CHANGE DETECTION ON SAR IMAGES BY A PARAMETRIC ESTIMATION OF THE KL-DIVERGENCE BETWEEN GAUSSIAN MIXTURE MODELS

Qian Xu and Lina J. Karam

School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ

ABSTRACT

In the context of multi-temporal synthetic aperture radar (SAR) images for earth monitoring applications, one critical issue is the detection of changes occurring after a natural or anthropic disaster. In this paper, we propose a new similarity measure for automatic change detection using a pair of SAR images acquired at different dates. This measure is based on the evolution of the local statistics of the image between two dates. The local statistics are modeled as a Gaussian Mixture Model (GMM), which approximates the probability density function in the neighborhood of each pixel in the image. The degree of evolution of the local statistics is measured using the Kullback-Leibler (KL) divergence. One analytical expression for approximating the KL divergence between GMMs is given and is compared with the Monte Carlo sampling method. The proposed change detector is compared to the classical mean ratio detector and also to other recent model-based approaches. Tests on the real data show that our detector outperforms previously suggested methods in terms of the rate of missed detections and the total error rates.

***Index Terms*—change detection, Gaussian mixture models, Kullback-Leibler (KL) divergence, multitemporal synthetic aperture radar (SAR) images.**

I. INTRODUCTION

Detecting temporal changes occurring on the earth surface by observing them at different times is one of the most important applications of remote sensing technology. In the case of synthetic aperture radar (SAR) imagery, change detection has many applications, such as environmental monitoring, agricultural surveys, urban studies, and forest monitoring [1]. One critical issue in multitemporal SAR images is the detection of changes occurring after a natural or anthropic disaster. Since the changes produced by these events are abrupt and seldom predictable, they are often difficult to model, even for the same kind of change. For example, an earthquake can have different features depending on when or where it happens. Also, the changes of interest are all mixed up with normal changes if the time gap between the two acquisitions is too long [2].

Many change detection methods have been proposed to solve this problem. According to the data sources, the existing methods fall into two categories: bi-temporal change detection and image time series change detection [3], [4]. Furthermore, most bi-temporal change detection techniques can be classified into supervised [5] and unsupervised change detection methods [6]. In our paper, we only consider the unsupervised change detection process for two images acquired at different time. In general, most unsupervised change detection methods include three steps (Figure 1): 1) preprocessing, such as despeckling and image registration, 2) image comparison to generate a difference image, and 3) thresholding the difference image to compute the final binary change detection map [7]–[9]. In this paper, we choose to focus on the second step, where the objective is to find a good detector to measure the degree of the similarly of each pixel between two image data. For the thresholding method, the method in [7] is adopted to generator the final edge map.

For image comparison, several detectors have been proposed. The classical detectors include differencing and ratioing techniques [1], which are carried out by pixel-by-pixel comparison. Comparing with the difference operator, the ratio operator is more robust to illumination variation, speckle noise and calibration errors. However, the ratio operator, also known as the mean ratio detector introduced by Ulaby [10], assumes that the texture is a zero-mean multiplicative contribution. As a result, it cannot detect changes taking place at the texture level. In recent years, promising methods based on information measures have been proposed, where the local probability density functions (pdfs) of the neighborhood of pixels of the pair images are compared, instead of a pixel-by-pixel comparison. In [11], the Gaussian model has been used to approximate the local pdf. However, SAR-intensity statistics are not typically normally distributed. Therefore, the Pearson system, which is composed of eight types of distributions, and one-dimensional Edgeworth series expansion techniques were proposed to estimate local statistics in [2]. However, the computation complexity of these two methods is very expensive since they have to use fourth order statistics to estimate the parameters of the local pdf. The method proposed in [12] extends the information measures-based method to the wavelet domain by using generalized Gaussian and Gamma distributions to model the subband coefficient magnitudes.

Recently, Gaussian Mixture Models (GMM) have been widely used in image similarity measure and recognition [13] since they can approximate a variety of distributions, and only second-order statistical parameters are needed to be estimated to obtain the GMM. For SAR images, the GMM has also been used in [14] to model the difference image for thresholding. In our paper, the GMM is used to approximate the

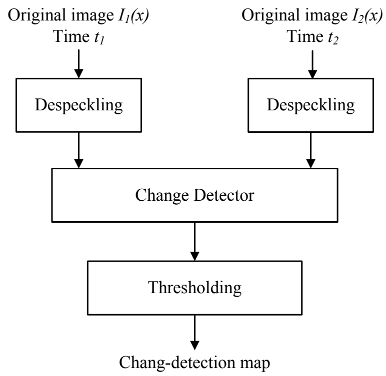


Fig. 1. General block diagram of the change-detection.

local statistical distribution of the SAR image. The degree of similarity between the local statistics of the pair images is measured using the Kullback Leibler (KL) divergence. Results from experiments that are conducted with real SAR data are shown to illustrate the performance of the proposed change detection method.

This paper is organized as follows. Section II presents the proposed algorithm, including the description of KL-divergence, GMMs and the methods to compute KL-divergence of GMMs. Section III presents the results on real data using the proposed detector. Finally, conclusions are drawn in Section IV.

II. ALGORITHM

Let us consider two coregistered SAR intensity images I_X and I_Y acquired over the same geographical area at two different times t_X and t_Y, respectively. Our aim is to generate a change detection map that represents changes that occurred on the ground between the acquisition dates. This change detection problem can be modeled as a binary classification problem where 1 represents change pixels and 0 represents unchanged pixels.

The proposed change-detection algorithm analyzes the difference of the local statistics of each pixel's neighborhood between the two acquired image data. A pixel will be considered as a changed pixel if its local statistical distribution changes from one to the other. In order to quantify this change, the KL divergence [15] between two probability density functions is used. In order to estimate the KL distance, the local pdf of each pixel in the two images needs to be estimated. In this paper, the local pdfs are approximated by the Gaussian Mixture Model. More details about the proposed algorithm are given as follow.

II-A. Kullback-Leibler divergence

The KL-divergence is used as a measure of similarity between two density distributions. Let f_x and g_y be two probability density functions of the random variables x and y, respectively. The KL-divergence between f_x and g_y, also known as the relative entropy, is given by [15]:

$$KL(f_x \| g_y) = \int f_x(x) log \frac{f_x(x)}{g_y(x)} dx.$$

It can be easily proved that $KL(f_x \| g_y) \neq KL(g_y \| f_x)$, but a symmetric version called the KL distance [2] may be defined as:

$$D(f_x, g_y) = D(g_y, f_x) = KL(f_x \| g_y) + KL(g_y \| f_x)$$

In order to compute the KL distance, the pdfs of two variables have to be known. As described in the following subsections, in this paper a Gaussian Mixture Model is used to estimate the local pdfs.

II-B. Image Modeling by GMM

A Gaussian mixture model (GMM) is a weighted sum of K component Gaussian distributions as given by: $p(\mathbf{x}) = \sum_{i=1}^{K} w_i \mathcal{N}(\mathbf{x}; \mu_i, \Sigma_i)$ where $\mathbf{x}$ is the measurement or feature vector, $w_i, i = 1 \ldots K$, are called mixing coefficients which must fulfill $\sum_{i=1}^{K} w_i = 1$, and $\mathcal{N}(\mathbf{x}; \mu_i, \Sigma_i)$ is a Gaussian density with mean vector μ_i and covariance matrix Σ_i. As a result, a GMM is parameterized by $\{w_i, \mu_i, \Sigma_i\}$. The problem of GMM parameter estimation is to estimate $\{w_i, \mu_i, \Sigma_i\}$ given training vectors. An efficient method for GMM parameter estimation is to use Expectation Maximization (EM) [16] which is an iterative refinement algorithm used for finding maximum likelihood estimates of parameters in probabilistic models.

Here, we propose the use of GMM to model the SAR image and use the EM algorithm to estimate the parameters of the GMM. In Figure 2, we compare the performance of Gaussian approximation and GMM approximation of the pdf of a window extracted from a considered SAR image. The top three figures show the fitting results for the SAR image before abrupt change. In this case, although the Gaussian approximation gives fairly good fitting performance, the GMM approximation fits the data much better. The bottom three figures show the results for the SAR image after abrupt change. In this case, there are more than one peak in the histogram, thus Gaussian approximation fails to fit the data. However, the GMM approximation still provides a good fitting performance as shown in Figure 2 (f). The Pearson system [2] is another widely used method in SAR image modeling and it can be used to represent 8 different types of distributions. Compared to the Pearson system, the GMM can be used to model any type of distributions which broaden its application range. As a result, we use a GMM to model the local probability distribution of SAR images.

II-C. KL divergence of the GMM

Using GMM, the normalized histogram of SAR images can be represented as follows: $f_x(x) = \sum_i \alpha_i \mathcal{N}(x; \mu_i, \Sigma_i)$; $g_y(x) = \sum_j \beta_j \mathcal{N}(x; \mu_j, \Sigma_j)$, where $f_x(x)$ is the normalized histogram before abrupt change and $g_y(x)$ is the normalized histogram after abrupt change. As a result, our objective is to find the similarity of these two GMM densities. For two

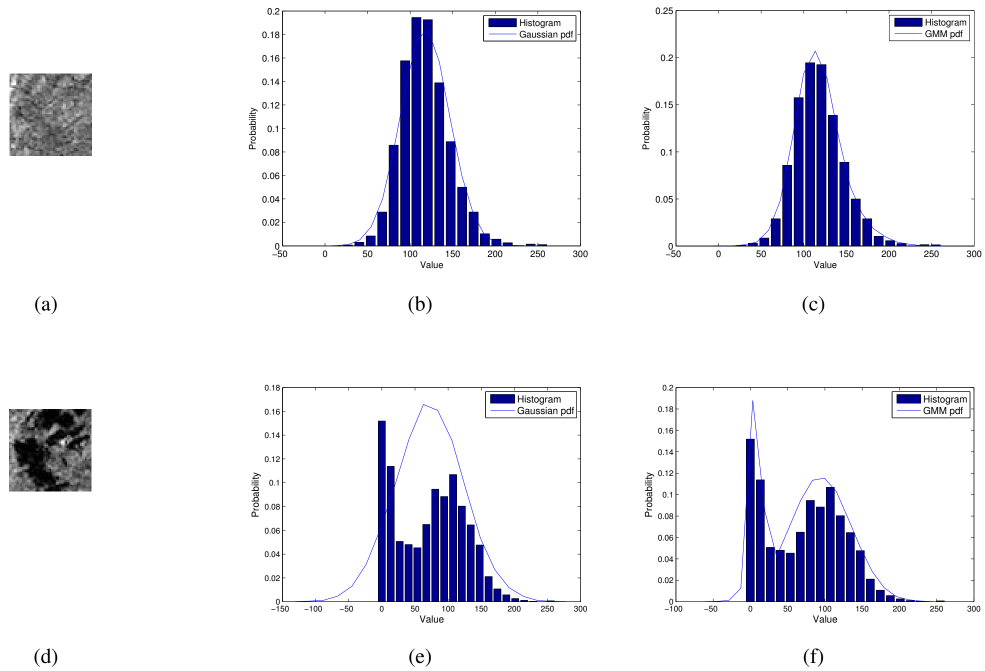


Fig. 2. Approximation of a histogram coming from a 50×50 window:(a) Window extracted from the SAR image before abrupt change. (b) Gaussian fitting of the histogram of (a). (c) GMM fitting of the histogram of (a). (d) Window extracted from the SAR image after abrupt change. (e) Gaussian fitting of the histogram of (d). (f) GMM fitting of the histogram of (d).

Gaussian densities $\hat{f}_x$ and $\hat{g}_y$, the KL-divergence has a closed form expression [17] given by:

$$\begin{aligned} KL(f_x \| g_y) &= \frac{1}{2}\left[log\frac{|\Sigma_{\hat{g}}|}{|\Sigma_{\hat{f}}|} + Tr[\Sigma_{\hat{g}}^{-1}\Sigma_{\hat{f}}]\right. \\ &\left. +(\mu_{\hat{f}} - \mu_{\hat{g}})^T \Sigma_{\hat{g}}^{-1}(\mu_{\hat{f}} - \mu_{\hat{g}})\right] \end{aligned} \quad (1)$$

However, there is no closed form expression for the KL-divergence between two GMMs. Monte Carlo simulation can estimate the KL-divergence with arbitrary accuracy [17]. Using Monte Carlo simulation, the KL-divergence of two GMM distributions $f(x)$ and $g(x)$ can be approximated as

$$KL_{MC}(f_x \| g_y) = \frac{1}{N}\sum_{i=1}^{N} log\, f_x(x_i)/g_y(x_i), \quad (2)$$

where $\{x_i\}_{i=1}^N$ are i.i.d samples drawn from the GMM $f_x(x)$. This can be achieved by first drawing a discrete sample a_i according to the mixing probability w_a, then drawing a continuous sample x_i from the resulting gaussian component $\mathcal{N}(x; \mu_{a_i}, \sigma_{a_i})$. As the number of samples $N \to \infty$, $KL_{MC}(f_x \| g_y) \to KL(f_x \| g_y)$. As a result, the KL distance is given by

$$D_{MC}(f_x, g_y) = KL_{MC}(f_x \| g_y) + KL_{MC}(g_y \| f_x). \quad (3)$$

The Monte Carlo (MC) method is a convergent method. However, the number of samples required for high accuracy approximation is very large which may cause a significant increase in computational complexity. The matching based approximation [13] can be used to reduce the computational complexity. The matching KL-divergence (MKL) approximation algorithm can be briefly described as follows:

1) *Components matching*: Define the matching function π which matches each components of $f_x(x)$ to the components of $g_y(x)$ as:

$$\pi(i) = argmin_j \left(KL(f_i \| g_j) - log\beta_j\right). \quad (4)$$

where $KL(f_i \| g_j)$ is the closed form KL-divergence between the ith Gaussian component of f_x and jth Gaussian component of g_y which can be calculated using (1).

2) *GMM KL Approximation*: Based on π, approximate the KL-divergence between two GMMs f_x and g_y as:

$$KL_{mat}(f_x \| g_y) = \sum_i \alpha_i\big(KL(f_i \| g_{\pi(i)}) + log\frac{\alpha_i}{\beta_{\pi(i)}}\big). \quad (5)$$

As a result, the closed form expression of the KL distance for the matching based method is given by

$$D_{mat}(f_x, g_y) = KL_{mat}(f_x \| g_y) + KL_{mat}(g_y \| f_x). \quad (6)$$

In our experiments, both MC and MKL methods are applied to approximate the KL-divergence between two SAR images. In order to compare the approximation performance, we use the Monte Carlo method with one million samples as the "ground truth". Table I shows the approximation results in terms of root mean square error (RMSE) which can be calculated as:

Table I. Comparison of KL approximation methods.

Algorithms	MC(100)	MC(1K)	MC(10K)	MKL
RMSE	0.21	0.13	0.06	0.25

$RMSE = \sqrt{\sum_{l=1}^{L}(\hat{KL} - KL_{true})^2}$, where L is the total number of pixels in the SAR image.

We can see from Table I that as the number of samples increase, the approximation performance of MC improves. However, this will also increase the computational complexity. The MKL approximation can reduce the computation complexity with minor performance decrease.

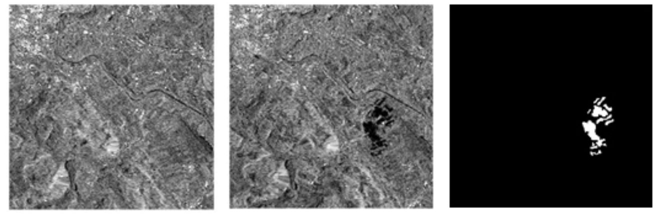
III. RESULTS WITH REAL DATA

In order to assess the effectiveness of the proposed approach, multi-temporal SAR dataset is considered. As shown in Figures 3 (a) and (b), this SAR dataset includes two images, which are two regions from two SAR images acquired by the European Remote Sensing 2 (ERS2) satellite SAR sensor over an area near the city of Bern, Switzerland, in April and May 1999, respectively [7]. Between the two acquisition dates, the river Aare flooded parts of the cities of Thun and Bern and the airport of Bern entirely. Therefore, the Aare valley between Bern and Thun was selected as a test site for detecting flooded areas. The available ground truth change between images in Figures 3 (a) and (b) is shown in Figure 3 (c).

The change-detection maps obtained by different change detectors are shown in Figure 4. The change detection map shown in Figure 4 (a) is obtained by using the mean-ratio detector (MDR) and the thresholding method in [7]; The change detection maps shown in Figures 4 (b) and (c) are obtained by using the Gaussian KL detector (GKLD) and the Pearson-based KL detector (PKLD) in [2], respectively. The result of our proposed GMMKLD method is shown in Figure 4 (d). From Figure 4, it is clear that PKLD and the proposed GMMKLD exhibit much better performance than the MRD and GKLD. In order to compare the Pearson-based KL detector and the proposed GGM-based detector, the false detections error, missed detection error and total error are measured by using the obtained binary change detection mask together with the ground truth change detection map. Table II summarizes those three errors for each detector. From Table II, it can be seen that the proposed detector produces the lowest total error rate of 0.35%.

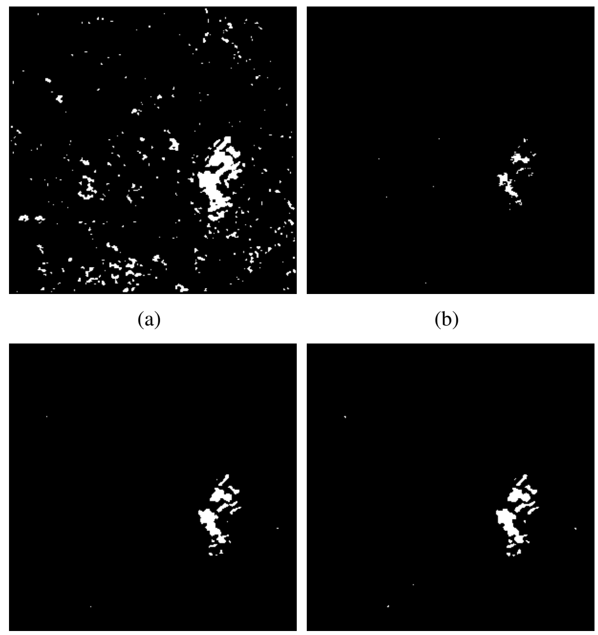
Table II. False detections, missed detections, and total errors (in number of pixels and percentage) resulting from different change detection algorithms on SAR image dataset.

Detector	False detections		Missed detections		Total errors	
	Pixels	%	Pixels	%	Pixels	%
MRD [7]	2683	3%	31	3%	2714	3%
GKLD [2]	7	0.008%	807	70%	814	0.9%
PKLD [2]	42	0.05%	315	27%	357	0.4%
GMMKLD	80	0.09%	243	21%	323	0.35%



(a) (b) (c)

Fig. 3. Multitemporal ERS2 SAR images used in the experiments: (a) image acquired in April 1999 before the flooding; (b) image acquired in May 1999 after the flooding; (c) the ground truth change map used as reference in the experiments.



(a) (b)

(c) (d)

Fig. 4. Change detection results from different algorithms on SAR image dataset: (a) MRD; (b) GKLD; (c) PKLD; (d) GMMKLD.

IV. CONCLUSION

In this paper, a novel change detector specifically is oriented to the analysis of multitemporal SAR images is proposed. This detector is based on an analysis of the local distribution of SAR images using a Gaussian Mixture Model and the KL distance. The Gaussian Mixture Model can approximate the local statistical distribution of SAR image very well, especially when modeling the local pdf where an abrupt change has occurred. Compared to existing detectors with higher order statistics, the proposed method has lower computational complexity with better change detection performance.

V. REFERENCES

[1] E. J. M. Rignot and J. J. van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 4, pp. 896 –906, July 1993.

[2] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432 –1445, May 2007.

[3] B. P. Salmon, J. C. Olivier, K. J. Wessels, W. Kleynhans, F. van den Bergh, and K. C. Steenkamp, "Unsupervised land cover change detection: Meaningful sequential time series analysis," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 4, no. 2, pp. 327C335, 2011.

[4] K. S. Chen, H. W.Wang, C. T.Wang, and W. Y. Chang, "A study of decadal coastal changes on western taiwan using a time series of ERS satellite SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 4, no. 4, pp. 826–835, 2011.

[5] L. Bruzzone and D. F. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection multitemporal remote-sensing images," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 452C466, Apr. 2002.

[6] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1171C1182, May 2000.

[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," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 874–887, Apr. 2005.

[8] T. Celik, "A bayesian approach to unsupervised multiscale change detection synthetic aperture radar images," *Signal Process.*, vol. 90, no. 5, pp. 1471–1485, May 2010.

[9] Y. Bazi, L. Bruzzone, and F. Melgani, "Image thresholding based on the EM algorithm and the generalized gaussian distribution," *Pattern Recogn.*, vol. 40, no. 2, pp. 619–634, 2007.

[10] F. T. Ulaby, F. Kouyate, B. Brisco, and T. H. L. Williams, "Textural information in SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. GRS-24, no. 2, pp. 235 –245, Mar. 1986.

[11] V. Frost, J. Stiles, K. Shanmugan, and J. Holtzman, "A model for radar images and its application to adaptive digital filtering of multiplicative noise," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, no. 2, pp. 157–165, Mar. 1982.

[12] S. Cui and M. Datcu, "Statistical wavelet subband modeling for multi-temporal SAR change detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 4, pp. 1095–1109, 2012.

[13] J. Goldberger, S. Gordon, and H. Greenspan, "An efficient image similarity measure based on approximations of KL-divergence between two gaussian mixtures," in *Proceedings of ICCV*, 2003, vol. 1, pp. 487 –493.

[14] T. Celik, "Bayesian change detection based on spatial sampling and Gaussian mixture model," *Pattern Recogn. Lett.*, vol. 32, no. 12, pp. 1635–1642, 2011.

[15] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Stat.*, vol. 22, no. 1, pp. 79–86, Mar. 1951.

[16] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1C38, 1977.

[17] J. R. Hershey and P. A. Olsen, "Approximating the kullback leibler divergence between gaussian mixture models," in *ICASSP*, 2007, vol. 4, pp. 317 –320.