Statistical Properties of Geodesic Distances between Samples and Elementary Backscatterers in PolSAR Imagery

Danilo Fernandes, Alejandro C. Frery,
Laboratório de Computação Científica e Análise Numérica
Universidade Federal de Alagoas
Maceió, Brazil
Email: dfc@laccan.ufal.br, acfrery@laccan.ufal.br

***Abstract*—PolSAR data are usually represented by complex scattering or covariance matrices. Another representation is by Kennaugh matrices, which are real and preserve the backscatter information. This approach allows measuring distances and, consequently, the dissimilarity between PolSAR data and known backscatters. In this report, we analyze the statistical properties of this dissimilarity measure between PolSAR data samples and elementary backscatters using a Geodesic Distance.**

***Index Terms*—PolSAR, Kennaugh matrix, Geodesic Distance.**

I. Introduction

In Polarimetric SAR, a radar target is characterized by a scattering matrix $\boldsymbol{S}$ that describes the dependence of its scattering properties on the polarization. It is defined as

$$\boldsymbol{S} = \begin{bmatrix} S_{\text{HH}} & S_{\text{HV}} \\ S_{\text{VH}} & S_{\text{VV}} \end{bmatrix},$$

where H and V denote, respectively, horizontal and vertical polarization.

This information can be projected and then visualized using the Pauli vector representation: $\boldsymbol{k} = 2^{-1/2}\begin{bmatrix} S_{\text{HH}} + S_{\text{VV}} & S_{\text{HH}} - S_{\text{VV}} & 2S_{\text{HV}} \end{bmatrix}^T$, where T denotes transposition. Another important matrix in PolSAR theory is the coherency $\boldsymbol{T}$ obtained by:

$$\boldsymbol{T} = \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{k}_\ell \mathbf{k}_\ell^{*T},$$

where $*$ denotes the complex conjugate, and L is the number of looks.

The Kennaugh matrix $\boldsymbol{K}$ can be obtained from the coherency matrix $\boldsymbol{T}$:

$$\begin{bmatrix} \frac{T_{11}+T_{22}+T_{33}}{2} & \Re(T_{12}) & \Re(T_{13}) & \Im(T_{23}) \\ \Re(T_{12}) & \frac{T_{11}+T_{22}-T_{33}}{2} & \Re(T_{23}) & \Im(T_{13}) \\ \Re(T_{13}) & \Re(T_{23}) & \frac{T_{11}-T_{22}+T_{33}}{2} & -\Im(T_{12}) \\ \Im(T_{23}) & \Im(T_{13}) & -\Im(T_{12}) & \frac{-T_{11}+T_{22}+T_{33}}{2} \end{bmatrix}.$$

The Geodesic Distance between two Kennaugh matrices $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ is [1]:

$$\text{GD}(\boldsymbol{K}_1, \boldsymbol{K}_2) = \frac{2}{\pi}\cos^{-1}\left(\frac{\text{Tr}(\boldsymbol{K}_1^T\boldsymbol{K}_2)}{\sqrt{\text{Tr}(\boldsymbol{K}_1^T\boldsymbol{K}_1)}\sqrt{\text{Tr}(\boldsymbol{K}_2^T\boldsymbol{K}_2)}}\right),$$

where Tr denotes the trace. It ranges between $[0, 1]$. This leads to defining a measure of similarity between PolSAR data: $f(\boldsymbol{K}_1, \boldsymbol{K}_2) = 1 - \text{GD}(\boldsymbol{K}_1, \boldsymbol{K}_2)$. This measure of dissimilarity has been used to classify images [1], in the proposal of a new generalized volume scattering model [2], and for the analysis of built-up areas [3]. A measure of similarity, often more convenient for classification purposes [4], is $1 - \text{GD}$.

In this work, we are interested in the similarity between samples and elementary backscatterers: trihedral, dihedral, random volume, narrow dihedral, cylinder, dipole, left helix, right helix, $+1/4$-wave and $-1/4$-wave, whose Kennaugh matrices are, respectively:

$$K_a = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, K_b = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$K_{rv} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, K_{nd} = \begin{bmatrix} 5/8 & 3/8 & 0 & 0 \\ 3/8 & 5/8 & 0 & 0 \\ 0 & 0 & -1/2 & 0 \\ 0 & 0 & 0 & 1/2 \end{bmatrix},$$

$$K_c = \begin{bmatrix} 5/8 & 3/8 & 0 & 0 \\ 3/8 & 5/8 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & -1/2 \end{bmatrix}, K_d = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$K_{lh} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}, K_{rh} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix},$$

$$K_{+1/4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, K_{-1/4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}.$$

Table I presents the geodesic distances between these elementary backscatterers. Fig. 1 shows a graphical representation of the relationships between them. The size of the edges in this complete graph is proportional to the distance between each pair, as presented in Table I.

TABLE I
THE GEODESIC DISTANCE BETWEEN ELEMENTARY BACKSCATTERERS

	K_a	K_b	K_{rv}	K_{nd}	K_c
K_a	0.000	1.000	0.391	0.936	0.287
K_b	1.000	0.000	0.732	0.287	0.936
K_{rv}	0.391	0.732	0.000	0.703	0.434
K_{nd}	0.936	0.287	0.703	0.000	0.765
K_c	0.287	0.936	0.434	0.765	0.000
K_d	0.666	0.666	0.580	0.871	0.871
K_{lh}	1.000	0.666	0.732	0.702	0.968
K_{rh}	1.000	0.666	0.732	0.702	0.968
$K_{+1/4}$	0.666	0.666	0.580	0.666	0.666
$K_{-1/4}$	0.666	0.666	0.580	0.666	0.666

	K_d	K_{lh}	K_{rh}	$K_{+1/4}$	$K_{-1/4}$
K_a	0.666	1.000	1.000	0.666	0.666
K_b	0.666	0.666	0.666	0.666	0.666
K_{rv}	0.580	0.732	0.732	0.580	0.580
K_{nd}	0.871	0.702	0.702	0.666	0.666
K_c	0.871	0.968	0.968	0.666	0.666
K_d	0.000	0.839	0.839	0.666	0.666
K_{lh}	0.839	0.000	1.000	0.839	0.839
K_{rh}	0.839	1.000	0.000	0.839	0.839
$K_{+1/4}$	0.666	0.839	0.839	0.000	1.000
$K_{-1/4}$	0.666	0.839	0.839	1.000	0.000

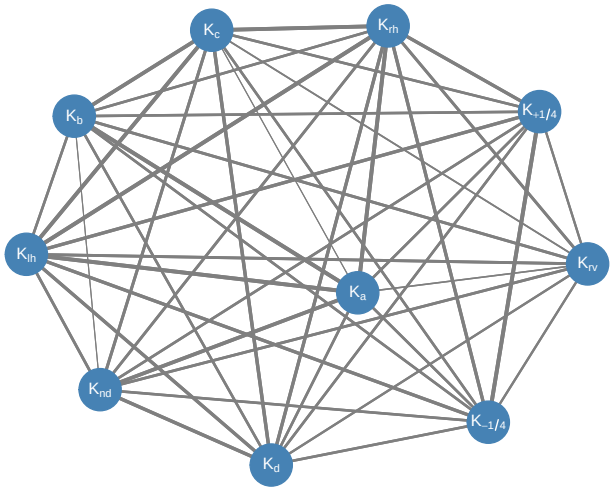


Fig. 1. Representation of the relationships among elementary backscatterers.

The forthcoming section presents an empirical analysis of the similarities between the aforementioned elementary backscatters and PolSAR data from samples of vegetation and bare soil. The data is from UAVSAR image over Sierra del Lacandon National Park, Guatemala.

II. Analysis of Distances

We selected samples of size 50×50 pixels over forest and bare soil regions, and computed the similarities of the observation to each elementary backscatterer. We also fitted a Beta distribution, with density given by

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1},$$

with $x \in [0,1]$, $\alpha, \beta > 0$. Some of the similarities did not span the $[0,1]$ interval; these cases were linearly mapped into this interval, and then fitted by the Beta law.

The mean is given by

$$\mu = \frac{\alpha}{\alpha+\beta}. \tag{1}$$

We estimated α and $\beta > 0$ by maximum likelihood and, thus, obtained models for these similarities.

We selected visually homogeneous regions of size 50×50 pixels from forested and bare soil areas. They are shown, in Pauli decomposition, in Figs. 2 and 3, respectively.

The histograms of the similarity of these data to each elementary backscatterer are shown in Figs. 4(a) to 4(j).

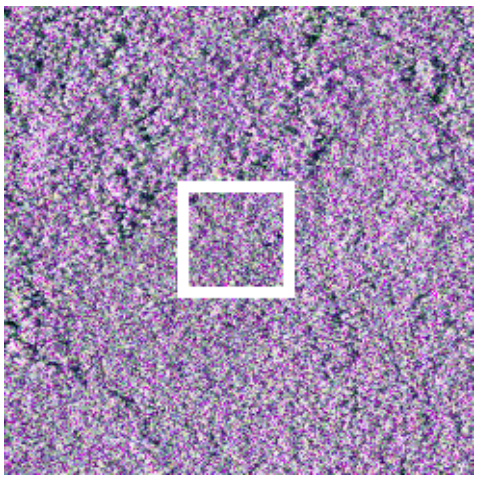
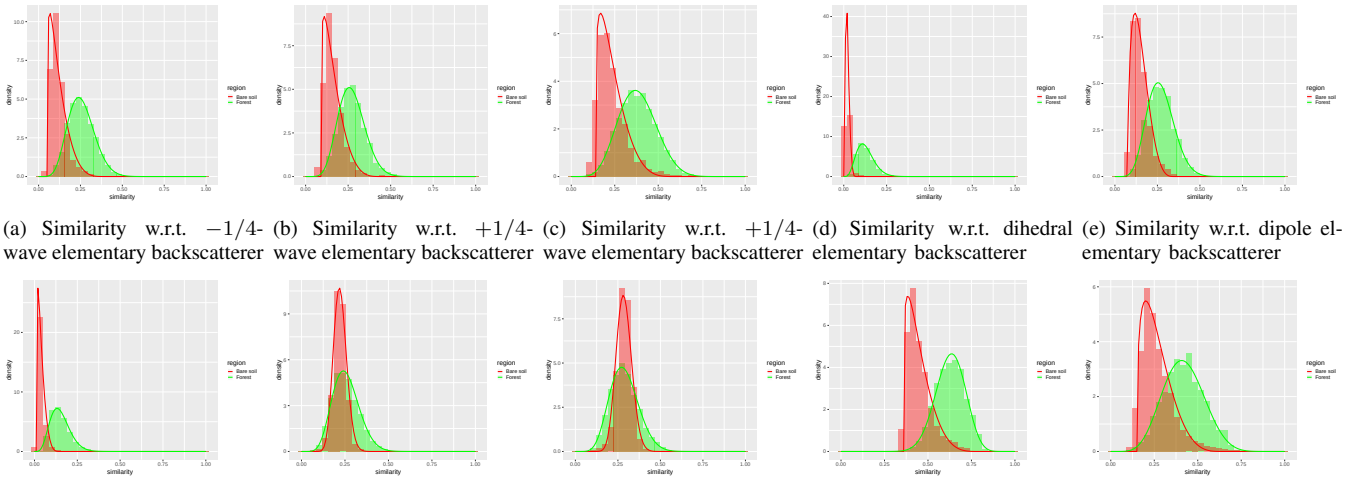


Fig. 2. Visualization of PolSAR data from the forest region through Pauli decomposition.

Fig. 3. Visualization of PolSAR data from the bare soil region through Pauli decomposition.



(a) Similarity w.r.t. $-1/4$-wave elementary backscatterer (b) Similarity w.r.t. $+1/4$-wave elementary backscatterer (c) Similarity w.r.t. $+1/4$-wave elementary backscatterer (d) Similarity w.r.t. dihedral elementary backscatterer (e) Similarity w.r.t. dipole elementary backscatterer

(f) Similarity w.r.t. narrow dihedral backscatterer (g) Similarity w.r.t. left helix backscatterer (h) Similarity w.r.t. right helix backscatterer (i) Similarity w.r.t. random volume backscatterer (j) Similarity w.r.t. trihedral backscatterer

Fig. 4. Histograms of similarities between forest and bare soil samples with respect to ten elementary backscatterers; Forest in green, Bare Soil in red, overlap in brown.

TABLE II
ESTIMATES OF THE BETA DISTRIBUTION, MINIMA, MAXIMA, AND ESTIMATED MEAN

	min	max	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\mu}$
			$-1/4$-wave		
Forest	0.000	1.000	7.830	22.758	0.255
Bare soil	0.055	0.400	1.127	4.872	0.119
			$+1/4$-wave		
Forest	0.000	1.000	8.681	23.277	0.271
Bare soil	0.090	0.450	1.200	4.800	0.162
			Cylinder		
Forest	0.000	1.000	7.500	12.165	0.381
Bare soil	0.140	0.600	1.243	4.756	0.235
			Dihedral		
Forest	0.000	1.000	5.380	36.870	0.127
Bare soil	0.009	0.070	1.327	4.672	0.022
			Dipole		
Forest	0.000	1.000	8.358	22.658	0.269
Bare soil	0.075	0.350	1.625	4.374	0.149
			Narrow dihedral		
Forest	0.000	1.000	5.890	33.198	0.150
Bare soil	0.016	0.150	1.119	4.880	0.041
			Left helix		
Forest	0.000	1.000	27.408	96.013	0.222
Bare soil	0.000	1.000	8.380	24.286	0.256
			Right helix		
Forest	0.000	1.000	28.522	71.238	0.285
Bare soil	0.000	1.000	8.316	20.574	0.287
			Random volume		
Forest	0.000	1.000	20.074	11.910	0.627
Bare soil	0.360	0.800	1.218	4.781	0.449
			Trihedral		
Forest	0.000	1.000	7.197	9.996	0.418
Bare soil	0.150	0.650	1.408	4.592	0.267

TABLE III
p-VALUES OF THE KOLMOGOROV-SMIRNOV GOODNESS-OF-FIT TEST OF THE SIMILARITY W.R.T. ELEMENTARY BACKSCATTERERS

	$-1/4$-wave	$+1/4$-wave	Cylinder	Dihedral	Dipole
Forest	0.979	0.808	0.763	0.733	0.975
Bare soil	0.361	0.893	0.264	0.443	0.475
	Left helix	Narrow dihedral	Random volume	Right helix	Trihedral
Forest	0.959	0.787	0.589	0.344	0.582
Bare soil	0.099	0.206	0.480	0.072	0.127

The differences are clear, denoting the expressiveness of this measure of similarity.

Table II shows the estimated parameters of the Beta distribution when used as model for all similarities. The table also presents the minima and maxima of each sample, and the estimate of the mean value computed with (1).

The elementary backscatterers which are closest to the samples are highlighted in red. It is noteworthy that both Forest and Bare Soil are closest to the Random Volume model.

Through those, the probability density functions of the figures 4(a) to 4(j) were adjusted to the histograms of the similarities. In addition, this table contains the estimate of the hope that results from the estimation of these parameters.

Table III presents the p-values of the Kolmogorov-Smirnov goodness-of-fit test of the similarities when expressed by Beta models. These tests were performed over 400 randomly selected observations (without replacement), and they show a good adequacy: the smallest p-value is 0.07.

III. CONCLUSIONS

This is a first step towards a statistical analysis of measures of similarities between PolSAR observations and elementary

backscatterers in the space of Kennaugh representations.

The Beta distribution is an acceptable model for describing the similarities.

REFERENCES

[1] D. Ratha, A. Bhattacharya, and A. C. Frery, "Unsupervised classification of PolSAR data using a scattering similarity measure derived from a geodesic distance," *IEEE Geoscience and Remote Sensing Letters*, pp. 151–155, Jan. 2018.

[2] D. Ratha, D. Mandal, V. Kumar, H. McNairn, A. Bhattacharya, and A. C. Frery, "A generalized volume scattering model based vegetation index from polarimetric SAR data," *IEEE Geoscience and Remote Sensing Letters*, p. in press, 2019.

[3] D. Ratha, P. Gamba, A. Bhattacharya, and A. C. Frery, "Novel techniques for built-up area extraction from polarimetric SAR images," *IEEE Geoscience and Remote Sensing Letters*, p. in press, 2019.

[4] A. C. Frery, A. H. Correia, and C. C. Freitas, "Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3098–3109, 2007.