PolSAR Image Segmentation by Mean Shift Clustering in the Tensor Space

WANG Ying-Hua[1] HAN Chong-Zhao[1]

Abstract We present an unsupervised segmentation algorithm for fully polarimetric synthetic aperture radar (PolSAR) data by using the mean shift clustering. The previous work using the span values of the PolSAR data as the features in the mean shift clustering, however, does not sufficiently exploit the full information contained in the polarimetric covariance matrix. When considering the polarimetric covariance matrices as the feature vectors, the traditional mean shift clustering in the Euclidean space is not applicable anymore, since these matrices do not form a Euclidean space. We first show that by regarding each Hermitian positive definite polarimetric covariance matrix at per pixel as a tensor, the tensor space can be represented as a Riemannian manifold. Then, the mean shift clustering is extended to the Riemannian manifold to explain the theoretical meanings of the tensor clustering and a practical segmentation algorithm based on the metric lying on the manifold is proposed. Experimental results using the real fully PolSAR data and simulated data verify the effectiveness of the proposed method.

Key words Polarimetric synthetic aperture radar (PolSAR), image segmentation, mean shift clustering, Riemannian manifold, tensor

DOI 10.3724/SP.J.1004.2010.00798

The fully polarimetric synthetic aperture radar (PolSAR) involves transmission and reception of both the horizontally and vertically polarized radar pulses, thus provides data containing the complete polarimetric scattering information. Therefore, these data has drawn more attention in recent years. Among various applications of PolSAR images, the unsupervised segmentation is an important step toward the automatic understanding of the data. The unsupervised segmentation schemes for PolSAR images can be categorized into two groups as suggested in [1]: the first one is based on the analysis of the polarimetric information (physical scattering mechanisms); the other applies the conventional image segmentation techniques to the PolSAR data.

For the first group of approaches, the polarimetric target decomposition theorems[2] provide powerful tools for extracting polarimetric parameters from the data. These parameters are used as features to segment the images. Take for instance the well-known H/α[3] and H/A/α[4] decomposition of PolSAR data. The segmentation is carried out by dividing the H/α plane into eight zones or the H/A/α space into sixteen zones. However, such division of the feature space is somewhat arbitrary.

In the second group, the clustering algorithms are most widely explored for the unsupervised segmentation of PolSAR images. The clustering of data aims at finding a natural grouping of clusters in the feature space. In the following, the various clustering algorithms are simply divided into the parametric ones and the non-parametric ones.

The parametric clustering usually relies upon a priori knowledge about the number of clusters or the distribution of the feature vectors. One example is the H/α Wishart classifier[5]. It uses the H/α decomposition results to get an initial segmentation into eight clusters, then the K-mean clustering is implemented by considering the polarimetric covariance matrices as the feature vectors, which are assumed to follow the complex Wishart distribution[6]. However, when the feature space becomes more complex, a predefined number of clusters based on the polarimetric parameters may be less meaningful. Moreover, for some cases, the Wishart distribution may not be suitable to characterize the polarimetric covariance matrices[7].

The nonparametric clustering does not impose any embedded assumptions, so it may be more appropriate for analyzing the complicated and arbitrarily structured feature space. The mean shift clustering proves to be a robust density-based clustering algorithm and has been applied to the color image filtering and segmentation[8]. In [9], the mean shift clustering is employed to segment the PolSAR data; nevertheless, only the span value is used as the feature for each pixel, and the full polarimetric information carried by the polarimetric covariance matrix or coherency matrix is not explored sufficiently. References [10−11] show that the use of full polarimetric information provides better classification results. Therefore, it is expected that using the polarimetric covariance matrices or coherency matrices as the feature vectors will improve the results, but these Hermitian positive definite matrices do not form a Euclidean space. Thanks to the work[12−13], where the mean shift clustering was extended to the analytic manifolds. Moreover, in [14] each real symmetric positive definite matrix is also called a tensor, and the tensor space is represented as a Riemannian manifold.

In the light of these new techniques, we present a new segmentation method for the PolSAR data by using the mean shift clustering, in which each 3×3 Hermitian positive definite polarimetric covariance matrix or coherency matrix at per pixel is used as the feature vector. The rest of the paper is organized as follows. The mean shift clustering and segmentation will be briefly reviewed in Section 1. In Section 2, we show that the PolSAR

Received March 12, 2009; in revised form August 24, 2009
Supported by National Basic Research Development Program of China (973 Program)(2007CB311006) and National Natural Science Foundation of China (60602026)
1. Institute of Integrated Automation, Xi′an Jiaotong University, Xi′an 710049, P. R. China

data space can be represented as a Riemannian manifold. In Section 3, the mean shift clustering algorithm is extended to the Riemannian manifold, and the new segmentation method for PolSAR data is presented. The evaluation of segmentations is described in Section 4. Experimental results and conclusions are provided in Sections 5 and 6, respectively.

1 Mean shift clustering and segmentation

The mean shift clustering algorithm is based on the kernel density estimation methods such as the Parzen window technique[15]. By estimating the density gradient, the mean shift vector is derived, which always points toward the direction of maximum increase in density. The local maxima of the density, i.e., the density modes in the feature space can thus be located by computing the mean shift vector in an iterative way.

Given n data points $\boldsymbol{x}_i$, $i = 1, 2, \cdots, n$, in the d-dimensional Euclidean space $\mathbf{R}^d$, the kernel density estimator at point $\boldsymbol{x}$ with kernel $K(\boldsymbol{x})$ and bandwidth h is given by

$$\hat{f}_{h,K}(\boldsymbol{x}) = \frac{1}{nh^d}\sum_{i=1}^{n} K\left(\frac{\boldsymbol{x}-\boldsymbol{x}_i}{h}\right) \tag{1}$$

Usually, the radially symmetric kernels are used which satisfy

$$K(\boldsymbol{x}) = c_{k,d}k(\|\boldsymbol{x}\|^2) \tag{2}$$

where $k(x)$ is called the profile of the kernel and $c_{k,d}$ is the normalization constant. Thus, the density estimator in (1) can be rewritten as

$$\hat{f}_{h,K}(\boldsymbol{x}) = \frac{c_{k,d}}{nh^d}\sum_{i=1}^{n} k\left(\left\|\frac{\boldsymbol{x}-\boldsymbol{x}_i}{h}\right\|^2\right) \tag{3}$$

According to [8], the gradient of $\hat{f}_{h,K}(\boldsymbol{x})$ can be expressed as

$$\bigtriangledown\hat{f}_{h,K}(\boldsymbol{x}) = \hat{f}_{h,G}(\boldsymbol{x})\frac{2c_{k,d}}{h^2c_{g,d}}m_{h,G}(\boldsymbol{x}) \tag{4}$$

where G is another kernel and its profile $g(x) = -k'(x)$; $m_{h,G}(\boldsymbol{x})$ is the mean shift vector and can be expressed as

$$m_{h,G}(\boldsymbol{x}) = \frac{\sum_{i=1}^{n}\boldsymbol{x}_i g(\|\frac{\boldsymbol{x}-\boldsymbol{x}_i}{h}\|^2)}{\sum_{i=1}^{n} g(\|\frac{\boldsymbol{x}-\boldsymbol{x}_i}{h}\|^2)} - \boldsymbol{x} \tag{5}$$

For each data point $\boldsymbol{x}$, its convergence point is computed by the following mean shift procedures:

1) Initialize $\boldsymbol{y}_0$ with $\boldsymbol{y}_0 = \boldsymbol{x}$.

2) Update $\boldsymbol{y}_j$ by $\boldsymbol{y}_{j+1}$ until the difference between $\boldsymbol{y}_{j+1}$ and $\boldsymbol{y}_j$ is small enough, where

$$\boldsymbol{y}_{j+1} = \boldsymbol{y}_j + m_{h,G}(\boldsymbol{y}_j) = \frac{\sum_{i=1}^{n}\boldsymbol{x}_i g(\|\frac{\boldsymbol{y}_j-\boldsymbol{x}_i}{h}\|^2)}{\sum_{i=1}^{n} g(\|\frac{\boldsymbol{y}_j-\boldsymbol{x}_i}{h}\|^2)} \tag{6}$$

After we get the convergence points of all the data points based on the above procedure, the local maxima are detected as the density modes. The data points visited by all the mean shift procedures converging to the same mode form a cluster of arbitrary shape. Thus, the mean shift clustering is completed. Based on mean shift clustering, the mean shift segmentation algorithm is proposed in [8]. Usually, the joint spatial-range domain is considered, which implies adding the pixel coordinates into the feature vector $\boldsymbol{x}_i$. More details can be found in [8].

2 PolSAR data space

We first review some main results for representing the space of real symmetric positive definite matrices as a Riemannian manifold. Then, we present that the PolSAR data space can also be represented as a Riemannian manifold.

2.1 The space of real symmetric positive definite matrices

It is sometimes encountered in image processing that the pixel value is a real symmetric positive definite matrix. For instance, the pixel value used for pedestrain detection in [16] is the covariance matrix. Another example is that each pixel value in the image obtained by the diffusion tensor image technique[17] is a real symmetric positive definite matrix. In [18], the 2×2 real symmetric positive definite matrices are used to exemplify that such matrices form a non-Euclidean space. Given a 2×2 real symmetric positive definite matrix $A = [a \quad b; b \quad c]$, where $ac - b^2 > 0$ and $a > 0$, if we consider the matrix A as a point $(a, b, c) \in \mathbf{R}^3$, the whole set of such matrices form a cone. The distance of two such matrices should be the length of the geodesic that connects the two points along the cone surface, but not the length of the straight line connecting the two points in $\mathbf{R}^3$.

When each pixel in an image is composed of a real symmetric positive definite matrix, such matrix space is not a Euclidean space. In [14], each symmetric positive definite matrix is also termed as a tensor, and the tensor space is represented as a Riemannian manifold $\mathcal{M}$. Some properties of the Riemannian manifold described in [14] 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: $T_X\mathcal{M} \to \mathcal{M}$ maps each tangent 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: $\mathcal{M} \to T_X\mathcal{M}$, which maps each point $Y \in \mathcal{M}$ to the tangent 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 point X.

For each point X, the Riemannian metric assigns to its tangent space $T_X\mathcal{M}$ an inner product as

$$\langle y, z\rangle_X = \mathrm{tr}\left(X^{-\frac{1}{2}}yX^{-1}zX^{-\frac{1}{2}}\right) \tag{7}$$

which varies smoothly from point to point. With this inner product, the norm of the tangent vector y in the tangent space $T_X\mathcal{M}$ can be computed by

$$\|y\|_X^2 = \langle y, y\rangle_X \tag{8}$$

The associated Riemannian exponential map and loga-

rithm map are defined as

$$exp_X(y) = X^{\frac{1}{2}} \exp\left(X^{-\frac{1}{2}} y X^{-\frac{1}{2}}\right) X^{\frac{1}{2}} \tag{9}$$

$$log_X(Y) = X^{\frac{1}{2}} \log\left(X^{-\frac{1}{2}} Y X^{-\frac{1}{2}}\right) X^{\frac{1}{2}} \tag{10}$$

where $\exp(\cdot)$ and $\log(\cdot)$ are the matrix exponential and natural logarithm, respectively.

The Riemannian distance of two points $\mathrm{dist}(X, Y)$ is defined as the length of the geodesic connecting the points X and Y on the manifold surface. This distance is given by $\|y\|_X$ in the tangent space, where $y = log_X(Y)$. In [14], the Riemannian affine invariant distance of two tensors X and Y can be computed by

$$\begin{aligned}\mathrm{dist}^2(X, Y) = \|y\|_X^2 = \|log_X(Y)\|_X^2 = \\ \left\|X^{-\frac{1}{2}} log_X(Y) X^{-\frac{1}{2}}\right\|_I^2 = \left\|\log\left(X^{-\frac{1}{2}} Y X^{-\frac{1}{2}}\right)\right\|_2^2\end{aligned} \tag{11}$$

where I is the identity matrix and $\|\cdot\|_2$ is the Euclidean norm. Note that $\|\cdot\|_X$ and $\|\cdot\|_I$ are respectively the norm defined in $T_X\mathcal{M}$ and $T_I\mathcal{M}$, and they can be computed by (7) and (8). Equation (11) indicates that from a certain reference point X, the distance of another tensor Y can be computed by the Euclidean norm of the vector $\log\left(X^{-\frac{1}{2}} Y X^{-\frac{1}{2}}\right)$ in the tangent space.

Under this Riemannian framework, the mean value $\bar{\mu}$ of a set of tensors $X_1, X_2, \cdots, X_n \in \mathcal{M}$ is defined as the tensor that minimizes the sum of the squared distances: $\bar{\mu} = \arg\min_{X \in \mathcal{M}} \sum_{i=1}^{n} \mathrm{dist}^2(X, X_i)$. Thus, it can be computed by the gradient descent algorithm as follows:

$$\begin{aligned}\bar{\mu}_{t+1} = exp_{\bar{\mu}_t}\left(\frac{1}{n}\sum_{i=1}^{n} log_{\bar{\mu}_t}(X_i)\right) = \\ \bar{\mu}_t^{\frac{1}{2}} \exp\left(\frac{1}{n}\sum_{i=1}^{n} \log\left(\bar{\mu}_t^{-\frac{1}{2}} X_i \bar{\mu}_t^{-\frac{1}{2}}\right)\right) \bar{\mu}_t^{\frac{1}{2}}\end{aligned} \tag{12}$$

Similarly, the weighted mean of a set of tensors is defined as the tensor that minimizes the sum of the weighted squared distances, and can be computed by the following iterative algorithm:

$$\begin{aligned}\bar{\mu}_{t+1} = exp_{\bar{\mu}_t}\left(\sum_{i=1}^{n} \omega_i log_{\bar{\mu}_t}(X_i)\right) = \\ \bar{\mu}_t^{\frac{1}{2}} \exp\left(\sum_{i=1}^{n} \omega_i \log\left(\bar{\mu}_t^{-\frac{1}{2}} X_i \bar{\mu}_t^{-\frac{1}{2}}\right)\right) \bar{\mu}_t^{\frac{1}{2}}\end{aligned} \tag{13}$$

where ω_i, $i = 1, 2, \cdots, n$, are the weighting parameters.

2.2 PolSAR data space represented as a Riemannian manifold

The PolSAR data is usually obtained as a set of 3×3 Hermitian positive definite polarimetric covariance matrices or coherency matrices. It has already been mentioned in Subsection 2.1 that the real symmetric positive definite matrices do not form a Euclidean space. Similarly, the Hermitian positive definite matrices also do not from a Euclidean space. In the following, we show the Riemannian framework for the Hermitian positive definite matrices.

The Hermitian positive definite matrix space can also be represented as a Riemannian manifold as proposed in [19]. The Riemannian metric we choose has the same form as proposed in [14], but in this work, all the corresponding computations are extended to the complex matrices. The associated exponential map, logarithm map, and Riemannian distance of two tensors also have the same expressions as shown in Subsection 2.1, however the corresponding computations will be carried out on complex matrices now, but not the real matrices. These computations are based on the matrix exponential, matrix logarithm, and matrix power, which are well defined for the real symmetric matrices in [14]. According to [20], these computations can also be well defined for the Hermitian matrices. Thus, the Riemannian framework for the Hermitian positive definite matrices is formed.

Henceforth, each pixel value in the PolSAR data, i.e., the polarimetric covariance matrix or the coherency matrix, is also termed as a tensor, and the tensor space is represented as a Riemannian manifold.

3 Mean shift clustering in the tensor space

Since the PolSAR data space can be represented as a Riemannian manifold, in Subsection 3.1, the original mean shift clustering algorithm is extended to the Riemannian manifold. The derived iterative algorithm has clear theoretical meanings, however, it involves extensive computational effort, which hinders its practical usefulness. To obtain similar results in practice while decreasing the computation burden, a new segmentation method is put forward in Subsection 3.2.

3.1 Extending the mean shift clustering to the Riemannian manifold

In [13], the original mean shift clustering is extended to the analytic manifolds. In [21], the mean shift clustering is extended to the space of symmetric positive definite matrices, and the extended algorithm is applied to the filtering. In the same way, the mean shift clustering is extended to the Hermitian positive definite matrix space in this section.

According to [13], the kernel density estimator in (3) can be extended to the manifolds by using the manifold distance as

$$\hat{f}_{h,K}(X) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^{n} k\left(\frac{\mathrm{dist}^2(X, X_i)}{h^2}\right) \tag{14}$$

The gradient of (14) is expressed as

$$\begin{aligned}\bigtriangledown \hat{f}_{h,K}(X) = \\ \hat{f}_{h,G}(X) \frac{c_{k,d}}{h^2 c_{g,d}} \left(-\frac{\sum_{i=1}^{n} \bigtriangledown \mathrm{dist}^2(X, X_i) g\left(\frac{\mathrm{dist}^2(X, X_i)}{h^2}\right)}{\sum_{i=1}^{n} g\left(\frac{\mathrm{dist}^2(X, X_i)}{h^2}\right)}\right)\end{aligned} \tag{15}$$

where G is another kernel and its profile $g(x) = -k'(x)$.

Based on the Riemannian distance in (11), for a fixed X_i,

$$\bigtriangledown \mathrm{dist}^2(X, X_i) = \bigtriangledown \|x_i\|_X^2 = -2x_i \tag{16}$$

as suggested in [22]. Equation (16) shows that $\bigtriangledown \mathrm{dist}^2(X, X_i)$ is a tangent vector. Introducing (16) into (15) yields

$$\bigtriangledown \hat{f}_{h,K}(X) = \hat{f}_{h,G}(X) \frac{2c_{k,d}}{h^2 c_{g,d}} \left(\frac{\sum\limits_{i=1}^{n} x_i g\left(\frac{\mathrm{dist}^2(X, X_i)}{h^2} \right)}{\sum\limits_{i=1}^{n} g\left(\frac{\mathrm{dist}^2(X, X_i)}{h^2} \right)} \right) \tag{17}$$

Thus, the mean shift vector for tensors over the Riemannian manifold is

$$m_{h,G}(X) = \frac{\sum\limits_{i=1}^{n} x_i g\left(\frac{\mathrm{dist}^2(X, X_i)}{h^2} \right)}{\sum\limits_{i=1}^{n} g\left(\frac{\mathrm{dist}^2(X, X_i)}{h^2} \right)} \tag{18}$$

It is known that $g\left(\frac{\mathrm{dist}^2(X, X_i)}{h^2} \right)$ is a scalar. Thus, $m_{h,G}(X)$ is a weighted sum of the tangent vectors, which implies $m_{h,G}(X)$ is also a tangent vector itself in $T_X\mathcal{M}$. Given the j-th kernel center Y_j, the next kernel center Y_{j+1} on the manifold surface is obtained by moving Y_j along the mean shift vector direction. Therefore, the updating relies on the Riemannian exponential map as:

$$\begin{aligned} &Y_{j+1} = \\ &exp_{Y_j}(m_{h,G}(Y_j)) = \\ &Y_j^{\frac{1}{2}} \exp\left(\frac{\sum\limits_{i=1}^{n} g\left(\frac{\mathrm{dist}^2(Y_j, X_i)}{h^2} \right) \log\left(Y_j^{-\frac{1}{2}} X_i Y_j^{-\frac{1}{2}} \right)}{\sum\limits_{i=1}^{n} g\left(\frac{\mathrm{dist}^2(Y_j, X_i)}{h^2} \right)} \right) Y_j^{\frac{1}{2}} \end{aligned} \tag{19}$$

In this way, the mean shift clustering for tensors can be interpreted as the alternate computation in the tangent space and the tensor space. For each point $X \in \mathcal{M}$, the mean shift vector in $T_X\mathcal{M}$ is computed first. Then, X is moved along the geodesic on the manifold surface corresponding to that tangent vector. By initializing $Y_0 = X$, the convergence point of (19) provides the mode that X belongs to.

3.2 The mean shift segmentation of PolSAR data

In [12−13], the mean shift clustering over the analytic manifolds is applied to the 3-D motion segmentation problems. Similarly, it is expected that the mean shift clustering algorithm for tensors derived in Subsection 3.1 can be directly applied to the PolSAR image segmentation. Nevertheless, it is observed that the mean shift iterative computation in (19) requires calculating the matrix powers, logarithm, and exponential for each data point in each iteration. This involves extensive computational effort even for a small image.

Now, we focus on a new segmentation method that decreases the computation burden but yields similar results in practice. In [23], the framework using the Log-Euclidean metric and the Riemannian framework mentioned in Subsection 2.1 are compared experimentally, and the results show that the Log-Euclidean metric can yield similar results in practice but with much less computations. The Log-Euclidean distance is defined as

$$\mathrm{dist}^2(X, X_i) = \|\log(X) - \log(X_i)\|_2^2 \tag{20}$$

which is equivalent to map all the tensors to the tangent space of the I matrix, i.e., the logarithm map log_I: $\mathcal{M} \to T_I\mathcal{M}$ is used. By introducing the Log-Euclidean distance and the logarithm map log_I into (19), a much simplified algorithm can be obtained.

The computational burden of each mean shift iteration by using (19) or using the simplified algorithm based on the Log-Euclidean metric are compared in the following. We have n data points $\{X_i\}_{i=1}^n$. For a certain data point, the j-th kernel center is known to be Y_j. For both methods, the computation of Y_{j+1} contains four steps, which are depicted in Table 1. From Table 1, we can note that Step 3 is the same for both methods. We only need to compare the computational effort of Steps 1, 2, and 4. All the computations involved in Steps 1, 2, and 4 for both methods can be regarded as the combinations of some matrix operations, including the matrix multiplication, matrix addition, matrix power, matrix logarithm, matrix exponential, and matrix Euclidean norm. For $P \times P$ matrices, the computation complexity of matrix multiplication, matrix addition, and matrix Euclidean norm are $\mathrm{O}(P^3)$, $\mathrm{O}(P^2)$ and $\mathrm{O}(P^2)$, respectively. By using the method proposed in [14] to compute the matrix power, matrix exponential, and matrix logarithm, their respective computation complexity is all approximately $\mathrm{O}(2P^3)$. Knowing the computation

Table 1 The procedures for updating the kernel center by using (19) and the Log-Euclidean metric

Using (19)	Using Log-Euclidean metric
Step 1. For each X_i, $V_i = Y_j^{\frac{1}{2}} \log(Y_j^{-\frac{1}{2}} X_i Y_j^{-\frac{1}{2}}) Y_j^{\frac{1}{2}}$	Step 1. For each X_i, $V_i = \log(X_i)$
Step 2. For each X_i, $\mathrm{dist}^2(Y_j, X_i) = \left\| \log(Y_j^{-\frac{1}{2}} X_i Y_j^{-\frac{1}{2}}) \right\|_2^2$	Step 2. For each X_i, $\mathrm{dist}^2(Y_j, X_i) = \|\log(Y_j) - V_i\|_2^2$
Step 3. $m(Y_j) = \dfrac{\sum\limits_{i=1}^{n} V_i g\left(\frac{\mathrm{dist}^2(Y_j, X_i)}{h^2} \right)}{\sum\limits_{i=1}^{n} g\left(\frac{\mathrm{dist}^2(Y_j, X_i)}{h^2} \right)}$	Step 3. $m(Y_j) = \dfrac{\sum_{i=1}^{n} V_i g\left(\frac{\mathrm{dist}^2(Y_j, X_i)}{h^2} \right)}{\sum_{i=1}^{n} g\left(\frac{\mathrm{dist}^2(Y_j, X_i)}{h^2} \right)}$
Step 4. $Y_{j+1} = Y_j^{\frac{1}{2}} \exp\left(Y_j^{-\frac{1}{2}} m(Y_j) Y_j^{-\frac{1}{2}} \right) Y_j^{\frac{1}{2}}$	Step 4. $Y_{j+1} = \exp(m(Y_j))$

complexity for each matrix operation, we evaluate the computational effort by the total number of each matrix operation. The computation complexity for each matrix operation and the total number of matrix operations used for Steps 1, 2 and 4 of the two methods are listed in Table 2.

Table 2 The computational effort evaluated by the number of matrix operations for Steps 1, 2, and 4 of the two methods shown in Table 1

	Using (19)	Using Log-Euclidean metric
Matrix multiplication ($\mathrm{O}(P^3)$)	$6n+4$	0
Matrix addition ($\mathrm{O}(P^2)$)	0	n
Matrix power ($\approx \mathrm{O}(2P^3)$)	$6n+4$	0
Matrix logarithm ($\approx \mathrm{O}(2P^3)$)	$2n$	$n+1$
Matrix exponential ($\approx \mathrm{O}(2P^3)$)	1	1
Matrix Euclidean norm ($\mathrm{O}(P^2)$)	n	n

Table 2 verifies that using the Log-Euclidean metric decreases the computation burden. Therefore, we propose a new segmentation method for PolSAR data using the mean shift clustering based on this metric. By observing the four steps using Log-Euclidean metric shown in Table 1, we can note that the $\exp(\cdot)$ operation in Step 4 can counteract the $\log(\cdot)$ operation in Step 2. Thus, by using the Log-Euclidean metric, the tensor space is in fact converted into a vector space as proposed in [23]. In this vector space, the original mean shift clustering can be used. The proposed new algorithm is depicted in the following. For each pixel $s \in S$, where S is the pixel set, the input is the 3×3 coherency matrix T_s, and the output is the label value L_s.

Algorithm 1. The new mean shift based segmentation algorithm for PolSAR data.

Input: T_s, $s \in S$.

Step 1. For each pixel s,

1) Map T_s into the tangent space using $V_s = log_I(T_s)$;

2) Change the 3×3 matrix V_s into a 6×1 vector $\boldsymbol{v}_s$ using the minimal representation[14];

3) if the joint domain is considered, add the pixel coordinates into the vector $\boldsymbol{v}_s$;

4) Compute the convergence point $\boldsymbol{y}_s$ corresponding to $\boldsymbol{v}_s$ using (6).

Step 2. Obtain the clustering results according to [8]

1) if only the range domain is considered, each cluster is formed by the pixels whose $\boldsymbol{y}_s$ values are closer than h_r in the range domain;

2) if the joint domain is considered, each cluster is formed by the pixels whose $\boldsymbol{y}_s$ values are closer than h_r in the range domain and h_s in the spatial domain.

Step 3. Eliminate the clusters containing less than M pixels, which results in N_c clusters $\{C_p\}_{p=1}^{N_c}$.

Step 4. For each pixel s, assign $L_s = p$ if $\boldsymbol{y}_s \in C_p$.

Output: L_s, $s \in S$.

[1]These data are downloaded from http://earth.esa.int/polsarpro/datasets.

4 Evaluation of segmentations

Similar to [9], the segmentation algorithms are evaluated in the way proposed in [24]. 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

$$OS(Q,G) = \frac{|Q \bigcap G|}{|Q \bigcup G|} \tag{21}$$

The best spatial score (BSS) is the maximum normalized overlap score of a group of segments with respect to a ground truth region G. It measures how well the best segment covers the region G.

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

5 Experimental results

The fully PolSAR data[1] obtained from German Aerospace Centre (DLR) E-SAR L-band are used for experiments. The original images have $1\,540 \times 2\,816$ pixels. The span image is shown in Fig. 1. Two sites from the image are tested first. Then, a simulated image constructed from this data set is tested.

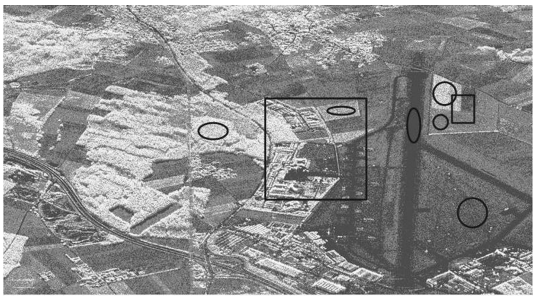


Fig. 1 The span image of the original data set (The test sites used are chosen as the areas in the two boxes. The data in the six ellipses are used to construct the six class simulated images.)

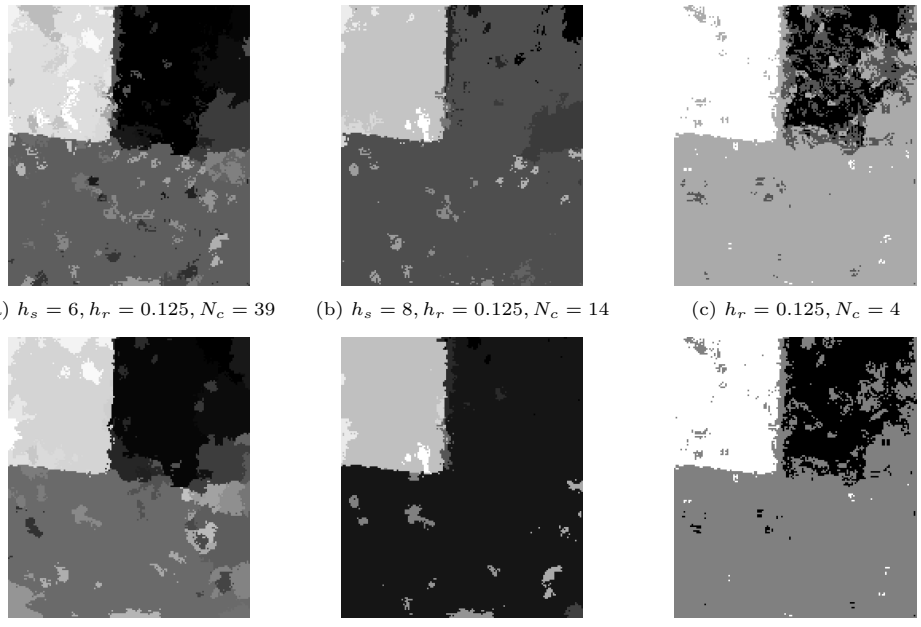
Before all our segmentation experiments, the speckle filtering is implemented using Lee′s filter[25]. Several segmentation algorithms are tested and compared in our experiments, including the H/α Wishart classifier, the H/A/α Wishart classifier[4], the mean shift clustering using the span image (denoted by "MSS") proposed in [9], and our proposed method using the mean shift clustering of tensors (denoted by "MST"). For the last two methods, the uniform kernel is used in all the experiments. For the parameter M contained in the algorithm shown in Algorithm 1, the values

20, 40, 60, and 80 are tested, and it is found that $M = 40$ yields good results for this data set. And for the spatial domain bandwidth h_s and the range domain bandwidth h_r, different values are tested in the experiments, which brings on different N_c in the final results. If the h_s value is not given for one result, it means that this result only considers the range domain. The label values of the segmentation results are first sorted according to the average span values of all the clusters, then the results are rendered using the color code shown in Fig. 2. For example, the cluster with label value 1 has the lowest average span value and is in black color, while the cluster with label value N_c has the highest average span value and is in white color. This color code is used in all our segmentation results.

Fig. 2 The color code for the segmentation results

The first test site is the small image with 176×152 pixels in the small rectangle in Fig. 1. We select this test site as in [26] since it is obviously composed of three classes. Thus, we could check if it is rightly segmented into three clusters as expected. Moreover, it will help us to select the appropriate h_s and h_r values, which will then be fixed when segmenting the other test sites. Because the H/α and H/A/α Wishart classifiers will segment the image into $N_c = 8$ or $N_c = 16$ classes, we only compare the segmentation results using MSS and MST. For the MSS results, $h_r = 0.125, 0.25, 0.375, 0.5$ are tested, while for the MST results, h_r is varied with values 0.5, 1, 1.5, and 2. For both MSS and MST, if the spatial domain is also considered, then $h_s = 4, 6, 8, 16, 64$ are tested. We select some results for exposition, which are comparatively good. The segmentation results using MSS are shown in Fig. 3 and the segmentation results using MST are shown in Fig. 4. From Fig. 3, we can observe that with $h_r = 0.25$, the image is rightly segmented into 3 classes as shown in Fig. 3 (f). Meanwhile, from Fig. 4, when $h_r = 1$, the image is also rightly segmented into 3 classes as shown in Fig. 4 (c) or Fig. 4 (e). Let us focus on Fig. 3 (f) and Fig. 4 (c). Fig. 4 (c) using MST is smoother than Fig. 3 (f) using MSS. For example, many pixels in the top-left homogenous region in Fig. 3 (f) are misclassified into the gray class, but in Fig. 4 (c), almost all the pixels in this homogenous region are rightly classified into the same class (the white class). This phenomenon can be interpreted as follows. It is observed from the span image in Fig. 1 that, even in a homogeneous region, the speckle effects cause the span values of the pixels vary greatly. Consider two pixels belonging to the same cluster, and the difference between their span values is big. In the MSS algorithm, the feature values used are their span values. Therefore, these two pixels can be easily segmented into different classes. However, in the MST algorithm, the feature values are computed from their polarimetric covariance matrices or coherency matrices. These matrices contain the full polarimetric information, which includes not only the span values but also other information. Thus, the MST feature values provide more complete and accurate descriptions for the pixels. And it is possible that these two pixels will be rightly classified into the same cluster by MST. The experimental results in Figs. 3 and 4 also provide information for choosing the appropriate parameter values for the mean shift algorithm. For the following test sites, only the range domain is considered in the mean shift clustering. For MSS, h_r is set to 0.25 and, for MST, h_r is set to 1.

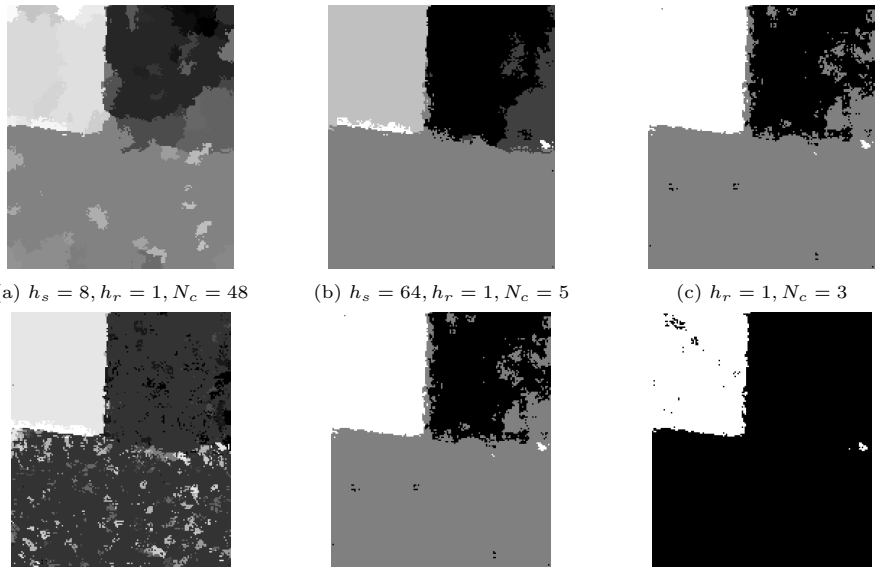
The second test site is selected as the image enclosed by the big rectangle in Fig. 1 with 512×512 pixels. This test site is chosen because its scene is quite complex, and it contains plenty of classes. Moreover, some salient objects shown in this test site as the buildings and the landing field



(a) $h_s = 6, h_r = 0.125, N_c = 39$ (b) $h_s = 8, h_r = 0.125, N_c = 14$ (c) $h_r = 0.125, N_c = 4$

(d) $h_s = 4, h_r = 0.25, N_c = 42$ (e) $h_s = 6, h_r = 0.25, N_c = 13$ (f) $h_r = 0.25, N_c = 3$

Fig. 3 The segmentation results using MSS for the test site enclosed by the small rectangle in Fig. 1



(a) $h_s = 8, h_r = 1, N_c = 48$ (b) $h_s = 64, h_r = 1, N_c = 5$ (c) $h_r = 1, N_c = 3$

(d) $h_r = 0.5, N_c = 11$ (e) $h_r = 1, N_c = 3$ (f) $h_r = 2, N_c = 2$

Fig. 4 The segmentation results using MST for the test site enclosed by the small rectangle in Fig. 1

can help us find the corresponding optical image, which is shown in Fig. 5 (a). The optical image illustrates that this test site has enough different classes, including buildings, different kinds of grassy ground, smooth ground, forests and so on. The results of the H/A/α Wishart classifier, MSS, and MST are compared in the following. Fig. 5 (b) is the segmentation result using the H/A/α Wishart classifier; Fig. 5 (c) is the segmentation result using MSS; and Fig. 5 (d) is the segmentation result using MST. First, the MST result in Fig. 5 (d) is compared with the H/A/α Wishart classifier result in Fig. 5 (b). It is observed that many homogeneous regions in the original image are segmented into less clusters in Fig. 5 (d) than that in Fig. 5 (b). For example, the grassy ground in the top-right part is rightly delineated as a single cluster in Fig. 5 (d) while several clusters in Fig. 5 (b). And the forest area in the top-left part is almost rightly delineated as one cluster (however two clusters with different gray colors are detected) in Fig. 5 (d), but a lot of meaningless small regions of different classes in Fig. 5 (b). Then, we focus on the MST result in Fig. 5 (d) and the MSS result in Fig 5 (c). We compare both of them with the optical image. By comparing with the optical image in Fig. 5 (a), some regions belonging to different classes are segmented into the same cluster in Fig. 5 (c). For instance, the grassy ground in the top-right part and the forest area in the top-left part are segmented into the same cluster in Fig. 5 (c). But these two areas are well segmented into two clusters in Fig. 5 (d). Also, the buildings are segmented into the same cluster as the grassy ground in the gray color in Fig. 5 (c). But in Fig. 5 (d), the buildings and the grassy ground are rightly delineated, i.e., they are segmented into clusters with different colors. When two distinct regions have close span values, they cannot be discriminated by the MSS since only the span value is used as the feature. However, by using the MST, two distinct classes with close span values can be well differentiated.

Finally, to evaluate the segmentation algorithms quantitatively, a simulated data set is constructed in the way similar to that in [27]. A ground-truth six class label image is first designed as shown in Fig. 6 (a). Then, some homogeneous regions are manually selected from the real E-SAR data set as shown in the ellipses in Fig. 1. The data contained in the six ellipses are finally used to fill in the six classes in the simulated PolSAR images. The span image of the derived simulated data set is shown in Fig. 6 (b). Since the image contains six classes, the results of the H/α Wishart classifier, MSS, and MST are compared in the following. For the MSS and MST, the same h_r values as before are used since this synthetic image is constructed from the same E-SAR data set. The H/α Wishart classification result is shown in Fig. 6 (c). Fig. 6 (d) is the segmentation result using MSS. The result using MST is shown in Fig. 6 (e). It is clearly observed that the MST result in Fig. 6 (e) has better visual quality than that in Fig. 6 (c) or Fig. 6 (d). Table 3 shows the number of clusters N_c and the mean BSS values for the three results in Fig. 6 (c) $\sim$ (e). The MST result in Fig. 6 (e) has the highest mean BSS value, which verifies the effectiveness of the proposed method and its performance improvement compared with the other two existing approaches.

Table 3 The mean BSS values for the segmentation results of the simulated data set

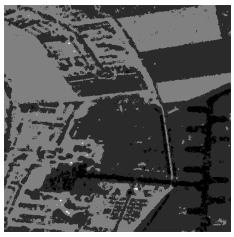
	N_c	Mean BSS value
H/α Wishart	8	0.7717
MSS	5	0.6849
MST	6	0.9277

6 Conclusion and perspective

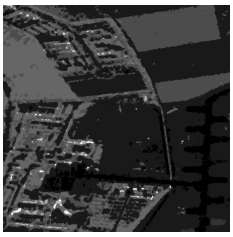
The mean shift clustering is employed to fulfill the segmentation of PolSAR data in this paper, and the polarimetric covariance matrix or coherency matrix is used as the feature for each pixel, but not the span value as in some existing approaches. To fulfill this purpose, we first

(a) The optical image of the test site from Google Earth ©2008 Google^{TM}, Image ©2008 Geocontent, ©2008 Tele Atlas[28]

(b) H/A/α Wishart classifier result with $N_c = 16$



(c) MSS result with $h_r = 0.25$ and $N_c = 7$

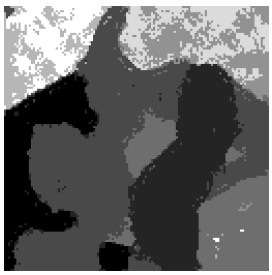


(d) MST result with $h_r = 1$ and $N_c = 11$

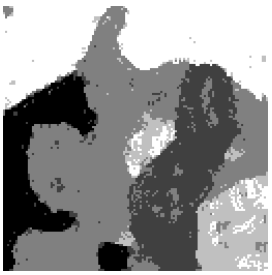
Fig. 5 The segmentation results for the test site enclosed by the big rectangle in Fig. 1

(a) The ground-truth label image

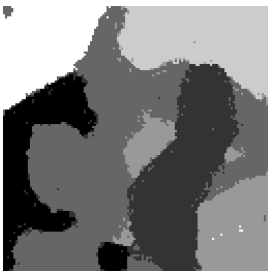
(b) The span image of the simulated data set



(c) H/α Wishart classifier result with $N_c = 8$



(d) MSS result with $h_r = 0.25$ and $N_c = 5$



(e) MST result with $h_r = 1$ and $N_c = 6$

Fig. 6 The experimental results for the simulated data set

show that the PolSAR data space can be represented as a Riemannian manifold. Then, the mean shift clustering is extended to the Riemannian manifold to explain the theoretical meanings of the tensor clustering. Finally, a new segmentation algorithm based on the Log-Euclidean metric is proposed. The efficiency of the proposed method is substantiated by the experiments on the real fully PolSAR data and synthetic images. Performance improvement is observed compared with several existing approaches.

The future work includes the following two aspects. 1) The mean shift clustering results depend greatly on the band width, which is determined empirically in this paper. The proposed method may be improved by introducing an adaptive band width based on analyzing the local structure of the PolSAR data or evaluating the clustering quality under different bandwidth values. 2) In the results of the proposed segmentation method, each class does not have any specific physical meanings. The proposed approach can be further improved so that each class in the results can be associated to some specific terrain types, thus to fulfill the terrain classification of PolSAR data.

Acknowledgement

The authors would like to thank ESA (http://earth.esa.int/polsarpro/datasets.html) for providing the data.

References

1 Maitre H. *Traitement des Images de Radar à Synthèse d'Ouverture*. Paris: Hermès Science Publication, 2001

2 Cloude S R, Pottier E. A review of target decomposition theorems in radar polarimetry. *IEEE Transactions on Geoscience and Remote Sensing*, 1996, **34**(2): 498−518

3 Cloude S R, Pottier E. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 1997, **35**(1): 68−78

4 Pottier E, Lee J S. Application of the H/A/α polarimetric decomposition theorem for unsupervised classification of

fully polarimetric SAR data based on the Wishart distribution. In: Proceedings of the Committee on Earth Observing Satellites SAR Workshop. Toulouse, France: ESA, 1999. 26−29

5 Lee J S, Grunes M R, Ainsworth T L, Du L J, Schuler D L, Cloude S R. Unsupervised classification using polarimetric decomposition and the complex Wishart classifier. *IEEE Transactions on Geoscience and Remote Sensing*, 1999, **37**(5): 2249−2258

6 Lee J S, Hoppel K W, Mango S A, Miller A R. Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 1994, **32**(5): 1017−1028

7 Bombrun L, Beaulieu J M. Fisher distribution for texture modeling of polarimetric SAR data. *IEEE Geoscience and Remote Sensing Letters*, 2008, **5**(3): 512−516

8 Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 603−619

9 He W, Jager M, Hellwich O. Comparison of three unsupervised segmentation algorithms for SAR data in urban areas. In: Proceedings of the International Geoscience and Remote Sensing Symposium. Boston, USA: IEEE, 2008. 241−244

10 Kong J A, Swartz A A, Yueh H A, Novak L M, Shin R T. Identification of terrain cover using the optimal polarimetric classifier. *Journal of Electromagnetic Waves and Applications*, 1988, **2**(2): 171−194

11 Lim H H, Swartz A A, Yueh H A, Kong J A, Shin R T, van Zyl J J. Classification of earth terrain using polarimetric synthetic aperture radar images. *Journal of Geophysical Research*, 1989, **94**(6): 7049−7057

12 Tuzel O, Subbarao R, Meer P. Simultaneous multiple 3D motion estimation via mode finding on Lie groups. In: Proceedings of the 10th International Conference on Computer Vision. Beijing, China: IEEE, 2005. 18−25

13 Subbarao R, Meer P. Nonlinear mean shift for clustering over analytic manifolds. In: Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2006. 1168−1175

14 Pennec X, Fillard P, Ayache N. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 2006, **66**(1): 41−66

15 Duda R O, Hart P E. *Pattern Classification and Scene Analysis*. New York: Wiley, 1973

16 Tuzel O, Porikli F, Meer P. Pedestrian detection via classification on Riemannian manifolds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(10): 1713−1727

17 Bihan D L, Mangin J, Poupon C, Clark C, Pappata S, Molko N. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 2001, **13**(4): 534−546

18 Fletcher P T, Joshi S. Principal geodesic analysis on symmetric spaces: statistics of diffusion tensors. In: Proceedings of the Conference on Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis. Prague, Czech Republic: Springer, 2004. 87−98

19 Bhatia R. *Positive Definite Matrices*. Princeton: Princeton University Press, 2007. 201−210

20 Gallier J H. *Geometric Methods and Applications: for Computer Science and Engineering*. New York: Springer-Verlag, 2001. 385−386

21 Subbarao R, Meer P. Discontinuity preserving filtering over analytic manifolds. In: Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition. Minneapolis, USA: IEEE, 2007. 1−6

22 Pennec X. Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 2006, **25**(1): 127−154

23 Arsigny V, Fillard P, Pennec X, Ayache N. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 2006, **56**(2): 411−421

24 Malisiewicz T, Efros A. Improving spatial support for objects via multiple segmentations. In: Proceedings of the British Machine Vision Conference. Warwick, UK: BMVA, 2007. 1−10

25 Lee J S, Grunes M R, de Grandi G. Polarimetric SAR speckle filtering and its implication for classification. *IEEE Transactions on Geoscience and Remote Sensing*, 1999, **37**(5): 2363−2373

26 Cao F, Hong W, Wu Y, Pottier E. An unsupervised segmentation with an adaptive number of clusters using the SPAN/H/α/A space and the complex Wishart clustering for fully polarimetric SAR data analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, **45**(11): 3454−3467

27 Yang Y, Han C Z, Han D Q. A Markov random field model-based fusion approach to segmentation of SAR and optical images. In: Proceedings of the International Geoscience and Remote Sensing Symposium. Boston, USA: IEEE, 2008. 802−805

28 Google Inc. Google earth [Online], available: http://earth.google.com/, October 11, 2008

WANG Ying-Hua Ph. D. candidate at the Institute of Integrated Automation, School of Electronics and Information Engineering, Xi′an Jiaotong University. Her research interest covers polarimetric synthetic aperture radar (PolSAR) image segmentation, classification and building detection. Corresponding author of this paper.
E-mail: yinghua.wang@stu.xjtu.edu.cn

HAN Chong-Zhao Professor, Ph. D. in the Department of Automation, School of Electronics and Information Engineering, Xi′an Jiaotong University. His research interest covers multi-source information fusion, stochastic control and adaptive control, and nonlinear spectral analysis.
E-mail: czhan@mail.xjtu.edu.cn