DISSIMILARITY MEASUREMENTS FOR PROCESSING AND ANALYZING POLSAR DATA: A SURVEY

Wen Yang[1], *Hui Song*[1], *Gui-Song Xia*[2] *and Carlos López-Martínez*[3]

[1] *School of Electronic Information, Wuhan University, 430072 Wuhan, China*
[2] *State Key Lab. LIESMARS, Wuhan University, 430079 Wuhan, China*
[3] *Universitat Politècnica de Catalunya (UPC), 08034 Barcelona, Spain*

ABSTRACT

Measuring the pairwise similarity/dissimilarity of data is of central importance for processing and analyzing PolSAR images. In the literature, a large variety of measurements have been used, however, it is still not clear how to choose appropriate similarity measures for a given task. This paper presents a brief summary and discussion of the dissimilarity measurements used for interpreting PolSAR Images.

Index Terms— Polarimetric SAR, dissimilarity measure, distance

1. INTRODUCTION

Polarimetric Synthetic Aperture Radar (PolSAR) systems provide new and more powerful ways to characterize the scattering properties of Earth's surface. The currently available satellite fully polarimetric SAR sensors are able to continually provide high quality data to support accurate Earth observations and topographic measurements. This in turn requires a number of efficient techniques for understanding these data, such as speckle reduction, segmentation, classification, object recognition and change detection [1, 2, 3, 4]. Among these tasks, measuring the similarity/dissimilarity is of central importance and lots of work have been devoted to it in the past decades. However, it is still not clear yet how to choose appropriate similarity measures for a given task. In this survey, we attempt to summarize and analyze the existing PolSAR dissimilarity measures, as well as to provide some hints for the direction of future efforts.

As the most common way of representing polarimetric information in a PolSAR data set, in this survey we mainly concern on dissimilarities defined for covariance matrix data.

2. PRELIMINARIES

2.1. Polarimetric SAR Data

For PolSAR measures of targets, the polarimetric information can be represented by a complex vector $\mathbf{k} = \{k_1, k_2, \cdots, k_p\}^T$, where ^{T} represents vector transposition. It can be depicted by a p-dimensional circular complex Gaussian distribution,

$$p(\mathbf{k}|\mathbf{\Sigma}) = \frac{1}{\pi^d|\Sigma|} \exp\left(-\mathbf{k}^{*T}\mathbf{\Sigma}^{-1}\mathbf{k}\right), \tag{1}$$

where $\Sigma = E[\mathbf{k}\mathbf{k}^{*T}]$ is a $p \times p$ complex covariance matrix, and E is the expectation operator.

For multilook processed PolSAR data, each data point is represented by a $p \times p$ covariance matrix, which is Hermitian (semi-) positive definite (HPD),

$$\mathbf{C} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{k}_i\mathbf{k}_i^{*T}, \tag{2}$$

where f^* denotes the complex conjugation of f.

The covariance matrix $\mathbf{C}$ can be modeled by a complex Wishart distribution, in homogeneous areas with fully developed speckle [5]. Let $\mathbf{\Sigma} = E\{\mathbf{k}\mathbf{k}^{*T}\}$, the probability density function (PDF) for $\mathbf{C}$ is

$$p_C(\mathbf{C}|n,\Sigma) = \frac{n^{np}\,|\mathbf{C}|^{n-p}}{\Gamma_d(n)\,|\mathbf{\Sigma}|^n} \exp\left\{-n{\cdot}\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{C}\right)\right\}, \tag{3}$$
$$\Gamma_p(n) = \pi^p \prod_{i=1}^{p}\Gamma(n-i+1),$$

where $\mathrm{tr}(\cdot)$ is the trace operator on a matrix. n is the number of looks and $\Gamma_p(n)$ is a normalization factor.

2.2. Dissimilarity Measures for PolSAR data

The dissimilarity measures are used for evaluating the dissimilarities between samples, in many cases, they are roughly called as distances. Mathematically, a distance is a function $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$, and when a distance function meets all the following properties [6], then it is a metric.

- $d(X,Y) \geq 0$ (nonnegativity).
- $d(X,Y) = 0 \Longleftrightarrow X = Y$ (definiteness).
- $d(X,Y) = d(X,Y)$ (symmetry).

- $d(X,Y) \leq d(X,Z) + d(Z,Y)$ (triangle inequality).

In practice, a well-defined dissimilarity measure should be at least nonnegative, definite and symmetric.

3. A REVISIT OF DISSIMILARITY MEASURES IN POLSAR DATA PROCESSING

The most widely used dissimilarity measures for PolSAR data classification and segmentation is the Wishart distance [1], a point to set distance defined between a sample covariance C and the center Σ_m of class m, which is derived from the log-likelihood of Wishart distribution,

$$d_W(C, \Sigma_m) = \log|\Sigma_m| + \mathrm{tr}(\Sigma_m^{-1} C). \quad (4)$$

The Wishart distance d_W was first used in the supervised maximum likelihood (ML) classifier [1], and the derived iterated Wishart Classifier has become an important component in many classification approaches [1, 7]

However, as pointed out by Anfinisen *et al.* [6], Wishart distance is not a well defined distance, as it do not meet any of the four conditions for metric. In fact, from the log-likelihood a more suitable distance can be obtained. Since the complex Wishart model belongs to exponential families, its negative log-likelihood can be written as the sum of a uniquely determined Bregman divergence and a function without distribution parameters [8],

$$-\log(p_{(\Psi,\theta)}(x)) = d_\phi(x, \mu(\theta)) - \log(b_\phi(x)). \quad (5)$$

where x is the sufficient statistic, Ψ is a convex function known as cumulant function, θ is the natural parameter. ϕ is the convex conjugate of Ψ and μ is the expectation parameter. It has $d_\phi(x,\mu(\theta)) \geq 0$ and $d_\phi(x,\mu(\theta)) = 0 \Longleftrightarrow x = \mu$. In the case of complex Wishart model, with a unknown constant parameter n to be estimated, Bregman divergence $d_\phi(x,\mu(\theta))$ can be written as

$$d_\phi(C, \Sigma_m) = n\left(\log\frac{|\Sigma_m|}{|C|} + \mathrm{tr}(\Sigma_m^{-1} C) - p\right). \quad (6)$$

Except for the dissimilarities directly derived from the log-likelihood, there are other ways to define matrices dissimilarities for PolSAR data, and many of them have been successful used in many tasks, such as filtering, segmentation, change detection, *etc*.

3.1. Dissimilarities based on Hypothesis Testing

One way to define dissimilarities is using hypothesis testing to measure the equality of two covariance matrices, which is derived by Conradsen *et al.* [9] in the task of PolSAR data change detection. The null and alternative hypothesis are $H_0: \Sigma_x = \Sigma_y$ and $H_1: \Sigma_x \neq \Sigma_y$, respectively. the resulting maximum-likelihood ratio test has a reduced form [9, 10],

$$Q = \frac{(n+m)^{p(n+m)}}{n^{pn} m^{pm}} \frac{|X|^n |Y|^m}{|X+Y|^{n+m}} \quad (7)$$

where X/n and Y/m are sample covariances and $X \sim W_C(p, n, \Sigma_x)$, $Y \sim W_C(p, n, \Sigma_y)$. The two dissimilarity measures derived from Q are the Bartlett distance, with the assumption that $n = m$,

$$d_B(\Sigma_x, \Sigma_y) = -\log(Q)/n = 2\log\frac{|(\Sigma_x + \Sigma_y)/2|}{\sqrt{|\Sigma_x||\Sigma_y|}} \quad (8)$$

and the revised Wishart distance, further assuming that Σ_y is known:

$$d_{RW}(\Sigma_x, \Sigma_y) = \log\frac{|\Sigma_x|}{|\Sigma_y|} + \mathrm{tr}(\Sigma_x^{-1}\Sigma_y) - p. \quad (9)$$

Note that d_B meets the first three conditions for metric and $\sqrt{d_B}$ also satisfies the last one [11], *i.e.*, it is a metric. However d_{RW} just satisfies the first two. A symmetric version of revised Wishart distance (SRW) has the form [6],

$$d_{SRW}(\Sigma_x, \Sigma_y) = \frac{1}{2}\left(\mathrm{tr}(\Sigma_x^{-1}\Sigma_y) + \mathrm{tr}(\Sigma_y^{-1}\Sigma_x)\right) - p \quad (10)$$

Kersten *et al.* [10] incorporated d_B and d_{RW} into fuzzy clustering. In the framework of spectral clustering [6, 12], symmetric distance are desirable to capture pairwise similarities. Anfinsen *et al.* [6] utilized d_B and d_{SRW} and demonstrated the way to transform the pairwise distance into affinity matrix. In [12], d_{SRW} was used to define the polarimetric cue.

In the region-based or hierarchical clustering approaches, we need to measure the dissimilarities between two clusters or two regions. Following Conradsen's paradigm, Cao *et al.* [7] derived the likelihood-ratio test statistic of two clusters in their agglomerative hierarchical clustering scheme. The distance between two sets (clusters) was defined as

$$\begin{aligned} d_m(S_i, S_j) &= -\frac{1}{n}\log\Lambda_{ij} \\ &= (N_i + N_j)\log|\hat{C}| - N_i \log|\hat{C}_i| - N_j \log|\hat{C}_j|. \quad (11) \end{aligned}$$

where $\hat{C}$, $\hat{C}_i$ and $\hat{C}_j$ are the estimated average covariances from sample sets $S_i \bigcup S_j$, S_i and S_j. In the patch-based filtering approaches [13, 2], d_B was adopted to define patch similarities. In the iterative bilateral filtering [14], d_{SRW} is used to define the weight coefficients between matrices data.

In the work of Alonso-Conzàlez *et al.* [15, 16], the construction of binary partition tree (BPT) also involves measuring the dissimilarities in the region model space. The proposed dissimilarities are based on two region features: the polarimetric information contained in covariance matrices and the region size. For example, SRW in this case has the following form:

$$d_{SRW'}(A, B) = \left(\mathrm{tr}(\Sigma_A^{-1}\Sigma_B + \Sigma_B^{-1}\Sigma_A)\right)(N_A + N_B) \quad (12)$$

3.2. Dissimilarities based on Information Theory

Information theory divergences have also been used in PolSAR data processing for a long time. For example, the revised

Wishart divergence d_{SRW} is actually the Kullback-Leibler (KL) distance between two Wishart distributed covariance matrics. In [17] Goudail *et al.* used KL and Bhattacharyya distance to quantify the dissimilarity between circular complex Gaussian distributions. Erten *et al.* [18] derived a coherent similarity using the mutual information for temporal multichannel scene characterization. Frery *et al.* [19] introduced the information divergence of $h-\phi$ family to define stochastic distances. The $h-\phi$ divergence between densities f_X and f_Y is defined as

$$D_{\phi}^{h}(X,Y)=h\left(\int\phi\left(\frac{f_X(C;\theta_1)}{f_Y(C;\theta_2)}\right)f_Y(C;\theta_2)dZ\right) \quad (13)$$

where $h:(0,\infty)\rightarrow[0,\infty)$ is a strictly increasing function with $h(0)=0$ and $\phi:(0,\infty)\rightarrow[0,\infty)$. By carefully choosing functions h and ϕ, some well-known divergences arise, such as the KL, Rènyi, Bhattacharyya and Hellinger [19]. To keep the symmetry property, the following expression is suggested,

$$d_{\phi}^{h}(X,Y)=\frac{D_{\phi}^{h}(X,Y)+D_{\phi}^{h}(Y,X)}{2} \quad (14)$$

By this way, all the derived distances keep the first three properties of a metric.

3.3. Dissimilarities based on Matrix Geometry

Another way to define the matrices dissimilarities is using the geodesic distances. In [10], Kersten *et al.* compared ℓ_p-norm defined in Euclidean vector space with several Wishart derived distances in fuzzy clustering, and the results show poor performances of the ℓ_p-norm based distances. Since the covariance matrices belong to the cone of HPD matrices, which forms a Riemannian maniflold, simple vectorization neglects their geometric structures. A natural choice is introducing the distances defined in Riemannian manifold, e.g. [14, 20]. In [14], D'Hondt *et al.* suggested two Riemannian distances for defining filtering weights, *i.e.* the affinity invariant Riemannian metric (AIRM)

$$d_{AI}(\Sigma_i,\Sigma_j)=\left\|\log\left(\Sigma_i^{-1/2}\Sigma_j\Sigma_i^{-1/2}\right)\right\|_F \quad (15)$$

and the log-Euclidean Riemannian metric (LERM):

$$d_{LE}(\Sigma_i,\Sigma_j)=\|\log(\Sigma_i)-\log(\Sigma_j)\|_F \quad (16)$$

where $\log(\cdot)$ is the matrix logarithm and $\|\cdot\|_F$ the Frobenius norm. AIRM is one of the most widely used similarity measure for covariance matrices, however, it is computationally expensive. LERM is an alternative of AIRM. The calculation of LERM involves matrix logarithms, is also computationally expensive, however, they could be evaluated off-line. The Jensen-Bregman LogDet divergence (JBLD) (the Bartlett distance, see Equation (8)) is another alternative of AIRM. As shown in [11, 21], JBLD also preserves several good properties of AIRM, but it can be evaluated more efficiently.

In the work of Alonso-Gonzales et al. [16], they compared several distance measures in the frame of BPT, and their results show that Riemannian distances had the best performances among all the tested distances.

3.4. Other Dissimilarities

Most of the dissimilarities described above require that the matrices should be positive-definite, *i.e.*, full-ranked. However, in some cases, e.g., filtering of PolSAR data, these matrices may be not full-ranked, and pre-filtering is needed. To address this problem, Alonso-Conzàlez *et al.* [15] introduced several dissimilarities that only employ the diagonal elements of the estimated covariance matrices, so pre-filtering will not be needed. However, these measures are not sensitive to the off-diagonal components, which reflect the correlations between channels.

Besides the covariance matrices data, vector form of polarimetric features is also important for PolSAR data processing. For feature vectors, the common used dissimilarities include the Euclidean distances, the cosine distances and the generalized squared Euclidean distance, *i.e.*, the Mahalanobis distance, etc.

Due to inherent imaging mechanism of PolSAR images, a local region maybe better described by the statistical distributions of features, rather than by individual feature vectors. As nonparametric estimations of empirical feature distributions, feature histograms are also very useful [22], the corresponding dissimilarities include the Minkowski-form distance, the Histogram Intersection (HI), the Kolmogorov-Smirnov distance (KS), the χ^2-statistic and the Earth Movers Distance (EMD), *etc*.

4. CONCLUSIONS

In this paper we have reviewed the main dissimilarity distances used in the literature and the main techniques concerning their applications. We would like to conclude with perspectives for further work:

1) The processing of PolSAR images usually involves calculating dissimilarity distances for three cases: point to point, point to set and set to set, such as the tasks of segmentation and classification. We need to seek a unified and robust dissimilarity distance for these three cases.

2) To analyze multitemporal PolSAR data, like filtering, classification or change detection in time series PolSAR data, we need to extend current dissimilarity measures or exploit new dissimilarity measures for simultaneously processing a series of $k>2$ PolSAR images.

3) Current dissimilarity distances are mostly developed under the well-established complex Wishart distribution, which is mainly used for describing the multilook complex

PolSAR images with Gaussian signal statistics. However, non-gaussian models usually give better performance, especially at high resolution. Therefore, we should investigate corresponding dissimilarities/distances for these non-gaussian models, such as the matrix-variate $\mathcal{K}$, $\mathcal{G}^0$ and the $\mathcal{U}$ distributions [23].

4) Handcrafting good dissimilarity measures for specific problems is generally difficult. An interesting way is to automatically learn dissimilarity measures for PolSAR images by using metric learning [24, 25], which can learn a distance function tuned to a particular task.

5. ACKNOWLEDGEMENTS

The research was partially supported by the National Key Basic Research and Development Program of China under Contract 2013CB733404 and the Chinese National Natural Sciences Foundation grants (No.61271401, No.91338113, No.61331016).

6. REFERENCES

[1] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*, CRC Press Inc., 2009.

[2] C.-A. Deledalle, L. Denis, G. Poggi, F. Tupin, and L. Verdoliva, “Exploiting patch similarity for sar image processing: the nonlocal paradigm,” *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 69–78, 2014.

[3] A.D. Nascimento, M.M. Horta, A.C. Frery, and R.J. Cintra, “Comparing edge detection methods based on stochastic entropies and distances for polsar imagery,” *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 7, no. 2, pp. 648–663, 2014.

[4] V. Alberga, “Similarity measures of remotely sensed multi-sensor images for change detection applications,” *Remote Sensing*, vol. 1, no. 3, pp. 122–143, 2009.

[5] J.S Lee, M.R. Grunes, and R. Kwok, “Classification of multi-look polarimetric SAR imagery based on complex wishart distribution,” *Int. J. Remote Sens.*, vol. 15, no. 11, pp. 2299–2311, 1994.

[6] S.N. Anfinsen, R. Jenssen, and T. Eltoft, “Spectral clustering of polarimetric SAR data with Wishart-derived distance measures,” in *Proc. POLInSAR 2007, Esrin, Italy, 22-26 January 2007*, Esrin, Italy, 2007.

[7] F. Cao, W. Hong, Y.-R. Wu, and E. Pottier, “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 Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3454–3467, 2007.

[8] A. Banerjee, S. Merugu, I.S. Dhillon, and J. Ghosh, “Clustering with bregman divergences,” *J. Mach. Learn. Res.*, vol. 6, pp. 1705–1749, Dec. 2005.

[9] K. Conradsen, A.A. Nielsen, J. Schou, and H. Skriver, “A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 4–19, 2003.

[10] P.R. Kersten, J.S. Lee, and Thomas L. Ainsworth, “Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and em clustering,” *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 519–527, 2005.

[11] A. Cherian, S. Sra, A. Banerjee, and N. Papanikolopoulos, “Jensen-bregman logdet divergence with application to efficient similarity search for covariance matrices,” *IEEE Trans. on Patt. Ana. and Mac. Int.*, p. 1, 2012.

[12] K. Ersahin, I.G. Cumming, and R.K. Ward, “Segmentation and classification of polarimetric SAR data using spectral graph partitioning,” *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 164–174, 2010.

[13] C.-A. Deledalle, L. Denis, and F. Tupin, “Iterative weighted maximum likelihood denoising with probabilistic patch-based weights,” *IEEE Trans. Image Process.*, vol. 18, no. 12, pp. 2661–2672, 2009.

[14] O. D’Hondt, S. Guillaso, and O. Hellwich, “Iterative bilateral filtering of polarimetric sar data,” *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 6, no. 3, pp. 1628–1639, 2013.

[15] A. Alonso-González, C. López-Martínez, and P. Salembier, “Filtering and segmentation of polarimetric sar data based on binary partition trees,” *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 593–605, 2012.

[16] A. Alonso-González, *Multidimensional and temporal SAR data representation and processing based on binary partition trees*, Ph.D. thesis, Universitat Politècnica de Catalunya, 2014.

[17] F. Goudail and P. Réfrégier, “Contrast definition for optical coherent polarimetric images,” *IEEE Trans. on Patt. Ana. and Mac. Int.*, vol. 26, no. 7, pp. 947–951, 2004.

[18] E. Erten, A. Reigber, L. Ferro-Famil, and O. Hellwich, “A new coherent similarity measure for temporal multichannel scene characterization,” *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2839–2851, 2012.

[19] A.C. Frery, A.D. Nascimento, and R.J. Cintra, “Analytic expressions for stochastic distances between relaxed complex wishart distributions,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1213–1226, 2014.

[20] N. Xue, G.-S. Xia, and L. Zhang, “Anisotropic diffusion on complex tensor fields for polsar image filtering,” in *IEEE Geoscience and Remote Sensing Symposium, IGARSS 2014*, 2014, pp. 3702–3705.

[21] H. Song, W. Yang, X. Xu, and M.S. Liao, “Unsupervised PolSAR Imagery Classification Based On Jensen-Bregman LogDet Divergence,” in *EUSAR 2014*. VDE, 2014, pp. 1–4.

[22] Y. Rubner, J. Puzicha, C. Tomasi, and J.M. Buhmann, “Empirical evaluation of dissimilarity measures for color and texture,” *Comp. Vis. Imag. Under.*, vol. 84, no. 1, pp. 25–43, 2001.

[23] S.N. Anfinsen and T. Eltoft, “Application of the matrix-variate mellin transform to analysis of polarimetric radar images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2281–2295, 2011.

[24] B. Kulis, “Metric learning: A survey,” *Foundations and Trends in Machine Learning*, vol. 5, no. 4, pp. 287–364, 2012.

[25] A. Bellet, A. Habrard, and M. Sebban, *Metric Learning*, Morgan & Claypool Publishers, 2015.