building structure prediction is to decompose the building into discrete points and identify its visibility along the RLOS.
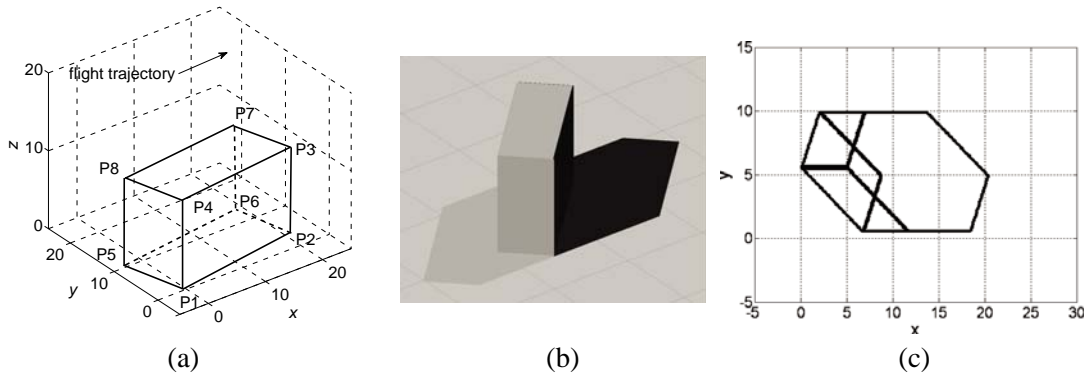
A building can be decomposed into a series of facets,

$$
B \triangleq \{ f_i^B, i = 1, 2, \dots, N_f^B \}
\tag{9}
$$

where N_f^B is the number of the facets. Each facet f_i^B can further be decomposed into a series of discrete points,
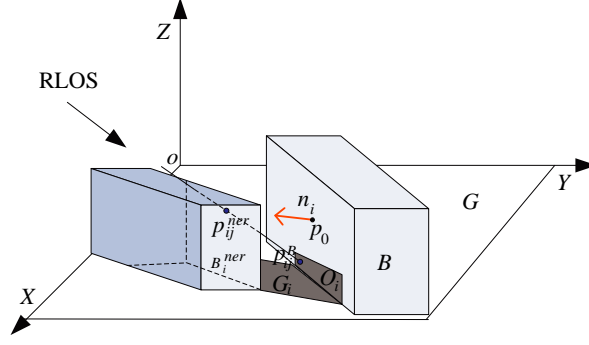
$$
f_i^B \triangleq \{ \mathbf{p}_{ij}^B, j = 1, 2, \dots, C_i \}
\tag{10}
$$

where $\mathbf{p}_{ij}^B = (x, y, z)_{ij}^B$ is the jth discrete point on facet f_i^B, and C_i is the total point number.

(a)   (b)   (c)

Figure 8. Mapping of the building to the slant image plane. (a) Building in the 3D coordinates. (b) Visible and invisible facets in the scene. (c) Corresponding geometrical structure in the slant image plane.

Figure 9. Visible facets identification.

The visible facets along the RLOS can be easily discriminated by the following formula,

$$\begin{cases} f_i \text{ is visible} & \text{if } \mathbf{R} \cdot \mathbf{n}_i < 0 \\ f_i \text{ is invisible} & \text{otherwise} \end{cases} \tag{11}$$

where "·" is the inner product operator, $\mathbf{R}$ the unit vector along the RLOS, and $\mathbf{n}_i$ the norm vector of the facet f_i. The ground facets $\{f_i^G\}$ which also have contribution surrounding the building B can be identified by the facets $\{f_i^B\}$ through the following linear equation,

$$\mathbf{p}_{ij}^G = \mathbf{p}_{ij}^B \mathbf{T} \tag{12}$$

where $\mathbf{p}_{ij}^G = (xyz)_{ij}^G$ is the point corresponding to $\mathbf{p}_{ij}^B$ in f_i^G, T is the transform matrix,

$$T = \begin{cases} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\cot\beta & 0 \end{bmatrix} & \text{if } \mathbf{R} \cdot \mathbf{n}_i < 0 \\[20pt] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \tan\beta & 0 \end{bmatrix} & \text{otherwise} \end{cases} \tag{13}$$

For the isolated building, by substituting these building points into Eq. (8), the corresponding geometrical structure can be obtained. However, for the partially occluded building, the visible facets $\{f_i^B\}$ and $\{f_i^G\}$ will not remain intact due to the partial occlusion caused by its adjacent buildings and thus the exact visible points should be further re-identified before mapping by Eq. (8). Denoting $\{B_i^{ner}, i = 1, \ldots, N_{ner}^B\}$ as the adjacent buildings of the building B which do not meet the minimum distance condition in Eq. (3), then the invisible parts of B occluded by the ith neighbor building B_i^{ner} can be identified through the sensor illumination geometry, as illustrated in Figure 9,

$$\begin{cases} \mathbf{p}_{ij}^B = \mathbf{p}_{ij}^{ner} + (0, \sin\beta \times t, -\cos\beta \times t) \\ \mathbf{n}_i^B \cdot (\mathbf{p}_{ij}^B - \mathbf{p}_0^B) = 0 \end{cases} \tag{14}$$

where $\mathbf{p}_{ij}^{ner} = (x, y, z)_{ij}^{ner} \in B_i^{ner}$, $\mathbf{n}_i^B = (u, v, k)$ is the norm vector of $f_i^B \in B$, $\mathbf{p}_{ij}^B$ the invisible point occluded by $\mathbf{p}_{ij}^{ner}$, and $\mathbf{p}_0^B = (x_o, y_o, z_o) \in f_i^B$ the arbitrary known point on f_i^B. Accordingly, the exact visible facet f'^B_i of f_i^B after occlusion by the adjacent buildings $\{B_i^{ner}\}$ is

$$f'^B_i = f_i^B \setminus \bigcup_{i=1}^{N_{ner}} \bigcup_{j=1}^{C_i} \{\mathbf{p}_{ij}^B\} \tag{15}$$

where "\" and "∪" are the set difference and union operator respectively. The exact visible parts of the ground facet for the partially occluded building can also be identified in a similar way by Eqs. (14)–(15), whose unit vector $\mathbf{n}_i \triangleq (0, 0, 1)$.

Based on the above analysis, the prediction of the building geometrical structure can be summarized as follows:

Step 1: Discriminate the isolated buildings from the partially occluded buildings through Eq. (3).
Step 2: Decompose the building into discrete points by Eq. (9)–Eq. (11). For the isolated building, directly calculate the interrelated points on the ground according to Eq. (12). For the partially occluded building, refine its visible points in advance by Eq. (14)–Eq. (15) and then identify the interrelated ground points by Eq. (12).
Step 3: Map the points obtained from **Step 2** onto the image plane through Eq. (8). For convenience, the mapped area on the slant image plane is denoted as a label image, which characterizes the building shape.

4.3. Evaluation Function

As discussed in Section 2, the contours of the building behave as salient edges in the SAR image, thus we design the evaluation function based on the edge strength along the building contours. Denote L_B as the label image which characterizes the geometrical structure of building B, and then by masking L_B onto the observed SAR image I, an image chip $P_I(L_B)$ containing the building contours is obtained.

$$P_I(L_B) = \left\{ p_i^{L_B}, i = 1 \ldots M \right\} \tag{16}$$

where $p_i^{L_B} \in I$ is the ith pixel along the building contours and M the total pixel number.

Considering the influence of speckle noise, the ratio of exponentially weighted averages (ROEWA) [20] is introduced for the edge strength calculation. Therefore, the evaluation function $L\{\cdot\}$ for Eq. (5) is formulated as the sum of ROEWA edge strength along contours of the predicted building geometrical structure in the following Eq. (17), where $R(\cdot)$ is the ROEWA operator.

$$L\left\{ I'(\mathbf{h}), I \right\} = \sum_{i=1}^{M} R\left(p_i^{L_B} \right) \tag{17}$$

4.4. Maximization

In order to get the optimal height parameter $\mathbf{h}$, the evaluation function $L\{\cdot\}$ in Eq. (17) should be jointly maximized for $\mathbf{h}$ and translation parameter $\mathbf{S}$. However, due to the lack of concave-convex property of $L\{\cdot\}$, the usefulness of the deterministic optimization techniques are limited. Therefore, the genetic algorithm is adopted here, which is specialized in finding a global maximum among several local maximums through a stochastic way. The fitness of the genetic algorithm is the evaluation function in Eq. (17). Other related parameters for the genetic algorithm are listed in Table 1.

Table 1. Parameter configuration for the genetic algorithm.

| Population type	Selection	Crossover	Mutation	Stopping Criteria
Double vector	Stochastic uniform	Arithmetic	Non-uniform	Maximum generation

5. EXPERIMENTAL RESULTS AND DISCUSSION

5.1. Evaluation by the Simulated Image

Aiming at the quantitative performance assessment for the proposed method, a dataset containing 176 simulated SAR images with azimuth and range resolution $0.2\,\text{m} \times 0.2\,\text{m}$ is constructed according to the Lambertian reflection model, which characterizes the salient relationship between the scattering intensity and the incident angle. Different acquisitions and geometry attributes for the isolated and
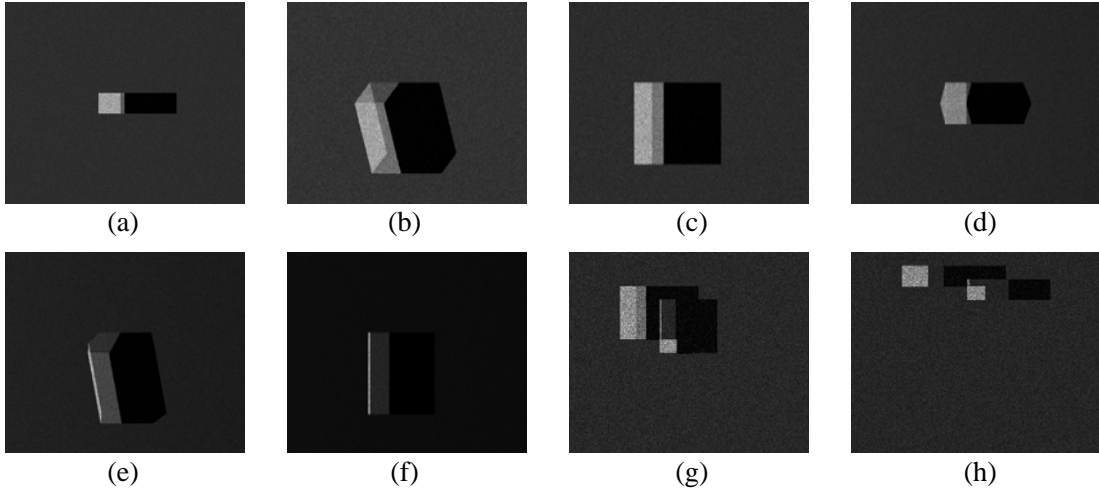
Figure 10. Examples of the simulated image data set. (a), (c) are isolated flat roof buildings with aspect angle at 0°, 60° and 90°. (d), (f) are isolated gable roof building with aspect angle at 0°, 60° and 90°. (g), (h) are partially occluded buildings containing two adjacent buildings with aspect angle at 0°, 90° and distance 8 m.
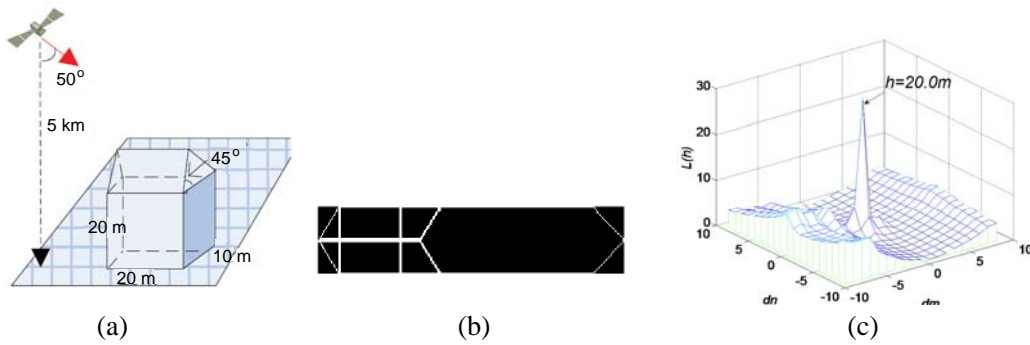

Figure 11. Example of the gable roof building height estimation. (a) 3D building in the scene. (b) Predicted building geometrical structure corresponding to h_0. (c) Surface of evaluation function $L(h)$.

partially occluded building cases are taken into account. For the isolated building cases, two common building types — the flat and the gable roof buildings, are considered. Their aspect angles vary at 0°, 20°, 30°, 45°, 60°, 80° and 90°, while the SAR depression angles are chosen at 30° and 51°. For the partially occluded building cases, a scene containing two adjacent flat roof buildings is considered for simplicity. Their aspect angles vary equally from 0° to 90° by step 10°. The distance between the buildings are set as 6 m, 8 m and 10 m respectively, while the SAR depression angles are chosen at 40° and 50°. The speckle effect is also considered. A multiplicative noise with variances of 0.1 and 0.2 are added to the whole dataset, which leads to their equivalent number of looks (ENL) as 3 and 1 respectively. Examples of the simulated images in the dataset are shown in Figure 10.

An example of the isolated gable roof building with geometric model (20 m, (20 m, 10 m, 0°, 45°)) is shown in Figure 11(a). The corresponding simulated image with multiplicative noise variance 0.2 is shown in Figure 10(d). The predicted building geometrical structure corresponding to the initial height parameter hypothesis $h_0 = 16.0$ m is shown as a label image in Figure 11(b). The "prediction-matching" process convergences after about 20 iterations by the genetic algorithm, where a maximum evaluation function value = 29.68 is found at = 20.0 m, as is shown in Figure 11(c).

The proposed algorithm is tested over the simulated image dataset. For comparison, the existing image simulation based algorithm [12, 16] and the projection based algorithm [15] have also been
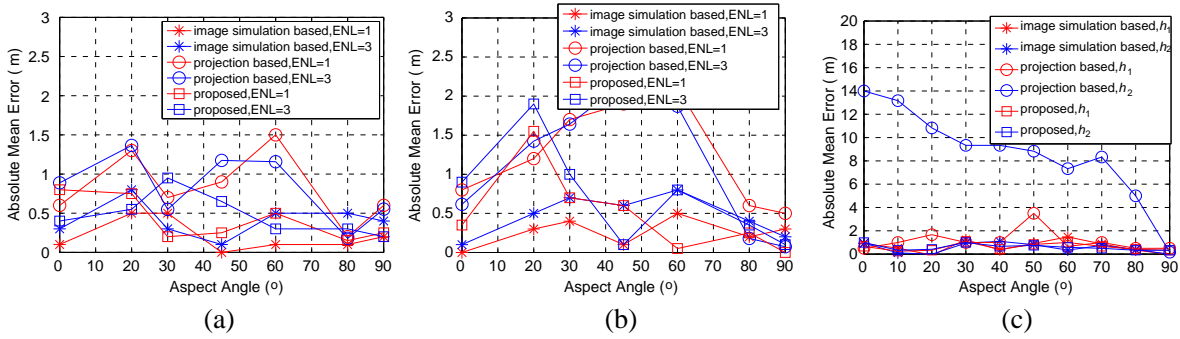
Figure 12. Comparison of height estimation results. (a) Isolated flat roof building. (b) Isolated gable roof building. (c) Adjacent two flat roof buildings.

investigated. Due to the fact that the coherent SAR images simulation techniques (such as PO, GO, ray tracing) usually entail complex electromagnetic computation, therefore, a non-coherent simulation technique [14] is adopted here for image simulation, which has a comparatively low computation cost and is prone to implement.

The absolute mean error (AME) of the above-mentioned algorithms with various aspect angles and noise variances are shown in Figure 12. Figure 12(a) is the result for isolated flat building case and Figure 12(b) is the isolated gable roof building case. From Figures 12(a)–12(b) we can see that both three algorithms can estimate the isolated building height parameter with good precision and their AME curves are comparatively low. For the image simulation based algorithm, a maximum AME of 0.5 m for isolated flat roof building and 0.8 m for the gable roof building is obtained. Similar results have been attained for the projection based algorithm and the proposed algorithm, which have a slightly high maximum AME of 1.5 m for isolated flat building and 2.1 m for the isolated gable roof building. Besides, their AME curves are also comparatively stable as the aspect angle varies, which means that their performances are not sensitive to the building aspect angle parameter.

Figure 12(c) refers to the result of the partially occluded building height estimation over the dataset, in which h_1 represents the building height at the near sensor range and h_2 is the building height at the far sensor range. For both the image simulation based and the proposed algorithm, their AME curves maintain low and smooth, with an average error of 1.6 m and a maximum error of 3.8 m. However, for the projection based method, the estimation error of h_2 becomes large, with an average error of 9.0 m and a maximum error of 14.0 m at the aspect angle 0°. The AME curves in Figure 12(c) indicates that the performance of the projection based algorithm deteriorates as the occlusion becoming worse, while the image simulated based algorithm and the proposed algorithm maintain effective.

The computation times of the above three algorithms during an individual "prediction-matching" procedure are also enumerated, as listed in Table 2. From Table 2 we can see that the image simulation based algorithm takes the longest computation time during the rendering phase, while the projection-based algorithm consumes the shortest time. Actually, both the projection based algorithm and the proposed algorithm use the range Doppler equation for the building mapping. However, the elapsed times of the projection-based algorithm are about 5 times shorter than the proposed algorithm for the partially occluded case, because only the ridges of the building are taken into account.

5.2. Evaluation by Airborne SAR Data

The X-band airborne SAR data with a spatial resolution of 1 m in both azimuth and range is used in the experiment, which is provided by East China Research Institute of Electronic Engineering. Figure 13(a) shows an image chip acquired over Hefei city in China, which contains a flat roof building. The corresponding optical image from the same area is shown in Figure 13(b). The ground truth of the building parameter in the image chip is (10.5 m, (13.1 m, 51.0 m, 40.4°)). The final geometrical structure obtained by the proposed algorithm corresponding to the optimal height estimation = 9.8 m is visually enhanced in gray line and superimposed onto the airborne SAR image chip in Figure 13(c), which is close to the ground truth.

Table 2. Computation time of an individual "prediction-matching" procedure.

Eclipsed time / Algorithms	Prediction (sec)		Matching (sec)		Total (sec)	
	isolated	partially occluded	isolated	partially occluded	isolated	partially occluded
image simulation based algorithm	1.032	3.661	0.012	0.012	1.044	3.673
projection-based algorithm	0.004	0.152	0.022	0.035	0.026	0.187
proposed algorithm	0.013	1.012	0.031	0.042	0.044	1.054

(a)   (b)   (c)
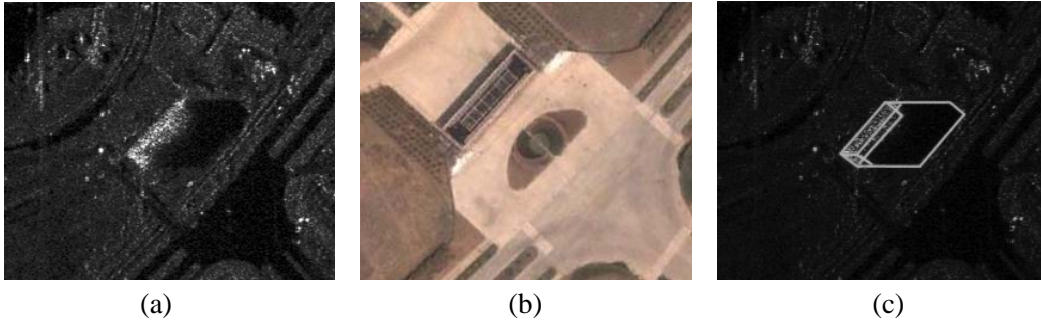
Figure 13. Airborne SAR data experiment. (a) Airborne SAR image chip. (b) Optical image (©Google Earth). (c) Overlaid with the predicted building geometrical structure.

(a)   (b)   (c)
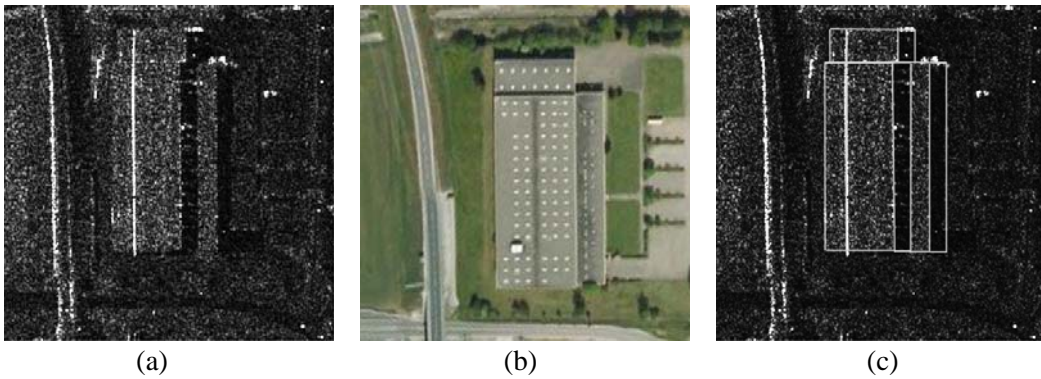
Figure 14. Spaceborne SAR data experiment. (a) TerraSAR-X image chip (©Infoterra GmbH). (b) Optical image (©Google Earth). (c) Overlaid with the predicted building geometrical structure.

5.3. Evaluation by TerraSAR-X Data

As reported, TerraSAR-X was successfully launched in June 2007, which is capable of acquiring meter and sub-meter resolution SAR images. The HH-polarization SAR image acquired over Nordlingen, Germany from ascending orbit is used, which contains three neighbor buildings, as shown in Figure 14(a). The optical image of the same district is shown in Figure 14(b).

Due to the constraints of field inspection, the ground truth of the building footprint is obtained by manual inspection over the optical image from Google Earth with a resolution of 0.6 m, which are (40.2 m, 24.2 m, 90°), (97.7 m, 40.2 m, 0°) and (97.7 m, 13.4 m, 0°) for buildings from top left to bottom right respectively. Although the actual building heights are unknown, the experiment results

on the simulated images in Section 5.1 suggest that the accuracy of the height estimation is closely correlated with the coherence degree between the predicted geometrical structure and the building areas. Thus, the estimated geometrical structure is overlaid on the origin image chip for an indirect and qualitative analysis, as shown in Figure 14(c), which is quiet in accordance with the building areas. The experimental results on the TerraSAR-X SAR data demonstrate the validity of the proposed method.

6. CONCLUSIONS

In this paper, a new building height estimation method is proposed, which is capable of retrieving the building even with partial occlusion from single VHR SAR images. The backscattering characteristics of two typical buildings are analyzed in detail. Base on that, the range Doppler equation is employed for the prediction of the building structure in the slant range plane. Then, the predicted building structure is masked onto the real SAR image for evaluation by the matching function. Finally, by maximizing the matching function through the genetic algorithm, the building height is retrieved. Experiments with the simulated and real SAR data show that the proposed method has a comparable performance with existing two algorithms for the isolated building case, and yields better results for the partially occluded building case. Further work for more complicate shaped man-made structures interpretation is currently in progress.

ACKNOWLEDGMENT

REFERENCES

1. Soergel, U., *Radar Remote Sensing of Urban Areas*, 1st edition, Springer, Berlin, 2010.
2. Ok, A. O., C. Senaras, and B. Yuksel, "Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, Vol. 51, No. 3, 1701–1717, 2013.
3. Izadi, M. and P. Saeedi, "Three-dimensional polygonal building model estimation from single satellite images," *IEEE Trans. Geosci. Remote Sens.*, Vol. 50, No. 6, 2254–2272, 2012.
4. Stankov, K. and D. He, "Building detection in very high spatial resolution multispectral images using the hit-or-miss transform," *IEEE Geosci. Remote Sens. Lett.*, Vol. 10, No. 1, 86–90, 2013.
5. Elbakary, M. I. and K. M. Iftekharuddin, "Shadow detection of manmade buildings in high-resolution panchromatic satellite images," *IEEE Trans. Geosci. Remote Sens.*, Vol. 52, No. 9, 5374–5386, 2014.
6. Balz, T., D. Perissinb, U. Soergel, L. Zhang, and M. S. Liao, "Post-seismic infrastructure damage assessment using high-resolution SAR satellite data," *Proceeding of 2nd International Conference on Earth Observation for Global Change*, 180–191, Chengdu, China, 2009.
7. Blacknell, D., R. D. Hill, and C. P. Moate, "Estimating building dimensions from synthetic aperture radar image sequences," *IET Radar, Sonar and Navigation*, Vol. 2, No. 3, 189–199, 2008.
8. Tupin, F., "Extraction of 3D information using overlay detection on SAR images," *The 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, 72–76, Berlin, Germany, 2003.
9. Meric, S., F. Fayard, and E. Pottier, "A multi-window approach for radargrammetric improvements," *IEEE Trans. Geosci. Remote Sens.*, Vol. 49, No. 10, 3803–3810, 2011.
10. Simonetto, E., H. Oriot, and R. Garello, "Rectangular building extraction from stereoscopic airborne radar images," *IEEE Trans. Geosci. Remote Sens.*, Vol. 43, No. 10, 2386–2395, 2005.
11. Guida, R., A. Iodice, and D. Riccio, "Height retrieval of isolated buildings from single high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, Vol. 48, No. 7, 2967–2979, 2010.

12. Brunner, D., 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, 1487–1504, 2010.

13. Wang, T. L. and Y. Q. Jin, "Postearthquake building damage assessment using multi-mutual information from pre-event optical image and post-event SAR image," *IEEE Geosci. Remote Sens. Lett.*, Vol. 9, No. 3, 452–456, 2012.

14. Zhao, L. J., *Building Extraction from High Resolution SAR Imagery*, National University of Defense Technology, Changsha, China, 2009.

15. Sportouche, H., 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. 9, 3932–3946, 2011.

16. Matzner, S. A., *Model-based Information Extraction from Synthetic Aperture Radar Signals*, Portland State University, Portland, USA, 2011.

17. Franceschetti, G., R. Guida, A. Iodice, D. Riccio, G. Ruello, and U. Stilla, "Building feature extraction via a deterministic approach: Application to real high resolution SAR images," *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, 2681–2684, Barcelona, Spain, 2007.

18. Gao, G., X. X. Qin, and S. L. Zhou, "Modeling SAR images based on a generalized gamma distribution for texture component," *Progress In Electromagnetics Research*, Vol. 137, 669–685, 2013.

19. Ferro, A., 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, 612–616, 2011.

20. Fjortoft, R., A. Lopes, P. Marthon, and E. Cubero-Castan, "An optimal multiedge detector for SAR image segmentation," *IEEE Trans. Geosci. Remote Sens.*, Vol. 36, No. 3, 793–802, 1998.