

# MULTIVARIATE STATISTICAL MODELING FOR MULTI-TEMPORAL SAR CHANGE DETECTION USING WAVELET TRANSFORMS

*Bouhlef Nizar<sup>(1,2)</sup>, Guillaume Ginolhac<sup>(2)</sup> and Abdourrahmane Atto<sup>(2)</sup>*

(1) IUT d'Annecy, University Savoie Mont Blanc, 74940 Annecy-le-Vieux

(2) LISTIC - Polytech' Annecy-Chambry, BP 80439, 74944 Annecy le Vieux Cedex, France

## ABSTRACT

In this paper, we propose a new method for automatic change detection in multi-temporal SAR images based on statistical wavelet subband modeling. The image is decomposed into multiple scales using wavelet transform and the probability density function of the sliding windows coefficients of each subbands is assumed to be multivariate Gaussian distribution. Kullback-Leibler similarity measures are computed between two corresponding subbands of the same scale and used to generate the change map. The multivariate statistical model is considered here to better model the spatial information given by texture than that given by a univariate statistical model. The proposed method is compared to the classical method based on univariate Gaussian distribution. Test on real data show that our approach outperforms the conventional approach.

**Index Terms**— Change Detection, Kullback-Leibler (KL) divergence, Multivariate Gaussian distribution, Multi-temporal synthetic aperture radar (SAR) images, Wavelet Transform.

## 1. INTRODUCTION

The multi-temporal SAR image change detection becomes very important for earth monitoring, earth observation, damage assessment and land cover dynamics. Change detection is a process that analyzes multitemporal remote sensing images acquired on the same geographical area for identifying changes occurred between the considered acquisition dates [1]. The result is a generation of a change detection map in which changed area are explicitly identified. In the literature, many methods are proposed to deal with the multi-temporal SAR change detection. These methods are classified into two categories according to the data: bi-temporal change detection and image time series change detection [2]. We consider here the case of bi-temporal change detection for study. In addition, bi-temporal change detection can be performed in supervised [3] or unsupervised way [4]. In supervised change detection, available training samples are selected and used to train a classifier. The latter will be used to classify each pixel as changed or unchanged pixel. In unsupervised change detection, the first step is to compare some features of the two images by some similarity metrics resulting in a change map and then threshold the change map to produce two classes associated with changed and unchanged pixels.

In the literature, several unsupervised change detection approaches have been proposed and can be classified in two classes: methods based on pixel intensity and methods based on local statistics [5]. The first group is based on the pixel intensity and the neighbouring of the pixel. They include image differencing, mean ratio/log-ratio measures [6, 7], Gauss log-ratio [8], etc. One widely used technique for SAR image change detection is the ratio of the local means in the neighborhood of each pair of co-located pixels

[9]. This ratio is robust but it is limited to the comparison of the first order statistics and if the change preserve the mean value but modifies the local textures, it will not be detected. While, the first group of measures is limited to the comparison of first order statistics, the second group uses the local statistics estimated by considering some local probability density functions (pdfs) of the neighborhood of homologous pixels of the pair of images used for the change detection. These distributions have been chosen particularly to model adequately the statistics of SAR images.

The estimation of pdfs can be made with different approaches: parametric and non-parametric approaches. For non-parametric estimation of pdfs, one can use the kernel density estimation (KDE) [10]. The authors in [9] use one dimensional Edgeworth series expansion to estimate the local probability density functions (pdfs) of the neighborhood of homologous pixels of the pair of images used for the change detection. For the parametric estimation of pdfs, one can pick up one distribution from a family of probability distributions called Pearson system and estimate its parameters. The Gaussian distribution is one of the distribution found in the Pearson system. It gives a quite good approximation of the probability distributions of a small region [9] and when some Gaussianity are introduced into the data when the images were resampled and filtered during the pre-processing step. When the SAR intensity statistics are not normally distributed, the radar intensity follows a Gamma distribution in the absence of texture. In the presence of texture, the intensity distribution follows a K-distribution. Once the parameters of the chosen distribution are estimated, their comparison can be performed using different criteria and the most usual one is the Kullback-Leibler divergence [9, 11].

Until now, all the pdfs used for local statistical model are univariate distributions. To the best of our knowledge, there aren't any papers dealing with a multivariate distribution of a random vector modeling the image for SAR image change detection. The idea is to model the spatial interaction and information found in the sub-image by multivariate Gaussian distribution and to exploit this to compute the Kullback-Leibler divergence and then generate a change detection map. Furthermore, we think that implementation of our approach in a decomposition scheme using wavelet transformation can enhance SAR change detection. Indeed, Texture can be easily represented and discriminated in wavelet domain. Wavelet transform is used to decompose the image into multiple scales. Probability density function of each sliding windows of the coefficient magnitudes of each subband is assumed to be multivariate Gaussian distribution. The paper is organized as follows. In section II, we introduce some statistical models used for our study. There are univariate and multivariate Gaussian distribution. Change detection using Kullback-Leibler divergence are presented in section III, where the former is calculated for three cases: one-dimensional Gaussian distribution (1D GD),  $k$ -dimensional multivariate Gaussian distribution

( $k$ D MGD) and  $k$ D MGD in wavelets domain. Real data used for evaluation and experimental setting are described in section IV. Finally, the discussion, and some concluding remarks close up this paper.

## 2. STATISTICAL MODELING

In this section, two statistical models univariate Gaussian distribution referred as 1D GD and Multivariate Gaussian distribution referred as  $k$ D MGD are used. As mentioned before, the sliding window manner is applied for change detection and the parameter estimation methods is applied in each sliding window.

### 2.1. 1D

In the community of radar image processing, a wide of statistical model distributions are used to characterize SAR images: Gamma distribution, generalized gamma distribution, K-distribution, etc. In our study, we simply use the Gaussian distribution which is given as follows

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R} \quad (1)$$

The Gaussian model may be justified where the real data were subject to transformations during the pre-processing step.

### 2.2. $k$ D

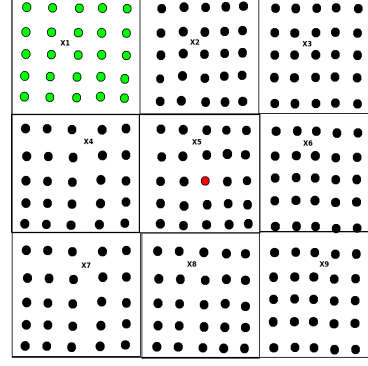
We model the sliding window or the sub-image as the realization of a random vector  $\mathbf{X} = (X_1, X_2, \dots, X_k)^t$ , where  $X_i$  are random variables. The joint density distribution function is given by  $f_{\mathbf{X}}(\mathbf{x})$ , where the vector  $\mathbf{x} = (x_1, x_2, \dots, x_k)^t$  is the realization of the random vector  $\mathbf{X}$ . The random  $k$ -vector  $\mathbf{X}$  has the  $k$ -variate Gaussian distribution with mean  $k$ -vector  $\boldsymbol{\mu}$  and positive-definite, symmetric ( $k * k$ ) covariance matrix  $\boldsymbol{\Sigma}$  and the density function is given by the

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{k}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}(\mathbf{x} - \boldsymbol{\mu})\right), \quad \mathbf{x} \in \mathbb{R}^k \quad (2)$$

Fig.1 shows an example of sliding window organized in  $3*3$  blocks each one is constituted by  $n * n$  pixels (here  $5*5$ ). One block is a realization of vector component. The random vector  $\mathbf{X}$  is composed here by  $k = 9$  components ( $X_i, i = 1..9$ )

### 2.3. $k$ D in wavelet domain

The previous models are performed in spatial domain. However, statistical models in wavelet domain are stable compared with models in spatial domain for SAR images. In addition, it helps to justify the use of MGD for modeling. After wavelet decomposition, we will have at the end  $3L + 1$  images where  $L$  is the number of scales. These images correspond to the horizontal detail, vertical detail, diagonal detail respectively at scale  $i$  and the approximation at scale  $L$ . Then, the sliding windows of each wavelet subbands coefficient are modeled using the  $k$ D multivariate Gaussian distribution. These sub-images are represented by  $\mathcal{X} = \{\mathbf{X}_{H_i}, \mathbf{X}_{V_i}, \mathbf{X}_{D_i}, \mathbf{X}_{A_L}\}$  where  $i \in \{1, \dots, L\}$  and  $\mathbf{X}_{H_i} = (X_{1,H_i}, X_{2,H_i}, \dots, X_{k,H_i})^t$ ,  $\mathbf{X}_{V_i} = (X_{1,V_i}, X_{2,V_i}, \dots, X_{k,V_i})^t$ ,  $\mathbf{X}_{D_i} = (X_{1,D_i}, X_{2,D_i}, \dots, X_{k,D_i})^t$ ,  $\mathbf{X}_{A_L} = (X_{1,A_L}, X_{2,A_L}, \dots, X_{k,A_L})^t$  are random  $k$ -vectors representing the sub-image horizontal detail, vertical detail and diagonal detail at scale  $i$  respectively, and finally the sub-image approximation at scale  $L$  and are distributed according  $k$ D MGD.



**Fig. 1.** An example of sliding window organized in  $3*3$  blocks each one is constituted by  $n * n$  pixels (here  $5*5$ ). One block is a realization of vector component. The random vector  $\mathbf{X}$  is composed here by  $k = 9$  components ( $X_i, i = 1..9$ )

## 3. CHANGE DETECTION BASED ON KULLBACK-LEIBLER

### 3.1. Kullback-Leibler divergence

To quantify a change detection between two acquisition dates we need to analyze the modification of the statistics of each pixel's neighborhood. Several approaches can be taken: the mean square error between the two distribution, the norm of a vector of moments, etc [9]. Symmetric Kullback-Leibler divergence is a good similarity indicator for change detection [2]. In our study, we choose to use the Kullback-Leibler divergence as a similarity measure. If the statistics of the two sliding windows are the same the symmetric Kullback-Leibler divergence is small. Let  $X^1$  and  $X^2$  be two random variables with probability density functions  $f_{X^1}$  and  $f_{X^2}$ . The Kullback-Leibler divergence from  $X^2$  to  $X^1$  is given by

$$KL(X^2 || X^1) = \int \log\left(\frac{f_{X^1}(x)}{f_{X^2}(x)}\right) f_{X^1}(x) dx, \quad (3)$$

The symmetric KL similarity measure between  $X^1$  and  $X^2$  is

$$KL(X^1, X^2) = KL(X^2 || X^1) + KL(X^1 || X^2). \quad (4)$$

In the case of 1D GD, if the  $X^1$  and  $X^2$  are distributed according to a Gaussian distribution with mean  $\mu_1$  and  $\mu_2$  and variance  $\sigma_1$  and  $\sigma_2$ , the symmetric version of the Kullback-Leibler divergence has the following form

$$KL(X^1, X^2) = \frac{\sigma_1^4 + \sigma_2^4 + (\mu_1 - \mu_2)^2(\sigma_1^2 + \sigma_2^2)}{2\sigma_1^2\sigma_2^2} \quad (5)$$

In the case of  $k$ D MGD, if  $\mathbf{X}^1$  and  $\mathbf{X}^2$  are two random  $k$ -vectors with joint density functions  $f_{\mathbf{X}^1}$  and  $f_{\mathbf{X}^2}$ , respectively, and are distributed according to the multivariate Gaussian distribution with  $k$ -dimensional mean vector  $\boldsymbol{\mu}_1$  and  $\boldsymbol{\mu}_2$  and  $k * k$  covariance matrix  $\boldsymbol{\Sigma}_1$  and  $\boldsymbol{\Sigma}_2$ , then the symmetric version of the Kullback-Leibler divergence has the following form

$$KL(\mathbf{X}^1, \mathbf{X}^2) = \frac{1}{2} [tr(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1) + tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2) - 2k + (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^t \boldsymbol{\Sigma}_1^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) + (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^t \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)] \quad (6)$$

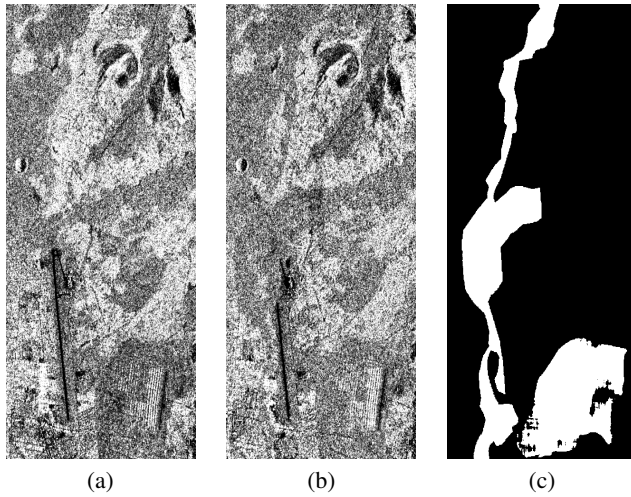
In the case of  $k$ D MGD in wavelet domain, the subbands are assumed independent and the total similarity of two blocks or two sliding widows are defined as the sum of similarity measures of each subband

$$KL(\mathcal{X}^1, \mathcal{X}^2) = KL(\mathbf{X}_{A_L}^1, \mathbf{X}_{A_L}^2) + \sum_{i=1}^L KL(\mathbf{X}_{H_i}^1, \mathbf{X}_{H_i}^2) + KL(\mathbf{X}_{D_i}^1, \mathbf{X}_{D_i}^2) + KL(\mathbf{X}_{V_i}^1, \mathbf{X}_{V_i}^2) \quad (7)$$

Where  $KL(\mathbf{X}_{H_i}^1, \mathbf{X}_{H_i}^2)$ ,  $KL(\mathbf{X}_{D_i}^1, \mathbf{X}_{D_i}^2)$ ,  $KL(\mathbf{X}_{V_i}^1, \mathbf{X}_{V_i}^2)$  and  $KL(\mathbf{X}_{A_L}^1, \mathbf{X}_{A_L}^2)$  are calculated using the Eq.(6)

#### 4. EXPERIMENTS WITH REAL DATA

To evaluate the proposed method, a pair of Radarsat images acquired before and after the eruption of the Nyiragongo volcano are used in this study. The volcano occurred in January 2002. Fig.(2) shows the two images before and after change and a binary change map produced using ground measures [9]. It is worth noticing that these images have undergone a series of image pre-processing such as filtering, re sampling, causing the modification of the local statistics of the image. The ground truth data in Fig.(2).c is not perfect due to the presence of a severe mis-registration caused by the lack of a proper digital terrain model [9]. To evaluate and to obtain reliable results,



**Fig. 2.** Data and ground truth for the Nyiragongo volcanic eruption of January 2002.

the input image is decomposed into  $L = 3$  scales using discrete stationary wavelet transform (SWT) with a Daubechies filter bank. For each coefficient magnitudes of each scale, different sliding windows with size (31, 37, 41, 47, 51, 59) are used in this study for performance evaluation. The two models both 1D GD and ( $k = 9$ )D MGD are estimated with different window size in spatial domain. The third method referred as SWT 9D is applied at each scale of the wavelet domain. To evaluate the accuracy of the change map independent of the thresholding algorithm, the receiver operating characteristic (ROC) curve is used and the area under ROC curve (AUC) is computed as a performance measure. The ROC curve is the evolution of the true positive rate (TPR) as function of false positive rate (FPR) [2]. The area under curve (AUC) is a good indicator of change. The larger the area the better the performance [2].

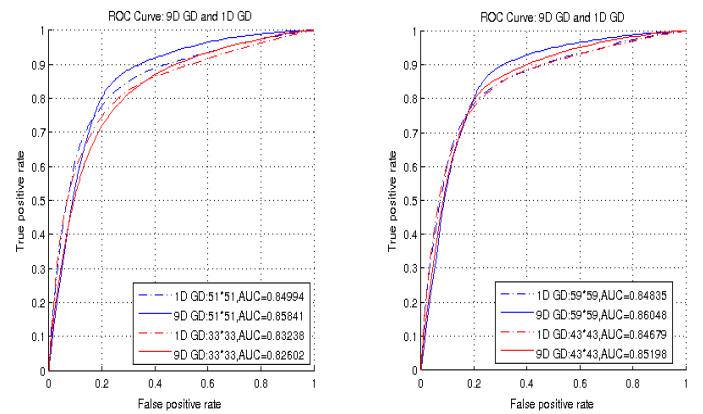
## 5. RESULTS

The area under the ROC curve are shown in Table 1. At each window size, the best value values are marked using red color. We can see clearly that the SWT9D is always the best for any window size. On the other hand, 1D gives the worst when the window size bigger than 33\*33. As the window size increases, the AUC increases until it reaches a maximum value and it decreases after. Based on this table, we conclude that the best window size for 1D, 9D, SWR1D and SWT9D are 51\*51, 59\*59, 51\*51 and 59\*59. For the sake of

Method	33	39	43	47	51	59
SWT1D	0.836	0.850	0.854	0.857	0.857	0.856
SWT9D	0.841	0.858	0.864	0.858	0.861	0.866
1D	0.832	0.842	0.846	0.849	0.849	0.848
9D	0.826	0.844	0.851	0.851	0.858	0.860

**Table 1.** The Area Under Curve (AUC) for different window size and different scales are measured for 1D, 9D, SWT1D and SWT9D. The best values are marked by red color and the worst by green color.

comparison, the ROC curves for 1D and 9D are presented in Fig.3 for different window size. We can see that the ROC Curve of 1D is above the ROC curve of 9D for low FPR values and is below for important FPR values. This behavior is the same for all the window size. It is worth noticing that the difference between the two curves is significant when the sliding window size increases especially for important FPR values. The same behavior is shown in Fig.4 comparing the ROC curve of SWT1D and SWT9D. Fig.5 shows a comparison between SWT9D and 9D. As we can see, the difference is minimal for important sliding windows 51\*51 and 59\*59. For the other window sizes, the ROC curve of SWT9D is always above the ROC curve of 9D.



**Fig. 3.** ROC curve comparison between 1D and 9D

## 6. DISCUSSION AND CONCLUSION

In this paper, an original method for SAR change detection in wavelet domain is proposed. It is based on multivariate Gaussian distribution for modeling the coefficient magnitude of the wavelet subbands. Wavelet transform is used to decompose the image into

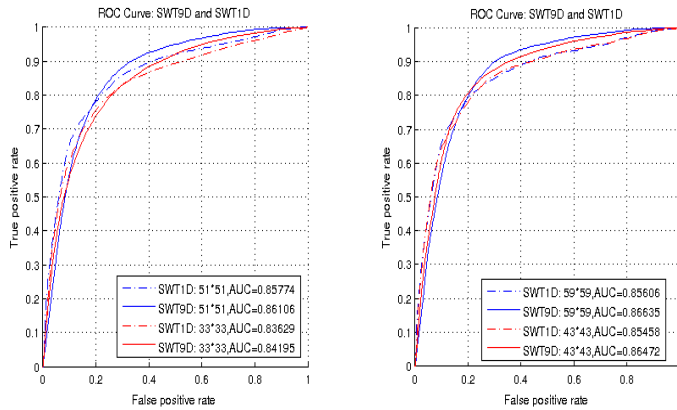


Fig. 4. ROC curve comparison between SWT1D and SWT9D

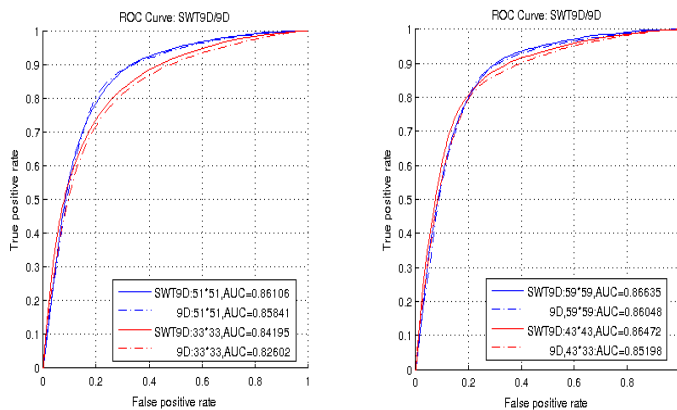


Fig. 5. ROC curve comparison between SWT9D and 9D

multiple scales. Probability density function of each sliding windows of the coefficient magnitudes of each subband is assumed to be multivariate Gaussian distribution. The Kullback-Leibler divergence between two corresponding subbands from the same scale is calculated. The total Kullback-Leibler divergence is the sum of the Kullback-Leibler of each subband. Our approach is evaluated using different windows size and compared with the univariate Gaussian distribution. Through the study, the multivariate Gaussian distribution in wavelet domain shows promising results comparing to the conventional approach as the univariate Gaussian distribution. Although the method realize good performance, improvement can be achieved by including other multivariate distributions as the multivariate Gamma distribution or multivariate K distribution.

## 7. REFERENCES

- [1] F. Bovolo and L. Bruzzone, "A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 6, pp. 1658–1670, June 2007.
- [2] Shiyong Cui and M. Datcu, "Statistical wavelet subband modeling for multi-temporal sar change detection," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 4, pp. 1095–1109, Aug 2012.
- [3] Chunlei Huo, Zhixin Zhou, Hanqing Lu, Chunhong Pan, and Keming Chen, "Fast object-level change detection for vhr images," *Geoscience and Remote Sensing Letters, IEEE*, vol. 7, no. 1, pp. 118–122, Jan 2010.
- [4] L. Bruzzone and D.F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 38, no. 3, pp. 1171–1182, May 2000.
- [5] S. Hachicha and F. Chaabane, "Multi-temporal sar change detection an monitoring," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, September 2011, pp. 293–298.
- [6] E.J.M. Rignot and J.J. van Zyl, "Change detection techniques for ers-1 sar data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 31, no. 4, pp. 896–906, Jul 1993.
- [7] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized gaussian model to automatic change detection in multitemporal sar images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 4, pp. 874–887, April 2005.
- [8] Biao Hou, Qian Wei, Yaoguo Zheng, and Shuang Wang, "Unsupervised change detection in sar image based on gauss-log ratio image fusion and compressed projection," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 7, no. 8, pp. 3297–3317, Aug 2014.
- [9] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal sar images and its extension to multiscale change analysis," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 5, pp. 1432–1445, May 2007.
- [10] Emanuel Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1065–1076, 09 1962.
- [11] A.M. Atto, E. Trouve, Y. Berthoumieu, and G. Mercier, "Multitide divergence matrices for the analysis of sar image time series," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 4, pp. 1922–1938, April 2013.