Robust low-rank change detection for SAR image time series

Ammar Mian\textsuperscript{1,2}, Arnaud Breloy\textsuperscript{3}, Guillaume Ginolhac\textsuperscript{2}, Jean-Philippe Ovarlez\textsuperscript{1,4}

\textsuperscript{1}: SONDRA, CentraleSupélec \textsuperscript{2}: LISTIC, Université Savoie Mont Blanc
\textsuperscript{3}: LEME, Université Paris Panthéon \textsuperscript{4}: ONERA, Palaiseau

Tuesday, July 30
Contents of the presentation

1 Motivations
2 Data
3 Statistical framework
4 Proposed Approach
5 Experimental results

Sources available at:
- Slides: https://ammarmian.github.io/igarss_slides_2019.pdf
- Code: https://github.com/AmmarMian/Robust-Low-Rank-CD
Motivations
Monitoring natural disasters:

PolSAR images of Ishinomaki and Onagawa areas [Sato et al., 2012], Nov.2010 (left), Apr.2011 (right).
Problems to consider

Huge increase in the number of available acquisitions:

- Sentinel-1: 12 days repeat cycle, since 2014
- TerraSAR-X: 11 days repeat cycle, since 2007
- UAVSAR, ...

Detect changes

- Massive amount of data  →  Automatic process
- Unlabeled data  →  Unsupervised detection
Data
Synthetic aperture radar (SAR)

**Principle of SAR**

Advantages:
- All weather and illumination conditions (active technology)
- Very high-resolution (sub-meter) imaging

Comparison of optical and image
Multivariate data: natural or pre-processing

Feature selection

- Leverage *diversity* to improve the detection
- Requires to process *multivariate* pixels
SAR image time series representation

Local patches

\[ X_t = \{ x_{k}^t | k \in [1, K] \} \]

t = 1

t = T

Figure 1: Sliding windows, the gray pixel corresponds to the test pixel
**Change detection (CD) problem (T=2)**

For each patch, decide if a change occurred between $X_1$ and $X_2$.

![Change detection images](image_url)

**Statistical detection framework**

- Can handle the multivariate aspect of the data
- Can account for physical modelling of the data/noise
- Strong theoretical guarantees from statistical literature
Statistical framework
Parametric change detection

Parametric probability model:

$$X_t \sim \mathcal{L}(X_t; \theta_t).$$

Change detection $\rightarrow$ Hypothesis test:

$$\begin{align*}
\text{H}_0 & : \theta_1 = \theta_2 \quad (\text{no change}) \\
\text{H}_1 & : \theta_1 \neq \theta_2 \quad (\text{change})
\end{align*}$$
Generalized likelihood ratio test (GLRT)

Statistical decision test derived as:

$$\max_{\theta_1, \theta_2} \frac{L(\{X_1, X_2\} ; \{\theta_1, \theta_2\})}{\max_{\theta_0} L(\{X_1, X_2\} ; \theta_0)} \quad H_1 \geq \lambda_{\text{GLRT}} \quad H_0.$$

Problems

- Specify $L$ and $\theta$ to model the data
  - Good fit
  - Robust to a large class of distributions and outliers
- Handy model to compute the ratio efficiently (closed form or optimization)
Seminal work [Conradsen et al., 2003]

Gaussian model

Assuming $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_p, \Sigma)$:

$$\theta = \Sigma$$

$$\mathcal{L}(\mathbf{X}; \Sigma) \propto |\Sigma|^{-K_{\text{etr}}} \left\{ -\mathbf{X}^H \Sigma^{-1} \mathbf{X} \right\}.$$  

Detection test:

$$\begin{cases} 
\text{H}_0 : & \Sigma_1 = \Sigma_2 \quad \text{(no change)} \\
\text{H}_1 : & \Sigma_1 \neq \Sigma_2 \quad \text{(change)} 
\end{cases}$$

Corresponding GLRT$^a$

$$\hat{\Lambda}_G = \left| \frac{1}{2} \left( \hat{\Sigma}_1 + \hat{\Sigma}_2 \right) \right|^2 \frac{1}{\left\| \hat{\Sigma}_1 \right\| \left\| \hat{\Sigma}_2 \right\|} \quad \text{H}_1 \gtrsim \text{H}_0 \lambda,$$

where

$$\forall t, \hat{\Sigma}_t = \mathbf{X}_t \mathbf{X}_t^H / K.$$  

$^a$ Other Gaussian/Covariance methods [Ciouonzo et al., 2017, Nascimento et al., 2019].
Non-Gaussian models in CD [Mian et al., 2019a]

Robust model: Compound-Gaussian distributions

Assuming $x_k \sim CN(0_p, \tau_k \Sigma)$.

$$\theta = \{\Sigma, \{\tau_k\}\}$$

$$\mathcal{L}(X; \Sigma, \{\tau_k\}) \propto \prod_{k=1}^K |\tau_k \Sigma|^{-1} \exp \left\{ -\frac{x_k^H \Sigma^{-1} x_k}{\tau_k} \right\}.$$ 

Corresponding GLRTs in [Mian et al., 2019a].
Structured covariance models in CD [Ben Abdallah et al., 2019]

Low-rank structured covariance

Assuming $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_p, \Sigma_R + \sigma^2 \mathbf{I})$.

$\theta = \Sigma_R$, with $\text{rank}(\Sigma_R) = R$

$\mathcal{L}(\mathbf{X}; \Sigma_R) \propto |\Sigma_R + \sigma^2 \mathbf{I}|^{-K} \text{etr} \left\{ -\mathbf{X}^H (\Sigma_R + \sigma^2 \mathbf{I})^{-1} \mathbf{X} \right\}$

Corresponding GLRTs in [Ben Abdallah et al., 2019].
Proposed Approach
Proposed CD test

Low-rank Compound-Gaussian model

Assuming \( x_k \sim \mathbb{C}\mathcal{N}(0_p, \tau_k(\Sigma_R + \sigma^2 I)) \).

\[
\theta = \{\Sigma_R, \{\tau_k\}\} \quad \text{with} \quad \text{rank}(\Sigma_R) = R
\]

\[
\mathcal{L}(X; \Sigma, \{\tau_k\}) \propto \prod_{k=1}^{K} |\tau_k(\Sigma_R + \sigma^2 I)|^{-1} \exp \left\{ -\frac{x_k^H(\Sigma_R + \sigma^2 I)^{-1}x_k}{\tau_k} \right\}.
\]

Recalling our problems

- Specify \( \mathcal{L} \) and \( \theta \) to model the data (✓)
- Compute the ratio efficiently (?)
Proposed block coordinate descent (BCD) algorithms

**Algorithm 1** BCD for MLEs under $H_1$

**Input:** $\{x^t_k\}$ with $t \in \{1, 2\}$

**repeat**

$\tau^t_k = \left( (x^t_k)^H \Sigma_t^{-1} x^t_k \right) / p$

$\Sigma_t = T \left\{ \frac{1}{K} \sum_{k=1}^{K} x^t_k (x^t_k)^H \right\}$

**until** convergence

**Output:** $\{\hat{\Sigma}_t, \{\hat{\tau}^t_k\}\}$

**Algorithm 2** BCD for MLE under $H_0$

**Input:** $\{x^1_k, x^2_k\}$

**repeat**

$\tau^0_k = \left( (x^1_k)^H \Sigma_0^{-1} x^1_k + (x^2_k)^H \Sigma_0^{-1} x^2_k \right) / 2p$

$\Sigma_0 = T \left\{ \frac{1}{K} \sum_{k=1}^{K} x^1_k (x^1_k)^H + x^2_k (x^2_k)^H \right\}$

**until** convergence

**Output:** $\{\hat{\Sigma}_0, \{\hat{\tau}^0_k\}\}$

**Low-rank Compound-Gaussian GLRT**

$$\frac{\mathcal{L}_{H_1} \left( \{X_1, X_2\}; \{\hat{\Sigma}_1, \hat{\Sigma}_2, \{\hat{\tau}^1_k\}, \{\hat{\tau}^2_k\}\} \right)}{\mathcal{L}_{H_0} \left( \{X_1, X_2\}; \{\hat{\Sigma}_0, \{\hat{\tau}^0_k\}\} \right)} \begin{cases} \equiv H_1 \quad \text{for} \quad \lambda_{GLRT}. \\ H_0 \end{cases}$$

Ammar Mian, 30 July 2019
Experimental results
Motivations

Data

Statistical framework

Proposed Approach

Experimental results

Dataset

Description

- Polarimetric data $\rightarrow$ wavelet decomp. [Mian et al., 2017] $\rightarrow$ $p = 12$ dim. pixels
- Image size: 2360px×600px
- Resolution: 1.67 m (Range) and 0.60 m (Azimuth)
- CD ground truth from [Nascimento et al., 2019]
Recall of the considered CD methods

| Gaussian                     | Low-rank Gaussian               | Low-rank Compound-Gaussian      |
|------------------------------|---------------------------------|---------------------------------|
| \( \mathbf{x} \sim \mathcal{CN}(\mathbf{0}_p, \Sigma) \) | \( \mathbf{x} \sim \mathcal{CN}(\mathbf{0}_p, \Sigma_R + \sigma^2\mathbf{I}) \) | \( \mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}_p, \tau_k(\Sigma_R + \sigma^2\mathbf{I})) \) |
| \( \theta = \Sigma \)        | \( \theta = \Sigma_R \), with rank(\(\Sigma_R\)) = R | \( \theta = \{\Sigma_R, \{\tau_k\}\}, \) with rank(\(\Sigma_R\)) = R |

| Compound-Gaussian          |                                |                                |
|---------------------------|--------------------------------|--------------------------------|
| \( \mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}_p, \tau_k \Sigma) \) |                                |                                |
| \( \theta = \{\Sigma, \{\tau_k\}\} \) |                                |                                |

Side parameters

- Rank \( R \) and noise floor \( \sigma^2 \) estimated on the whole datacube
Results with a $5 \times 5$ sliding windows: Gaussian detectors
Results with a $5 \times 5$ sliding windows: Robust detectors
**Performance curves**

**Figure 2:** Probability of detection $P_D$ versus probability of false alarm $P_{FA}$ with $(p = 12, N = 25, R = 3)$

**Figure 3:** $P_D$ versus the size of window at $P_{FA} = 5\%$ with $(p = 12, R = 3)$
Thanks for your attention!
Ben Abdallah, R., Mian, A., Breloy, A., Korso, M. N. E., and Lautru, D. (2019). *Detection methods based on structured covariance matrices for multivariate SAR images processing.* *IEEE Geoscience and Remote Sensing Letters.*

Ciuonzo, D., Carotenuto, V., and Maio, A. D. (2017). *On multiple covariance equality testing with application to SAR change detection.* *IEEE Transactions on Signal Processing,* 65(19):5078–5091.

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 Transactions on Geoscience and Remote Sensing,* 41(1):4–19.
Mian, A., Ginolhac, G., Ovarlez, J.-P., and Atto, A. M. (2019a).  
**New robust statistics for change detection in time series of multivariate SAR images.**  
*IEEE Transactions on Signal Processing*, 67(2):520–534.

Mian, A., Ovarlez, J.-P., Atto, A. M., and Ginolhac, G. (2019b).  
**Design of new wavelet packets adapted to high-resolution SAR images with an application to target detection.**  
*IEEE Transactions on Geoscience and Remote Sensing.*

Mian, A., Ovarlez, J.-P., Ginolhac, G., and Atto, A. M. (2017).  
**Multivariate change detection on high resolution monovariate SAR image using linear time-frequency analysis.**  
In *2017 25th European Signal Processing Conference (EUSIPCO)*, pages 1942–1946.
Nascimento, A. D. C., Frery, A. C., and Cintra, R. J. (2019). Detecting changes in fully polarimetric SAR imagery with statistical information theory. *IEEE Transactions on Geoscience and Remote Sensing*, to appear:1–13.

Sato, M., Chen, S., and Satake, M. (2012). Polarimetric sar analysis of tsunami damage following the march 11, 2011 east japan earthquake. *Proceedings of the IEEE*, 100(10):2861–2875.
Figure 4: $P_D$ versus the size of window at $P_{FA} = 10\%$
As studied in [Mian et al., 2017], it is possible to increase detection performance by increasing data diversity using wavelet decomposition. By doing a $2 \times 2$ decomposition, we obtain vectors of dimension $p = 12$. 
Wavelet decomposition pre-processing

As studied in [Mian et al., 2017], it is possible to increase detection performance by increasing data diversity using wavelet decomposition. By doing a $2 \times 2$ decomposition, we obtain vectors of dimension $p = 12$. 
Wavelet decomposition pre-processing

As studied in [Mian et al., 2017], it is possible to increase detection performance by increasing data diversity using wavelet decomposition. By doing a $2 \times 2$ decomposition, we obtain vectors of dimension $p = 12$. 
Wavelet decomposition pre-processing

As studied in [Mian et al., 2017], it is possible to increase detection performance by increasing data diversity using wavelet decomposition. By doing a $2 \times 2$ decomposition, we obtain vectors of dimension $p = 12$. 
As studied in [Mian et al., 2017], it is possible to increase detection performance by increasing data diversity using wavelet decomposition. By doing a $2 \times 2$ decomposition, we obtain vectors of dimension $p = 12$. 