

Adaptive Neighborhood Selection Semi-supervised Discriminative Locality Alignment Based Urban Building Areas Extraction from High-Resolution SAR Images

Bo Cheng and Shiai Cui

Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China

Email: {chengbo, cuisai}@radi.ac.cn

Abstract—Currently, the majority of manifold learning algorithms applied to SAR image feature extraction are unsupervised. The semi-supervised manifold learning becomes a research trend for it can make full use of class information and be more coincident with actual data. Considering the non-uniform distribution of high-resolution SAR data and the problem of manually set neighbor values, adaptive neighborhood selection was introduced into the Semi-supervised Discriminative Locality Alignment (SDLA) and adaptive neighborhood selection semi-supervised discriminative locality alignment (ANSSDLA) was proposed for building extraction of high-resolution SAR data. Then, single polarization TerraSAR-X data and fully polarization RADARSAT-2 data were taken as experiment data to validate the ANSSDLA algorithm. It was found that the ANSSDLA algorithm has strong adaptability by comparing the results of ANSSDLA and SDLA. In addition, the result of comparative experiments of multi-polarization data shows that cross polarization data is more suitable for building extraction.

Index Terms—synthetic aperture radar, manifold learning, adaptive neighborhood selection, semi-supervised discriminative locality alignment

I. INTRODUCTION

Synthetic Aperture Radar (SAR) has the unique advantages in its all-time and all-weather observation capabilities to compensate defects of optical sensors that unable to obtain valid data in cloudy, foggy and rainy weather, and it has become an important technique of remote sensing information extraction [1]. With the increase of resolution, urban SAR data has more and more abundant and unique features. The SAR technique is playing an increasingly important role in building area extraction, which provided the basic data for urban planning, land use monitoring and population density survey [2]. However, the high resolution SAR image of city is very complex, and its high dimensional nonlinear characteristics also increase the difficulty in automatic extraction. As a machine learning method, manifold learning can find the inherent characteristics of data and

is good at processing nonlinear data. It can improve the accuracy of target recognition by applying manifold learning to feature extraction of high-dimensional nonlinear SAR image.

By using its local geometric structure, manifold learning finds intrinsic manifold structure of a data set to achieve the purpose of efficient dimension reduction, e.g., Locally Linear Embedding (LLE) [3], Isometric feature mapping (ISOMAP) [4], Local Tangent Space Alignment (LTSA) [5], Laplacian Eigenmap (LE) [6] and so on. Currently, most manifold learning methods are unsupervised. In order to make full use of the class information, supervised manifold learning was proposed. However, it is not easy to obtain a large number of data with categories information. Therefore, the semi-supervised manifold learning method which uses a small amount of class information becomes a research trend and has extensive application value. Semi-supervised Discriminative Locality Alignment (SDLA) is based on Discriminative Locality Alignment (DLA) [7]. DLA can deal with nonlinear samples, enhance the importance of marginal samples to learn low dimensional representations and avoid the matrix singularity problem. The application of SDLA to feature extraction of high-resolution SAR image can improve the accuracy of target recognition. Considering that the distribution of SAR data is not uniform and the neighborhood selection of unlabeled samples in SDLA affects extraction result, the adaptive neighborhood selection technique was introduced into the SDLA algorithm and the adaptive neighborhood selection semi-supervised discriminative locality alignment (ANSSDLA) was proposed for building extraction of high-resolution SAR in this paper. Section II introduces the proposed method, which contains data preprocessing, high-dimensional feature cube construction and ANSSDLA algorithm. Section III introduces the experimental data. The experiment results of this algorithm are presented in Section IV. Section V gives the conclusion and discuss.

II. METHODOLOGY

The framework of building extraction is composed of filtering, feature set, dimensional reduction of ANSSDLA,

classification and post-processing. Firstly, the Lee filter was applied to high-resolution SAR image. Secondly, eight texture features were obtained by using the classic second order probability statistical method (gray level co-occurrence matrix, GLCM), and combined with gray feature to construct original high-dimensional feature set of SAR image. Thirdly, the ANSSDLA algorithm reduced the data set of training samples and the projection matrix was obtained to deal with testing samples respectively to reduce dimension and achieve new features. Finally, the new features were used to extract building areas as the input of simple classifier. Post-processing was conducted to achieve the final results.

A. The Feature Set

The building area has rich texture features different from non-building area. The classical texture extraction algorithm is GLCM proposed by Haralick [8] in 1973. Texture features were acquired through statistical calculation of the gray level co-occurrence matrix. Angular Second Moment, Contrast, Correlation, Dissimilarity, Energy, Homogeneity, Mean and Variance were commonly used.

B. Adaptive Neighborhood Selection Semi-supervised Discriminative Locality Alignment (ANSSDLA)

SDLA is a semi-supervised manifold learning algorithm based on DLA, only part of the samples is labeled. The unlabeled samples and the labeled samples can be expressed as:

$$X = [X_i, \dots, X_N, X_{N+1}, \dots, X_{N+N_U}] \in R^{m \times N_u}$$

where the front part are labeled samples, the rest N_U samples are unlabeled. The aim of the feature extraction is to get $Y \in R^{d \times N_u}$ ($d < m$), which is the projection of sample X in the low-dimensional feature space, and to get a projection matrix A , expressed as $Y = A^T X$. The algorithm mainly includes part optimization and whole alignment.

In the first part, we construct local patch structure for each sample of data set and design objective function to keep local discrimination information. Since each single sample can be seen as part of the integral data set, this stage is called patch optimization.

For the labeled sample X_i , the local patch $X_i = [X_{i1}, X_{i1}, \dots, X_{ik_1}, X_{i1}, \dots, X_{ik_2}]$ was built by searching k_1 's neighbor samples from the same class and k_2 's neighbor samples from different classes. Each local patch corresponding to output in low-dimensional feature space should be $Y_i = [y_{i1}, y_{i1}, \dots, y_{ik_1}, y_{i1}, \dots, y_{ik_2}]$. In the low-dimensional space, the best case is that the distance between X_i and neighbor samples from the same class is shortest, and the distance between X_i and neighbor samples from different classes is longest.

Since the patch consisting of neighbor samples can be seen as approximate Euclidean space, and it constructs local optimization function:

$$\arg \min_{Y_i} \left(\sum_{s=1}^{k_1} \|y_i - y_{is}\|^2 - \beta \sum_{t=1}^{k_2} \|y_i - y_{it}\|^2 \right) \quad (1)$$

where β is a scaling parameter, valuing in $[0,1]$, used to weight the within-class distance and the between-class distance. Define a coefficient vector as follows:

$$\omega_i = \begin{bmatrix} \overbrace{1, \dots, 1}^{k_1}, \overbrace{-\beta, \dots, -\beta}^{k_2} \end{bmatrix}^T \quad (2)$$

Then (1) can be rewritten as:

$$\arg \min_{Y_i} \text{tr}(Y_i L_i Y_i^T) \quad (3)$$

$$L_i = \begin{bmatrix} \sum_{s=1}^{k_1+k_2} (\omega_i)_s & -\omega_i^T \\ -\omega_i & \text{diag}(\omega_i) \end{bmatrix}$$

For those unlabeled samples x_i , $i = N+1, \dots, N+N_U$, search k_s 's neighbor in training samples, construct patch, where include both labeled samples and unlabeled samples. The patch optimization function is:

$$\arg \min_{Y_i} \sum_{j=1}^{k_s} \|y_i - y_{ij}\|^2 = \arg \min_{Y_i} \text{tr}(Y_i L_i^U Y_i^T) \quad (4)$$

where, $L_i^U = \begin{bmatrix} k_s & -e_{k_s}^T \\ -e_{k_s} & I_{k_s} \end{bmatrix}$, $e_{k_s} = [1, \dots, 1]^T \in R^{k_s}$,

I_{k_s} is the identity matrix.

Then, the whole alignment is to obtain the full optimization by summing all the patch optimizations of samples. $Y_i = [y_{i1}, y_{i1}, \dots, y_{ik_1}, y_{i1}, \dots, y_{ik_2}]$ was actually selected from $Y = [y_1, \dots, y_{n_1}]$, we define:

$$Y_i = Y S_i, (S_i)_{pq} = \begin{cases} 1, & \text{if } p = F_i(q) \\ 0, & \text{else.} \end{cases} \quad (5)$$

The whole alignment:

$$\arg \sum_{i=1}^N \min_{Y_i} \text{tr}(Y_i m_i L_i Y_i^T) + \gamma \arg \sum_{i=N+1}^{N+N_U} \min_{Y_i} \text{tr}(Y_i L_i^U Y_i^T) \quad (6)$$

$$= \arg \min_Y \text{tr}(Y L^S Y^T)$$

where, m_i is a margin degree parameter, through an iterative sum we have: $L(F_i, F_i) \leftarrow L(F_i, F_i) + m_i L_i$, $F_i = \{i, i1, \dots, ik_1, i1, \dots, ik_2\}$ is the patch index vector for samples. In the end, to get projection matrix, (6) can be rewritten as:

$$\arg \min_A \text{tr}(A^T X L X^T A) \text{ s.t. } A^T A = I_d \quad (7)$$

The hypothetical condition for the unlabeled samples in SDLA is the uniform distribution of data, and each sample has same neighborhood size set by manual

experiences. It is not applicable to SAR data with speckle noise, which likely shortens the divergence between classes and cannot recognize the target. Accordingly, the way to adaptively selecting neighborhood was introduced into SDLA. The procedure is listed as follows:

1. Select the initial k nearest neighbors of x_i ;
2. Compute the average Euclidean distance and manifold distance, then achieve the adaptive parameter k_i ;
3. Compared k_i and k , to increase or decrease the corresponding number of neighbors [9].

III. DATA

The experiment was conducted on two sets of SAR data from single polarization TerraSAR-X data and full polarization RADARSAT-2 data. The polarization mode of TerraSAR data is HH and spatial resolution is 1.25m. The data format of RADARSAT-2 is SPG and the spatial resolution is 3.125m. Two sub images of 1000×1000 were selected from TerraSAR-X data (Fig. 1) to validate the proposed algorithm. Fig. 1(a) was taken as training image and Fig. 1(b) was test image. The sub image of 1000×1000 from RADARSAT-2 (Fig. 2) was to explore the application of multi-polarization SAR in building extraction. And the experimental images were all filtered.



Figure 1. TerraSAR-X images.

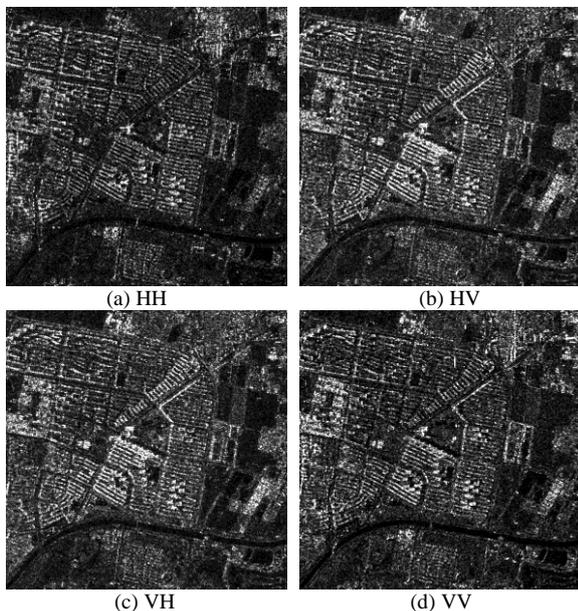


Figure 2. RADARSAT-2 images.

IV. RESULTS

A. The Feature Set

Feature set is composed of texture feature and gray feature. Eight texture features were extracted with GLCM, containing different information. The filtered images present gray feature. GLCM is mainly influenced by direction, step length, window size and image gray level [10]. The greatest influence on texture extraction of images is window size, which should be chosen according to the actual situation. Since building area, plant and water is main objects in urban, we analyzed the trend of each texture feature of different objects to determine the optimal window size.

Taking Mean as example (Fig. 3), Mean was calculated in different window sizes with interval of 4 between the window size of 3×3 and 39×39 . As shown in Fig. 3, the value of Mean in building area, plant and water constantly increases from the window size of 3 to that of 23, and slightly decreases at the window size of 27, then picks up. The increasing trend slows down between the window size of 31 and that of 39. Moreover, with the increase of the window size, the edge of texture images become more and more fuzzy. Therefore, the window size of 31×31 is more appropriate. Fig. 4 shows the feature set.

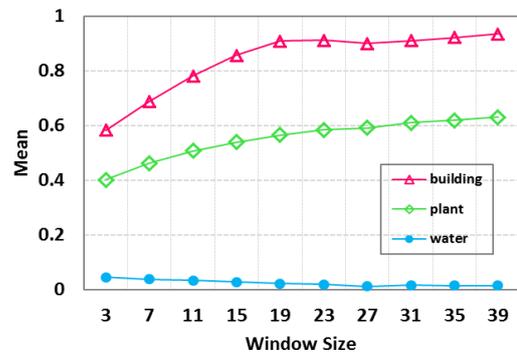


Figure 3. Variation trend of Mean with the window size

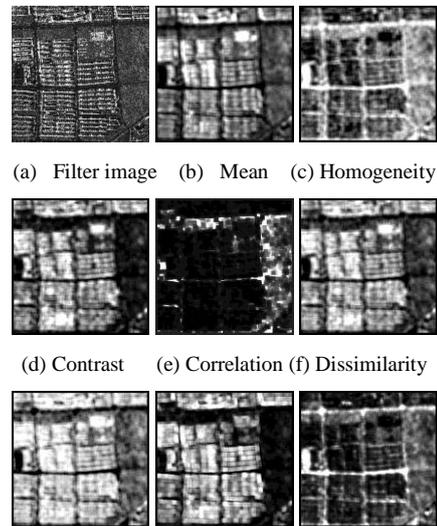


Figure 4. Filtered image and texture feature images.

B. Dimensionality Reduction of Feature Set

To validate the ANSSDLA algorithm, both experiments with SDLA algorithm and ANSSDLA algorithm were conducted on the base of TerraSAR-X data (Fig. 1). Fig. 1(a) was training image, and part of samples in training image had labeled information. A projection transformation matrix was respectively obtained through these two algorithms. Then a new feature was achieved by the operation between this matrix and the high-dimensional feature set (Fig. 5). The new features of the whole shows a gray stone based on SDLA algorithm, and the contrast between building and non-building is not obvious. However, building in the new feature based on ANSSDLA are in bright color and has a clear outline. The new feature is more conducive to extract building area.

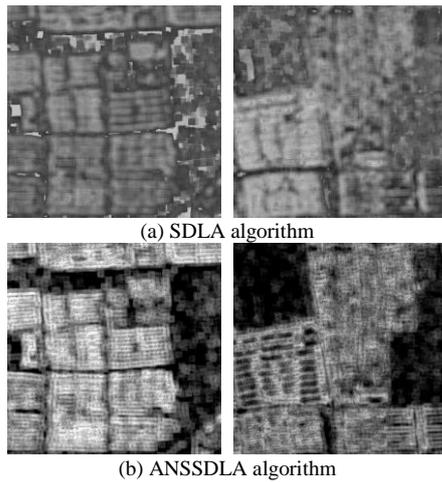


Figure 5. New features extracted by ANSSDLA algorithm and SDLA algorithm

C. Extraction of Building Areas and Accuracy Evaluation

The Otsu [11] method on the new feature image was used to extract building area. Then post-processing was made to achieve the final building areas (Fig. 6). Based on the principle of confusion matrix, Detection Rate (DR) [12], False Alarm Rate (FAR), Missing Alarm Rate (MAR) and Overall Accuracy (OA) were calculated to evaluate the accuracy of building extraction.

There are obvious differences between ANSSDLA and SDLA algorithm in Fig. 6 and Table I. The ANSSDLA shows a better result. The detection rate (DR) of the image achieved by ANSSDLA algorithm is higher than that achieved by SDLA algorithm. The overall accuracy (OA) of ANSSDLA method is more than 80%, and it shows that the new feature based on ANSSDLA is in favor of building extraction. The ANSSDLA algorithm can keep spatial geometry of feature set and improve recognition accuracy. For the test image, the DR reach 95%. Thus, ANSSDLA algorithm has good generalization ability. The projection matrix obtained by training image can be directly applied to feature set of test image and obtain good discriminated new feature.

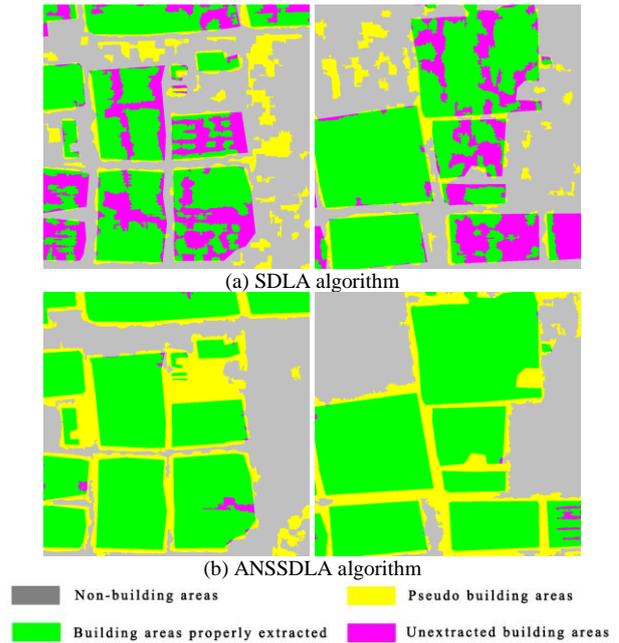


Figure 6. The extraction results of urban building areas

TABLE I. VALUES OF ACCURACY EVALUATION INDEXES (UNITS: %)

Algorithm	Experimental images	Detection Rate	False Alarm Rate	Missing Alarm Rate	Overall Accuracy
ANSSDLA	Training image	98.63	26.99	1.37	81.21
ANSSDLA	Test image	99.23	22.07	0.77	84.31
SDLA	Training image	66.89	23.98	33.11	73.8
SDLA	Test image	73.97	15.55	27.03	78.47

D. Application of the Multi-polarization SAR Data in Building Extraction

With the development of SAR, the multi-polarization data is gradually increasing. The full-polarization SAR data is rich in information. To study application of polarization SAR to extract building area with manifold learning, nine groups of contrast experiments were designed using the sub image of RADARSAT-2 (Fig. 2). In the first four groups, each polarization image followed the framework of building extraction based on ANSSDLA algorithm, and each polarization image generated a new feature. In the other six groups, two arbitrary polarization images combined and the classifier chose Random Forest (RF) [13]. The comparative experiment used random forest directly to obtain feature set to extract building. The last experiment was four polarization images using random forest classier.

As we can see from Table II and Table III that all indexes of single polarization data based on the ANSSDL is higher than that using random forest directly. Therefore, it verifies the result in the last section. In Table II and Table III, the OA of cross polarization data is higher than one of co-polarization.

TABLE II. VALUES OF ACCURACY EVALUATION INDEXES OF SINGLE POLARIZATION DATA (UNITS: %)

Experimental images	HH	HV	VH	VV
DR	100	99.36	99.44	100
FAR	46.42	30.52	30.62	47.54
MAR	0	0.64	0.56	0
OR	56.14	77.58	77.5	54.14

TABLE III. VALUES OF ACCURACY EVALUATION INDEXES OF SINGLE POLARIZATION DATA (UNITS: %)

Experimental images	HH+	HV+	VH+	VV+
	RF	RF	RF	RF
DR	77.87	99.2	98.19	87.64
FAR	54.97	49.37	49.62	52.19
MAR	22.12	0.80	1.80	12.35
OR	40.69	50.62	50.12	45.31

Table IV shows the results of two arbitrary polarization data in the ANSSDLA framework. The best combination is that of cross polarization data and vertical polarization data. HH+HV has the lowest value of OA. While all the full polarization data were used to extracted building area, the value of OA was less than that of HV+VH, VH+VV, and HV+VV. The result may be related to the quality of the image itself. (See Table V and Table VI).

TABLE IV. VALUES OF ACCURACY EVALUATION INDEXES OF DOUBLE POLARIZATION DATA (UNITS: %)

Experimental images	HH+	HH+	HH+	VH+	VH+	HV+
	HV	VH	VV	HV	VV	VV
DR	39.39	44.30	89.00	95.15	95.22	96.13
FAR	15.52	15.64	32.95	27.37	25.05	25.26
MAR	60.61	55.7	11.00	4.85	4.78	3.89
OR	65.66	67.65	72.3	79.4	81.47	81.6

TABLE V. VALUES OF ACCURACY EVALUATION INDEXES OF DOUBLE POLARIZATION DATA (UNITS: %)

Experimental images	HH+	HH+	HH+	VH+	VH+	HV+
	HV+ RF	VH+ RF	VV+ RF	HV+ RF	VV+ RF	VV+ RF
DR	99.35	99.23	84.64	97.92	98.76	98.51
FAR	49.33	49.37	52.76	49.69	49.48	49.54
MAR	0.65	0.77	15.36	2.08	1.24	1.49
OR	50.71	50.64	44.37	49.99	50.41	50.30

TABLE VI. VALUES OF ACCURACY EVALUATION INDEXES OF FULL POLARIZATION DATA (UNITS: %)

Experimental images	Full polarization+ANSSDLA	Full polarization+RF
DR	84.39	92.24
FAR	29.21	51.34
MAR	15.6	7.76
OR	74.47	46.81

V. CONCLUSIONS

In this paper, a novel semi-supervised manifold learning was introduced for building extraction in urban area. The image was filtered to reduce noise as preprocessing. Eight texture features by GLCM and filter image is both stacked into the input feature set. Then, the ANSSDLA algorithm was adopted to obtain a low-dimensional projection in which the discriminative information of some training samples can be best preserved. Finally, the recognized results were achieved by using Otsu in the reduced subspace. Two sub images of TerraSAR-X data took experiment to valuate. The experimental results show that the introduced ANSSDLA has the potential to keep the intrinsic geometry features of SAR, which could deal with out-of-sample problem. The projection matrix can be applied to test image and obtained a good discriminative feature. Also, comparative experiments were conducted between single polarization and multi-polarization. The result shows that the accuracy of building extraction with cross polarization SAR data is higher than that of full Polarization SAR data. The performance of HH image is the lowest, which is caused by the image itself. It is expected to further study in later time. Meanwhile, experiment data only chose SPG full polarization SAR including few polarization information. Our future work will be focused on the polarization classification of urban area.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation of China under Grant No. 61372189.

REFERENCES

- [1] F. L. Zhang and Y. Shao, "urban target monitoring using high resolution SAR data," *Remote Sensing Technology and Application*, vol. 25, no. 3, pp. 415-422, 2010.
- [2] J. Xu, "Buildings extraction based on scattering features for polarimetric SAR images," M.S. thesis, China University of Mining and Technology, China, 2013.
- [3] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323-2326, 2000.
- [4] J. B. Tenenbaum, V. D. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319-2323, 2000.
- [5] Z. Y. Zhang and H. Y. Zha, "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment," *SIAM Journal of Scientific Computing*, vol. 26, no. 1, pp. 313-338, 2005.
- [6] M. Belkin and P. Niyogi, "Laplacian Eigen maps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373-1396, 2003.
- [7] T. Lin and H. Zha, "Riemannian manifold learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 5, pp. 796-809, 2007.
- [8] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610-621, 1973.
- [9] S. Roweis, L. K. Sual, and G. E. Hinton, "Global coordination of local linear models," presented at Neural Information Processing Systems Conference, Vancouver, British Columbia, Canada, Dec. 9-14, 2002. pp. 889-896.
- [10] F. Wu, C. Wang, and H. Zhang, "Residential areas extraction in high resolution SAR image based on texture features," *Remote Sensing Technology and Application*, vol. 20, no. 2, pp. 148-153, 2005.

- [11] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
- [12] J. A. Shufelt, "Performance evaluation and analysis of monocular building extraction from aerial imagery," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 311-326, 1999.
- [13] Y. Liu, P. J. Du, H. Zheng, *et al.*, "Classification of China small satellite remote sensing image based on random forests," *Science of Surveying and Mapping*, vol. 37, pp. 194-196, 2012.



Bo Cheng received the Ph.D. degrees in cartography and geographic information engineering from China University of Geoscience (Beijing) in 2003.

Since then, he has been working for 10 years in photogrammetry and remote sensing fields at Institute of Remote Sensing and Digital Earth (RADI), Chinese Academy of Sciences (CAS). He has actively conducting research using remote sensing and GIS in support of natural resource, urban land-use monitoring, and environmental applications.



Shiai Cui studied in Lanzhou University for a bachelor's degree from 2009 to 2013. She attended Institute of Remote Sensing and Digital Earth Chinese Academy of Sciences in 2013 for a master's degree and her direction is remote sensing image processing.