A JOINT DENSITY OF INTERFEROMETRIC AND/OR POLARIMETRIC IMAGES: APPLICATION TO CHANGE DETECTION

E. Erten[1,2], **A. Reigber**[2], **and R. Zandoná-Schneider**[2]

[1]*Computer Vision and Remote Sensing, Technical University of Berlin, 10587, Berlin, Germany*
[2]*Microwaves and Radar Institute, German Aerospace Centre (DLR), 82234 Oberpfaffenhofen, Germany.*

ABSTRACT

Polarimetric data of distributed scatterer can be fully characterised by the (3×3) Hermitian positive definite matrix which follows a complex Wishart distribution under Gaussian assumption. A second observation in time will also follow Wishart distribution. Then, these observations are correlated or uncorrelated process over time related to the monitored objects. To not to make any assumption concerning their independence, the (6×6) matrix which is also modelled as a complex Wishart distribution is used in this study to characterise the behaviour of the temporal polarimetric data. In particular, we derive a closed-form expression of the joint probability density function of two polarimetric data thus enabling the exact evaluation of the change detection performances. Regarding the proposed temporal polarimetric data distribution, we propose a new algorithm for evaluating the change detection with KL-divergence test. The KL-divergence is a distance measurement between two probability distributions. In our case, also the case of mutual information, it measures the dependency of two variables by calculating the distance between the probability density of joint distribution of polarimetric data and their marginal probability densities. We illustrate this new change detection algorithm is independent from the dimension of the system that it can be easily implemented to lower or higher multi-channel SAR systems.

Key words: multi-channel SAR; PolInSAR; change detection; KL-divergence test.

1. CHARACTERISATION OF MULTI-CHANNEL TEMPORAL SAR IMAGES

To be precise, with the same notations in [1], let $\mathbf{k} = [\mathbf{k}_1, \mathbf{k}_2]^T$ be a complex target vector distributed as multivariate complex Gaussian $\mathcal{N}^C(0, \Sigma)$ that consist of two target vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ obtained from temporal images at time t_1 and t_2. Since true covariance matrix of target vector Σ hasn't known, it is estimated by n samples incoherent averaging and it has a complex Wishart distribution with n degrees of freedom. Here, the number of elements in one of the target vector $\mathbf{k}_i$ is represented by m, and the target vector $\mathbf{k}$ has the dimension of $p = 2 \times m$.

The n look covariance matrix $\mathbf{A} = \frac{1}{n}\sum_{j=1}^{n} \mathbf{k}_j \mathbf{k}_j^\dagger$ is the estimation of the true covariance matrix Σ, and it summaries whole (joint and marginal) information from both images. If $\mathbf{A}$ is portioned as $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}$, the conditioned on $\mathbf{A}_{11}$, the joint density of element $\mathbf{A}_{22}$ follows the complex Wishart distribution

$$\mathbf{A}_{11|22} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$$
$$p(\mathbf{A}_{11}|\mathbf{A}_{22}) = \mathcal{W}_q^C(n-q, \mathbf{\Sigma}_{11|22}),$$

and it is independent from $\mathbf{A}_{12}$ and $\mathbf{A}_{22}$. Then, using the well known rule that the conditional distribution of correlation matrix $\mathbf{A}_{12}$ given $\mathbf{A}_{22}$ is a complex normal distribution

$$p(\mathbf{A}_{12}|\mathbf{A}_{22}) \sim \mathcal{N}_{q\times q}^C(\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{12}^{-1}\mathbf{A}_{22},\ \mathbf{\Sigma}_{11.2} \otimes \mathbf{A}_{22})$$

where $\otimes$ indicates Kronecker products and the *theorem 10.3.2* in [2] , the conditional distribution of $\mathbf{R}^2 = \mathbf{A}_{12}\mathbf{A}_{11}^{-1}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$ on $\mathbf{A}_{22}$ $\left(p(\mathbf{R}^2|\mathbf{A}_{22})\right)$ is a non central Wishart distribution. Since $p\left(\mathbf{A}_{11|22}, \mathbf{A}_{22}, \mathbf{R}^2\right) = p\left(\mathbf{A}_{11|22}\right) p\left(\mathbf{R}^2|\mathbf{A}_{22}\right) p\left(\mathbf{A}_{22}\right)$, after transforming $\mathbf{A}_{11|22}$ into $\mathbf{A}_{11}(\mathbf{I} - \mathbf{R}^2)$, the joint density $p\left(\mathbf{A}_{11}, \mathbf{A}_{22}, \mathbf{R}^2\right) = p\left(\mathbf{A}_{11|22}\right) p\left(\mathbf{R}^2|\mathbf{A}_{22}\right) p\left(\mathbf{A}_{22}\right)$ can be calculated.[1] After integrating this density function over $\mathbf{R}^2$, the joint distribution of two complex Wishart distribution $p(\mathbf{A}_{11}, \mathbf{A}_{22})$ can be written in the following form

$$p\left(\mathbf{A}_{11}, \mathbf{A}_{22}\right) = etr\left(-n\frac{\mathbf{\Sigma}_{22}^{-1}\mathbf{A}_{22} + \mathbf{\Sigma}_{11}^{-1}\mathbf{A}_{11}}{\mathbf{I} - \mathbf{P}^2}\right)$$
$${}_0\tilde{F}_1\left(m, m^2\mathbf{A}_{11}^{1/2}\mathbf{\Sigma}_{11.2}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{A}_{22}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\Sigma_{11.2}^{-1}\mathbf{A}_{11}^{1/2}\right)$$
$$\frac{n^{2mn}|\mathbf{A}_{11}\mathbf{A}_{22}|^{n-q}}{|\mathbf{\Sigma}_{11}\mathbf{\Sigma}_{22}|^n|I - P^2|^n\tilde{\Gamma}_q(n)\tilde{\Gamma}_q(n)}. \quad (1)$$

Here, $\mathbf{P}^2 = \mathbf{\Sigma}_{11}^{-1/2}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1/2}$, ${}_0\tilde{F}_1(n, \Psi)$ is the complex hypergeometric function of matrix argument [4]

[1]The proof of this distribution for the real case can be found in [2] and [3].

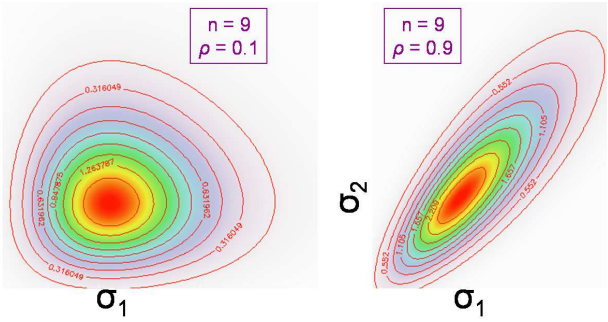

Figure 1. Comparison of theoretical pdf (1) with histogram of simulated data with different correlation and the number of samples scenario. Images are obtained from simulated data and the over plotted counters are obtained form the simulated data.

$$ {}_0\tilde{F}_1(n,t) = \frac{\|t_j^{(m-n+i-1)/2} I_{n-m+i-1}(2\sqrt{t_j})\|_{i,j=1}^{m}}{\prod_{i<j}^{m}(t_i - t_j)} \times (-1)^{m(m-1)/2} \prod_{k=1}^{m}(n-k)!, \quad (2) $$

where t are the positive eigenvalues of the complex hermitian matrix of $\mathbf{\Psi}$ and $\tilde{\Gamma}_q(n)$ is a complex gamma function

$$ \tilde{\Gamma}_q(n) = \pi^{q(q-1)/2} \prod_{i=1}^{q} \Gamma(n-i+1). \quad (3) $$

It is clear that (1) is valid for $\mathbf{0} < \mathbf{P}^2 < \mathbf{I}$, which means that both $\mathbf{P}^2$ and $\mathbf{I} - \mathbf{P}^2$ are positive definitive. When $\mathbf{\Sigma}_{12} = 0$, then $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ are independent. Regarding the dimension of multi-temporal images it is not easy two visualise the proposed distribution cause we will demonstrate (1) in the case of $m = 1$, see Figure 1. It can be proved from [5] that for $m = 1$ joint probability density $p(\mathbf{A}_{11}, \mathbf{A}_{22})$ converges to bigamma distribution.

2. CHANGE DETECTION DECISION STATISTICS

In previous section, to investigate the temporal behaviour of polarimetric data, the joint density functions of polarimetric temporal images were derived in the context of multi-channel SAR systems. In this section, previous results are used with the aim of change detection.

In probability theory, KL-divergence test is used as a measure of the difference between two probability distributions with following algorithm

$$ \int log\left[\frac{p(A_{11}, A_{22})}{p(A_{11}), p(A_{22})}\right] p(A_{11}, A_{22}) d\overrightarrow{A}. \quad (4) $$

Here, $p(\mathbf{A}_{11})$ and $p(\mathbf{A}_{22})$ are marginal densities of the m^2 complex element vector obtained by stacking the columns of $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ respectively. After substituting (1) and the marginal distributions of $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ in (4), we can obtain the following decision statisting D_n regarding to KL-divergence test.

$$ D_n(\Sigma_{11}, \Sigma_{22}, \Sigma_{12}) = \langle log({}_0\tilde{F}_1(n,t)) \rangle - n \log(|1-P|) - tr\left(-\frac{2nP}{I-P}\right) \quad (5) $$

where $\langle . \rangle$ and $|.|$ indicate the expectation and the determinant operator respectively. Regarding (5), it can be seen that decision statistics is directly related correlation information between multi-channel images.

As expected, the proposed change detection algorithm is a decreasing function regarding the dimension of the multi-channel SAR system. Which can be a advantage in comparing the different system configuration performances in change detection, especially in polarimetry, we can see the contribution of the different channel in change detection analysis.

3. COMPARISON WITH OTHER CHANGE DETECTORS

In this section, the performance of the proposed algorithm is compared with other two well-known change detection statistics known as ML Ratio [6] and the Ratio-Edge statistics [7]. These two change detection statistics are very powerful and easy to implement techniques.

For example, the probability of the detection that there is a change is a function of the detection threshold $\mathcal{T}$, that can be obtained as

$$ p(D_n \leq \mathcal{T}) = \int_0^{\mathcal{T}} p_{D_n}(x) \mathrm{d}x \quad (6) $$

with the following binary hypothesis test:

$$ H_o \text{ (presence of change)} : D_n \leq \mathcal{T} $$
$$ H_1 \text{ (absence of change)} : D_n > \mathcal{T}. $$

Expressions for the probability of detection (p_D) and of false alarm (p_{FA}) can be defined by two following probabilities

$$ p_D = p[\text{accepting } H_0 | H_0 \text{ is true}] $$
$$ p_{FA} = p[\text{accepting } H_0 | H_1 \text{ is true}]. $$

Thus, for each value of $\mathcal{T}$, there exists a pair (p_{FA}, p_D). The curves of p_D versus p_{FA} called Receiver Operating Characteristic (ROC) curves are analysed in the following example for three change statistics algorithm called as KL-divergence (the proposed technique, D_n), Ratio-Edge and ML-Ratio statistics. However, it is difficult to compare the performance of different techniques cause the ROC curves are related to the selected threshold. Because of this amplitude normalisation (unity intensity) should be applied to simulated data.

In Fig. 2, ROC curves are represented for three different techniques with different number of looks. For small number of looks, the proposed detector has a better performance than other two detectors. Moreover, it is interesting to see that these three approaches have similar performances for larger number of samples[2]. Regarding simulated data, this can be explained by that increasing the number of samples causes more strong estimation.

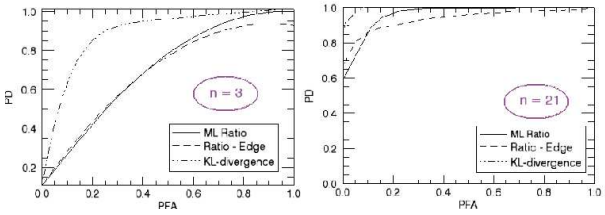

Figure 2. Comparison of three decision statistics with different number of samples.

For a probability of error of ROC curves, $P_e = \frac{1-P_D+P_{FA}}{2}$, below 0.2, it is more interesting to use KL-divergence test statistics instead of other two detectors, see Figure 3. It is interesting point of view to evaluate the threshold to be applied to obtain the best tradeoff between detection and false alarms. The best value for threshold is to be found at the minimum of the curves. In this case, it is interesting to see that KL-divergence and the ML-ratio detectors have a similar behaviour for selected threshold, whereas KL-divergence and the Ratio-Edge detectors have very different behaviour versus threshold.

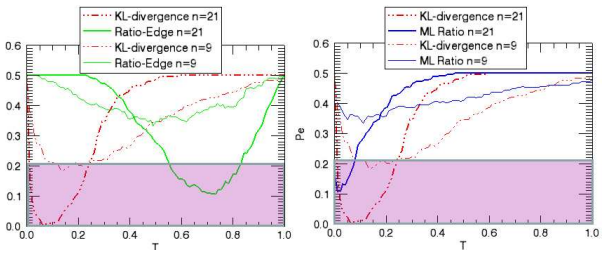

Figure 3. Probability of error for different detectors regarding selected threshold.

Unfortunately, in real data as indicated in [8], the performance of the detectors can be shown with different performances cause threshold is very sensitive to noise and fluctuations.

[2]However, increasing the number of sample can not guaranteed better estimation in real SAR images if the local stationary condition is not supplied.

4. CONCLUSION

A new joint distribution and a change detection decision statistic has been proposed for multi-channel temporal SAR Images. A new algorithm for change detection, which is based on KL-divergence test, has been introduced. The proposed algorithm uses mutual information of temporal images to evaluate their dependence of each other. The proposed detector has been compared to the classical change detector regarding simulated data and has been shown to have a more robust behaviour than the classical algorithms. However, true evaluation of the proposed algorithm will be done with real data, and is an area of future research of proposed techniques.

REFERENCES

[1] Ferro-Famil, L.P. and Lee, E.J.S., 2001, Unsupervised classification of multifrequency and fullypolarimetric SAR images based on the H/A/Alpha-Wishart classifier, IEEE Trans. on Geoscience and Remote Sensing, vol 39, No 11, Nov., pp 2332-2342

[2] Muirhead, R. J., 1982, Aspects of multivariate statistic theory, John Wiley & Sons

[3] Lliopoulos, G., 2066, UMVU estimation of the ratio of powers of normal generalized variances under correlation, Journal of Multivariate Analysis

[4] Smith, P. J. and Garth, L. M., 2007, Distribution and characteristic functions for correlated complex Wishart matrices, Journal of Multivariate Analysis, vol 98, no 4, pp 661-677

[5] Chatelain, F., Tourneret, J. Y., Inglada, J. and Ferrari, A., 2007, Bivariate gamma distributions for image registration and change detection, IEEE Trans. on Image Processing, Vol 16, No 7, Jul., pp 1796-1806

[6] Conradsen, K., Nielsen, A. A., Schou, J. and Skriver, H., 2003, A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data, IEEE Trans. on Geoscience and Remote Sensing, vol 41, Jan., pp 4-19

[7] Touzi, R., Lopes, A. and Bousquet, P., 1988, A statistical and geometrical edge detector for SAR images, IEEE Trans. on Geoscience and Remote Sensing, vol 26, No 6, Nov., pp 764-773

[8] Kersten, P. R., Lee, J. S. and Ainsworth, T. L., 2005, A comparison of change detection statistics in POLSAR images, Geoscience and Remote Sensing Symposium, IGARSS'05. Proceedings, vol 7