

# Building detection from high-resolution PolSAR data at the rectangle level by combining region and edge information

Yinghua Wang<sup>a,b</sup>, Florence Tupin<sup>b</sup>, Chongzhao Han<sup>a</sup>

<sup>a</sup>*Institute of Integrated Automation, Xi'an Jiaotong University, Xi'an 710049, China  
yinghua.wang@stu.xjtu.edu.cn, dodgerbluester@gmail.com*

<sup>b</sup>*Institute TELECOM, TELECOM ParisTech, LTCI CNRS, 46, rue Barrault 75013  
Paris, France  
florence.tupin@telecom-paristech.fr*

---

## Abstract

We propose a new approach at the rectangle feature level to extract buildings from high-resolution polarimetric synthetic aperture radar (PolSAR) data, using both region-based and edge-based information. The first step employs low-level detectors to provide raw region and edge information of the scene. In the second step, the rectangle features are initially extracted from the edge detection results, and further optimized to best fit the rough region-based building detection results. In the last step, a novel Markov random field (MRF) framework for rectangles is proposed, in which the data energy term of rectangles is defined from the region information while the *smoothness term* is defined according to the contextual prior knowledge about the buildings. Under this framework, the building rectangles are identified from the optimized rectangle candidates by minimizing the total energy. The effectiveness of the proposed method is demonstrated using the real *fully* PolSAR data.

*Key words:* high-resolution synthetic aperture radar (SAR), polarimetric SAR, building detection, rectangle, Markov random field (MRF), region, edge

---

## 1. Introduction

Building detection is a crucial issue for many practical applications of the *fully* polarimetric synthetic aperture radar (PolSAR) data, such as the 3-D

reconstruction, the environment or urban planning, etc.. The full information carried by PolSAR data is of great benefit to the building detection problem. These information permits to explore the special physical scattering nature of buildings. In the low and middle resolution images, buildings can only be detected as a certain class at the pixel level since no details about building edges or shapes are visible. The building class *can be identified based on the scattering mechanism classification*. For example, the method proposed in (Guillaso et al., 2005) *classified the pixels into three basic categories: single-bounce, double-bounce and volume scattering. Then buildings were detected as the double-bounce scattering pixels*.

Another group of methods to detect buildings from the PolSAR data are based on the study of the target backscattering behavior during the synthetic aperture radar (SAR) image integration. For example, the Time-Frequency (TF) approaches have been proposed in (Ferro-Famil and Pottier, 2007) *to analyze the signal stationarity and coherence in the time-frequency domain. They were used to study dense urban areas and detect man-made targets (Ferro-Famil and Pottier, 2007; Reigber et al., 2007)*. These approaches are also at the pixel level and region-based results are obtained.

With the improved resolution of SAR data, more details emerge in the building areas, such as roofs, the bright double-reflection line, the shadows, the layover areas, etc.. Such details allow to detect buildings based on the edges *of the bright strips/blocks. These bright strips/blocks in the image mostly comes from the radar echoes of the building facades, roofs, edges and wall-ground corners*. One example is the approach proposed in (Xu and Jin, 2007). It first implemented the edge detection using the constant false alarm rate (CFAR) edge detector for PolSAR data (Schou et al., 2003); then parallel line segments were extracted from the edges using the extended Hough transform techniques (Xu and Jin, 2007); finally buildings were localized as *some certain* rectangles formed by these parallel line segments.

The low-level (pixel level) detectors do not solve the problem well enough when they come to the high-resolution images, since in these images buildings are not only some nearby pixels, but also some objects with certain edges and shapes, especially rectangle shaped objects as assumed in (Simonetto et al., 2005; Xu and Jin, 2007). Thus high-level features, like rectangle features for instance, should be exploited. Moreover, the region-based and edge-based approaches employ two kinds of different yet complementary information: *the region-based approaches detect buildings as the groups of pixel having the similar responses in the image, such as pixels with double-bounce scattering*

(Guillaso et al., 2005) or high coherence value (Ferro-Famil and Pottier, 2007); while the edge-based approaches indirectly detect buildings by retrieving the characteristic edges of buildings, such as the pairs of parallel line segments (Xu and Jin, 2007), and then deducing the building regions from these edges. It is expected that the coupling of these two kinds of information will improve the results.

In this paper, we propose a new approach at the rectangle feature level, combining the region-based and edge-based information to detect buildings from the high-resolution PolSAR data. Some low-level detectors are used in the first step to provide raw region or edge information of the scene, which offers the entries for the next step. In the second step, the rectangle features are initially extracted from the edge detection results, and further optimized to fit the low-level region-based building detection results. Finally a novel Markov random field (MRF) framework of rectangles is proposed, under which the rectangles belonging to buildings are identified from the optimized rectangle candidates by minimizing the total energy of the MRF *framework*.

The rest of the paper is organized as follows. Section 2 presents the proposed approach. Experimental results are provided in Section 3 and conclusions in Section 4.

## 2. The proposed approach

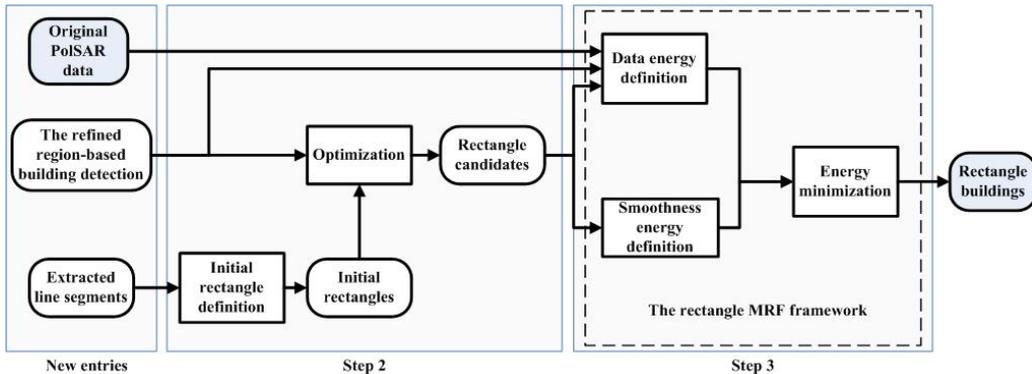


Figure 1: The diagram of the proposed approach.

The diagram of the proposed approach is shown in Figure 1. The new entries include the original PolSAR data and the low-level results obtained in the first step. In Step 2, rectangle features are directly retrieved from the

line segments and further optimized under the criterion proposed in (Simonetto et al., 2005), resulting in the optimized rectangle candidates that are located near the true positions of the buildings. Finally in Step 3 a novel MRF framework of rectangles is defined. The data energy term in this MRF framework is defined by the region information while the *smoothness energy* term is defined according to the contextual information. By minimizing the total energy, the rectangles belonging to buildings are identified from all the optimized candidates. *In Section 2.1, the low-level detectors used in Step 1 are briefly described.* The optimized rectangle candidate definition in Step 2 is introduced in Section 2.2. The new MRF framework of rectangles in Step 3 will be presented in Section 2.3. The relevant parameter setting of the MRF framework is analyzed in Section 2.4.

### 2.1. Low level detector description

Several low-level detectors are used in the first step to retrieve the raw region-based and edge-based information. The region-based detectors are introduced in Section 2.1.1. The edge-based detectors are described in Section 2.1.2. The new entries obtained are *listed* in Section 2.1.3.

#### 2.1.1. The region-based detectors

Two region-based detectors are used: the MRF based segmentation of PolSAR data and the TF coherence approach (Ferro-Famil and Pottier, 2007). Moreover the results of these two detectors are fused to obtain a rough region-based building detection result.

- *The MRF based segmentation. This segmentation is obtained by imposing smoothness constraint on the  $H/\alpha$  Wishart classification (Lee et al., 1999a) result. Let us introduce the  $H/\alpha$  Wishart classification first. An initial scattering mechanism classification is derived by dividing the  $H/\alpha$  plane into eight zones and eight classes (Cloude and Pottier, 1997), in which entropy  $H$  and  $\alpha$  angle computed from the polarimetric covariance matrix are used to characterize the scattering mechanism of each class. The initial cluster centers are defined as the average polarimetric covariance matrices of each class. This initial classification is further refined by considering the statistical properties, which is achieved by the  $K$ -means algorithm and the revised Wishart distance measure (Kersten et al., 2005) here. We denote the  $H/\alpha$  Wishart classification result by  $r^{[0]} = \{r_s^{[0]} \in \{1, 2, \dots, L_R\}, s \in S\}$ , where  $r_s^{[0]}$  is the class label of pixel  $s$ ,  $S$  is the pixel set, and  $L_R$  is the number of classes, i.e.,  $L_R = 8$ . Then,  $r^{[0]}$  is used to estimate the parameters contained in the energy*

definitions in the MRF framework. The data energy in the MRF framework is defined by the observation log-likelihood while the smoothness energy is defined based on the  $L_R \times L_R$  class distance matrix  $\mathbf{D}$ . The total energy is minimized to achieve the final regularized segmentation  $r$ . More details can be found in (Wang et al., 2008). The MRF based segmentation provides a region map of the whole image. In this region map, each homogeneous region is composed of pixels with similar statistical properties. Interpretation for each class is also provided based on scattering mechanisms characterized by  $H$  and  $\alpha$ .

- *The TF coherence approach.* This approach is based on the correlation properties of the targets across different sup-aperture images. The man-made targets usually appear correlated among the sub-apertures, thus they can be detected as pixels with high coherence values. By analyzing the signal coherence in the time-frequency domain, a coherence image is obtained, in which each pixel's coherence value is denoted by  $\rho_s$ ,  $0 \leq \rho_s \leq 1$ . Then a threshold is imposed on this coherence image, in order to obtain the binary rough building labeling  $x^{[0]} = \{x_s^{[0]} \in \{0, 1\}, s \in S\}$ , where  $x_s^{[0]} = 1$  if  $\rho_s$  is bigger than the threshold, otherwise  $x_s^{[0]} = 0$ .  $x_s^{[0]} = 1$  means pixel  $s$  belongs to buildings.

- *The fusion of the results from the above two detectors.* Under the fusion framework proposed in (Jodoin et al., 2007), the segmentation  $r$  and the rough building labeling  $x^{[0]}$  are fused to achieve  $\hat{x}_{\text{opt}}$ .  $\hat{x}_{\text{opt}}$  is close to  $x^{[0]}$ , however is adapted to fit the regions in  $r$ . The two label fields are assumed to form a joint MRFs. The fusion is fulfilled by minimizing the total energy based on this joint MRFs with the iterated conditional modes (ICM) algorithm (Besag, 1986). Yet we improve the original fusion framework by minimizing the total energy using the  $\alpha$ - $\beta$ -swap algorithm (Boykov et al., 2001). One could refer to (Wang et al., 2008) for more details.

### 2.1.2. The edge-based detectors

Speckle filtering with Lee's filter (Lee et al., 1999b) is first implemented. Then we use the CFAR edge detector (Schou et al., 2003) to get an edge intensity image. *In this detector, a likelihood-ratio test is used to test whether the two mean polarimetric covariance matrices estimated on each side of the central pixel are equal. Edges are detected when the hypothesis of equality is rejected. And the edge intensity of the central pixel can be computed from the likelihood-ratio value.* This edge intensity image is further filtered by means of the ridge filter (Xu and Jin, 2007) to provide the binary edge image. Finally

line segments are extracted from this binary edge image using the approach proposed in (Venkateswar and Chellappa, 1992). Thus a label image with line segments is achieved, in which the background has label value 0 and each extracted line segment has a distinct label value. *All the significant line segments in the image are extracted, including those belonging to the edges of buildings.*

### 2.1.3. New entries

With the low-level detectors, new region-based and edge-based inputs are now available: the rough region-based building detection result and the edge-based label image with line segments.

## 2.2. Optimized rectangle candidate definition

*In Step 2 of our method, the initial rectangle definition will be introduced in Section 2.2.1. These initial rectangles are further optimized in Section 2.2.2.*

### 2.2.1. Initial rectangle definition

In the label image with the extracted line segments, each line segment has a certain label value. All the pixels on one line segment are regarded as a group of samples, then a straight line segment can be obtained to fit these samples. In this way, the extracted line segments from the new entries are substituted by a *set* of straight line segments.



Figure 2: (a) *The pair of parallel line segments that can be used to define a rectangle.* (b) *The pair of parallel line segments that can not be used to define a rectangle.*

For each straight line segment, the line segments adjacent to it and parallel or orthogonal to it are sought out first. Then each pair of parallel or orthogonal straight line segments may define an initial rectangle. *For a pair of parallel line segments, there are four endpoints. If there exists at least one endpoint that satisfies the following condition, a rectangle can be defined*

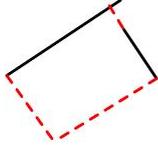


Figure 3: The pair of orthogonal line segments and the defined rectangle.

based on this pair of parallel line segments. Let the two parallel line segments be  $L1$  and  $L2$ . We assume the considered endpoint belongs to  $L1$ . Then there is a unique line, denoted by  $L3$ , that passes through this endpoint and is orthogonal to  $L2$ . The condition is that the intersection point of  $L3$  and  $L2$  is exactly on  $L2$ . If this condition is satisfied, the rectangle can be defined as shown in Figure 2 (a). In contrast, the pair of parallel line segments as shown in Figure 2 (b) can not be used to define a rectangle since the condition is not satisfied. For each pair of orthogonal line segments, a rectangle can be always defined. The intersection point of the two orthogonal lines provides the first vertex of the rectangle. On each line segment, there is an endpoint that is farther from the intersection point than the other. These two far endpoints provide the other two vertices of the rectangle. Then the rectangle can be defined as shown in Figure 3.

Let  $S_R$  be the set of the defined initial rectangles. We denote each rectangle  $R_i \in S_R$  by its four vertices  $\{R_i^k\}_{k=1}^4$  as  $R_i = R_i^1 R_i^2 R_i^3 R_i^4$ , where  $R_i^k$  is the  $k$ th vertex of  $R_i$ . In addition, each vertex is composed of one pixel in the image, thus it can be further represented by  $R_i^k = (R_{ix}^k, R_{iy}^k)$ , given  $R_{ix}^k$  and  $R_{iy}^k$  being the column and row indexes of this pixel in the image.

### 2.2.2. Rectangle optimization

For  $i$ th rectangle  $R_i$ , a mask  $M_i$  is defined, which is a box surrounding  $R_i$ . Similarly, we denote  $M_i$  by its vertices as  $M_i = M_i^1 M_i^2 M_i^3 M_i^4$ .  $M_i$  is obtained by moving the corresponding vertices of  $R_i$  along their respective off-center directions by an offset  $(N_x, N_y)$ , measured in the number of pixels, where  $N_x$  is the offset in the column direction while  $N_y$  denotes that along the row direction in the image. Here we set  $N_x = N_y = 15$ . With respect to  $R_i$ , the following four numbers must be computed in advance, based upon the *binary rough region-based building detection result obtained in Step 1 of our method and the mask  $M_i$* : the number of building pixels inside  $R_i$ , denoted by  $n_{1i}$ , that outside  $R_i$  but inside  $M_i$  by  $n_{2i}$ , and the total number of pixels inside  $R_i$ , say  $N_{1i}$ , as well as the total number of pixels outside  $R_i$  but inside  $M_i$ ,

by  $N_{2i}$ . The optimized  $\hat{R}_i$  inside  $M_i$  can be obtained according to (Simonetto et al., 2005), in which the criterion to be optimized is given by

$$\hat{R}_i = \max_{R_i|M_i} \{n_{1i}/N_{1i} \times (1 - n_{2i})/N_{2i}\}, \quad (1)$$

that is, to find the optimum  $\hat{R}_i$  that covers to the most the number of building pixels while leaving as few of them as possible within the domain between  $R_i$  and  $M_i$ . Each initially defined rectangle can be optimized in this way. The detailed optimization procedures were outlined in (Wang et al., 2008).

It is noteworthy in Eq. (1) that the accuracy of the optimized  $\hat{R}_i$  is restricted by the mask  $M_i$ . Thus the performance may deteriorate if we simply design the mask according to the initial value of  $R_i$  and keep it fixed during the optimization. To improve this, we suggest the optimization algorithm depicted in Table 1. This algorithm may always have the mask adapted if necessary.

Table 1 The optimization algorithm to obtain the rectangle candidates from the initial rectangles.

---

**Input:** Initial rectangle set  $S_R$ .

1. Initialize  $\hat{S}_R = \emptyset$ . Set  $i = 1$  and define  $N_R$  the total number of rectangles in  $S_R$ .
2. Get the candidate  $R_i \in S_R$ . Let  $R_i^{(0)} = R_i$  and initialize the iteration index with  $t = 1$ .
3. For  $R_i^{(t-1)}$  we define its mask  $M_i^{(t)}$ .
4. Find the optimized  $\hat{R}_i^{(t)}$  based on  $M_i^{(t)}$  under the optimization criterion in Eq. (1).
5. If the difference between  $\hat{R}_i^{(t)}$  and  $R_i^{(t-1)}$  is small enough or a maximum number of iterations has been reached:
  - if the  $[n_{1i}/N_{1i} \times (1 - n_{2i})/N_{2i}]$  value of  $\hat{R}_i^{(t)}$  is not 0, update the set by  $\hat{S}_R = \hat{S}_R \cup \{\hat{R}_i^{(t)}\}$ , break;
 otherwise:
  - let  $R_i^{(t)} = \hat{R}_i^{(t)}$  and  $t := t + 1$ . Then go to Step 3.
6. Increase  $i$  by 1. If  $i \leq N_R$  go to Step 2.

**Output:** Optimized rectangle candidate set  $\hat{S}_R$ .

---

In Step 5 of the optimization algorithm depicted in Table 1, the difference between  $\hat{R}_i^{(t)}$  and  $R_i^{(t-1)}$  is defined as the maximum value of the Euclidean distances between all the vertices of  $\hat{R}_i^{(t)}$  and their counterparts of  $R_i^{(t-1)}$ . It is noticed that in Step 5, if the  $[n_{1i}/N_{1i} \times (1 - n_{2i})/N_{2i}]$  value of one rectangle

is 0, this rectangle is not added into the optimized rectangle set  $\hat{S}_R$ , i.e., it is removed. Thus after this optimization algorithm, the final total number of rectangles is decreased.

### 2.3. The rectangle MRF framework

We introduce the proposed MRF framework for rectangles in this section. For each rectangle  $\hat{R}_i \in \hat{S}_R$ , its label value is denoted by  $L_i$ ,  $L_i \in \{0, 1\}$ , where  $L_i = 1$  means the rectangle  $\hat{R}_i$  belongs to buildings otherwise not. Now let  $N_R$  be the total number of elements in  $\hat{S}_R$ . If we model the label field  $\mathcal{L} = \{L_i, i \in \{1, 2, \dots, N_R\}\}$  with the MRF, the optimized configuration of  $\mathcal{L}$  can be achieved by minimizing the energy of the form (Szeliski et al., 2008)

$$E(\mathcal{L}) = E_{data}(\mathcal{L}) + \lambda E_{smooth}(\mathcal{L}). \quad (2)$$

In Eq. (2), the data energy term  $E_{data}(\mathcal{L})$  measures the disagreement between  $\mathcal{L}$  and the observations; the smoothness energy term  $E_{smooth}(\mathcal{L})$  contains the clique potentials of the MRF and describes the contextual information; and  $\lambda$  is the regularization parameter. A new definition for the *data* energy term will be given in Section 2.3.1. And the *smoothness* term is defined in Section 2.3.2. The minimization of the total energy in Eq. (2) will be described in Section 2.3.3, as well as the postprocessing.

#### 2.3.1. Data energy

For each pixel  $s$  in the original PolSAR images, the observed value is a complex polarimetric covariance matrix  $\mathbf{T}_s$ . If this  $3 \times 3$  sample covariance matrix has  $n_L$  number of looks (Maître, 2001), it is known that  $\mathbf{Z}_s = n_L \mathbf{T}_s$  follows the complex Wishart distribution (Lee et al., 1994)  $W_c(n_L, \Sigma)$ , i.e.,  $\mathbf{Z}_s \sim W_c(n_L, \Sigma)$ , where  $\Sigma$  is the expectation of the sample covariance matrices. A combined use of the rough binary building detection result and the original PolSAR data allows to estimate the Wishart distribution parameters  $\Sigma_1$  and  $\Sigma_0$  for the building class and the background class. Under the Maximum Likelihood (ML) criterion, a new label image  $\{L_s \in \{0, 1\}, s \in S\}$  is obtained, where  $L_s = 1$  means pixel  $s$  belongs to buildings otherwise background. Here the ML criterion can be formulated as

$$L_s = \begin{cases} 1, & \text{if } P(\mathbf{Z}_s | \Sigma_1) \geq P(\mathbf{Z}_s | \Sigma_0), \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

where  $P(\cdot)$  is the conditional probability.

Based on the new label image, for each  $\hat{R}_i \in \hat{S}_R$ , with the corresponding mask  $M_i$ , the data energy of  $\hat{R}_i$  is defined as

$$E_{data}(\hat{R}_i|L_i = 1) = -n_{1i}/N_{1i} \times (1 - n_{2i})/N_{2i}, \quad (4)$$

$$E_{data}(\hat{R}_i|L_i = 0) = -[1 - n_{1i}/N_{1i} \times (1 - n_{2i})/N_{2i}], \quad (5)$$

where  $n_{1i}$ ,  $n_{2i}$ ,  $N_{1i}$  and  $N_{2i}$  are given in Section 2.2.2. Then  $E_{data}(\mathcal{L})$  in Eq. (2) is given by

$$E_{data}(\mathcal{L}) = \sum_{i=1}^{N_R} E_{data}(\hat{R}_i|L_i). \quad (6)$$

### 2.3.2. smoothness energy

The *smoothness energy* term can be defined by the clique potentials of the MRF. Before defining the clique potentials, let us define the neighborhood of a rectangle first. For one vertex  $\hat{R}_i^k$  of the current rectangle  $\hat{R}_i$ , the  $N \times N$  neighborhood of  $\hat{R}_i^k$  is denoted by  $\mathcal{N}_{\hat{R}_i^k}$ . The neighbor rectangle set of  $\hat{R}_i$  is then defined as

$$\mathcal{N}_{\hat{R}_i} = \{\hat{R}_j \in \hat{S}_R | \exists (k', k) \in \{1, 2, 3, 4\}^2, \hat{R}_j^{k'} \in \mathcal{N}_{\hat{R}_i^k}, j \neq i\}. \quad (7)$$

Here we set  $N = 65$ . Only considering the second-order cliques, we have the *smoothness energy* term of the form

$$E_{smooth}(\mathcal{L}) = \sum_{\hat{R}_i, \hat{R}_j | \hat{R}_i \in \mathcal{N}_{\hat{R}_j}} V(L_i, L_j). \quad (8)$$

The second-order clique potentials  $V(L_i, L_j)$  are defined based on our prior knowledge about the buildings: the rectangle buildings in a small neighborhood usually have the same orientation, i.e., they are parallel or orthogonal to one another; two different buildings in the neighborhood should not overlap, however, in images two *adjacent* buildings can be considered to have small overlapping area. *We have made the assumption that the rectangle buildings in a small neighborhood are parallel or orthogonal to one another. Although this assumption does not always hold in the real world and the usefulness of the proposed method may be hence limited, it holds in many urban areas in developed countries.*

Some parameters used in the clique potential definitions are introduced first. The angle between two rectangles  $\hat{R}_i$  and  $\hat{R}_j$  is denoted by  $\theta_{ij}$ ,  $0 \leq$

$\theta_{ij} \leq \pi$ . The overlapping degree of  $\hat{R}_i$  and  $\hat{R}_j$  is evaluated by the parameter  $O_{ij} = \max\{Q(\hat{R}_i \cap \hat{R}_j)/Q(\hat{R}_i), Q(\hat{R}_i \cap \hat{R}_j)/Q(\hat{R}_j)\}$ ,  $0 \leq O_{ij} \leq 1$ , where  $Q(\hat{R}_i)$ ,  $Q(\hat{R}_j)$  and  $Q(\hat{R}_i \cap \hat{R}_j)$  are the area of  $\hat{R}_i$ , the area of  $\hat{R}_j$ , and the overlapping area of  $\hat{R}_i$  and  $\hat{R}_j$ , respectively. With these parameters, the second-order clique potentials are defined as

$$\begin{aligned} & \text{if } \min(\theta_{ij}, \pi - \theta_{ij}) \leq |\theta_{ij} - \pi/2|, \\ & \quad V(L_i = L_j = 1) = K_\theta W(|\sin(\theta_{ij})|, |\sin(\theta_{thr1})|) + K_O W(O_{ij}, O_{\max}), \\ & \text{else,} \\ & \quad V(L_i = L_j = 1) = K_\theta W(|\cos(\theta_{ij})|, |\cos(\theta_{thr2})|) + K_O W(O_{ij}, O_{\max}), \end{aligned} \quad (9)$$

$$V(L_i = 0 \quad \text{or} \quad L_j = 0) = 0, \quad (10)$$

where  $K_\theta$  and  $K_O$  are two weighting parameters;  $\theta_{thr1}$ ,  $\theta_{thr2}$  and  $O_{\max}$  are three thresholds. Function  $W(x, x_{\max})$ , where  $0 \leq x \leq 1$  and  $0 \leq x_{\max} \leq 1$ , has the property: if  $x \leq x_{\max}$ ,  $W(x, x_{\max}) \leq 0$ ; otherwise,  $W(x, x_{\max}) > 0$ . This function is defined based on that used in (Ortner et al., 2008), as formulated in Eq. (11).

$$W(x, x_{\max}) = -\frac{1}{x_{\max}^2} \left( \frac{1 + x_{\max}^2}{1 + x^2} - 1 \right) \quad (11)$$

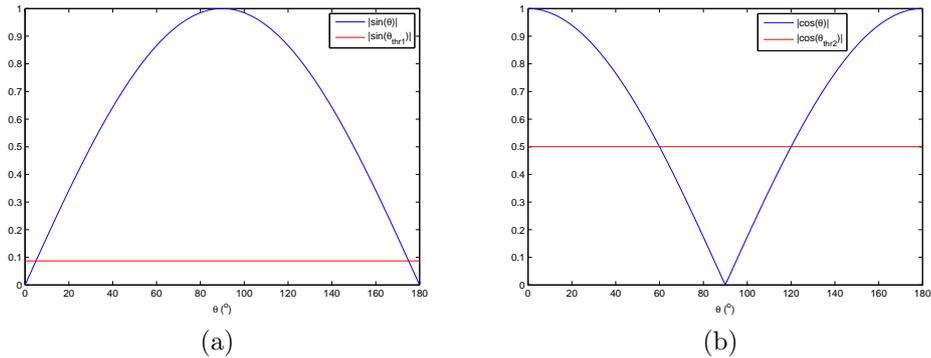


Figure 4: The curves of  $|\sin(\theta)|$  and  $|\cos(\theta)|$  versus  $\theta$ .

The  $W(O_{ij}, O_{\max})$  part in Eq. (9) is easy to understand: if the overlapping degree of the two rectangles in a clique is smaller than  $O_{\max}$ , these two rectangles are both encouraged to be labeled by 1. The curves of  $|\sin(\theta)|$  and  $|\cos(\theta)|$  shown in Figure 4 will help us to understand the meaning of

the  $W(|\sin(\theta)|, |\sin(\theta_{thr1})|)$  or  $W(|\cos(\theta)|, |\cos(\theta_{thr2})|)$  part in Eq. (9). As shown in Figure 4 (a), if  $\theta$  between the two rectangles is close to 0 or  $\pi$ , then  $|\sin(\theta)| < |\sin(\theta_{thr1})|$ , therefore  $W(|\sin(\theta)|, |\sin(\theta_{thr1})|) < 0$ ; otherwise,  $W(|\sin(\theta)|, |\sin(\theta_{thr1})|) \geq 0$ . This implies we favor the two rectangles in a second-order clique both to have label value 1 if they are parallel. Similarly, the two rectangles in a second-order clique are also favored to be both buildings if  $\theta$  between them is approximately  $\pi/2$ . To sum up, if two rectangles in a clique are parallel or orthogonal to each other and the overlapping part of them is small enough, the clique potentials defined in Eqs. (9) and (10) encourage the two rectangles both to be labeled by 1.

*Since during the image formation, the parallelism of two objects are well preserved,  $\theta_{thr1}$  should be set close to 0 or  $\pi$ . However, the angle between two orthogonal objects is usually no longer  $\pi/2$  in the obtained image due to different azimuth and range pixel spacing. Thus  $\theta_{thr2}$  should be set according to our prior knowledge about the angle between two orthogonal objects in the image. This threshold may be a little far from  $\pi/2$ . An example is given in Figure 4, in which  $\theta_{thr1} = 5/180 \times \pi$  and  $\theta_{thr2} = 60/180 \times \pi$ . Another threshold in our clique potential definition is  $O_{max}$ . We hope that the two detected buildings in a second-order clique are not overlapping too much, thus  $O_{max}$  should be set to a small value, such as 0.15.*

### 2.3.3. Energy minimization and postprocessing

The total energy in Eq. (2) is minimized by the ICM algorithm. Although the results may not be stable with different initializations, good initial values usually provide satisfactory regularization results. The estimation *minimizing only the data energy (the smoothness term is assigned to be 0)* is used as the initial value for the ICM algorithm in our experiments.

*After energy minimization, postprocessing is executed to combine two rectangles into one when the two rectangles satisfy the following condition: each long edge of the first rectangle and its counterpart of the second rectangle are on the same straight line, and the two rectangles are partly overlapping.*

### 2.4. Parameter setting

The parameters contained in the rectangle MRF framework are as follows: three thresholds  $\theta_{thr1}$ ,  $\theta_{thr2}$ , and  $O_{max}$ ; the weighting parameters  $K_\theta$  and  $K_O$ ; and the regularization parameter  $\lambda$ . *The setting of the three thresholds has been discussed at the end of Section 2.3.2. Then the left three parameters are*

$K_\theta$ ,  $K_O$  and  $\lambda$ . We analyze their influences on the regularization results in this section.

Two special situations shown in Figure 5 and Figure 6 are analyzed in the following to provide a general knowledge about how to choose the appropriate values. In order to introduce the *smoothness energy* influence, it is assumed that one rectangle in the clique already has label value 1. For brevity,  $f(\theta)$  is used to substitute  $|\sin(\theta_{ij})|$  or  $|\cos(\theta_{ij})|$ , and  $f(\theta_{thr})$  to substitute  $|\sin(\theta_{thr1})|$  or  $|\cos(\theta_{thr2})|$  in Eq. (9) henceforth.



Figure 5: The cliques in which one rectangle contains the other one.

In the first situation as shown in Figure 5, one rectangle contains the other one in the clique. If one rectangle has label value 1, we expect the other has label value 0, then, its energetic variation  $\Delta E$  between taking label value 1 and 0 should be positive. Since  $O_{ij} = 1$ ,  $\Delta E$  can be expressed as

$$\begin{aligned} \Delta E = & E_{data}(\hat{R}_i | L_i = 1) - E_{data}(\hat{R}_i | L_i = 0) \\ & + \lambda [K_\theta W(f(\theta), f(\theta_{thr})) + K_O \frac{1 - O_{\max}^2}{2O_{\max}^2}]. \end{aligned} \quad (12)$$

Let  $\Delta E_{data} = E_{data}(\hat{R}_i | L_i = 1) - E_{data}(\hat{R}_i | L_i = 0)$ . We have the following four cases.

1. If  $\Delta E_{data} < 0$  and  $W(f(\theta), f(\theta_{thr})) < 0$ ,  $\Delta E > 0$  implies

$$\lambda K_O \frac{1 - O_{\max}^2}{2O_{\max}^2} > |\Delta E_{data}| + \lambda K_\theta |W(f(\theta), f(\theta_{thr}))|. \quad (13)$$

2. If  $\Delta E_{data} < 0$  and  $W(f(\theta), f(\theta_{thr})) \geq 0$ ,  $\Delta E > 0$  means

$$\lambda K_\theta |W(f(\theta), f(\theta_{thr}))| + \lambda K_O \frac{1 - O_{\max}^2}{2O_{\max}^2} > |\Delta E_{data}|. \quad (14)$$

3. If  $\Delta E_{data} \geq 0$  and  $W(f(\theta), f(\theta_{thr})) < 0$ ,  $\Delta E > 0$  is equivalent to

$$|\Delta E_{data}| + \lambda K_O \frac{1 - O_{\max}^2}{2O_{\max}^2} > \lambda K_\theta |W(f(\theta), f(\theta_{thr}))|. \quad (15)$$

4. If  $\Delta E_{data} \geq 0$  and  $W(f(\theta), f(\theta_{thr})) \geq 0$ ,  $\Delta E > 0$  is always met.

It is noted that both Eqs. (14) and (15) hold if Eq. (13) is satisfied. With the knowledge that  $\max(|\Delta E_{data}|) = 1$ , and if  $W(f(\theta), f(\theta_{thr})) < 0$  then  $\max(|W(f(\theta), f(\theta_{thr}))|) = 1$ , that

$$\lambda K_O \frac{1 - O_{\max}^2}{2O_{\max}^2} > 1 + \lambda K_\theta \quad (16)$$

is sufficient for Eq. (13). Denote the inequality (16) by “Condition 1”.

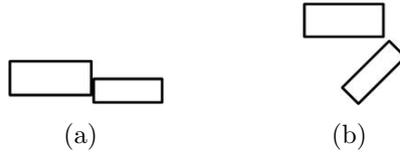


Figure 6: The cliques in which two rectangles are not overlapping at all.

In the second scenario as illustrated in Figure 6, two rectangles in the clique are not overlapping at all. If the label value of one rectangle is 1, we hope the labeling of the other depends only on its own data energy. Since  $O_{ij} = 0$ , its energetic variation  $\Delta E$  between taking label value 1 and 0 is

$$\Delta E = E_{data}(\hat{R}_i | L_i = 1) - E_{data}(\hat{R}_i | L_i = 0) + \lambda [K_\theta W(f(\theta), f(\theta_{thr})) - K_O]. \quad (17)$$

With the aforementioned definition for  $\Delta E_{data}$ , we have

1. If  $\Delta E_{data} \leq 0$  and  $W(f(\theta), f(\theta_{thr})) \leq 0$ ,  $\Delta E < 0$  is met by any means.
2. If  $\Delta E_{data} \leq 0$  and  $W(f(\theta), f(\theta_{thr})) > 0$ ,  $\Delta E < 0$  is equivalent to

$$K_\theta |W(f(\theta), f(\theta_{thr}))| - K_O < \frac{|\Delta E_{data}|}{\lambda}. \quad (18)$$

This is denoted by “Condition 2”.

3. If  $\Delta E_{data} > 0$  and  $W(f(\theta), f(\theta_{thr})) < 0$ ,  $\Delta E > 0$  implies

$$K_O + K_\theta |W(f(\theta), f(\theta_{thr}))| < \frac{|\Delta E_{data}|}{\lambda}, \quad (19)$$

which is termed by “Condition 3”.

4. If  $\Delta E_{data} > 0$  and  $W(f(\theta), f(\theta_{thr})) > 0$ ,  $\Delta E > 0$  yields

$$K_O - K_\theta |W(f(\theta), f(\theta_{thr}))| < \frac{|\Delta E_{data}|}{\lambda}, \quad (20)$$

i.e., “Condition 4”.

The effects of the three parameters  $\lambda$ ,  $K_O$  and  $K_\theta$  on the regularization results are analyzed based on the above four conditions. *For a certain image,  $|\Delta E_{data}|$  for each rectangle is fixed.* For simplicity, when we discuss the influence of one parameter, the other two are assumed to be fixed.

1. The effect of  $\lambda$ : Condition 1 implies

$$\lambda > \frac{1}{K_O \frac{1-O_{\max}^2}{2O_{\max}^2} - K_\theta}. \quad (21)$$

For Conditions 2, 3, and 4, it is noticed that the bigger  $\lambda$  is, the less possible that Conditions 2, 3, and 4 are met. So  $\lambda$  should not be too large, i.e., it should be around the value of  $1/[K_O(1 - O_{\max}^2)/(2O_{\max}^2) - K_\theta]$ .

2. The effect of  $K_O$ : Since Condition 1 is satisfied once  $\lambda$  is chosen according to Eq. (21), we only need to discuss the influence of  $K_O$  on Conditions 2, 3 and 4. The bigger  $K_O$  is, the more possible Condition 2 is met but the less possible Conditions 3 and 4 are met. This means it is more likely for the cliques shown in Figure 6 (a) and (b) to be detected as buildings.
3. The effect of  $K_\theta$ : Still, we only consider the influence of  $K_\theta$  on Conditions 2, 3 and 4. Condition 4 will be more easily satisfied given a bigger value for  $K_\theta$ , in contrast to an decreasing possibility for Conditions 2 and 3 to hold. That is, there is a higher possibility for the cliques as in Figure 6 (a) to be detected as buildings while opposite for the cliques in Figure 6 (b).

Generally, the effects of  $K_O$  and  $K_\theta$  on the regularization results are consistent with what we expect in Eq. (9). The bigger  $K_O$  is, the term relevant to the overlapping degree  $O_{ij}$  in the clique potentials casts more effects on the regularization results; *while given a bigger  $K_\theta$  value, the term dependent on the angle  $\theta_{ij}$  brings greater impact.*

### 3. Experimental results

The effectiveness of the proposed method is verified by experimental results on two different data sets. For each test site, the detection result using our method is compared with the region-based building detection result using the TF coherence approach and the edge-based result using the extended Hough transform.

#### 3.1. The first data set

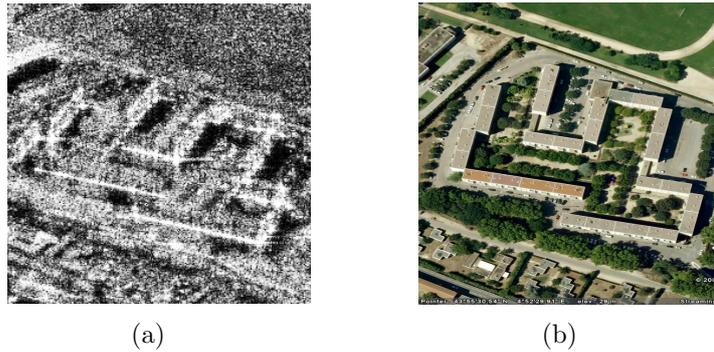


Figure 7: (a) The span image of the first data set. (b) The corresponding optical image from Google Earth ©2007 Google.

The first data set is a fully PolSAR data set with metric resolution. The images have  $256 \times 256$  pixels. The span image is shown in Figure 7 (a) and the corresponding optical image is shown in Figure 7 (b). In Figure 7 (a), the range is the top to the bottom while the azimuth is the left to the right. It is observed that in this test site, there are several buildings parallel or orthogonal to one another, located in the urban areas and surrounded by trees. The detection results in different steps of our method will be displayed in the following.

In Step 1 of our method, the obtained new inputs using the low-level detectors are shown in Figure 8. Figure 8 (a) is the binary rough region-based building detection result. Figure 8 (b) is the binary image with the extracted line segments. Here the label image with the extracted line segments is changed into a binary image for displaying purpose. In fact, each line segment in this label image has a distinct label value as mentioned before. These results are achieved with the following parameter values: the regularization parameter for the MRF based segmentation has the value of 3; the

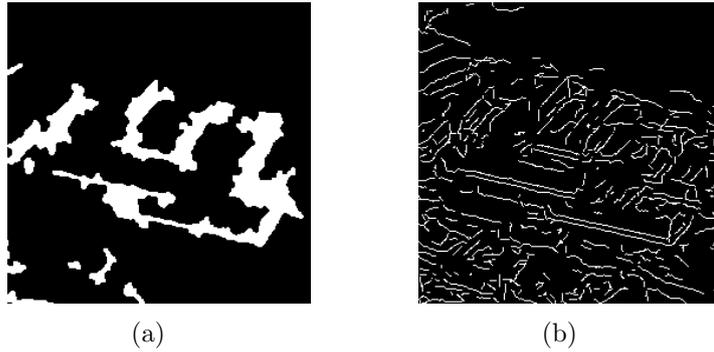


Figure 8: The low-level detectors' results for the first data set. (a) The binary rough region-based building detection result. (b) The extracted line segments.

threshold for the coherence image is 0.32; for the CFAR edge detector, the parameters are taken as  $7 \times 3$  window, 1 pixel gap, 4 directions and false alarm rate 0.01.

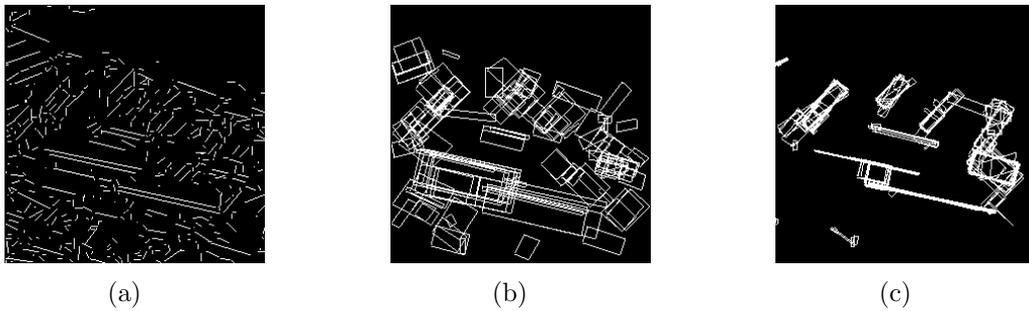


Figure 9: (a) The extracted straight line segments. (b) The initial rectangles. (c) The optimized rectangle candidates.

*Figure 9 displays the results obtained in Step 2 of our method. The extracted line segments [Figure 8 (b)] are substituted by a set of straight line segments [Figure 9 (a)]. Then initial rectangles are defined based on these straight line segments [Figure 9 (b)]. These initial rectangles are further optimized to provide the rectangle candidates. For the optimization algorithm depicted in Table 1, the maximum number of iterations in Step 5 is set to 5. The obtained rectangle candidates are shown in Figure 9 (c).*

*In Step 3 of our method, a new label image under the ML criterion is obtained in order to define the data term. The obtained new label image for*

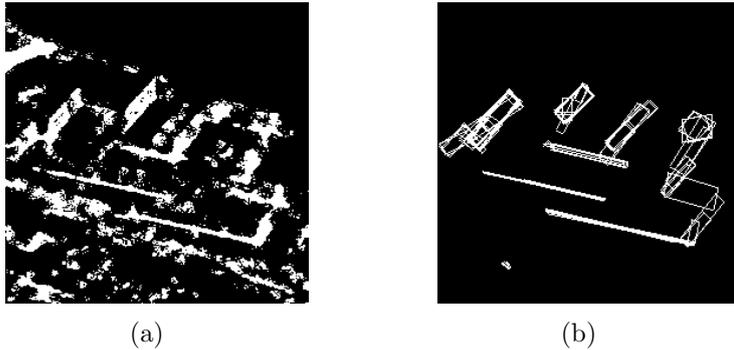


Figure 10: (a) The label image with two classes (the building class in white and the background class in black). (b) The detected buildings using only the data energy term.

the current test site is shown in Figure 10 (a). To verify the effectiveness of the data term, we neglect the smoothness term (no contextual information is considered) and assign  $L_i = 1$  when  $E_{data}(\hat{R}_i|L_i = 1) < E_{data}(\hat{R}_i|L_i = 0)$ . In this way the detected buildings are shown in Figure 10 (b). It is observed that most of the buildings in the scene are detected, which verifies the effectiveness of the proposed data term. Now we focus on the results after regularization. The three thresholds are set as  $\theta_{thr1} = 5/180 \times \pi$ ,  $\theta_{thr2} = 60/180 \times \pi$  and  $O_{max} = 0.15$  for this test site. In the following, the influences of  $K_\theta$ ,  $K_O$  and  $\lambda$  values on the regularization results are first illustrated. Then the final detection results are provided.

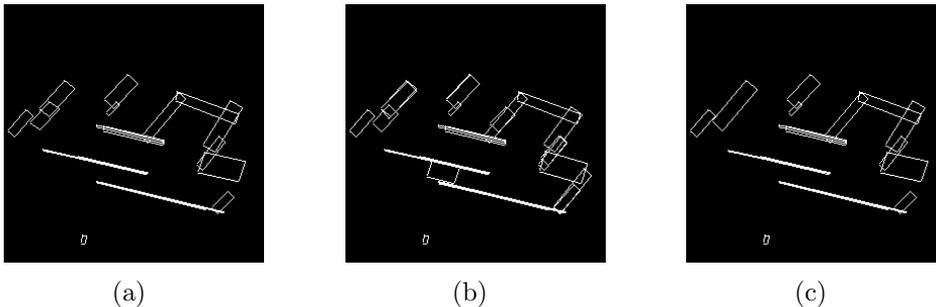


Figure 11: Regularization results for the first data set with different parameter values. (a)  $K_\theta = 1$ ,  $K_O = 1$  and  $\lambda = 0.05$ . (b)  $K_\theta = 1.5$ ,  $K_O = 1$  and  $\lambda = 0.05$ . (c)  $K_\theta = 1$ ,  $K_O = 1.5$  and  $\lambda = 0.035$ .

The effects of  $K_\theta$ ,  $K_O$  and  $\lambda$  values on the regularization results have been analyzed in Section 2.4. Based on these remarks, once the two weighting

parameters  $K_\theta$  and  $K_O$  are selected,  $\lambda$  can be set according to Eq. (21). Some realizations with different parameter values are displayed in Figure 11. Compared to the result in Figure 11 (a), more overlapping rectangles are detected as buildings in Figure 11 (b), since bigger  $K_\theta$  is used for the result in Figure 11 (b) and the overlapping term in the clique potentials affects the results less. When bigger  $K_O$  is used, *it is expected that less overlapping rectangles are detected. This is verified by Figure 11 (c), in which one less overlapping rectangle is detected as buildings compared to Figure 11 (a).*

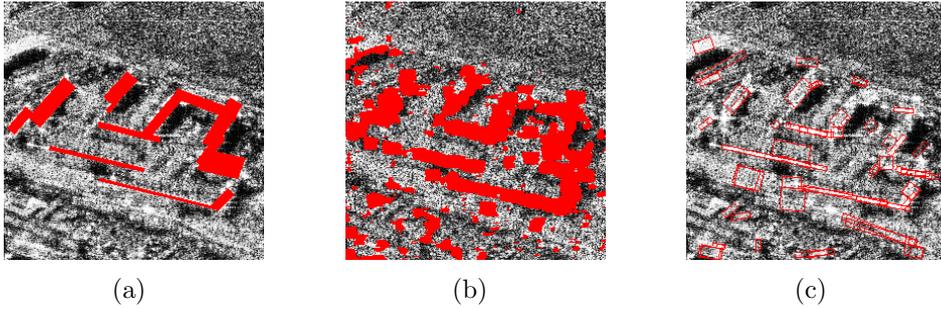


Figure 12: The final building detection results for the first data set. (a) The detection result of the proposed method. (b) The region-based detection result using the TF coherence approach. (c) The edge-based detection result using the extended Hough transform.

*The final detection result of the proposed method under the values  $K_O = 1$ ,  $K_\theta = 1$ , and  $\lambda = 0.05$  is displayed in Figure 12 (a). Compared with the region-based result in Figure 12 (b), the detected buildings in Figure 12 (a) have clearer shapes and edges since the fusion improves the region-based results and the edge constraints are considered. Meanwhile, there are less false detections in the non-building areas *in the result of our method [Figure 12 (a)] compared to the edge-based result in Figure 12 (c), which may profit from the region-based results. If we compare the detection result in Figure 12 (a) with the optical image in Figure 7 (b), it is observed that the main buildings in the scene are mostly detected. In addition, almost all of them have right locations and orientations. These results verify that the interaction of the region and edge information improves the building detection quality.**

*Due to the difficulty for the authors to collect the ground truth data, the detection results are quantitatively evaluated in a simple way based on our visual inspection. The probability of detection (POD) and the false alarm ratio (FAR) are used to evaluate the results. The actual number of buildings observed in the scene is denoted by  $N_{\text{build}}$ . For a real building region, if there*

is at least one detected rectangle overlapping with this region, this building is considered to be rightly detected. The number of rightly detected building is denoted by  $N_{right}$ . Then POD is given by  $POD = N_{right}/N_{build}$ . It measures the fraction of observed targets that are correctly detected. FAR is defined as follows. We denote the total number of detected rectangles by  $N_{rect}$  and the number of false detections by  $N_{false}$ . As long as one detected rectangle does not overlap with any real building region at all, it is regarded as a false detection. FAR is computed by  $FAR = N_{false}/N_{rect}$ . It gives the fraction of the detected targets that are observed to be non-targets. Since the region-based approach detects buildings as groups of pixels, we only compute the POD and FAR values of the results from the proposed method and the edge-based method. The POD and FAR values for the results of the first data set are shown in Table 2. From it, one can observe that the proposed method has a POD value of 92.86%, which is comparable to the POD value of the edge-based method (100%). While the FAR value of the proposed method (0%) is much lower than that of the edge-based method (44.26%). These values represent a good detection result of the proposed method.

Table 2 The POD and FAR values for the detection results of the first data set.

	The proposed method	The edge-based method
$N_{build}$	14	14
$N_{right}$	13	14
POD	92.86%	100%
$N_{rect}$	13	61
$N_{false}$	0	27
FAR	0%	44.26%

### 3.2. E-SAR data set

The second data set used for experiments is the German Aerospace Center (DLR) E-SAR L-band *fully* PolSAR data. *These data are obtained ...* The test site used is around the Dresden city (Germany), with  $512 \times 512$  pixels. The span image is shown in Figure 13, *in which the range is the left to the right while the azimuth is the top to the bottom.* This test site contains a lot of buildings. Most of them have the same orientations. However, exceptions exist, like the two bright vertical buildings in the middle of the image.

For this data set, the parameters are assigned with different values from the first data set. For the low-level detectors *in Step 1 of our method*, the

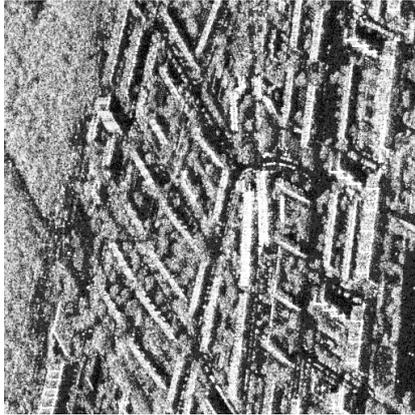
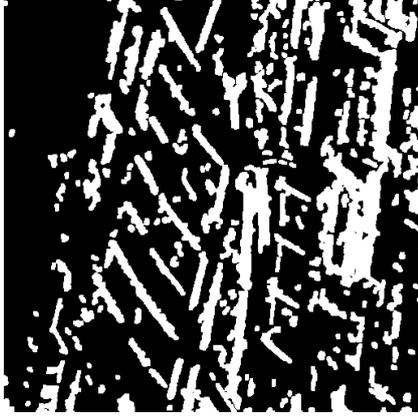


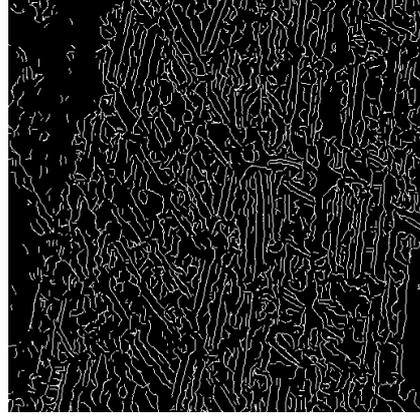
Figure 13: The span image of the test site from the E-SAR L-band fully PolSAR data around Dresden city ©DLR.

parameters relevant are set as the following values: the regularization parameter for the MRF based segmentation has the value of  $3\text{E-}6$ ; the threshold for the coherence image is 0.7; for the CFAR edge detector, the parameters are taken as  $7 \times 3$  window, 1 pixel gap, 4 directions and false alarm rate 0.01. *In Step 2, the maximum number of iterations for the optimization algorithm depicted in Table 1 is also set to 5.* For the parameters with respect to the MRF framework of rectangles *in Step 3*, the following values are used:  $\theta_{thr1}$  and  $\theta_{thr2}$  are set to  $15/180 \times \pi$  and  $60/180 \times \pi$ , respectively;  $O_{\max}$  is 0.05;  $K_{\theta}$ ,  $K_O$  and  $\lambda$  are 1, 1, and 0.0051, respectively.

The results obtained from the low-level detectors *in Step 1 of our method* are shown in Figure 14, *including the binary rough region-based building detection result [Figure 14 (a)] and the extracted line segments [Figure 14 (b)].* *In Step 2, we obtain the initial rectangles as shown in Figure 15 (a) and the optimized rectangle candidates as shown in Figure 15 (b).* *In Step 3, the label image used to define the data energy of the rectangles is shown in Figure 16 (a) while the detected buildings using only the data energy term are shown in Figure 16 (b).* *It is observed from Figure 16 (b) that almost all the rectangles located in the building areas are detected while those in the background areas are suppressed, which verifies the effectiveness of our data energy definitions.* *The final detection results of various methods are given in Figure 17. In the result of our method [Figure 17 (a)], most detected buildings have right orientations and locations. However, a few false detections also exist.* It is *noticed* that most of these falsely detected rectangles

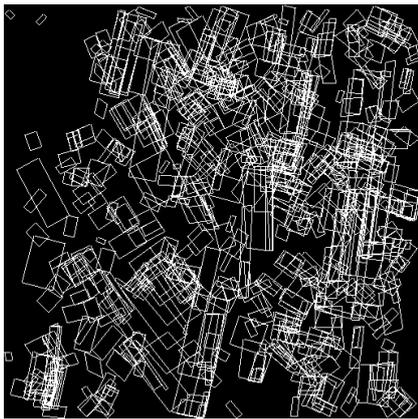


(a)

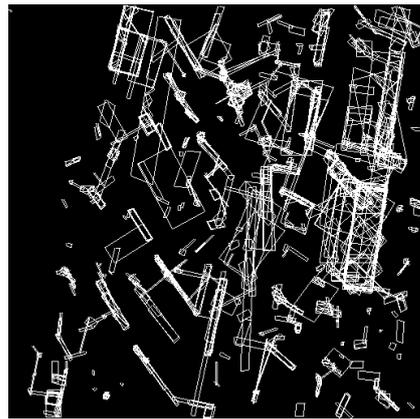


(b)

Figure 14: The low-level detectors' results for the E-SAR test site. (a) The binary rough region-based building detection result. (b) The extracted line segments.



(a)



(b)

Figure 15: (a) The initial rectangles. (b) The optimized rectangle candidates.

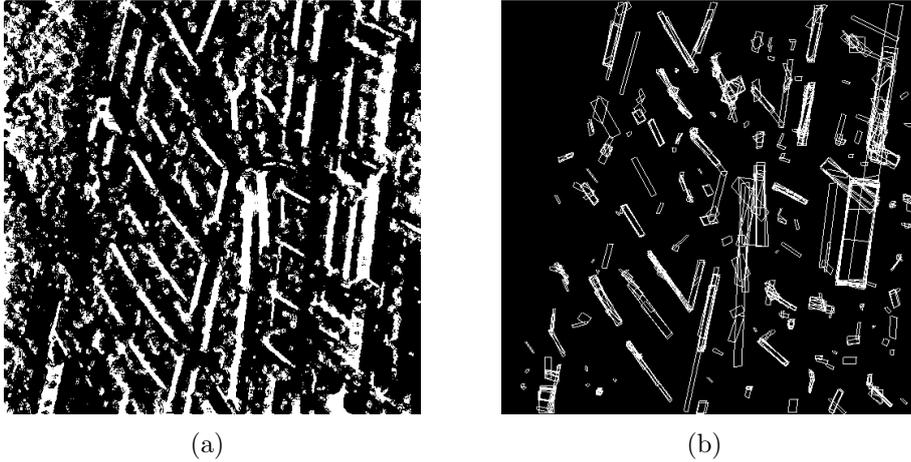


Figure 16: (a) The two class label image used to define the data energy of the rectangles. (b) The detected buildings using only the data energy term.

are not overlapping with their neighbor buildings and sharing the similar orientations. Also some buildings are missed in the result, which may due to small data energy values. The data energy values depend on the threshold used for the coherence image. When bigger threshold is used, fewer buildings will be detected finally. *On the contrary, if we decrease this threshold, there will be more false detections. Hence the appropriate threshold value should be selected empirically to reach a compromise between over-detection and incomplete detection. It is also observed that our method yields better performance than the approach using only the region information [Figure 17 (b)] or that using only the edge information [Figure 17 (c)].*

*The POD and FAR values for the results of this data set are displayed in Table 3. It is observed that the FAR value of the proposed method (32.43%) is comparable to that of the edge-based method (30.67%), but the proposed method has a higher POD value (82.61%) than the edge-based method (65.22%). These results verify the effectiveness of the proposed method and the usefulness of the combination of the region and edge information when detecting buildings.*

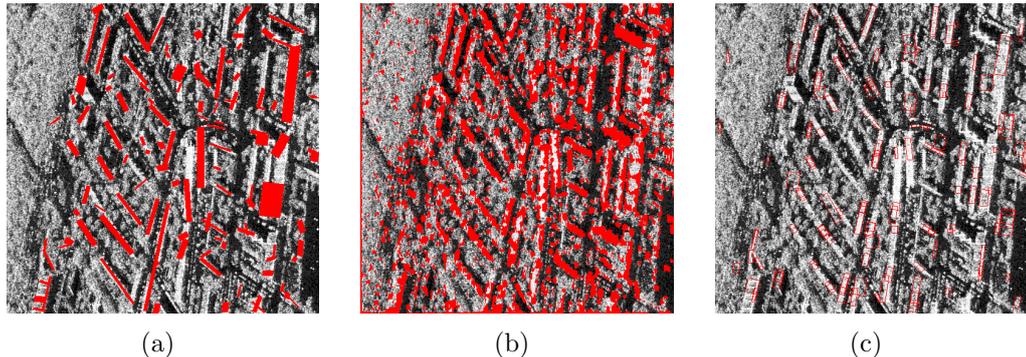


Figure 17: The final building detection results for the E-SAR test site. (a) The detection result of the proposed method. (b) The region-based detection result using the TF coherence approach. (c) The edge-based detection result using the extended Hough transform.

Table 3 The POD and FAR values for the results of the E-SAR data set.

	The proposed method	The edge-based method
$N_{build}$	69	69
$N_{right}$	57	45
POD	82.61%	65.22%
$N_{rect}$	111	150
$N_{false}$	36	46
FAR	32.43%	30.67%

#### 4. Conclusion

*In this paper, a new approach is proposed to detect buildings from the high-resolution PolSAR data. The rectangle features are extracted and analyzed from the high-resolution images directly. The edge-based information is employed to define the initial rectangles, which are further optimized based on the region information. The region-based information is then exploited to characterize the optimized rectangle candidates, i.e., to define the data energy terms of the rectangle candidates. And a novel MRF is defined on a set of rectangles. The smoothness term, composed of the second-order clique potentials of this MRF, takes into account the contextual information of buildings. The total energy is minimized to provide the final detection results. The experimental results using the real fully PolSAR data verifies the effectiveness of the proposed method. Even if we assign the smoothness term of the MRF*

framework to be 0, most buildings can be detected with right positions and orientations. However, many false detections also exist. After introducing the contextual information using the MRF of rectangles, the false detections are suppressed to some extent. The final detection results show that the combined use of the region and edge information improves the detection quality, compared to the results using only the region or edge information. When the buildings in one area are parallel or orthogonal to one another, the proposed method can rightly detect most of the buildings with few false detections.

It is noticed that in our MRF framework, there are six free parameters that should be set by the users. The robustness of the proposed method is influenced by these parameters, aslo by the initialization of the ICM algorithm used to minimize the total energy of the MRF framework. In order to get good detection results in practice, we proposed an analysis about how to choose the appropriate parameter values and also suggested a simple and comparatively good initialization for the ICM algorithm.

In the proposed approach, the final detected buildings are some rectangles selected from the optimized rectangle candidates. Therefore the accuracy of the optimized rectangle candidates affects the final detection efficiency greatly. To improve this, two aspects can be tracked in the future: one is to seek for a better optimization criterion; the other is to improve the region-based building detection results we used to fit the initial rectangles to.

The main drawback of our method is the assumption that the neighboring buildings in an area are parallel or orthogonal to one another. In fact, the orientations of buildings in the real world are versatile. The usefulness of the proposed method may be hence limited, *depending on the considered urban area.*

## 5. Acknowledge

The first data set used is provided by “Guichet Ramses, ONERA<sup>1</sup>”. The authors would like to thank ONERA, France for providing the images. The E-SAR data used are from the German Aerospace Center (DLR). The authors would like to thank DLR, Germany for providing these data. *This work was also supported in part by a grant from the State Key Program for Basic Research of China (973) (No. 2007CB311006). Moreover, the authors would*

---

<sup>1</sup>ONERA=Office National d’Etudes et de Recherches Aérospatiales

*like to thank Dr. Loïc Denis (from Ecole Supérieure de Chimie Physique Electronique de Lyon, France) and Dr. Laurent Ferro-Famil (from University of Rennes I, Rennes, France) for their help and suggestions. We are also grateful to the anonymous reviewers for their helpful comments that greatly improve the paper.*

## References

- Besag, J., 1986. On the statistical analysis of dirty pictures. *J. Roy. Statist. Soc.* 48 (3), 259–302.
- Boykov, Y., Veksler, O., Zabih, R., Nov. 2001. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (11), 1222–1239.
- Cloude, S. R., Pottier, E., Jan. 1997. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Trans. Geosci. Remote Sens.* 35, 68–78.
- Ferro-Famil, L., Pottier, E., 2007. Urban area remote sensing from L-band PolSAR data using Time-Frequency techniques. In: *Urban Remote Sensing Joint Event*. IEEE.
- Guillaso, S., Ferro-Famil, L., Reigber, A., Pottier, E., July 2005. Building characterization using L-band polarimetric interferometric SAR data. *IEEE Geoscience and Remote Sensing Letters* 2 (3), 347–351.
- Jodoin, P.-M., Mignotte, M., Rosenberger, C., Oct. 2007. Segmentation framework based on label field fusion. *IEEE Trans. Image Proces.* 16 (10), 2535–2550.
- Kersten, P. R., Lee, J.-S., Ainsworth, T. L., Mar. 2005. Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering. *IEEE Trans. Geosci. Remote Sens.* 43 (3), 519–527.
- Lee, J.-S., Grunes, M., Ainsworth, T., Du, L.-J., Schuler, D., Cloude, S., Sept. 1999a. Unsupervised classification using polarimetric decomposition and the complex Wishart classifier. *IEEE Trans. Geosci. Remote Sens.* 37 (5), 2249–2258.

- Lee, J.-S., Grunes, M. R., De Grandi, G., Sept. 1999b. Polarimetric SAR speckle filtering and its implication for classification. *IEEE Trans. Geosci. Remote Sens.* 37 (5), 2363–2373.
- Lee, J.-S., Hoppel, K. W., Mango, S. A., Miller, A. R., Sept. 1994. Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Trans. Geosci. Remote Sens.* 32 (5), 1017–1028.
- Maître, H., 2001. *Traitement des Images de Radar à Synthèse d’Ouverture*. Hermès Science Publication, Paris, France.
- Ortner, M., Descombes, X., Zerubia, J., Jan. 2008. A marked point process of rectangles and segments for automatic analysis of Digital Elevation Models. *IEEE Trans. Pattern Anal. Mach. Intell.* 30 (1), 105–119.
- Reigber, A., Jäger, M., He, W., Ferro-Famil, L., Hellwish, O., 2007. Detection and classification of urban structures based on high-resolution SAR imagery. In: *Urban Remote Sensing Joint Event*. IEEE.
- Schou, J., Skriver, H., Nielsen, A., Conradsen, K., Jan. 2003. CFAR edge detector for polarimetric SAR images. *IEEE Trans. Geosci. Remote Sens.* 41 (1), 20–32.
- Simonetto, E., Oriot, H., Garello, R., Oct. 2005. Rectangular building extraction from stereoscopic airborne radar images. *IEEE Trans. Geosci. Remote Sens.* 43 (10), 2386–2395.
- Szeliski, R., Zabih, R., Scharstein, D., Veksler, O., Kolmogorov, V., Agarwala, A., Tappen, M., Rother, C., June 2008. A comparative study of energy minimization methods for Markov random fields with smoothness-based priors. *IEEE Trans. Pattern Anal. Mach. Intell.* 30 (6), 1068–1080.
- Venkateswar, V., Chellappa, R., Nov. 1992. Extraction of straight lines in aerial images. *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (11), 1111–1114.
- Wang, Y.-H., Tupin, F., Han, C.-Z., Nicolas, J.-M., 2008. Building detection from the high-resolution PolSAR data by combining the region and edge information. In: *Proc. IGARSS*. IEEE.
- Xu, F., Jin, Y.-Q., July 2007. Automatic reconstruction of building objects from multiaspect meter-resolution SAR images. *IEEE Trans. Geosci. Remote Sens.* 45 (7), 2336–2353.