Building Extraction from Polarimetric SAR Data using Mean Shift and Conditional Random Fields

Wenju He, Marc Jäger, Andreas Reigber, Olaf Hellwich
Computer Vision and Remote Sensing, Berlin University of Technology, Germany

Abstract

This paper presents a classification framework for extracting buildings from polarimetric SAR (PolSAR) data. Buildings in SAR data are generally composed of layover and shadow regions. First, mean shift bottom-up segmentation approach divides a SAR image into small homogeneous patches. Then conditional random fields (CRF) framework is applied to classify the patches into layover, shadow and other regions. The spatial connectivity between layover and shadow regions is exploited to improve the accuracy of CRF shadow detection. Promising segmentation results of buildings are presented and compared to the results of a basic logistic regression classifier.

1 Introduction

Since high resolution SAR data are available nowadays, building characterization and extraction become possible. PolSAR data have great potential in view of building analysis. Buildings are evident in SAR data because of their strong backscattering. Segmentation methods for buildings have promising applications in 3D characterization, change detection, reconstruction, and so on. However, inherent characteristics of SAR imaging in urban areas and speckle make the segmentation problem very difficult. The statistical characteristics of SAR images in urban areas are still in investigations. Therefore, robust segmentation methods are required to enhance the quality of building extraction.

Integration of unsupervised segmentation and supervised classification is an effective way to localize objects from images. Unsupervised bottom-up segmentation, e.g. mean shift [1], generates many small homogeneous patches from an image. Low level information such as intensity, color and texture is used to group similar pixels. Bottom-up segmentation is ill-posed because low-level features are not enough to obtain meaningful segmentation. Supervision and other priori knowledge are necessary to assemble the patches into objects and background.

This paper puts forward a classification framework for building segmentation from PolSAR data. The framework chooses the mean shift approach for bottom-up segmentation. The assembling of the patches into building and non-building regions is performed using the CRF model [6], which incorporates data-dependent interactions in image classification. The object in this paper is the radar response of buildings, which usually consists of layover, roof and shadow regions. Roof regions is treated here as layover regions. The connectivity of layover and shadow is incorporated into a hybrid method for shadow detection. Experiments on fully PolSAR data in urban areas are conducted.

Section 2 describes the classification framework. In section 3, the experiments on polarimetric SAR data in urban areas are presented. Section 4 concludes the article.

2 Classification Framework

In image segmentation layover and shadow regions are treated separately. When extracting layover or shadow regions, we consider all other regions as background. Therefore, extracting building regions from SAR data is a foreground/background segmentation problem. The shadow regions considered in this paper are only those generated by buildings.

The first step is to use mean shift method to generate small segments from SAR images. Mean Shift iteratively searches for representative modes of the density in a joint spatial-feature space. Delineation of the modes results in many segments. A segment contains a set of pixels with similar intensities. The segments are the basic elements in the following classification. We expect to classify and combine them by some rules into layover, shadow and other regions.

The benchmark classifier used in the evaluation is the logistic regression classifier, which works well for object detection [6]. CRF classification is adopted in order to obtain smoother and more accurate results.

For the shadow regions, we intend to make use of the fact that the layover and shadow regions are connected. If a layover region is present at a certain place, its neighbors are probably shadow regions. This knowledge could help to improve the CRF classification and eliminate some shadow regions which are not caused by buildings.

2.1 Feature Extraction

PATCH FEATURES
Polarimetry
P1. Polarimetric entropy, anisotropy and α: mean
P2. Sublook coherence, entropy: mean
P3. Optimized coherence: mean
Intensity
I1. HH, VV and HV: mean
I2. Span: histogram (5 bins)
Texture
T1. Filter bank: texton histogram (15 bins)
Shape
S1. Area: normalized by total image area
S2. Lines: (number of line pixels)/sqrt(area)
S3. Lines: percent of nearly parallel pairs of lines

Table 1: Features computed for a segment. 'mean' means averaging over pixels in the segment.

We extract polarimetric and low-level image features for the small patches obtained using the mean shift method. The set of 38 features is described in Table 1.

PolSAR data reveal scattering mechanisms and scatter characteristics through combination of different polarizations. Polarimetric entropy H, anisotropy A and α angle are extracted from the coherency matrix based on Pauli basis. Coherent scatters detected by sublook coherence or sublook entropy [4] are often situated on man-made structures. Polarimetric coherence optimization helps to detect more coherent scatters. These polarimetric statistics could be efficient features, since coherent scatters are generally in layover regions rather than shadow regions.

We also use intensity, texture and line features. Log intensity is an important feature, because layover regions have very high intensities and shadow have very low intensities. To represent texture, we apply a filter bank to log span image. K-means algorithm clusters the filter responses into representative textons [5]. For a segment generated by mean shift method, we calculate a texton histogram over pixels within the region. We also measure the number of pixels in each patch normalized by total image area. Layover and shadow regions usually have regular quadrangle shapes and edges. We extract straight lines and compute the occurrence of parallel lines in each region using log span image. Detailed explanations of exacting these low-level image features are presented in [2].

2.2 Conditional Random Fields

CRF is a discriminative framework for labeling and segmenting data. Given input data $\mathbf{y}$ and labels $\mathbf{x}$, the CRF models the conditional distribution $P(\mathbf{x}|\mathbf{y})$. The distribution over $\mathbf{y}$ is not modeled. Rich and overlapping features can be integrated in the CRF framework. $\mathbf{y}_i$ is the observation of ith patch, and its label $x_i \in \{-1, 1\}$, where -1 indicates background and 1 indicates object.

In a variant of CRF model in [6], the distribution of the labels $\mathbf{x}$ given the patches $\mathbf{y}$ is given by

$$p(\mathbf{x}|\mathbf{y}) = \frac{1}{Z}\exp(\sum_{i\in S} A_i(x_i, \mathbf{y}) + \sum_{i\in S}\sum_{j\in N_i} I_{ij}(x_i, x_j, \mathbf{y})) \tag{1}$$

where S is the set of all patches and N_i is the set of neighbors of $\mathbf{y}_i$; log-partition function Z is a global normalization factor. The association potential A_i describes the degree of membership of $\mathbf{y}_i$ being object. The interaction potential I_{ij} measures the cost of assigning labels to two neighboring patches.

This paper uses the Ising model potentials:

$$A_i(x_i, \mathbf{y}) = \exp(x_i \boldsymbol{\omega}^T \mathbf{f}_i(\mathbf{y})) \tag{2}$$

$$I_{ij}(x_i, x_j, \mathbf{y}) = \exp(x_i x_j \boldsymbol{\nu}^T \boldsymbol{\mu}_{ij}(\mathbf{y})) \tag{3}$$

where $\boldsymbol{\omega}$ and $\boldsymbol{\nu}$ are the model parameters. $\mathbf{f}_i(\mathbf{y})$ is the feature vector for patch $\mathbf{y}_i$. The pairwise function $\boldsymbol{\mu}_{ij}(\mathbf{y})$ is defined as the absolute difference between the feature vectors of $\mathbf{y}_i$ and $\mathbf{y}_i$. The parameter learning and label inference using Broyden-Fletcher-Goldfarb-Shanno gradient descent and sum-product loopy belief propagation are presented in [6].

2.3 Hybrid Detection

We apply the CRF to detect layover and shadow respectively. Most of the layover regions are associated with buildings, but shadow regions might be generated by non-manmade objects. A hybrid method is proposed here to improve the accuracy of detecting shadow areas of buildings. CRF detection results can be integrated with the fact that layover and shadow are often connected: the presence of layover regions indicates that its neighbors mights be shadow regions.

For an unseen SAR image, CRF classification produces a probability map $P(O_S = 1|C)$ that suggests how likely the regions are shadow. Here $O = 1$ indicates the presence of object, subscript S represents shadow, and C means that the probability is provided by CRF.

We apply logistic classifier to detect layover regions. The output of the classifier $P(O_L = 1)$ is the probability of the region being layover, where subscript L represents layover. We simply assign to each region the highest probability of its neighboring regions and itself. This spreads the probability from a layover region to its neighboring shadow regions. The propagation results in likelihood map $P(O_S = 1|L)$ that gives evidence of the shadow regions.

Thus we obtain two probability maps, i.e. $P(O_S = 1|C)$ and $P(O_S = 1|L)$. Both of them are evidences of shadow regions, and can be viewed as two independent experts. There are several approaches to integrate experts to produce more reliable results [3], one of them is logistic regression:

$$P(O_S = 1|C, L) = \sigma(\boldsymbol{\theta}^T[1, P(O_S = 1|C), P(O_S = 1|L)]) \tag{4}$$

where σ is logistic function. The weight $\boldsymbol{\theta}$ is estimated using maximum likelihood on a separate validation dataset.

3 Experiments

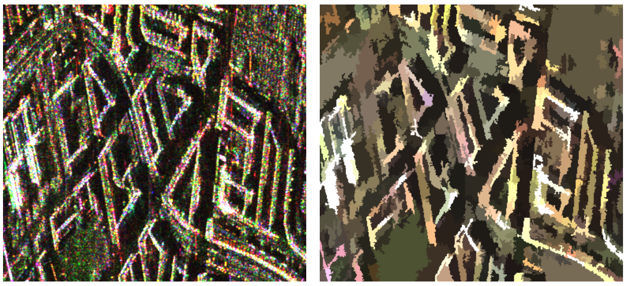

Figure 1: (Left) A test region of the PolSAR data; (Right) It's mean shift result.

Fully polarimetric SAR data of Copenhagen acquired by EMISAR are used in the experiments. The spatial resolutions in range and azimuth are 1.499 m and 0.748 m, respectively. The left image of Figure 1 is a sample image of the data. We extract a data set containing 98 images from the Copenhagen data. Each image is of size 384×352. We randomly select 78 images as training dataset, and the other 20 images for validation and accuracy evaluation.

3.1 Mean shift and ground truth

Single SAR intensity image can be modeled by Gamma distribution. For simplicity, we assume the span image of PolSAR data follows log-normal distribution. Log span image is the feature used in mean shift segmentation, since the Euclidean distance then makes sense for the smoothing in mean shift procedure. We use Gaussian kernel in the smoothing process. Building regions, which have higher variance than others, are smoothed to a relatively smaller extent. Thus patches composing layover and shadow regions are separated from other patches. The right image of Figure 1 shows the segmentation result of the left image.
For each image, layover and shadow masks are labeled manually indicating the locations and shapes of layover and shadow regions, respectively. Notice that there are errors in the masks, since sometimes it is difficult to identify building regions from obscured SAR images. For each patch generated by mean shift, the patch is identified as an object element if at least one-third of its pixels are labeled as objects in the ground truth. This labeling is then used for validation.

3.2 Layover detection

In this experiment we apply the logistic classifier and CRF to detect layover regions. Results of the sample image are shown in Figure 2. Comparisons of these methods are given in Table 2. Overall accuracy is the percentage of correctly labeled image patches. The ROC curves measures the relationship between detection rate and false positive rate under different thresholds, shown in the left side of Figure 3.

We also apply logistic classifier without mean shift preprocessing. In this case, the element in the classification is pixel rather than patch. Due to a large number of pixels in the dataset, we randomly select 10 images from the training dataset for training. Features P1$\sim$P3, I1, T1 and S2 in Table 1 are extracted for each pixel. Table 2 shows that logistic classifier with mean shift has a much higher detection rate than that without Mean Shift. One possible reason is that the spatial support provided by the small patches enables complex features to be extracted.

Using patch as an element in classification, CRF is compared to logistic classifier. CRF detects more layover regions and has a lower false positive rate.

The initial weight ν for interaction potential influences the strength of interaction factors. Higher initial weight tends to strengthen the influence of interaction potential, resulting in both higher detection rate and false positive rate.

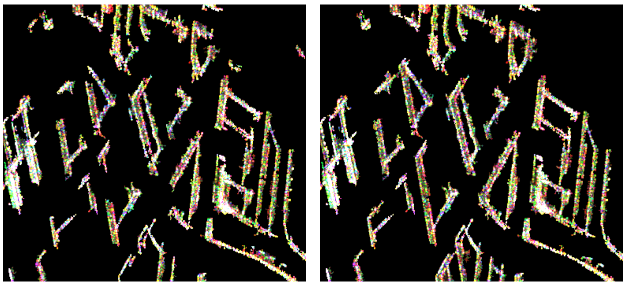

Figure 2: (Left) Logistic Classifier result; (Right) CRF result;

	Logistic^{-}	Logistic	CRF
Detection Rate (%)	44.85	76.45	79.69
False Positive Rate (%)	2.99	4.99	4.34
Overall Accuracy (%)	89.86	90.70	91.95

Table 2: Detection rates and false positive rates for layover detection. Superscript '-' indicates that no mean shift segmentation is used.

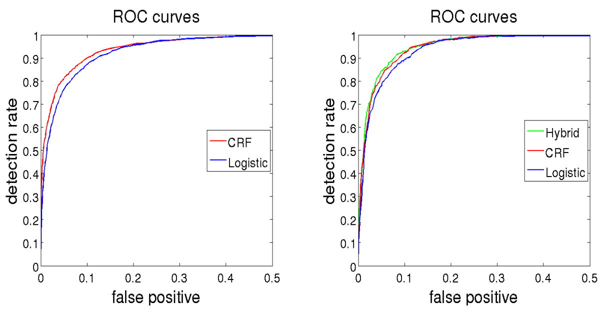

Figure 3: (Left) ROC curves for localizing layover; (Right) ROC curves for localizing shadow.

3.3 Shadow detection

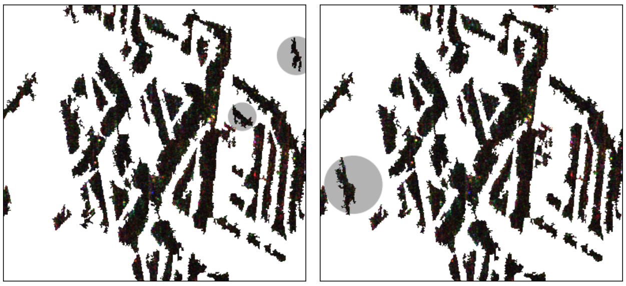

Figure 4: (Left) Shadows detected by CRF; (Right) Shadows detected by the proposed hybrid method.

	Logistic	CRF	Hybrid
Detection Rate (%)	82.56	86.36	84.81
False Positive Rate (%)	6.50	7.04	5.24
Overall Accuracy (%)	91.58	91.80	93.01

Table 3: Accuracies of Logistic, CRF and the proposed Hybrid method for shadow detection.

We split the test data set into two sets, each with 10 images. One set is used to training parameters $\boldsymbol{\theta}$, and the another set is used for accuracy evaluation. Logistic classifier, CRF and the proposed hybrid method are applied to detect shadow regions. Since the sensor is looking toward its right side, shadows are located on the right side of layovers. Therefore, in the hybrid method we only propagate the probability of each region in the map $P(O = 1|L)$ to its right neighboring regions. Figure 4 shows the classification results of the sample image. The ROC curves are displayed in the right side of Figure 3.

Table 3 shows that CRF has a higher false positive rate than logistic classifier. A probable reason can be that, CRF detects more shadow regions regardless of whether they are associated with buildings. Logistic classifier has the same problem although it outputs less false positives. But the hybrid method has a much lower positive rate. It demonstrates the improvement by incorporating connectivity information of layover and shadow regions. The hybrid method has the highest overall accuracy.

The results of CRF and the hybrid method in Figure 4 are almost the same except several minor differences. The inked shadow region at the top right detected by CRF is correctly eliminated by the hybrid method. The shadow region is probably associated with trees. The hybrid method detects several more shadow regions, e.g. the inked one at the mid-left of the left image of Figure 4. This shows the effectiveness of the hybrid method.

However, we can find that the inked shadow region in the middle of the left image disappears in the right image. The reason lies in the likelihood map $P(O = 1|L)$ and the fact that there is a roof region between the layover and shadow region. The roof region has a low probability to be detected as layover by the logistic classifier, and propagates the probability to the shadow region. The small probability prohibits the region from being detected by the hybrid method. Although the shadow region is detected by one expert (i.e. CRF) of the hybrid method, it is eventually discarded. So this is a weakness of the hybrid method. Frequently there are intermediate roof regions between layover and shadow regions. The hybrid method with one-step probability propagation does not work if the logistic classifier fails to identify the roof regions as layover. The probability from layover region to its corresponding shadow regions. Therefore, the hybrid method tends to detect less but purer building shadows than CRF does.

4 Conclusions

This paper proposes the use of conditional random fields for segmenting building areas from high resolution SAR data. mean shift method is applied in the preprocessing stage to generate small segments from SAR images. CRF classifies the segments into layover and shadow regions. The proposed hybrid method improves the accuracy of detecting shadow regions. The exaction of layover and shadow of buildings are efficient and promising for further building analysis and applications.

References

[1] Dorin Comaniciu, Peter Meer: *Mean Shift: A Robust Approach Toward Feature Space Analysis*, IEEE Transactions on Pattern Analysis and Machine Learning, Vol. 24, No. 5, pp. 603-619, 2002.

[2] Derek Hoiem, Alexei A. Erfos, Martial Hebert: *Recovering Surface Layout from an Image*, International Journal of Computer Vision, Vol. 75, No. 1, pp. 151-172, 2007.

[3] Kevin Murphy, Antonio Torralba, Daniel Eaton, William Freeman: *Object detection and localization using local and global features*, Towards Category-Level Object Recognition, LNCS, Vol. 4170, 2006.

[4] Rafael Z. Schneider, Konstantinos P. Papathanassiou, Irena Hajnsek, Alberto Moreira: *Polarimetric and Interferometric Characterization of Coherent Scatters in Urban Areas*, IEEE Transactions on Geoscience and Remote Sensing, Vol. 44, No. 4, pp. 971-984, 2006.

[5] Manik Varma, Andrew Zisserman: *A Statistical Approach to Texture Classification from Single Images*, International Journal of Computer Vision, Vol. 62, No. 1-2, pp. 61-81, 2005.

[6] S.V.N. Vishwanathan, Nicol N. Schraudolph, Mark Schmidt, Kevin Murphy: *Accelerated Training Conditional Random Fields with Stochastic Gradient Methods*, Proc. Intl. Conf. Machine Learning, pp. 969-976, ACM Press, New York, NY, USA, 2006.