Acta Automatica Sinica
Vol. 36, No. 6 June, 2010

Available online at www.sciencedirect.com
ScienceDirect

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

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
DOI: 10.1016/S1874-1029(09)60037-9

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

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

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

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

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: $T_X\mathcal{M} \rightarrow \mathcal{M}$ 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: $\mathcal{M} \rightarrow T_X\mathcal{M}$, 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-