Semantic Segmentation of Polarimetric SAR Imagery Using a Few Well-selected Training Samples

Wen Yang, Xun Zhang, Lijun Chen, Hong Sun

Signal Processing Lab, School of Electronic Information, Wuhan University, Wuhan, 430079, China

1. INTRODUCTION

During the last decade, multi-frequency and polarimetric SAR (PolSAR) imaging has been investigated with respect to classification of terrain types, many supervised and unsupervised segmentation and classification methods for PolSAR data have been proposed. However, it is still very difficult to get a reliable and consistent scene semantic segmentation for PolSAR imagery. Recently, with the introduction of Conditional Random Field (CRF), the use of discriminative models for semantic segmentation and classification tasks has become popular [1]. CRF directly models the conditional probability which is able to incorporate a rich set of arbitrary non-independent overlapping features of the observations. It has been shown to outperform the MRF based generative models in semantic segmentation of natural images [2]. In general, these systems require a large number of full labeled images to produce an effective segmenter. Unfortunately, it requires the labor-intensive and time-consuming works to label every pixel in training samples. Furthermore, due to the intrinsic speckle phenomenon, sometimes only experts of SAR image interpretation are qualified for making ground truth labels. This often limits the amount of training data available, which can lead to an inferior segmentation system. In this work we will investigate the semantic labeling of PolSAR imagery based on CRF with a few well-selected training samples, and an active learning process will be employed to address the above-mentioned problem by identifying which of the unlabeled images should be labeled [3].

2. METHODOLOGY

2.1. CRF model

The major challenge in the task of semantic segmentation is how to jointly consider image pixel properties and the context of a pixel at different scales. Of particular interest are methods based on discriminative models such as Conditional Random Fields (CRF) [4]. The CRF model are generally characterized by energy functions defined on the unary and the pairwise cliques as

$$E(\mathbf{x}) = \sum_{i\in\mathcal{V}} \Phi_i(x_i) + \sum_{i\in\mathcal{V}, j\in\mathcal{N}_i} \psi_i(x_i, x_j) \qquad (1)$$

Every possible configuration of the CRF defines a segmentation. The unary potential Φ_i of the CRF is defined as the negative log of the likelihood of a label being assigned to pixel i. It can be computed from the properties of the pixel or the patch. The pairwise term ψ_i of the CRF model generally takes the form of a contrast-sensitive Potts model for encouraging label consistency in adjacent pixels. Given the trained CRF model and its learned parameters, the final semantic segmentation is the most probable labeling that maximizes the conditional probability of the model. Firstly, an initially over-segmented subimage is obtained by meanshift algorithm to build a region adjacency graph (RAG). The over-segmentation provides the connectivity for the patches. Here we compute the over-segmentation with the Edge Detection and Image Segmentation (EDISON) System of Mean Shift [5] implementation on the HH, HV and VV channel amplitude data of POLSAR images. The parameters of the segmentation are chosen to mostly over-segment the subimages. In this way, descriptors extracted from coherent patches are less affected to noise deriving from co-presence of different categories, and object boundary detection accuracy is improved in comparison with methods based on rectangular partitioned patches. Then, the CRF model is defined over the RAG, and multiple diverse features, such as polarimtric signatures, texture, intensity, and image context are easily incorporated into the CRF-RAG model [6]. Surely, any other helpful features can also be integrated into our CRF-RAG model conveniently. Finally, the maximum margin learning method [7] with cutting plane algorithm is applied to efficiently train our CRF-RAG mode, and the approximate labeling via Graph-Cuts optimization method [8] is employed to infer the CRF-RAG model quickly.

2.2. Multiple feature descriptors

PolSAR is sensitive to the orientation and characters of target and thus yields many new polarimetric signatures which produce a more informative description of the scattering behavior of the imaging area. However, in earlier PolSAR image segmentation experiments, we found that using only polarimetric signature is not a very discriminative feature and fails to produce high accurate segmentations, especially for high resolution PolSAR images which include the abundant structural and textural information. To overcome the very high intraclass variability of PolSAR images, it is better to combine multiple diverse and complementary features based on different aspects, therefore we extract multiple polarimetric signatures and low-level image features to describe the over-segmented patches, that is, the unary potentials of our CRF-RAG model are constructed on combining the polarimetric signatures, texture and intensity features. The polarimetric signatures are obtained from our previous work [9], it is a concatenation of several normalized polarimetric statistics. For the texture descriptors, we use a multi-channel extention to the recently proposed multi-scale local pattern histogram (MLPH) which can capture both local and global structural information [10]. It is based on recognizing that there exist certain local patterns which are elementary properties of SAR image and their frequency histogram is proven to be a very powerful texture descriptor for SAR image. For the intensity features, we just use the mean value of the HH, VV and HV channel, and the Span histogram (5 bins) as done in [11].

2.3. Active learning for selecting good training samples

In the active learning framework, our aim is to obtain the best possible performance using the hand-labeled training samples as few as possible. Here we try to training on very few informative images to produce a better semantic segmenter as done in [3].

3. EXPERIMENTS AND DISCUSSIONS

The initial experiment is performed on RadarSat-2 fully polarimetric SAR images of Flevoland in Netherlands, with 12m$\times$8m resolution at fine quad-pol mode. The PolSAR scene to be labeled is of size 4000$\times$2400 pixels, which mainly contains four classes: woodland, cropland, water, building area. An additional “void” category in the label space accounts for observations that are not associated to any of the given semantic classes. We divide the PolSAR scene into 240 subimages, each subimage is 200$\times$200 pixels. We also implement two baseline classifiers for performance evaluation, one is the traditional maximum likelihood method based on Wishart distribution, the other baseline employs support vector machine (SVM) classifier with the radial basis function (RBF) kernel [12]. The primary experimental result is promising and demonstrates the effectiveness of our semantic segmentation framework, and also implies using a few well-selected labeled training images could produce an effective segmenter for POLSAR imagery.

4. REFERENCES

[1] S. Kumar and M. Hebert. “Discriminative random fields: A discriminative framework for contextual interaction in classification,” in Proc. Int. Conf. on Computer Vision, Oct. 2003, Vol. 2, pp.1150–1157.

[2] X. He, R.S. Zemel, and M.Á. Carreira-Perpinan. “ Multiscale Conditional Random Fields for Image Labeling,” in Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, 2004, Vol. 2, pp. 695-702.

[3] A. Farhangfar, R. Greiner, and C.Szepesvári, “Learning to segment from a few well-selected training images,” Proceedings of the 26th Annual International Conference on Machine Learning, 14-18 June,2009, Montreal, Quebec, Canada, pp.305-312

[4] T. Hofmann, “Unsupervised learning by probabilistic latent semantic anaylysis,” Machine Learning, 2001, Vol. 42, No. 1-2, pp.177-196.

[5] D. Comaniciu and P. Meer, “ Mean Shift: A Robust Approach toward Feature Space Analysis,” IEEE Trans. On Pattern Analysis and Machine Intelligence, May 2002, Vol. 24, No. 5, pp. 603-619

[6] D. D. Dai, W. Yang, and H. Sun, “Labeling SAR Images Using Conditional Random Field on Region Adjacency Graph ,” submitted to IET Radar, Sonar and Navigation.

[7] M. Szummer, P. Kohli, and D. Hoiem, “ Learning CRF using Graph Cuts,” in Proc. Tenth European Conf. Computer Vision, 2008.

[8] V. Kolmogorov and R. Zabih. “What energy functions can be minimized via graph cuts,” IEEE Trans. On Pattern Analysis and Machine Intelligence, Feb. 2004, Vol.26, No.2, pp.147-159

[9] T. Y. Zou, W. Yang, D. X. Dai, and H. Sun, “Polarimetric sar image classification using multi-features combination and extremely randomized clustering forests,” EURASIP Journal on Advances in Signal Processing, special issue on Advances in Multidimensional Synthetic Aperture Radar Signal Processing, to appear.

[10] D.D.Dai, W.Yang, and H.Sun, "Multi-scale Local Pattern Histogram for SAR Image Classification", submitted to IEEE Geoscience and Remote Sensing Letters.

[11] W. J .He, M. Jäger , A. Reigber, and O. Hellwich, “Building extraction from polarimetric sar data using mean shift and conditional random fields,” The 7th European Conference on Synthetic Aperture Radar (EuSAR), 2008.

[12] C.C. Chang and C.J. Lin, “LIBSVM : a library for support vector machines,” 2001.