第35卷第2期
2016年4月

红外与毫米波学报
J. Infrared Millim. Waves

Vol. 35, No. 2
April, 2016

文章编号：1001－9014(2016)02－0160－06

DOI:10.11972/j.issn.1001－9014.2016.02.007

Combining high-resolution optical and SAR data for rectangle building detection and height estimation in urban areas

ZHANG Yong-Hua*, WEN Xian-Bin

(Key Laboratory of Computer Vision and System of Ministry of Education, Tianjin University of Technology, Tianjin 300384, China)

Abstract: A new complete scheme for urban rectangle building detection and height estimation was proposed using one high-resolution optical image and one synthetic aperture radar (SAR) image. The scheme exploited adequately the complementarities of both data and consisted of a sequential processing chain: building boundary extraction from both data, registration from optical image to SAR image based on an integrated criterion and building height estimation based on a "hypothesis generation-simulation-matching" procedure. Especially, the proposed method performed well also for partially occluded buildings. The method was tested and validated over two different scenes.

Key words: building detection, height estimation, synthetic aperture radar (SAR), optical
PACS: 84.40.Xb

联合高分辨率光学和SAR数据的城市区域矩形建筑物检测与高度估计

张永华*, 温显斌
(天津理工大学 计算机视觉与系统省部共建教育部重点实验室,天津 300384)

摘要:提出了一种新的完整地利用高分辨率光学图像和SAR图像进行城市矩形建筑物检测和高度估计的方法.该方法充分利用了两种数据的互补性,由三个顺序执行的步骤组成:首先对两种数据分别进行建筑物检测,然后基于一种整合的准则对两种图像进行配准,最后利用一种"假设生成-仿真-匹配"的方法对建筑物高度进行估计.该方法对部分遮掩的建筑物也能取得好的效果.对两个不同场景的测试验证了该方法的有效性.

关 键 词:建筑物检测;高度估计;合成孔径雷达;光学

中图分类号:TP75 **文献标识码**:A

Introduction

With the advantages of all-weather and all-time, synthetic aperture radar (SAR) plays an important role in remote sensing field. Nowadays, abundant high-resolution (HR) image data from airborne or spaceborne SAR systems provide great support for scene analysis and understanding in urban areas, especially for urban building detection and height estimation.

Urban building detection and height estimation have a wide range of applications such as urban change detection, disaster assessment, urban planning. Many methods for building detection and height estimation have been developed based on single-aspect SAR image[1-2], InSAR images[3], polarimetric SAR images[4] and tomographic SAR (TomoSAR)[5].

However, some specific phenomena like layerover, occlusion, shadow and the speckle generated based on the physical principle of SAR sensors will seriously hamper the performance of these methods. Meanwhile, other imaging systems such as Quickbird, LiDAR also bring massive available data. This provides a new chance to resolve the problem by exploiting adequately the complementarities of multi-source data. Only few researches have focused on combining optical image with SAR image for building detection and height estimation[6-8].

When only one optical and one SAR image are a-

Received date: 2015 - 06 - 26, **revised date**: 2015 - 12 - 25 **收稿日期**:2015 - 06 - 26, **修回日期**:2015 - 12 - 25
Foundation items:Supported by National Natural Science Foundation of China(61472278, 60872064)
Biography:ZHANG Yong-Hua(1986-), Linzhou China. Ph.D. Research interest includes remote sensing image processing
* **Corresponding author**:E-mail: zhangyhuaee@163.com

vailable, building detection and height estimation are still challenging tasks. Sportouche *et al*.[9] presented a semiautomatic processing chain to perform the simple 3D reconstruction of building in urban areas from HR optical and SAR image. Building footprint was extracted from monoscopic optical image and then the homologous building features were projected and registered for building validation and height estimation based on a SAR criteria. However, this method only performed well for the isolated building not for partially occluded buildings in the scene. In fact, in urban areas, partial occlusions between neighboring buildings are existing widely due to the interference with each other. The appearance of partially occluded buildings in SAR image is obviously different with the isolated ones, which complicates the problem of building detection and height estimation.

To solve the problem of partially occluded building detection and height estimation in urban areas, in this paper, we proposed a novel complete scheme by combining one optical image and one SAR image. Unlike the most existing methods, the proposed method not only performs well for the isolated building but also has high accuracy for partially occluded buildings. We assume that our building models are parallelepiped (rectangular footprints and flat roof). The complete flowchart of our method is shown in Fig. 1. The proposed scheme consists of three main steps: 1) the detection of potential rectangle building footprints from optical image and the extraction of building regions from SAR image; 2) the registration of extracted building footprints from optical image to SAR image; 3) the height estimation of refined buildings. The whole scheme achieves satisfying results for two different scenes.

1 Building detection

1.1 Building detection from optical image

To get high precision of building detection from optical image, a two-phase process was proposed. First, a coarse boundary map is generated by two region extraction algorithms: automatic threshold and the morphological algorithm. Automatic threshold method is used to extract the shadows which are clearly visible on the optical image. Then, the extracted shadows are masked to make the detection easier. The differential morphological profile (DMP) introduced in Ref. [10] is used to extract rectangle regions. The extracted rectangle regions with the size below a specific threshold are removed to reduce the false alarms. Then, an optimization criterion[12] is provided to get the refined map.

1.2 Building regions extraction from SAR image

The conditional random field (CRF) framework[3] was used to extract the building regions from SAR image. Unlike extracting characteristic features for building in both optical and InSAR data in Ref. [3], we extracted features only in SAR data.

y denotes all labels (either -1 or 1), x denotes all data, i denotes an image site out of the set of all image sites, j denotes the 4-connectivity neighborhood N_i of site i, and $Z(x)$ is the partition function. The building detection can be modeled as:

$$P(y \mid x) = \frac{1}{Z(x)} \exp\left[\sum_{i \in S} A_i(x, y_i) + \sum_{i \in S} \sum_{j \in N_i} I_{ij}(x, y_i, y_j)\right] . \quad (1)$$

The association potential $A_i(x, y_i)$ determines how likely a site is labeled with given all data x. Data x in our case is the SAR data. A linear model is used to distinguish building and non-building sites.

$$A_i(x, y_i) = \exp(y_i w^{\mathrm{T}} h_i(x)) , \quad (2)$$

where vector $h_i(x)$ contains all described node features. w^{T} is the weights of the features in $h_i(x)$ that are tuned during the training process.

The interaction potential $I_{ij}(x, y_i, y_j)$ measures how two sites i and j should interact regarding the data x.

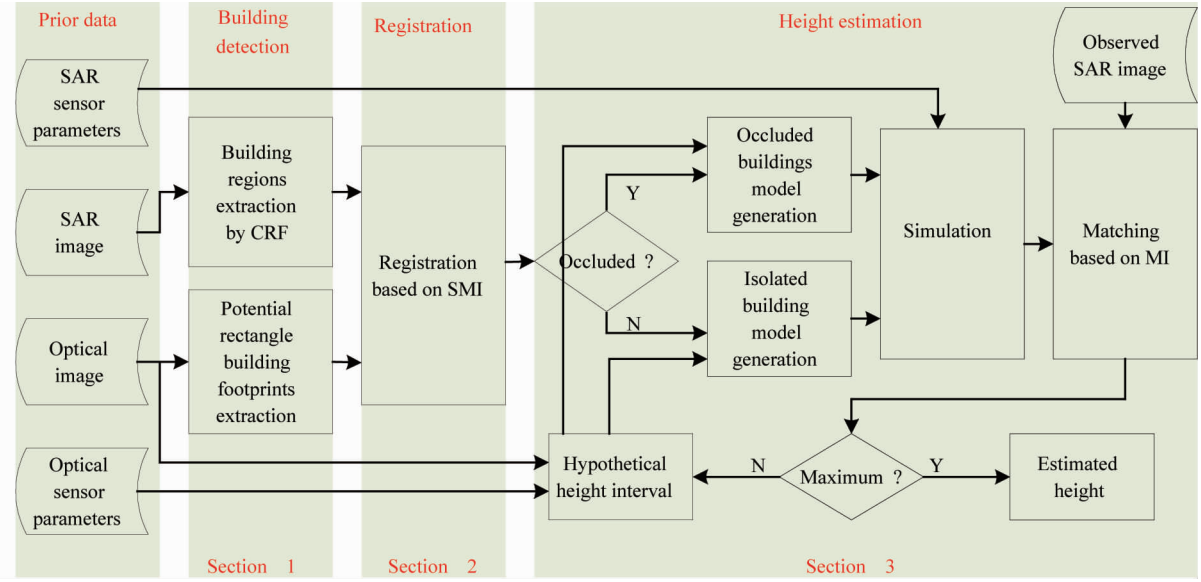


Fig. 1　Flowchart of the proposed method

图 1　所述方法的流程图

here,

$$I_{ij}(x,y_i,y_j) = \exp(y_i y_j v^{\mathrm{T}}\mu_{ij}(x)) \quad . \quad (3)$$

2 Registration

To get the accurate locations of the potential buildings in SAR image, we provided a registration procedure from optical image to SAR image. The choice of similarity metric is an important factor for effective registration process.

An integrated method (integration of area-based and feature-based methods) was used in registration[11]. The similarity metric consists of two components, the mutual information (MI) of image intensity and the spatial information (SI). The metric SMI can be defined as

$$\mathrm{SMI}(T_\alpha) = \mathrm{SI}_{\mathrm{A,B}}\mathrm{MI}_{\mathrm{A,B}} \quad . \quad (4)$$

The component SI of SMI is used to estimate near global optimal transformation and correct the bias of MI. The regions information is extracted from optical image and SAR image in the above step. Here, the centers of gravity of extracted building regions are used as the point sets. CP_{A} denotes the point set of optical image and CP_{B} denotes the point set of SAR image. The modified Hausdorff distance is used to measure the spatial relations of the point sets. SI is defined as:

$$\mathrm{SI}_{\mathrm{A,B}} = \sum_{cp_{\mathrm{A}}\in CP_{\mathrm{A}}} h\left(\min_{cp_{\mathrm{B}}\in CP_{\mathrm{B}}} \| cp_{\mathrm{A}} - cp_{\mathrm{B}} \|\right) \quad , \quad (5)$$

where

$$h(x) = \frac{1}{\sqrt{2\pi\sigma}}\exp\left(-\frac{x^2}{2\sigma^2}\right) \quad . \quad (6)$$

The component of MI is used to obtain the local optimal transformation with sub-pixel accuracy. MI is defined as the relevance of two random variables A and B. The formula of MI in this paper is defined as the following:

$$\mathrm{NMI}(A,B) = \frac{H(A)+H(B)}{H(A,B)} \quad , \quad (7)$$

where $H(A)$ and $H(B)$ is marginal entropy, $H(A,B)$ is joint entropy.

$$H(A,B) = -\sum_{i,j} p(i,j)\log p(i,j) \quad . \quad (8)$$

Only when both values of the two components of SMI are large will the SMI reach the maximum.

3 Height estimation

A "hypothesis generation-simulation-matching" procedure was proposed to estimate the height of both isolated buildings and partially occluded buildings in SAR image. Firstly, a series of height hypotheses are generated based on optical image. Then, the simulated images are generated by a SAR simulator using all known parameters. Finally, all the simulated images are matched with the observed SAR image. The height hypothesis for which the simulation best matches with the observed SAR image is the desired height.

3.1 Building geometric and backscattering characteristics

As shown in Fig. 2, the building geometric characteristic affects its appearance in the slant image plane due to the SAR sensor's active side-looking imaging. Figure 2(a) shows an example of the isolated flat roof building. The single bounce reflections from ground, front wall and roof generate the layerover area, denoted by a. The brighter line b denotes the dihedral corner. The area c is formed by the single bounce reflections from roof. The shadow area where no reflection happens is denoted by d. Figure 2(b) interprets the situation of partially occluded buildings. In Fig. 2(b), three cases are drawn for two neighboring buildings B_0 and B_1 with different heights. As shown in Fig. 2(b), only when the distance between two neighboring buildings is larger than a specific value will the occlusion not arise. This specific value, defined as the minimum distance $\Delta_{\min}$, can be calculated by:

$$\Delta_{\min} = h_1\tan(\beta) + h_0\cot(\beta) \quad , \quad (9)$$

where h_1 and h_0 are the heights of the buildings at the near and far range of the SAR sensor respectively, β is the depression angle.

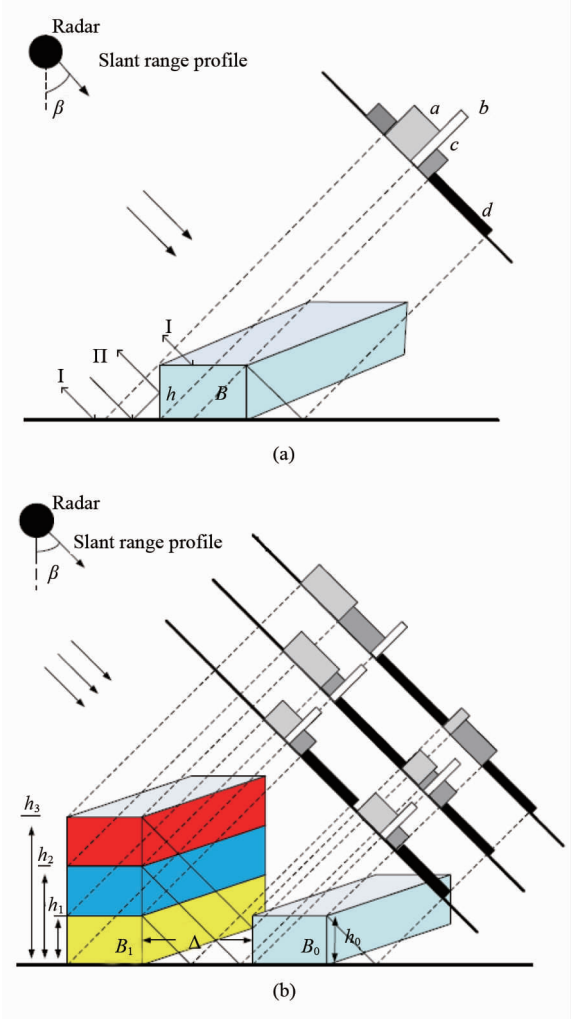


Fig. 2 Slant range profile of building for isolated building (a) and occluded buildings (b)

图2 建筑物斜距剖面(a)孤立的建筑物,(b)遮掩的建筑物

3.2 Problem statement of height estimation

A building B can be denoted as $B=(l,w,h,\phi,C)$, where w and l are the width and length of the building,

respectively, h is the height of building, ϕ is the aspect angle, C is the center of building footprint. β and H are the depression angle and the height of the SAR sensor, respectively, and are known. We denote $\theta = (l, w, \phi, C)$ and $\varphi = (H, \beta)$ in the vector form for simplicity.

From a model-based point, the problem of building height estimation can be regarded as searching for the best matched parameters with the observed SAR image. Based on this consideration, height estimation for the isolated building can be formulated as:

$$\hat{h} = \arg\max_{h} M\{X'(B(h,\theta) \mid \varphi), X\} \quad , \quad (10)$$

where X denotes the true SAR image, X' is the simulated SAR image at building height h, M is the matching function. Meanwhile, the heights and geometry parameters of the occluded buildings should be considered simultaneously, the formula in Eq. 10 is rewritten as:

$$\hat{\boldsymbol{h}} = \arg\max_{\boldsymbol{h}} M\{X'(B(\boldsymbol{h},\boldsymbol{\theta}) \mid \varphi), X\} \quad , (11)$$

where $\boldsymbol{h} = (h_1, h_2, \cdots, h_n)^{\mathrm{T}}$ is the height vector of the building and its occluded buildings, $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_n)^{\mathrm{T}}$ is the geometry parameter vector and n is the number of them. For $n = 1$, it is the representation of the isolated building, and for $n > 1$, it is the representation of the partially occluded buildings.

3.3 Solution of height estimation $\hat{h}$

The building height estimation problem has been formulated as the maximization function in Eq. 11. This section defines the generation of height hypothesis, the simulated scheme, matching function and the optimization algorithm to find the function's best solution.

3.3.1 Height hypothesis generation

In Eq. 11 only the height parameter $\boldsymbol{h}$ is variable and other parameters are known as prior. In order to reduce the number of simulation and improve the efficiency, the initial height $\tilde{h}$ of each building is estimated by the shadow from the optical image based on the geometrical imaging formula. Then, the interval of hypothetical height of each building is determined as $[\tilde{h} - 3\mathrm{m}; \tilde{h} + 3\mathrm{m}]$ sampled in a fixed step η. Thus, each building has $N = 6/\eta$ possible heights.

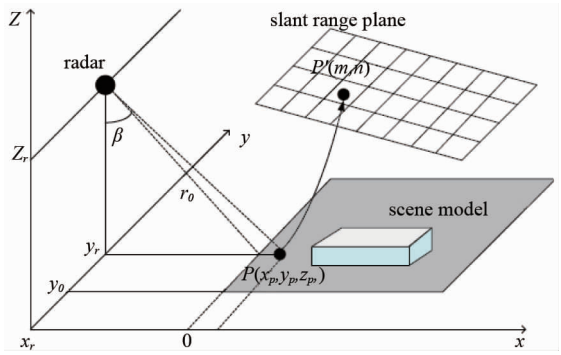


Fig. 3 Mapping from the scene coordinates to the slant range plane

图 3 场景坐标到斜距平面的映射

3.3.2 The simulated scheme

As shown in Fig. 3, let (x_p, y_p, z_p) be the coordinates of the point P in the scene, and (m, n) is the coordinates of its corresponding mapping point P' in the slant range plane, the mapping from P to P' follows the basic range Doppler equation. Thus, the coordinates of P' can be deduced by the equations:

$$\begin{cases} m = \dfrac{\sqrt{(x_p - x_r)^2 + (y_p - y_r)^2 + (z_p - z_r)^2} - r_0}{\Delta m} \\ n = \dfrac{y - y_0}{\Delta n} \end{cases} \quad , \quad (12)$$

where r_0 is the near slant range of the scene, (x_r, y_r, z_r) is the location of radar, Δm and Δn are the azimuth and range resolutions, respectively, y_0 denotes the initial azimuth coordinate.

Firstly, the isolated building and partially occluded buildings are distinguished based on Eq. 9 and the corresponding scene models are built respectively. Then, based on Eq. 12, the scene models are simulated in a line-by-line way. The backscattering of each point is calculated based on a radiation model with respect to local incident angle, which is the angle between radar wave and scene surface element. The backscattering is accumulated in the corresponding range unit. By repeating the process for all of the scanning lines, we can get the whole simulated image. In the simulated process, we also integrated the shadow regions and the effect of speckle noise to the final simulated image.

3.3.3 Definition of matching function

The matching function M was used to measure the similarity between the observed SAR image and the simulated image. The choice of matching function influences greatly the results of height estimation. Comparing the real SAR data to the simulated image, the geometries of the images are similar, but the local statistics and the radiometry are different. Considering the drawback of feature-based methods dependent on the effectiveness of the feature extraction process, we chose mutual information as the matching function M, defined in Sect. 2.

4 Experiments

To evaluate the performance of the proposed 3D reconstruction method, the results of the building detection and height estimation process for the test data are presented in this section.

4.1 The results for the isolated building

4.1.1 Description of the test data

A couple of Quickbird image (resolution of 0.68 m) and TerraSAR-X image (resolution of 1.1 m in range and 1.1 m in azimuth; incidence angle of 32°) on an urban area (Marseille, France) were used to verify the proposed method for isolated building. The method proposed by Hélène *et al*. [9] is a representative urban building 3D reconstruction method based on the fusion of optical and SAR images in recent years. Due to employing a non-symmetric scheme, we denoted this method as NSS for simplicity. We used the same data set for comparison. As shown in Fig. 4, the scene is composed of rectangular residential buildings located in an urban area.

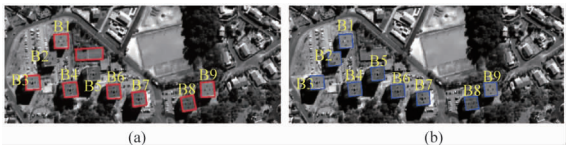
4.1.2 Building detection and registration results

Results of building footprint extraction from optical

(a) (b)

Fig. 4 Scene 1. (a) optical image, and (b) SAR image
图4 场景1(a)光学图像,(b)SAR图像

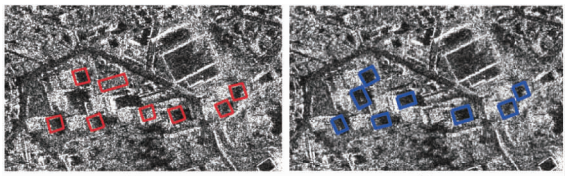
image are shown in Fig. 5. For scene 1 with nine actual buildings, eight potential footprints (seven good detections and one false alarm) were extracted using NSS, while nine ones (nine good detections and no false alarm) were proposed using our method.



(a) (b)

Fig. 5 Results of the building footprint extraction derived with NSS (a) and our method (b)
图5 建筑物轮廓提取结果(a)NSS,(b)本文方法

The results of the building footprint projection and registration are shown in Fig. 6. To obtain more accurate results, our method was combined with the results of CRF. It is worthy although a good detection (B6) is removed in this scene. While for NSS, the building footprints are registered simply with the SAR image no matter whether the potential buildings are good detection or false alarm.



(a) (b)

Fig. 6 Results of the building footprint projection and registration, superimposed to the SAR image. (a) NSS, and (b) our method
图6 建筑物轮廓提取和配准结果,显示于SAR图像上(a)NSS,(b)本文方法

4.1.3 Height estimation results

In addition to NSS, we also compared other two height estimation algorithms which were proposed in Ref. [2] to evaluate the performance of our method. One algorithm is based on the Sun shadow from optical image (denoted as SSO), which is the computational method for initial height. The other algorithm is based on the overlap of double-bounce line in SAR image and roof edge in optical image (denoted as OLE). The results of the height estimation are presented in Table 1 for the test data set. As shown in Table 1, it is clearly shown that our estimated heights of buildings are very approximate to the ground truth. The root mean square error (RMSE) of height estimation proves that our method obtains higher accuracy than other algorithms.

Table 1 The results of height estimation on scene 1
表1 场景1高度估计结果

scene 2	ground truth	estimated height				error			
building	h_{truth}	h_{NSS}	h_{SSO}	h_{OLE}	h_{our}	E_{NSS}	E_{SSO}	E_{OLE}	E_{our}
B1	11.2	11	13	12.1	11.3	−0.2	+1.8	+0.9	+0.1
B2	11	miss detected	12.6	11.6	10.7	–	+1.6	+0.6	−0.3
B3	19.9	20	21.2	18.7	19.5	+0.1	+1.3	−1.2	−0.4
B4	20.3	21	19.1	19.8	20.6	+0.7	−0.4	−0.5	+0.3
B5	20.6	miss detected	21	20.2	20.2	–	+0.4	−0.4	−0.4
B6	19.4	miss detected	19.2	18.8	19.3	–	−0.2	−0.6	−0.1
B7	20.6	22.5	20.9	21.3	20.1	+1.9	+0.3	+0.7	−0.5
B8	14.4	14	13.5	14.1	14.6	−0.4	−0.9	−0.3	+0.2
B9	15	15	14.6	14.5	15.4	0	−0.4	−0.5	+0.4
		RMSE:				0.85	0.99	0.68	0.33

4.2 The results for the partially occluded buildings

4.2.1 Description of the test data

A pair of optical image (from Google Earth) and COSMO-SkyMed SAR image (resolution of 1m in range and 1m in azimuth; incidence angle of 38°) on an university campus (Tor Vergata, Italy) were used to verify the proposed method for partially occluded buildings.

As shown in Fig. 7, the scene is composed of seven rectangular buildings and one non-regular building. Although distinguished obviously in the optical image, these buildings generate some mixtures in the SAR image. This makes many building detection and height estimation algorithms invalid because they only take the isolated building into consideration.

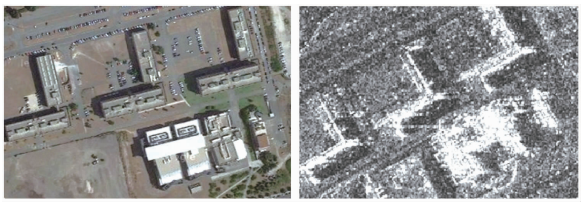


Fig. 7 Scene 2 on the optical image and SAR image
图7 场景2(a)光学图像,(b)SAR图像

4.2.2 Building detection and registration results

Figure 8 shows the results of building detection from optical image and SAR image. For the optical image, ten potential footprints (six good detections and four false alarms) were extracted with NSS, while nine ones (seven good detections and two false alarms) were proposed with our method. Detection for the complicated building C7 is beyond the scope of this paper and it will be our further research. For the SAR image, as shown in Fig. 8(c), the buildings (C1-C2, C3-C4 and C5-C6) are grouped as a new L-structure which can not be distinguished without additional information.

Figure 9 demonstrates that one false alarm is removed in the registration process with our method while all the false alarms are still reserved with NSS. Therefore, our registration results are more accurate than NSS.

4.2.3 Height estimation results

The results of partially occluded buildings with four algorithms (NSS, SSO, OLE and our method) are listed

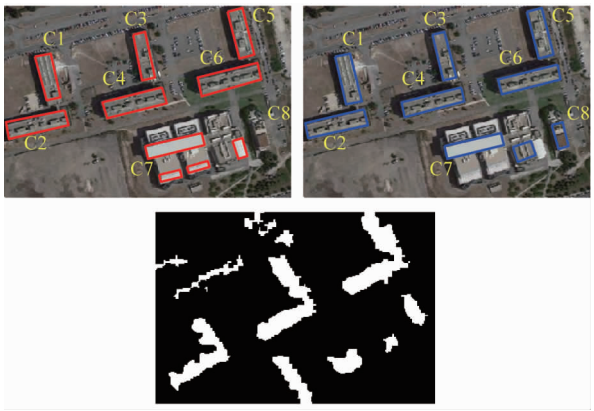


Fig. 8　The results of building footprint extration from optical image and SAR image. (a) NSS. (b) out method, And (c) the results of SAR image
图 8　建筑物轮廓提取结果(a)NSS,(b)本文方法,(c)SAR 图

in Table 2. The RMSE with NSS (3.60) is much higher than others. This indicates that NSS has a very poor ability to deal with partially occluded buildings. The performances of SSO and OLE are also barely satisfactory although they have a higher accuracy than NSS. While, our method brings exciting results (RMSE of 0.69) for partially occluded buildings.

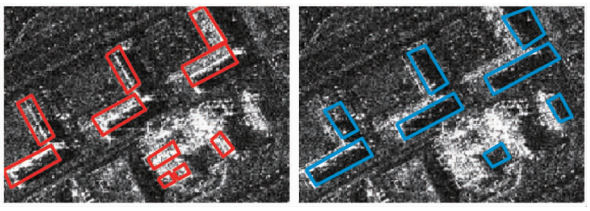


Fig. 9　The results of registration. (a) NSS. (b) our method
图 9　配准结果(a)NSS,(b)本文方法

Table 2　The results of height estimation on scene 2
表 2　场景 2 高度估计结果

scene 3	ground truth	estimated height				error			
building	h_{truth}	h_{NSS}	h_{SSO}	h_{OLE}	h_{our}	E_{NSS}	E_{SSO}	E_{OLE}	E_{our}
C1	7.2	5.1	6.3	6.1	6.5	−2.1	−0.9	−1.1	−0.7
C2	12	9.3	10.9	10.3	11.4	−2.7	−1.1	−1.7	−0.6
C3	7.2	10	5.4	8.6	8.3	+2.8	−1.8	+1.4	+1.1
C4	16	12.4	14.3	14.2	16.5	−3.6	−1.7	−1.8	+0.5
C5	10.5	7.8	13.2	13.6	11.4	−2.7	+2.7	+3.1	+0.9
C6	20	16.5	22.3	19.7	19.6	−3.5	+2.3	−0.3	−0.4
C7	9	15.2	6.9	7.4	Miss detected	+6.2	−2.1	−1.6	–
C8	8.1	Miss detected	9.5	9.3	7.8	–	+1.4	+1.2	−0.3
		RMSE:				3.60	1.84	1.70	0.69

5　Conclusion

A systematic scheme for urban building detection and height estimation from HR SAR and optical image was presented and illustrated on real scenes. Compared with other similar methods, our approach achieved more satisfying results for the isolated building. Especially, the approach also achieved good performance for the partially occluded buildings. The main limitation of the method is the exhaustive matching for all possible height hypotheses. In further works, we will focus in building with more complicated shapes (such as gable-roof, polyhedral and multi-structure), and a sophisticated technique is desired for such building detection and height estimation.

References

[1] Ferro A, Brunner D, Bruzzone L. Automatic detection and reconstruction of building radar footprints from single VHR SAR images [J]. *IEEE Trans. Geosci. Remote Sens.*, 2013. **51**(2): 935 −952.

[2] Wegner J D, Ziehn J R,. Soergel U. Combining high-resolution optical and InSAR features for height estimation of buildings with flat roofs [J]. *IEEE Trans. Geosci. Remote Sens.*, 2014, **52**(9): 5840 −5854.

[3] Wegner J D,. Haensch R, Thiele A, *et al*. Building detection from one orthophoto and high-resolution InSAR data using conditional random fields [J]. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 2011, **4**(1):83 −91.

[4] Xu F, Jin Y.-Q. Automatic reconstruction of building objects from multiaspect meter-resolution SAR images [J]. *IEEE Trans. Geosci. Remote Sens.*, 2007, **45**(7): 2336 −2353.

[5] Shahzad M,. Zhu X. Robust reconstruction of building facades for large areas using spaceborne TomoSAR point clouds [J]. *IEEE Trans. Geosci. Remote Sens.*, 2015, **53**(2): 752 −769.

[6] Denis L, Tupin F, Darbon J, *et al*. Joint regularization of phase and amplitude of InSAR data: Application to 3-D reconstruction [J]. *IEEE Trans. Geosci. Remote Sens.*, 2009, **47**(11): 3774 −3785.

[7] Wegner J D, Ziehn J R, Soergel U. Building detection and height estimation from high-resolution InSAR and optical data [C]. *IEEE IGARSS*, 2010, pp: 1928 −1931.

[8] J. Zhu, C. Ding, H. You et al. 3D reconstruction of building based on high-resolution SAR and optical images [C]. *IEEE IGARSS*, 2006, pp: 3794 −3797.

[9] Sportouche H, Tupin F, Denise L. Extraction and three-dimensional reconstruction of isolated buildings in urban scenes from high-resolution optical and SAR spaceborne images [J]. *IEEE Trans. Geosci. Remote Sens.*, 2011, **49**(10): 3932 −3946.

[10] Chanussot J, Benediktsson J A,. Fauvel M. Classification of remote sensing images from urban areas using a fuzzy possibilistic model. *IEEE Geosci. Remote Sens. Lett.*, 2006, **3**(1): 40 −44.

[11] Liang J, Liu X, Huang K, *et al*. Automatic registration of multisensor images using an integrated spatial and mutual information (SMI) metric [J]. *IEEE Trans. Geosci. Remote Sens.*, 2014, **52**(1): 603 −615.

[12] LeguizamÕn G, Coello A. A study of the scalability of ACOR for continuous optimization problems [C] Departamento de Computación, Mexico, Tech. Rep. 2010.