PolSAR data segmentation by combining tensor space cluster analysis and Markovian framework

Yinghua Wang, Chongzhao Han, Florence Tupin

Abstract—We present a new segmentation algorithm for the full polarimetric synthetic aperture radar (PolSAR) data by coupling the cluster analysis in the tensor space and the Markov random field (MRF) framework. The PolSAR data are usually obtained as a set of $3 \times 3$ Hermitian positive definite polarimetric covariance matrices, which do not form a Euclidean space. By regarding each matrix as a tensor, we show that the PolSAR data space can be represented as a Riemannian manifold. Firstly, a mean shift clustering algorithm extended to the manifold is presented to cluster the tensors. Then under the MRF framework, the data energy term is defined from the degree of membership of all tensors in all the clusters while the regularization energy term is defined according to the cluster overlap rate. These parameters regarding the cluster analysis are computed under the Riemannian framework. The total energy is minimized using a graph cut based optimization to achieve the segmentation results. The effectiveness of the proposed method is demonstrated using real full PolSAR data.

Index Terms—Polarimetric synthetic aperture radar (PolSAR), image segmentation, Markov random field (MRF), cluster analysis, Riemannian manifold, tensor.

I. INTRODUCTION

The segmentation of full polarimetric synthetic aperture radar (PolSAR) data is still a challenging problem due to speckle. By introducing contextual information, the speckle effects on the segmentation results can be suppressed to some extent. Markov random fields (MRFs) [1] have proven to be powerful tools to impose spatial regularity constraint on the segmentation. The MRF based segmentation is usually achieved via minimizing an energy including two terms: the data energy, which measures the disagreement between the segmentation labels and the observed data, and the regularization energy, which introduces the spatial contextual information.

For the data energy term definition, the parametric models assume the observed data distribution is known and defined by some parameters, then the data energy is defined by the observation log-likelihood. The non-parametric models, such as the approach in [2] using the kernel methods to estimate the observation log-likelihood, do not impose any observation distributions, thus they may be more appropriate for describing the complicated and arbitrary observations. However, it may be problematic if the kernel methods such as the Parsen window techniques [3] are directly applied to PolSAR images, since they often require to compute the Euclidean distances among the observations. For the PolSAR images, the observed value at each pixel is usually the $3 \times 3$ Hermitian positive definite polarimetric covariance matrix. All these matrices form a non-Euclidean space. By representing the matrix space as a Riemannian manifold [4], some non-parametric methods can be extended to the PolSAR data space.

For the definition of the regularization energy term, it is known that the widely used Potts model [5] may deform the details in the image due to uniform smoothness. To resolve this problem, [6] proposes to use an adaptive smoothness term depending on the class overlap degree in the feature space. In [7], the overlap rate (OLR) between two components of a mixture is computed using the probability density function (p.d.f.) values of the mixture. Therefore, the class overlap degree can be evaluated by analyzing the empirical p.d.f. in the feature space.

The contribution of this paper is twofold: 1) We represent the PolSAR data space as a Riemannian manifold. 2) We propose a practical algorithm. Firstly, the clustering of PolSAR data over the manifold is achieved by extending the original mean shift clustering algorithm [8]. Then a new MRF based segmentation method employing the cluster analysis is presented. Under the Riemannian framework [4], the degree of membership of all the polarimetric covariance matrices in all the clusters are computed and the OLRs between every two clusters are also calculated. These parameters are used to define the data energy and the regularization energy terms in the MRF framework. The total energy is minimized using a graph cut based optimization to achieve the regularized segmentations.

II. POLSAR DATA SPACE

It is often encountered in image processing that the pixel value is a real positive definite symmetric matrix. In [4], such matrix is also termed as a tensor. The tensor space, which is not a Euclidean space, can be represented as a Riemannian manifold $\mathcal{M}$. Some properties of this Riemannian framework are briefly reviewed in the following.

For each point $X \in \mathcal{M}$, its tangent space $T_X \mathcal{M}$ is the plane tangent to the surface of the manifold at that point. The exponential map $exp_X : T_X \mathcal{M} \rightarrow \mathcal{M}$ maps each vector $y \in T_X \mathcal{M}$ to the point $Y \in \mathcal{M}$. The inverse of the exponential map at point $X$ is the logarithm map $log_X : \mathcal{M} \rightarrow T_X \mathcal{M}$, which maps each point $Y \in \mathcal{M}$ to the vector $y \in T_X \mathcal{M}$. Thus the two maps provide a one-to-one mapping between the tensor space and the tangent space around the point $X$. For a certain

---

Y.-H. Wang and C.-Z. Han are with the Institute of Integrated Automation, Xi’an Jiaotong University, Xi’an 710049, China e-mail: yinghua.wang@stu.xjtu.edu.cn, dodgerbluester@gmail.com

F. Tupin is with the Institute TELECOM, TELECOM ParisTech, LTCI CNRS, 46, rue Barrault 75013 Paris, France e-mail: florence.tupin@telecom-paristech.fr
reference point \( \mathbf{X} \), if the norm \( \| \cdot \|_{\mathbf{X}} \) in the tangent space is defined, then the length of the geodesic connecting the points \( \mathbf{X} \) and \( \mathbf{Y} \) on the manifold surface is given by \( \| \mathbf{y} \|_{\mathbf{X}} \) in the tangent space.

The Riemannian exponential map and logarithm map are defined as
\[
\exp_{\mathbf{X}}(\mathbf{y}) = \mathbf{X} \tilde{z} \exp \left( \mathbf{X}^{-\frac{1}{2}} \mathbf{y} \mathbf{X}^{-\frac{1}{2}} \right) \mathbf{X} \tilde{z},
\]
\[
\log_{\mathbf{X}}(\mathbf{Y}) = \mathbf{X} \tilde{z} \log \left( \mathbf{X}^{-\frac{1}{2}} \mathbf{Y} \mathbf{X}^{-\frac{1}{2}} \right) \mathbf{X} \tilde{z},
\]
where \( \exp(\cdot) \) and \( \log(\cdot) \) are the matrix exponential and nature logarithm respectively. The Riemannian distance of two tensors \( \mathbf{X} \) and \( \mathbf{Y} \) is given by
\[
d^2(\mathbf{X}, \mathbf{Y}) = \| \log_{\mathbf{X}}(\mathbf{Y}) \|^2 = \| \log \left( \mathbf{X}^{-\frac{1}{2}} \mathbf{Y} \mathbf{X}^{-\frac{1}{2}} \right) \|^2_2,
\]
where \( \| \cdot \|_2 \) is the Euclidean norm.

The mean \( \bar{\mu} \) of a set of tensors \( \{\mathbf{X}_i\}_{i=1}^n \) in \( \mathcal{M} \) is defined as the tensor that minimizes the sum of the squared distances:
\[
\bar{\mu} = \arg \min_{\mathbf{X} \in \mathcal{M}} \sum_{i=1}^n d^2(\mathbf{X}, \mathbf{X}_i).
\]
Thus it can be computed by the gradient descent algorithm as follows:
\[
\bar{\mu}_{t+1} = \exp_{\bar{\mu}_t} \left( \frac{1}{n} \sum_{i=1}^n \log_{\mu_i}(\mathbf{X}_i) \right).
\]

The linear interpolation \( \mathbf{Z} \) of two tensors \( \mathbf{X} \) and \( \mathbf{Y} \) is on the geodesic joining these two tensors. It can be computed by
\[
\mathbf{Z} = \exp_{\mathbf{X}}(t \cdot \log_{\mathbf{X}}(\mathbf{Y})), t \in [0, 1].
\]

For the PolSAR data, although the observed value at each pixel is not a real matrix, the Riemannian framework depicted above can be directly extended to the PolSAR data space. Henceforth, the polarimetric covariance matrices are also called tensors.

### III. MEAN SHIFT CLUSTERING OF POLSAR DATA

The mean shift clustering algorithm is a robust density-based clustering algorithm. It seeks for the local maxima of the empirical p.d.f. (so-called density modes) in the feature space. We review some main results described in [8], to which the empirical p.d.f. (so-called density modes) in the feature space.

Henceforth, the polarimetric covariance matrices are also called tensors.

#### IV. THE MRF BASED SEGMENTATION OF POLSAR DATA EMPLOYING THE CLUSTER ANALYSIS

Given the PolSAR data observations \( \mathbf{T} = \{\mathbf{T}_s, s \in S\} \), the aim of segmentation is to estimate the labels \( \mathbf{L} = \{L_s, s \in S\} \). The MRF based segmentation is usually achieved via minimizing the following energy:
\[
E(\mathbf{L}) = \sum_{s \in S} D_s(L_s) + \sum_{s,t \in \mathcal{N}} V_{s,t}(L_s, L_t),
\]
where the data energy \( D_s(L_s) \) measures the disagreement between \( L_s \) and \( \mathbf{T}_s \); \( \mathcal{N} \) is the neighborhood definition and the regularization energy \( V_{s,t}(L_s, L_t) \) is the penalty for assigning the label pair \( (L_s, L_t) \) to the neighboring pixels \( s \) and \( t \). These two energy terms will be newly defined based on the cluster analysis in Sections IV-A and IV-B, respectively. The energy minimization will be introduced in Section IV-C.

#### A. Data energy

It is known that the fuzzy memberships in the Fuzzy c-means algorithm [10] describe the agreement between each observed value and each class. Based on this point, we define our data energy term by the memberships of all the tensors in all the clusters. Being different from the original membership definitions, our membership values are computed via the mean shift clustering results under the Riemannian framework.
With the mean shift clustering output \( L_s \) and \( Z_s \), \( s \in S \), the cluster centers \( \{ C_p \}_{p=1}^{N_c} \) are defined as:
\[
C_p = E_M[Z_s | L_s = p],
\]
where \( E_M[\cdot] \) is the mean of tensors computed by (4). The membership of \( T_s \) in the \( p \)th cluster is denoted by \( u_{s,p} \), which is defined based on the Riemannian metric in (3) as:
\[
u_{s,p} = \frac{1}{d^2(T_s,C_p)} \sum_{k=1}^{N_c} 1/d^2(T_s,C_k).
\]
The data energy \( D_s(L_s) \) is then defined as:
\[
D_s(L_s) = -u_{s,L_s}.
\]
Equation (13) does not impose any assumption about the observation distributions. Moreover, the number of classes is not required to be set beforehand. This number is derived from the parameters \( h \) and \( M \) in the mean shift algorithm, thus can be set more intuitively.

B. Regularization energy

The OLRs between every two classes are introduced to define an adaptive smoothness term depending on the class labels. The original OLR definition in [7] is extended to our tensor space.

With the cluster centers \( \{ C_p \}_{p=1}^{N_c} \), the OLR between the \( p \)th and \( q \)th clusters, denoted by \( OLR(p,q) \), is defined as:
\[
OLR(p,q) = \frac{P(saddle)}{\min[P(C_p), P(C_q)]},
\]
where \( P(saddle) \) is the p.d.f. value of the saddle point on the geodesic joining the two cluster centers \( C_p \) and \( C_q \); \( P(C_p) \) and \( P(C_q) \) are the p.d.f. values of the cluster centers \( C_p \) and \( C_q \), respectively. Each point \( X_s \) on the geodesic joining \( C_p \) and \( C_q \) can be computed by (5). The point with the lowest p.d.f. value is considered as the saddle point, which implies
\[
P(saddle) = \min P(X_s).
\]

\( P(X_s) \), \( P(C_p) \) and \( P(C_q) \) are all estimated by the kernel methods using the Riemannian metric in (3). For instance, \( P(X_s) \) is computed by
\[
P(X_s) = \frac{c_k d}{nV_d} \sum_{t \in S} k \left( \frac{d^2(X_s,T_t)}{h^2} \right),
\]
where \( k(\cdot) \) is the kernel profile, \( c_k d/(nV_d) \) is the normalization parameter. Based on the OLRs between every two clusters, \( V_{s,t}(L_s, L_t) \) is defined as:
\[
V_{s,t}(L_s, L_t) = \begin{cases} 0 & L_s = L_t \\ \lambda \max_{(p,q) \in \Omega} \frac{OLR(L_s, L_t)}{OLR(p,q)} & L_s \neq L_t \end{cases},
\]
which are the same as the Potts model if for all the \( (p,q) \in \{1,2,\ldots,N_c\}^2 \), we assign \( OLR(p,q) = 1 \).

Equation (17) provides adaptive smoothness depending on the class labels. For instance, if two clusters are well separated in the feature space, the OLR between them tends to be 0, thus no regularization is imposed. Oppositely, if two clusters are totally overlapping in the feature space, the OLR between them will become 1, thus strong regularization is needed since the two clusters compete for the observations and the label values can be selected according to the spatial relations.

C. Energy minimization

The recently proposed graph cut optimization methods [11], including the \( \alpha \)-expansion algorithm and the \( \alpha \)-\( \beta \)-swap algorithm, are fast energy minimization algorithms that converge to the approximate global minimum with guaranteed optimality bounds. Since \( V_{s,t}(L_s, L_t) \) in (17) is a semi-metric, the \( \alpha \)-\( \beta \)-swap algorithm is exploited to minimize the total energy.

V. EVALUATION OF SEGMENTATIONS

The segmentation results are evaluated by the way proposed in [12]. Given a segment \( Q \) and a ground truth region \( G \), \( Q \)'s degree of spatial support with respect to \( G \) is evaluated by the normalized overlap score \( OS(Q,G) \in [0,1] \), computed as
\[
OS(Q,G) = \frac{|Q \cap G|}{|Q \cup G|}.
\]
For each \( G \), the Best Spatial Score (BSS) is the maximum normalized overlap score of a group of segments with respect to \( G \). It measures how well the best segment covers \( G \).

For an entire data set, the segmentation performance can thus be evaluated by computing the mean BSS value across all the ground truth regions in the data set. The segmentation with higher mean BSS value implies better spatial support. Moreover, the number of segments is another index to evaluate the segmentations. A small number of segments will decrease the computation burden for further processing. To sum up, a high quality segmentation should have high BSS values and low number of segments.

VI. EXPERIMENTAL RESULTS

The German Aerospace Center (DLR) E-SAR L-band full PolSAR data \(^1\) are used for experiments. The original images have 1540 \times 2816 pixels. The span image is shown in Fig. 1.

Four MRF based segmentation algorithms are tested and compared. The first one is the traditional MRF based segmentation method, which uses the Wishart distribution [13]

\(^1\)These data are downloaded from http://earth.esa.int/polsarpro/datasets.html.
log-likelihood to define the data energy and the Potts model to define the regularization energy. The distribution parameters are estimated using the \(H/\alpha\) Wishart classifier \([14]\) or the \(H/A/\alpha\) Wishart classifier \([15]\) results. We denote this approach by “MRF-Wish-Potts”. For the left three approaches, the data energy terms are all defined by the proposed one using memberships. The regularization energy terms are respectively defined by 0, the Potts model, and the proposed one using OLRs. We denote these three approaches respectively by “MRF-Mem-0”, “MRF-Mem-Potts”, and “MRF-Mem-OLR”.

The speckle filtering is implemented beforehand using Lee’s filter \([16]\). The uniform kernel is used in both the mean shift clustering and the kernel methods to estimate the p.d.f. values of tensors. By varying the mean shift parameters \(h\) and \(M\), we find \(h = 1 \) and \(M = 40\) yield satisfactory results for this E-SAR data set. For the regularization parameter \(\lambda\) contained in the Potts model or (17), different values are tested. In each segmentation, the label values of all the classes are first sorted in an ascending way according to their average span values. Then each class is rendered using the color code shown in Fig. 2.

The test site is selected as the image in the red rectangle in Fig. 1 with 512 \(\times\) 512 pixels. The optical image is shown in Fig. 3 (a), which illustrates that this test site has plenty of classes, including buildings, different kinds of grass ground, smooth ground, forests, etc.. The segmentation results are shown in Fig. 3 (b)~(f). Fig. 3 (b) employs the \(H/A/\alpha\) Wishart classifier result to estimate the Wishart distribution parameters. Hence the image is segmented into \(N_c = 16\) classes. In Fig. 3 (c)~(f), the image is segmented into \(N_c = 11\) classes. Firstly, the result in Fig. 3 (c) is compared with Fig. 3 (b). Obviously Fig. 3 (c) has better visual quality. The MRF-Mem-0 with regularization energy assigned to be 0 even performs better than the traditional MRF-Wish-Potts, which demonstrates the effectiveness of the proposed data energy. Then, we compare the result in Fig. 3 (e) with that in Fig. 3 (d). With the same \(\lambda\) value, it is noted that the details are better preserved in Fig. 3 (e), such as the building inside the red rectangle and the fine structures contained in the red ellipses. When \(\lambda\) is set to 0.25 in MRF-Mem-OLR, we get the result in Fig. 3 (f), which provides comparative smoothness like that in Fig. 3 (d) while the details are still well preserved as in Fig. 3 (e).

Furthermore, to evaluate the segmentation algorithms quantitatively, a simulated data set is constructed in the way similar to that in \([17]\). A ground truth image with six class labels is first designed as shown in Fig. 4 (a). Then we manually select six homogeneous regions from the real E-SAR data set as shown in the ellipses in Fig. 1 and use them to fill in the six classes in the simulated PolSAR data. The span image of the derived simulated data set is shown in Fig. 4 (b). The segmentation results are shown in Fig. 4 (c)~(h). Fig. 4 (c) uses the \(H/\alpha\) Wishart classifier result to estimate the Wishart distribution parameters. Therefore the image is segmented into \(N_c = 8\) classes. In Fig. 4 (d)~(h), the image is rightly segmented into \(N_c = 6\) classes based on the mean shift clustering. The same \(h\) and \(M\) values as before are used here since this simulated data set is constructed using the real E-SAR data. Table 2 displays the BSS values of the six results in Fig. 4 (c)~(h). Firstly, we compare the mean BSS value of the traditional MRF-Wish-Potts with that of others. It is noticed that it has the lowest mean BSS value, which verifies that the proposed data energy defined by memberships outperforms the one defined by the Wishart distribution log-likelihood. Then, MRF-Mem-Potts is compared with MRF-Mem-OLR. When \(\lambda = 0.1\), their results are shown in Fig. 4 (e) and (g). Table 2 displays that although the mean BSS value of MRF-Mem-Potts is higher than that of MRF-Mem-OLR
due to stronger smoothness constraint for some classes, the Class 5 BSS value of MRF-Mem-Potts is lower. According to our color code, Class 5 is the class in the red color, to which the thin red line belongs. Hence we conclude that the fine structures are better preserved in Fig. 4 (g) than (e). When $λ = 0.45$, the corresponding results are shown in Fig. 4 (f) and (h). Table 2 shows that five classes of MRF-Mem-OLR result have higher BSS values than that of MRF-Mem-Potts. This implies MRF-Mem-OLR performs better, which can also be observed from Fig. 4 (f) and (h). To sum up, for the Potts model, smoother results may imply less details, yet by using certain $λ$ values, MRF-Mem-OLR can provide results with comparative smoothness while preserving most of the details.

![Fig. 4. The segmentation results for the simulated data set. (a) The ground truth label image. (b) The span image of the simulated data set. (c) MRF-Wish-Potts result with $λ = 0.1$. (d) MRF-Mem-0 result. (e) MRF-Mem-Potts result with $λ = 0.1$. (f) MRF-Mem-Potts result with $λ = 0.45$. (g) MRF-Mem-OLR result with $λ = 0.1$. (h) MRF-Mem-OLR result with $λ = 0.45$.](image)

### Table 2 The BSS values for the segmentation results of the simulated data set.

<table>
<thead>
<tr>
<th>Class</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Mean BSS value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRF-Wish-Potts, $λ = 0.1$</td>
<td>0.9028</td>
<td>0.8642</td>
<td>0.9006</td>
<td>0.8222</td>
<td><strong>0.5406</strong></td>
<td>0.5527</td>
<td>0.7638</td>
</tr>
<tr>
<td>MRF-Mem-0</td>
<td>0.9251</td>
<td>0.8792</td>
<td>0.8982</td>
<td>0.8203</td>
<td><strong>0.9061</strong></td>
<td>0.9691</td>
<td>0.8997</td>
</tr>
<tr>
<td>MRF-Mem-Potts, $λ = 0.1$</td>
<td>0.9523</td>
<td>0.9379</td>
<td>0.9390</td>
<td>0.9157</td>
<td><strong>0.9262</strong></td>
<td>0.9788</td>
<td>0.9416</td>
</tr>
<tr>
<td>MRF-Mem-Potts, $λ = 0.45$</td>
<td>0.9493</td>
<td>0.9125</td>
<td>0.9147</td>
<td>0.8858</td>
<td><strong>0.9148</strong></td>
<td>0.9671</td>
<td>0.9240</td>
</tr>
<tr>
<td>MRF-Mem-OLR, $λ = 0.1$</td>
<td>0.9211</td>
<td>0.9328</td>
<td>0.9323</td>
<td>0.8787</td>
<td><strong>0.9290</strong></td>
<td>0.9622</td>
<td>0.9260</td>
</tr>
<tr>
<td>MRF-Mem-OLR, $λ = 0.45$</td>
<td>0.8183</td>
<td>0.9128</td>
<td>0.9300</td>
<td>0.8914</td>
<td><strong>0.9362</strong></td>
<td>0.9734</td>
<td>0.9103</td>
</tr>
</tbody>
</table>

**VII. CONCLUSION AND PERSPECTIVE**

A new MRF based segmentation method for the PolSAR data has been presented in this paper. Its effectiveness was verified by experimenting with the real full PolSAR data. By regarding the polarimetric covariance matrices as feature vectors, the cluster analysis is carried out in the tensor space. Under the Riemannian framework for tensor computing, the feature space is analyzed without the prior knowledge about the observation distributions or the cluster numbers. The data energy term in the MRF framework defined based on the cluster membership performs better than that defined by the observation log-likelihood using the complex Wishart distribution. The regularization energy term coupling the OLRs of clusters into the Potts model outperforms the original Potts model, for preserving the details in the image after the regularization. The future work includes improving the proposed method by using adaptive bandwidth for the mean shift algorithm and applying it to the terrain classification.

**ACKNOWLEDGMENT**

The authors would like to thank the ESA (http://earth.esa.int/polsarpro/datasets.html) for providing the data. This work was supported in part by a grant from the State Key Program for Basic Research of China (973) (No. 2007CB311006).

**REFERENCES**