A New Statistical Similarity Measure for Change Detection in Multitemporal SAR Images and Its Extension to Multiscale Change Analysis

Jordi Inglada and Grégoire Mercier, *Member, IEEE*

***Abstract*—In this paper, we present a new similarity measure for automatic change detection in multitemporal synthetic aperture radar images. This measure is based on the evolution of the local statistics of the image between two dates. The local statistics are estimated by using a cumulant-based series expansion, which approximates probability density functions in the neighborhood of each pixel in the image. The degree of evolution of the local statistics is measured using the Kullback–Leibler divergence. An analytical expression for this detector is given, allowing a simple computation which depends on the four first statistical moments of the pixels inside the analysis window only. The proposed change indicator is compared to the classical mean ratio detector and also to other model-based approaches. Tests on the simulated and real data show that our detector outperforms all the others. The fast computation of the proposed detector allows a multiscale approach in the change detection for operational use. The so-called multiscale change profile (MCP) is introduced to yield change information on a wide range of scales and to better characterize the appropriate scale. Two simple yet useful examples of applications show that the MCP allows the design of change indicators, which provide better results than a monoscale analysis.**

***Index Terms*—Change detection, Edgeworth series expansion, Kullback–Leibler (KL) divergence, multiscale change profile (MCP), multitemporal synthetic aperture radar (SAR) images.**

I. Introduction

Remote-sensing imagery is a precious tool for rapid-mapping applications. In this context, one of the main uses of remote sensing is the detection of changes occurring after a natural or anthropic disaster. Since they are abrupt and seldom predictable, these events cannot be well temporally sampled—in the Shannon sense—by the polar orbit satellites, which provide the medium, high, and very high resolution imagery needed for an accurate analysis of the land cover. Therefore, rapid mapping is often produced by detecting the changes between an acquisition after the event and available archive data.

This change-detection procedure is made difficult due to the time constraints imposed by the emergency context. Indeed, the first available acquisition after the event has to be used, whatever its modality, which is more likely to be a radar image, due to weather and daylight constraints.

The kind of changes produced by the event of interest are often difficult to model. The same kind of event—a flood—can have different signatures, depending on where it happens—high-density built-up areas, agricultural areas, etc.—and on the characteristics of the sensor. Also, the changes of interest are all mixed up with normal changes, which can be the majority if the time gap between the two acquisitions is too long.

All these issues present us with a very difficult problem: detecting abrupt unmodeled transitions in a temporal series with only two dates.[1]

From this position of the problem, one can make the straightforward deduction that pixelwise comparison between the two images will not be robust enough.

In the case of radar acquisitions, the standard detector is based on the ratio of local means [3]. This detector is robust to speckle noise, but it is limited to the comparison of first-order statistics. The classical model for synthetic aperture radar (SAR) intensity introduced by Ulaby *et al.* [4] assumes that the texture is a zero-mean multiplicative contribution. Therefore, changes taking place at the texture level, which preserve the mean value, will not be detected by the mean-ratio detector (MRD). One can, thus, assume a miss-detection behavior of the detectors using only the mean pixel values. This remark invites a more accurate analysis of the local statistics of the images to be compared. Bujor *et al.* [5] did a very interesting work by analyzing the interest of higher order statistics for change detection in SAR images. They concluded that the ratio of means was useful for step changes and that the second- and third-order log-cumulants were useful for progressive changes appearing in consecutive images in multitemporal series. Since higher order statistics seem to be helpful, one may want to compare the local probability density functions (pdfs) of the neighborhood of the homologous pixels of the pair of images used for the change detection.

Of course, this assumes that the pdfs are known and that there exists a robust way to compare them. The estimation of pdfs can be made with different approaches, but the straightforward histogram method should be avoided due to the need of a high number of samples for the estimation. Indeed, small analysis window sizes are required to yield high-resolution change maps. In this paper, we will present several approaches for this estimation by using only a small number of samples for the local statistics estimation, up to the fourth order.

Manuscript received February 16, 2006; revised December 1, 2006.

J. Inglada is with the Centre National d'Études Spatiales, 31401 Toulouse, France.

G. Mercier is with the École Nationale Supérieure des Télécommunications de Bretagne, F-29238 Brest, France.

Digital Object Identifier 10.1109/TGRS.2007.893568

[1] In the case where a sequence of several images is to be processed, the approaches presented in [1] and [2] may be applied.

Once the pdfs are estimated, their comparison can also be performed using different criteria. Information theory shows that a good measure is the Kullback–Leibler (KL) divergence, which is also called information gain. We will use a symmetrical version of this measure and show that it is superior to the classical detector when the pdfs are correctly estimated.

Therefore, these measures will be based on the comparison of local neighborhoods where an analysis window for the computation of the local estimation of probabilities is used. The problem, which arises here, is the one of choice of the window size. Since we are facing unmodeled changes, we cannot choose the window size to fit the size of the expected changes. An inappropriate window size can produce miss- and overdetections: 1) When using a small window for a correlation analysis, no detection will be performed in a homogeneous area, which was globally changed to another homogeneous area, and 2) on the contrary, when using a larger window size, change areas have to be of larger size or strong in intensity (relative to the measure) to be detected. In these cases, it will produce a coarse-resolution change map. One way to overcome this problem is by applying a multiscale change-detection analysis.

Scale is to be understood in its geographic meaning, which is the spatial extent of the study area. It does not refer to the cartographic meaning of scale (the larger the scale, the more detailed is the information [6]; for an interesting discussion on scale issues in remote sensing, see [7]).

Image-processing techniques for multiscale analysis often use the cartographic meaning and apply low-pass filtering and possibly subsampling. For change-detection analysis, this filtering and subsampling can be justified in the case where the images are not perfectly registered [8]. In other cases, we think that it is better to use all the available information, that is, maximizing the number of available samples by using increasing window sizes. Nevertheless, pyramidal multiscale decompositions can also be useful in the case of phenomena characterizations (see, for example, [9]).

Therefore, the main point of the problem is how to choose the largest window size that robustly detects the changes, but which is small enough to preserve the resolution of the final map without missdetections.

We propose to use multiscale change profiles (MCP), which are defined as the change indicator for each pixel in the image as a function of the analyzing window size. The computation of the change detection for each window size can be very time-consuming. We present here a method for the computation of these profiles, which allows the change indicator at scale n to be computed from the value obtained at scale $n-1$ plus a correction term which takes into account the addition of new samples only. Analytical expressions are given for three different change indicators. This paper proposes the following three main contributions:

1) an information-theory-based similarity measure which uses full local statistics;
2) the use of cumulant-based series expansions of similarity measures, which allow a robust and fast computation by using a small number of samples;
3) the concept of MCP and its fast implementation using recurrence evaluations.

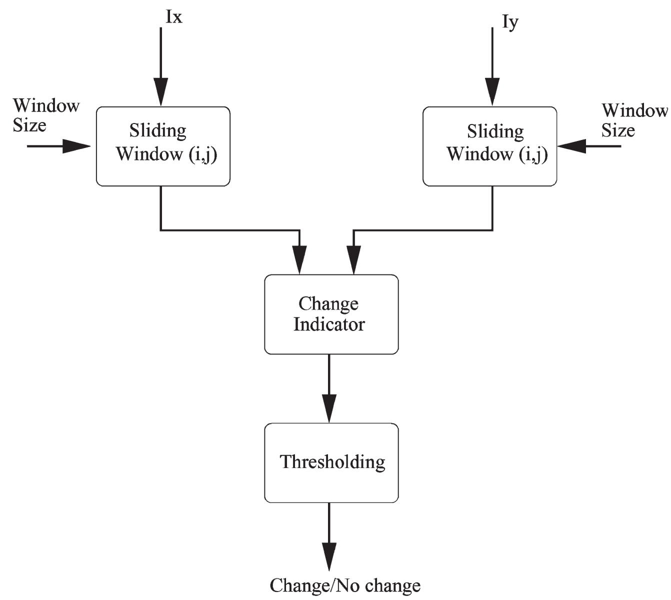


Fig. 1. Block diagram for a classical change-detection processing chain.

This paper is organized as follows. Section II presents the problem formulation; Section III introduces the measures used for the production of a change image; in Section IV, we introduce the concept of MCP and present the mathematical formulation, allowing its optimized computation; Sections V and VI present the results obtained on simulated and real data, respectively, and Section VII concludes this paper and proposes some directions for future work.

II. Problem Formulation

Let us consider two coregistered SAR intensity images I_X and I_Y acquired at two different dates t_X and t_Y, respectively. Our objective is to produce a map representing the changes occurring in the scene between t_X and t_Y. The final goal of a change-detection analysis is to produce a binary map corresponding to the two classes: change and no change. The problem can be decomposed into two steps: the generation of a change image and the thresholding of the change image in order to produce the binary change map. Fig. 1 shows a block diagram describing a classical change-detection processing chain.

The overall performance of the detection system will depend on both the quality of the change image and the quality of the thresholding. In this paper, we choose to focus on the generation of an indicator of change for each pixel in the image. For interesting approaches in the field of unsupervised change image thresholding, the reader can refer to the works of Bruzzone and Prieto [10], [11], Bruzzone and Serpico [12], and Bazi *et al.* [13]. The reader may note that some of these approaches need a statistical modeling of the detectors' response, which is not presented here.

The change indicator can also be useful by itself. Indeed, often the end user of a change map wants not only the binary information, given after thresholding, but also an indicator of the intensity of the change and, eventually, a confidence level. In order to evaluate the quality of a change image independently of the choice of the thresholding algorithm, the evolution of the detection probability $P_{\rm det}$ as a function of the false-alarm probability $P_{\rm fa}$ may be evaluated in the case where a set of constant thresholds is applied to the whole image. These are the so-called receiver operating characteristics (ROC), and the plots of $P_{\rm det}(P_{\rm fa})$ are called the ROC plots.

III. Distance Between Probability Densities

The main difficulty in the multitemporal analysis of SAR images is the presence of speckle noise. When moving away from interferometric configurations, the speckle is different from one image to the other, and it can induce a high number of false alarms in the change-detection procedure. Because of the multiplicative nature of speckle, the classical approach in SAR remote sensing involves using the ratio of the local means in the neighborhood of each pair of colocated pixels. The MRD is usually implemented as the following normalized quantity:

$$r_{\rm MRD} = 1 - \min\left\{\frac{\mu_X}{\mu_Y}, \frac{\mu_Y}{\mu_X}\right\} \tag{1}$$

where μ_X and μ_Y stand for the local mean values of the images before and after the event of interest, respectively. The logarithm of (1) may also be used. Nevertheless, this operation does not modify the performance of the detector, in terms of ROC, even if the contrast of the image of change indicator is modified. However, the logarithm is used since it modifies the initial pdf of the image of change indicator and then facilitates the development of Bayesian thresholding approaches [13].

This detector assumes that a change in the scene will appear as a modification of the local mean value of the image. If the change preserves the mean value but modifies the local texture, it will not be detected.

The change-detection algorithm proposed in this paper extends the MRD by analyzing the modification of the statistics of each pixel's neighborhood between the two acquisition dates. A pixel will be considered as having changed if its statistical distribution changes from one image to the other. In order to quantify this change, a measure, which maps the two estimated statistical distributions (one for each date at a colocated area) into a scalar change index is required. Several approaches could be taken into consideration: the mean-square error between the two distributions, the norm of a vector of moments, etc. We have chosen to use a measure derived from the information theory called the KL divergence [14].

A. KL Divergence

Let P_X and P_Y be two probability laws of the random variables X and Y. The KL divergence from Y to X, in the case where these two laws have the densities f_X and f_Y, is given by

$$K(Y|X) = \int \log \frac{f_X(x)}{f_Y(x)} f_X(x) dx. \tag{2}$$

The measure $\log(f_X(x)/f_Y(x))$ can be thought of as the information on x for the discrimination between the hypothesis $\mathcal{H}_X$ and $\mathcal{H}_Y$, if hypothesis $\mathcal{H}_X$ is associated with the pdf $f_X(x)$ and $\mathcal{H}_Y$ with $f_Y(x)$. Therefore, the KL divergence $K(Y|X)$ can be understood as the mean information for the discrimination between $\mathcal{H}_X$ and $\mathcal{H}_Y$ per observation. This divergence appears to be an appropriate tool to detect changes when we consider that changes on the ground induce different shapes on the local pdf.

Since the KL divergence can be understood as the entropy of P_X relative to P_Y, it is also called information gain. It can easily be proven that $K(Y|X) \geqslant 0$; $K(Y|X)$ vanishes only when the two laws are identical. $K(Y|X)$ can be used as a measure of the divergence from P_Y to P_X. This measure is not symmetric as it stands: $K(Y|X) \neq K(X|Y)$, but a symmetric version may be defined by writing

$$D(X,Y) = D(Y,X) = K(Y|X) + K(X|Y) \tag{3}$$

that will be called the KL distance.

In order to estimate the KL distance, the pdfs of the two variables to be compared have to be known. As stated in the Introduction, the processing of high-resolution change maps requires analysis windows of small size, which makes impossible the use of local histogram estimations. In the following sections, we will introduce several approaches, which allow the estimation of the pdfs by using a limited number of samples only. This requires some *a priori* information on the data, which can be introduced by using the models of local statistics.

B. Gaussian KL Detector (GKLD)

As shown in Section III, the classical detector of (1) uses first-order statistics only. Yet, second-order statistics are often used for SAR-image processing. For instance, many speckle-reduction filters [15]–[17] are based on the contrast coefficient σ_X^2/μ_X^2, that is, the ratio between the variance and the square of the mean value. If the local statistics have to be compared up to the second order, the local random variables X and Y may be assumed to be normally distributed (i.e., of Gaussian law). Then, the pdf of P_X can be written as

$$f_X(x) = \mathcal{G}(x; \mu_X, \sigma_X) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}}. \tag{4}$$

An analogous expression holds for $f_Y(x)$.

Fig. 2(b) shows the Gaussian approximation of the probability distribution of a small region of interest (ROI) [Fig. 2(a)] extracted from a SAR image.

If this Gaussian model is used in (3), it yields the GKLD

$$r_{\rm GKLD} = \frac{\sigma_X^4 + \sigma_Y^4 + (\mu_X - \mu_Y)^2(\sigma_X^2 + \sigma_Y^2)}{2\sigma_X^2\sigma_Y^2} - 1. \tag{5}$$

It can be seen that even in the case of identical mean values, this detector is able to underline the shading of textures, which is linked to the local variance evolution.

Nevertheless, the reader should note that the Gaussian model should not be used, since SAR-intensity values are always positive. However, this example has been given as a simple case of a parametric model, which takes into account second-order statistics. Since some Gaussianity may be introduced into

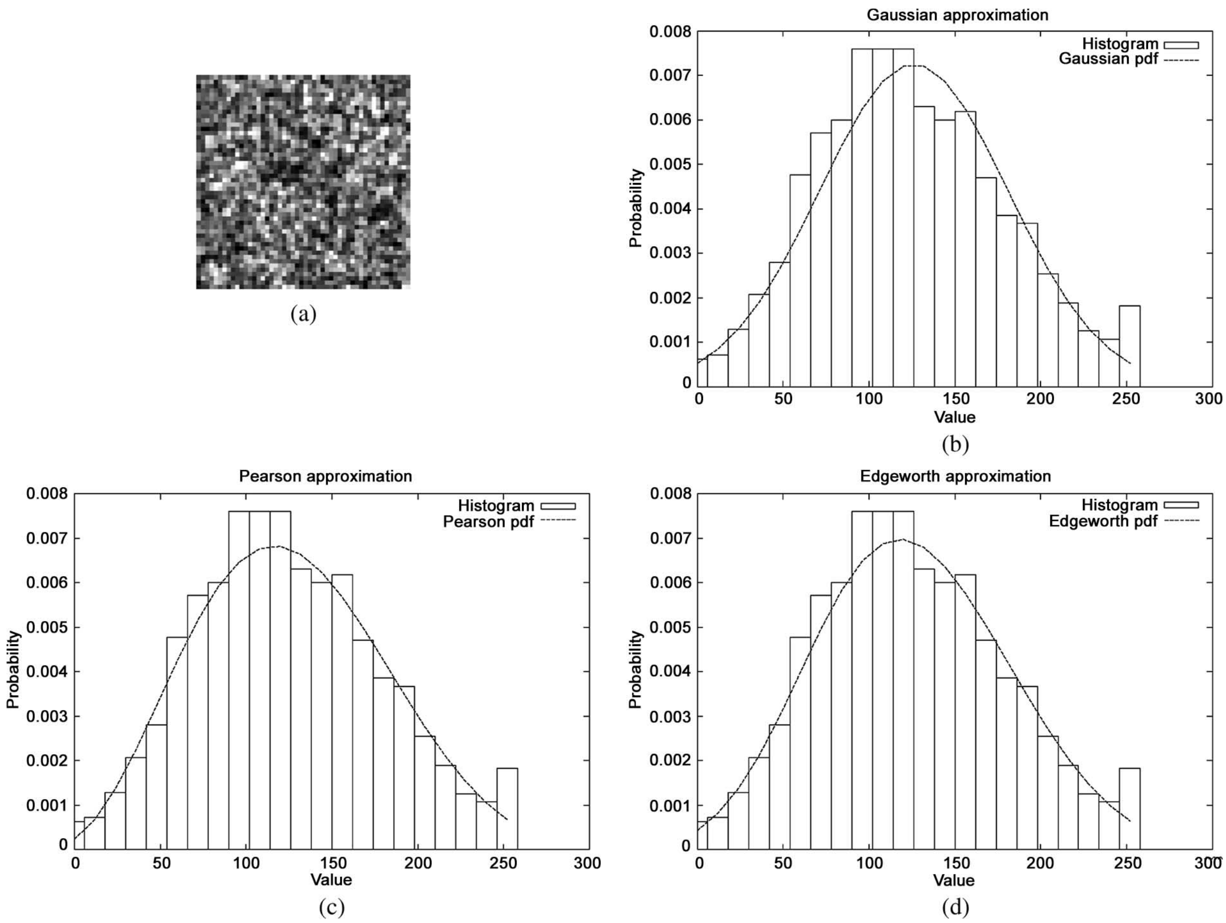


Fig. 2. Approximation of a histogram, coming from a 50 × 50 ROI, using three different strategies. The Pearson fitting yields a Beta distribution of the first type. (a) ROI extracted from a SAR image. (b) Gaussian fitting. (c) Pearson fitting: $\beta_1 = 2.51 \times 10^{-6}$ and $\beta_2 = 1.87$. (d) Edgeworth approximation.

the data when resampling and filtering the images during the preprocessing step, the Gaussian model may nevertheless be justified.

C. KLD Using the Pearson System

The drawback of the GKLD is that SAR-intensity statistics are not normally distributed, and the use of a bad model can induce bad performance of the detector, whatever the accuracy of the parameter estimation is. In the absence of texture, the radar intensity follows a Gamma distribution:

$$f_X(x) = \frac{1}{\Gamma(L)} \left(\frac{L}{\sigma_X}\right)^L e^{-\frac{Lx}{\sigma_X}} x^{L-1}. \tag{6}$$

The Gamma distribution is characterized by the following parameters: L is the number of looks, and σ_X is the square-root of the SAR-intensity image. $\Gamma(\cdot)$ is the Gamma function.

In the presence of texture, the local statistics can deviate from the Gamma distribution. For instance, if the texture is modeled by a Gamma distribution with a shape parameter ν, then the resulting intensity distribution follows a K-law [18]:

$$f_X(x) = \frac{2}{x}\left(\frac{L\nu x}{\mu_X}\right)^{L+\nu} \frac{1}{\Gamma(L)\Gamma(\nu)} K_{\nu-L}\left(2\left(\frac{L\nu x}{\mu_X}\right)^{1/2}\right) \tag{7}$$

where $K(\cdot)$ is the modified Bessel function of the second kind and μ_X is the mean of X.

More generally, it is now accepted that the statistics of SAR images can be well modeled by the family of probability distributions known as the Pearson system [19]. It is composed of eight types of distributions, among which the Gaussian and the Gamma distributions may be found. The Pearson system is very easy to use since the type of distribution can be inferred from the following parameters:

$$\beta_{X;1} = \frac{\mu_{X;3}^2}{\mu_{X;2}^3} \qquad \beta_{X;2} = \frac{\mu_{X;4}}{\mu_{X;2}^2}$$

where $\mu_{X;i}$ is the centered moment of order i of variable X. That means that any distribution from the Pearson system can be assessed from a given set of samples by computing the first four statistical moments. Any distribution, therefore, can be represented by a point on the $(\beta_{X;1}, \beta_{X;2})$ plane. For instance, the Gaussian distribution is located at $(\beta_{X;1}, \beta_{X;2}) = (0, 3)$, and the Gamma distributions lie on the $\beta_{X;2} = (3/2)\beta_{X;1} + 3$ line. Details about the theory of the Pearson system can be found in [20].

Fig. 2(c) shows an example of distribution estimation. The Pearson approximation fits the data better than the Gaussian one [Fig. 2(b)]. The example shown corresponds to a Beta distribution of the first type with parameters $\beta_1 = 2.51 \times 10^{-6}$, $\beta_2 = 1.87$.

The Pearson-based KLD (PKLD) was originally introduced in [21]. It does not have a unique analytic expression, since eight different types of distribution may be held. Therefore, 64 different possibilities for the couples of pdf exist. Once the couple of pdfs is identified, the detection can be performed by numerical integration

$$r_{\mathrm{PKLD}}(X,Y) = \int \Bigg[\log\left(\frac{f_X(x;\beta_{X;1},\beta_{X;2})}{f_Y(x;\beta_{Y;1},\beta_{Y;2})}\right) \times f_X(x;\beta_{X;1},\beta_{X;2}) + \log\left(\frac{f_Y(x;\beta_{Y;1},\beta_{Y;2})}{f_X(x;\beta_{X;1},\beta_{X;2})}\right) \times f_Y(x;\beta_{Y;1},\beta_{Y;2}) \Bigg] dx. \tag{8}$$

The correct way in proceeding to use the Pearson system is to choose a pdf using the estimated moments and, then, estimate the parameters of the distribution by maximum likelihood. While this can improve the results of the pdf estimation, the effect is not noticeable in terms of the estimation of the change indicator. This approach was not used in this paper in order to reduce the computation cost.

The reader should note that, in the case of single-look high-resolution SAR data (better than 10 m), other statistical models may be more appropriate, mainly on urban areas. Nicolas *et al.* have proposed a new model based on the log-statistics and a set of pdfs coming from the Fisher system of distributions [22], [23]. It has been applied to high-resolution SAR images on dense urban areas with promising results [24], [25].

D. Cumulant-Based KL Approximation

Instead of considering a parameterization of a given density, or set of densities, it may be of interest to describe the shape of the distribution. Such a description is based on quantitative terms that may approximate the pdf itself. The cumulants themselves do not provide such a pdf estimation directly but are necessary to describe its shape: For example, third-order (κ_3) is linked to the symmetry (i.e., skewness), while the fourth-order (κ_4) is linked to the flatness (i.e., kurtosis). The density is then estimated through a series expansion. In fact, the cumulant generating function is used for such an estimation. By definition, the cumulant generating function $\mathcal{K}_X(\cdot)$ of a random variable X is defined by

$$\mathcal{K}_X(\omega) = \ln \mathcal{M}_X(\omega) = \sum_n \kappa_{X;n} \frac{\omega^n}{n!}$$

where $\mathcal{M}_X(\cdot)$ is the moment-generating function defined by

$$\mathcal{M}_X(\omega) = \int e^{\omega x} f_X(x) dx = \int \left(1 + \omega x + \frac{\omega^2}{2} x^2 + \cdots + \right) f_X(x) dx.$$

For the case of the four first-order cumulants, the following expressions hold [26, p. 8]:

$$\begin{aligned}
\kappa_{X;1} &= \mu_{X;1} \\
\kappa_{X;2} &= \mu_{X;2} - \mu_{X;1}^2 \\
\kappa_{X;3} &= \mu_{X;3} - 3\mu_{X;2}\mu_{X;1} + 2\mu_{X;1}^3 \\
\kappa_{X;4} &= \mu_{X;4} - 4\mu_{X;3}\mu_{X;1} - 3\mu_{X;2}^2 + 12\mu_{X;2}\mu_{X;1}^2 - 6\mu_{X;1}^4.
\end{aligned} \tag{9}$$

Let us assume that the density to be approximated is not too far [27] from a Gaussian pdf (denoted as $\mathcal{G}_X$ to underline the fact that it has the same mean and variance as X), that is, with a shape similar to the Gaussian distribution. The difference between $\mathcal{K}_X(\cdot)$ and $\mathcal{K}_{\mathcal{G}_X}(\cdot)$ can be written in terms of the difference of the cumulants $\kappa_{X;n} - \kappa_{\mathcal{G}_X;n}$. By inversion, the density may be expressed by a formal Taylor-like series

$$f_X(x) = \mathcal{G}_X(x) + c_1 \frac{d\mathcal{G}_X}{dx} + c_2 \frac{d^2\mathcal{G}_X}{dx^2} + \cdots + .$$

Since a Gaussian density is used, it yields

$$f_X(x) = \sum_{r=0}^{\infty} c_r H_r(x) \mathcal{G}_X(x)$$

where $H_r(x)$ is known as the Chebyshev–Hermite polynomial of order r [27]. When choosing a Gaussian law so that its first and second cumulants agree with those of X, the number of terms of the series expansion is greatly reduced. This is the so-called Edgeworth series expansion. Its expression, when truncated to order of six, is the following:

$$f_X(x) = \left(1 + \frac{\kappa_{X';3}}{6} H_3(x) + \frac{\kappa_{X';4}}{24} H_4(x) + \frac{\kappa_{X';5}}{120} H_5(x) + \frac{\kappa_{X';6} + 10\kappa_{X';3}^2}{720} H_6(x)\right) \mathcal{G}_X(x). \tag{10}$$

It can be thought of as a model of the form $X = X_{\mathcal{G}} + X'$, where $X_{\mathcal{G}}$ is a random variable with Gaussian density with the same mean and variance as X, and X' is a standardized version of X [28] with

$$X' = (X - \kappa_{X;1})\kappa_{X;2}^{-1/2}.$$

Fig. 2(d) shows an example of such an approximation of a histogram.

The Edgeworth series expansion of the two pdfs f_X and f_Y may be introduced into the KL divergence (2). It yields an approximation of the KL divergence by Edgeworth series, truncated at a given order. In [29], such an approximation has been truncated to order of four by using the equality $(f_X/f_Y) = (f_X/\mathcal{G}_X)\ (\mathcal{G}_X/\mathcal{G}_Y)\ (\mathcal{G}_Y/f_Y)$, where $\mathcal{G}_X$ (respectively, $\mathcal{G}_Y$)

is a Gaussian density of the same mean and variance as f_X (respectively, f_Y). Then

$$\begin{aligned}
&\mathrm{KL}_{\mathrm{Edgeworth}}(X,Y)\\
&= \frac{1}{12}\frac{\kappa_{X';3}^2}{\kappa_{X;2}^2}\\
&+ \frac{1}{2}\left(\log\frac{\kappa_{Y;2}}{\kappa_{X;2}} - 1 + \frac{1}{\kappa_{Y;2}}\left(\kappa_{X;1} - \kappa_{Y;1} + \kappa_{X;2}^{1/2}\right)^2\right)\\
&- \left(\kappa_{Y';3}\frac{a_1}{6} + \kappa_{Y';4}\frac{a_2}{24} + \kappa_{Y';3}^2\frac{a_3}{72}\right)\\
&- \frac{1}{2}\frac{\kappa_{Y';3}^2}{36}\left(c_6 - 6\frac{c_4}{\kappa_{X;2}} + 9\frac{c_2}{\kappa_{Y;2}^2}\right)\\
&- 10\frac{\kappa_{X';3}\kappa_{Y';3}(\kappa_{X;1} - \kappa_{Y;1})(\kappa_{X;2} - \kappa_{Y;2})}{\kappa_{Y;2}^6}
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
a_1 &= c_3 - 3\frac{\alpha}{\kappa_{Y;2}}\\
a_2 &= c_4 - 6\frac{c_2}{\kappa_{Y;2}} + \frac{3}{\kappa_{Y;2}^2}\\
a_3 &= c_6 - 15\frac{c_4}{\kappa_{Y;2}} + 45\frac{c_2}{\kappa_{Y;2}^2} - \frac{15}{\kappa_{Y;2}^3}\\
c_2 &= \alpha^2 + \beta^2\\
c_3 &= \alpha^3 + 3\alpha\beta^2\\
c_4 &= \alpha^4 + 6\alpha^2\beta^2 + 3\beta^4\\
c_6 &= \alpha^6 + 15\alpha^4\beta^2 + 45\alpha^2\beta^4 + 15\beta^6\\
\alpha &= \frac{\kappa_{X;1} - \kappa_{Y;1}}{\kappa_{Y;2}}\\
\beta &= \frac{\kappa_{X;2}^{1/2}}{\kappa_{Y;2}}.
\end{aligned}$$

Finally, the cumulant-based KLD (CKLD) between two observations X and Y is written as

$$r_{\mathrm{CKLD}} = \mathrm{KL}_{\mathrm{Edgeworth}}(X,Y) + \mathrm{KL}_{\mathrm{Edgeworth}}(Y,X). \tag{12}$$

The reader should note the fact that, like for the Pearson-based detector, despite the apparent complexity of the formulas, and owing to (9), only the moments up to the order of four have to be computed.

IV. Multiscale-Change Profile

Scale plays a strategic role in image analysis and more especially in change-detection applications. In Section I, it has been shown how an inappropriate scale of analysis can produce mis- or overdetections. Bovolo and Bruzzone [30] stress the fact that the scale of analysis is a key parameter for better discrimination between change and no-change areas. Such a point of view is implemented by a wavelet transform of the log-ratio estimated with a window of a user-defined size.

Instead of applying a multiscale analysis of the change image, the purpose here is to produce a set of change indicators estimated at various scales. We will call it MCP.

As stated in the Introduction, the multiscale term refers here to the size of the analyzing window. The MCP will therefore involve computing the change indicator for a pixel by using neighborhoods of increasing sizes. The so-called profile corresponds to the sequence of change measures as a function of scale. We will restrict our formulation to the case of the CKLD. Given the fact that this detector needs the estimation of the statistical moments of the samples inside the analyzing window, we are interested in finding an approach which avoids the computation from scratch of the moments at every scale.

A. *Optimized Computation of the MCP*

Let us consider the following problem: how to update the moments when an $(N+1)$th observation x_{N+1} is added to a set of N observations $\{x_1, x_2, \ldots, x_N\}$ already processed? When considering raw moments of order r, the formulation comes easily as

$$\tilde{\mu}_{r,[N+1]} = \frac{N}{N+1}\tilde{\mu}_{r,[N]} + \frac{1}{N+1}x_{N+1}^r.$$

$\tilde{\mu}_{r,[N]}$ (respectively, $\tilde{\mu}_{r,[N+1]}$) stands for the raw moment of order r estimated with N samples (respectively, $N+1$ samples). Since the analyzing window may contain textured areas, the mean value itself may be modified by the increase in the number of samples. Therefore, by using simple binomial properties, it can be shown that central moments may be characterized by

$$\begin{aligned}
\mu_{1,[N]} &= \frac{1}{N}s_{1,[N]}\\
\mu_{r,[N]} &= \frac{1}{N}\sum_{\ell=0}^{r}\binom{r}{\ell}(-\mu_{1,[N]})^{r-\ell}s_{\ell,[N]}
\end{aligned} \tag{13}$$

where the notation $s_{r,[N]} = \sum_{i=1}^{N} x_i^r$ has been used.

Hence, when considering a new sample x_{N+1}, each moment may be updated directly by using updates of $s_{1,[N+1]}$ and, then, $s_{r,[N+1]}$ for increasing values of order r. The Edgeworth series is also updated by transforming moments to cumulants [by using (9)] to be introduced in (10) and, then, in (11).

Fig. 3 shows an example of a pdf estimation on a homogeneous area [shown in Fig. 2(a)] when the window increases from 9 × 9 to 17 × 17. In fact, the availability of updating the estimation of the distance between distributions from windows of any size without reprocessing the overall data is the most interesting point for multiscale change-detection purposes. This online multiscale moment estimation is the key for the operational use of the MCP concept.

For example, the computation of r_{CKLD} with windows of size ranging from 5 × 5 pixels to 51 × 51 pixels (22 different window sizes) takes only 42% additional time with respect to the computation of a single detection with a window of median size of 29 × 29 pixels (300 versus 210 s for a 800 × 400 pixel image).

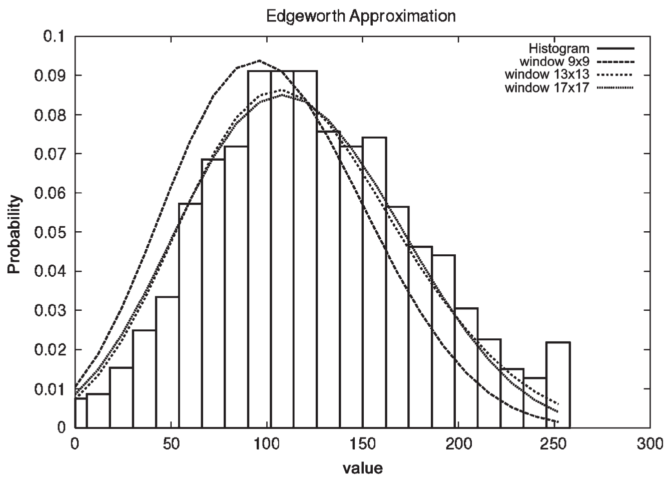


Fig. 3. Example of a pdf estimation update by increasing the sample set from a window of size 9×9 to 17×17. The histogram has been estimated with a 17×17 window.

B. MCP Exploitation

The MCP computation produces a multichannel image (one scale per channel) whose pixels have to be transformed into scalar values in order to provide a change indicator. In order to exploit the information available at all scales, two approaches may be investigated. The first one consists in choosing the best scale for each image pixel. The second one consists in fusing the information available at all scales in order to provide a single-change value.

The development of an optimal approach for the exploitation of the MCP may be application-dependent. Indeed, multiscale fusion approaches could be tuned to a particular type of change—shape, nature, etc. In this section, two simple, yet useful, choices will be proposed, which will yield an improvement in comparison to the performance of single scale detection: 1) In order to choose the best scale, we will choose the one which produces the highest KLD value. This assumes that this scale is the one that is associated with the largest window inside a homogeneous area with respect to the classes change and no change. 2) The fusion of the multiscale information will be performed by using the principal component analysis (PCA). The first principal component of the MCP multichannel image will be considered as the change indicator. This corresponds to a linear combination of all scales which maximizes the contrast of the final image.

V. Experiments With the Simulated Data

A. Data-Set Description

Simulations have been performed to better understand the behavior of the detectors relatively to a given kind of change and a given size of the change area. Since this study focuses on change detection on radar images, a speckle simulation is performed from a map of ground reflectivity. The simulated changes are applied on a small area, drawn as a circle, located in the center of the initial image.

The simulation procedure is based on the radar-image-formation mechanism. Each pixel is simulated with a given amplitude (coming from a SPOT near-infrared-band image, normalized to [0, 1]) and a thousand phases coming from independent uniform generations in $[0, 2\pi]$ to characterize elementary wave scatterers. Taking the square of the modulus of each pixel yields a one-look intensity image. A four-look intensity image is obtained by averaging and subsampling two adjacent pixels along lines and rows.

Each simulation of an change is applied to the initial image by using a change circle of a given size taken from $\{5, 10, 15, \text{ and } 20\}$. Once the speckle simulation is performed (independently from one image to another), the speckle-changed images are mosaicked on a 2×2 grid as shown on Fig. 4(b).

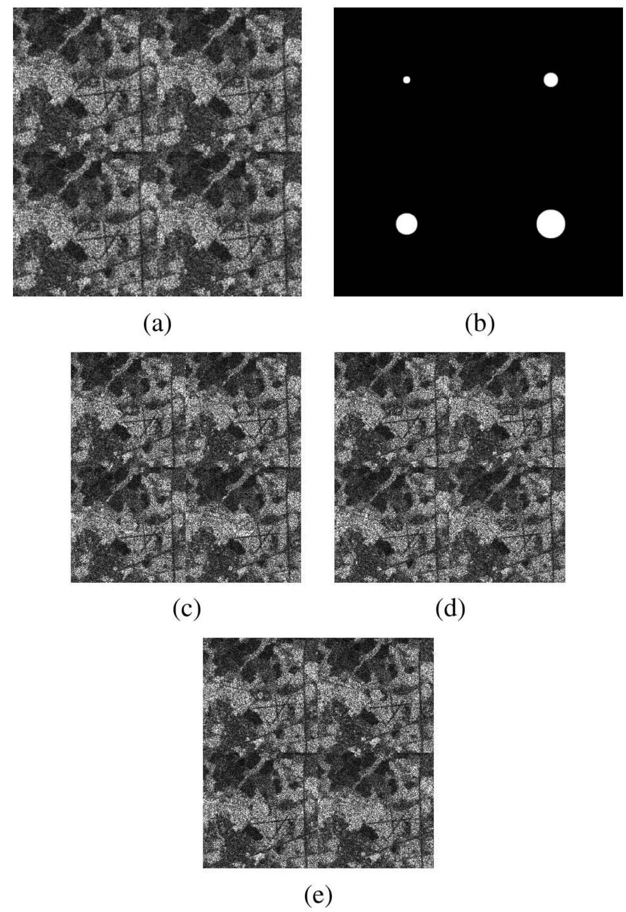


Fig. 4. Simulated data set. (a) Before. (b) Mask. (c) After offset. (d) After Gaussian. (e) After deterministic.

B. Simulation of Changes

Three kinds of change were considered.

1) Offset Change: The initial image is modified by applying an offset value (i.e., a shift) to the inital data [Fig. 4(c)]. This is a very simple type of change, which seldom occurs in reality, but is useful to characterize the behavior of the detectors.

2) Gaussian Change: The initial image is modified by applying a zero-mean Gaussian additive noise to the initial data [Fig. 4(d)]. This corresponds to a change in the state of the surface—field and vegetation. This is the main type of change that one can encounter in medium-resolution SAR images.

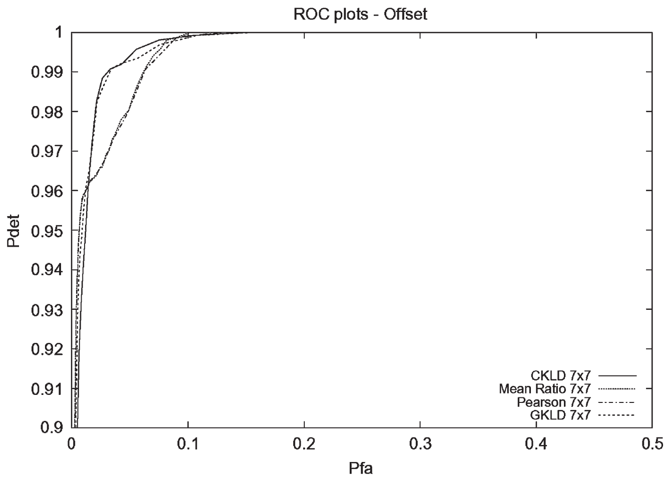


Fig. 5. ROC plot comparison of the four detectors for a simulated change consisting in an offset on reflectivity.

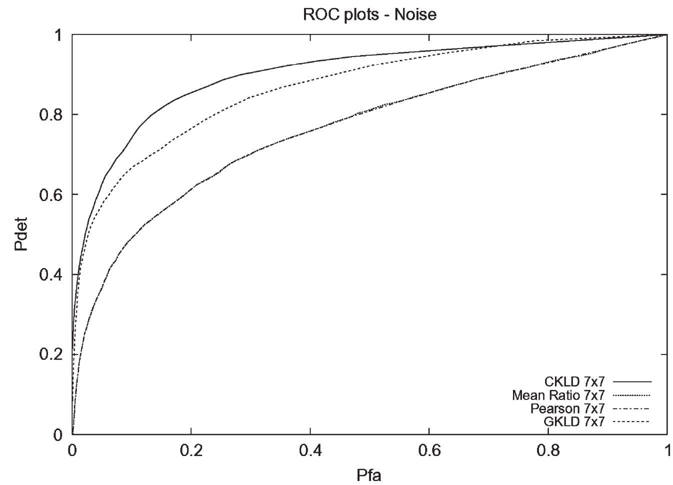


Fig. 6. ROC plot comparison of the four detectors for a simulated change consisting in a Gaussian random modification of the reflectivity.

3) Deterministic Change: The initial image is modified by pasting values copied from another area of the image itself [Fig. 4(e)]. This type of change can occur when there is a land-use change, anthropic activities, etc.

C. Results

1) Monoscale Detection: The results of the different detectors for a fixed analysis window size are analyzed.

Fig. 5 shows the ROC plots for the case where the change consists in a shift of the reflectivity value [Fig. 4(c)]. In this case, all four detectors are able to detect the changes with high accuracy. There is a slight difference in performance between the pairs CKLD–GKLD and PKLD–MRD, but it is difficult to infer the general behavior from this result. To draw a preliminary conclusion, for a simple change such as a reflectivity shift, the mean-value criterion is efficient enough for good discrimination in the changes, even on speckled images.

Fig. 6 shows the ROC plots in the case of a Gaussian change. The change is simulated by the addition of a Gaussian noise to the reflectivity (before speckle simulation). In this case, the mean value of the observed pixels remains approximately the same. It is difficult for this kind of change to be observed by a human operator. However, it is more likely to occur when the modifications affect the surface without changing its nature. In this case, even if all the detectors show bad performance in comparison to the offset case, the MRD and the PKLD are far below the GKLD and CKLD. The bad performance of the MRD is easy to understand, since the zero-mean Gaussian noise added to the reflectivity slightly changes the observed mean value. For the PKLD, it can be argued that the type of law in the Pearson system is not very different from the initial case, and the main difference is seen through the mean value, thus obtaining the same performance as the MRD. On the contrary, the GKLD assumes a simpler model than the PKLD and is able to take into account the mean and the variance modifications together. Finally, the ability of the CKLD to fit many different types of densities allows better detection for this difficult type of change.

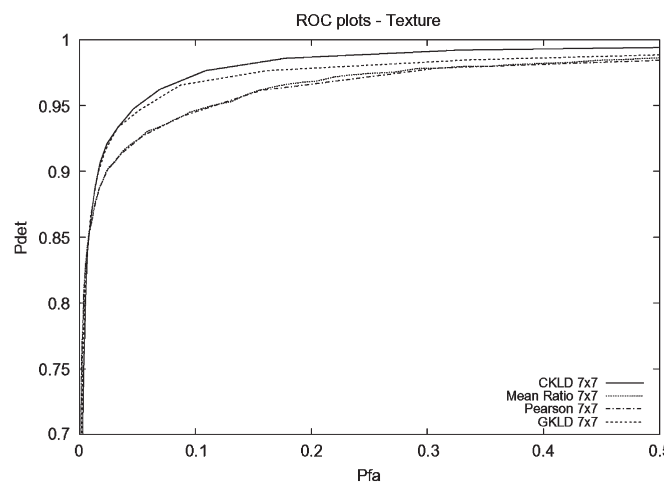


Fig. 7. ROC plot comparison of the four detectors for a simulated change consisting in a deterministic modification of the reflectivity.

The third type of change is that of a texture change, which can occur when there is a land-use change, anthropic activities, etc. In this case, as can be inferred from Fig. 7, the mean value of the regions may or may not change, and it is, therefore, interesting to analyze the shape of the density. The Pearson detector can be even worse than the MRD when the model does not fit the data, which is the case in presence of mixtures.

2) Analysis of the MCPs: Some collected MCPs, obtained by applying r_{CKLD} of (12) to our data set, are analyzed. Four different profiles are presented. They are extracted from a change area of the simulated data set for the case of a deterministic texture change and a radius of ten pixels. These profiles are labeled as follows: Far for the case where the analysis window is located 30 pixels from the center of the change area; Outside border for a distance of 15 pixels; Inside border for a distance of seven pixels; and Inside centered for a distance of zero pixels. Fig. 8(a) presents a diagram explaining how the profiles are extracted with respect to the change area, and Fig. 8(b) presents the profiles themselves.

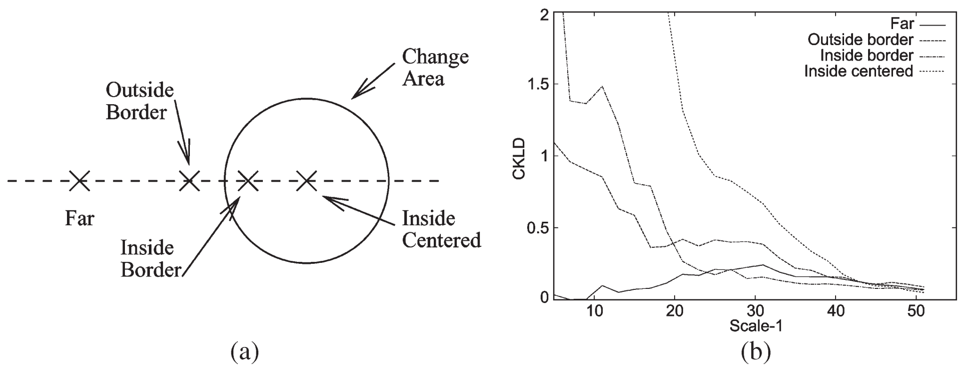


Fig. 8. Typical examples of MCPs obtained from the Edgeworth approximation of the KL distance. (a) Positions of the profiles. (b) Gaussian fitting.

The Far profile shows low values for small window sizes, and these values increase as the window size increases and it begins to include pixels from the change area. The values decrease for large window sizes, since the window stops including new change pixels while including no-change pixels present in all directions. The Outside border profile has a similar behavior, but the CKLD values are high for small scales since the pixel is nearer to the change area. The Inside border profile shows higher values for the change indicator for small window sizes. Finally, the Inside centered profile shows very high values of the detector for a large interval of window sizes. It is worth noting that the CKLD values are nearly the same for all detectors for the largest window sizes, since at this scale, all detectors include the same proportion of change and no-change pixels.

3) MCP Exploitation: In this section, the interest of the use of the MCP is illustrated with respect to the selection of a fixed scale of analysis (i.e., a fixed window size). The MCP allows the best scale to be selected for each pixel location in the images. Here, the maximum of the profile is used as a means to select the appropriate scale.

The maximum of the MCP and two different scales, 5×5 and 17×17, are compared. The small window size is used in order to detect small changes, but its main drawback is that the false alarms may increase in the presence of noise. The larger window size gives a lower false-alarm rate, since the noise is averaged and, therefore, its effect is reduced. But small changes can also be averaged, and therefore, the detection probability may be lowered. Also, false alarms may be increased in the neighborhood of the change areas.

The results of the comparison are presented in Figs. 9–11. As expected, small windows were able to give high detection rates. In the case of radiometric shift, the false-alarm rates are low for a given detection probability, since the type of change is easily detected by computing the mean value over a few pixels only. However, when more complex changes occur (Figs. 10 and 11), the false-alarm rate is very high at a given detection probability. Another interesting effect can be observed in Figs. 9 and 11, where for the large window sizes, the false-alarm rate increases without an increase in the detection probability. This is due to the fact that when the window is too large for the small changes and not as large as the larger changes [see the mask in Fig. 4(b)], the new detections induce false alarms only in the neighborhood of the small changes.

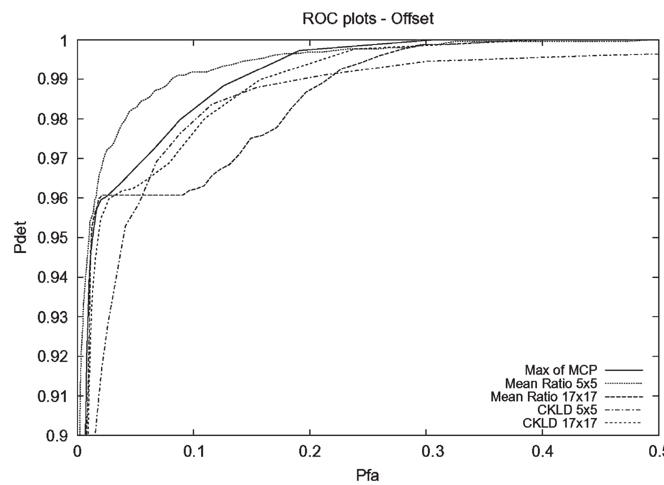


Fig. 9. ROC plot comparison between MRD—2 scales, CKLD—2 scales, and MCP—maximum of the profile, for a simulated change consisting in a reflectivity offset.

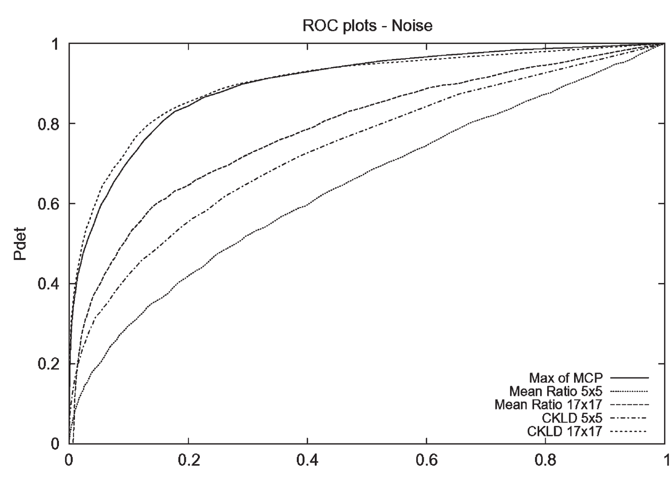


Fig. 10. ROC plot comparison between MRD—2 scales, CKLD—2 scales, and MCP—maximum of the profile, for a simulated change consisting in a Gaussian random modification of the reflectivity.

In addition, the MCP gives results which do not suffer from these drawbacks without the constraint of choosing a window size without prior information on the size of the changes in the images.

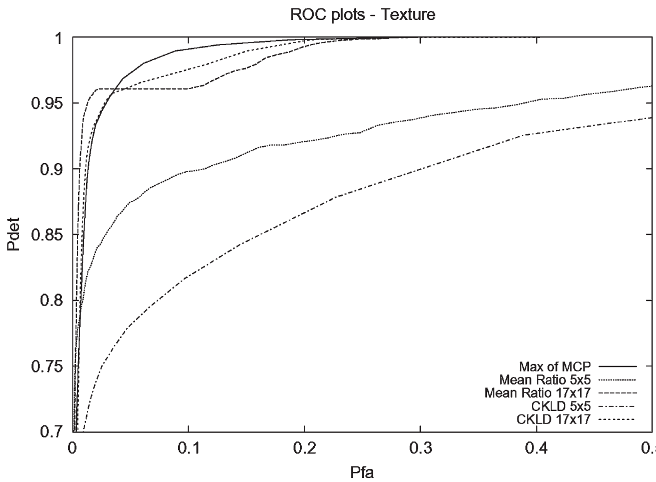
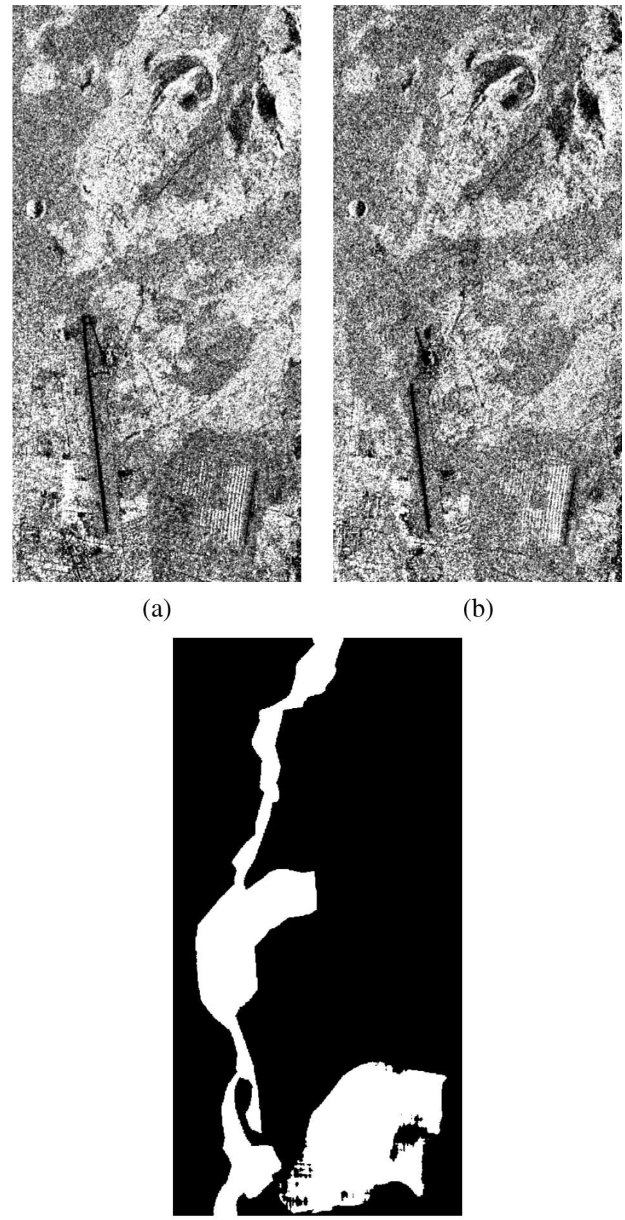


Fig. 11. ROC plot comparison between MRD—2 scales, CKLD—2 scales, and MCP—maximum of the profile, for a simulated change consisting in a deterministic modification of the reflectivity.



(a) (b) (c)

Fig. 12. Data and ground truth for the Nyiragongo volcanic eruption in January 2002. (a) Before. (b) After. (c) Mask.

VI. Experiments With Real Data

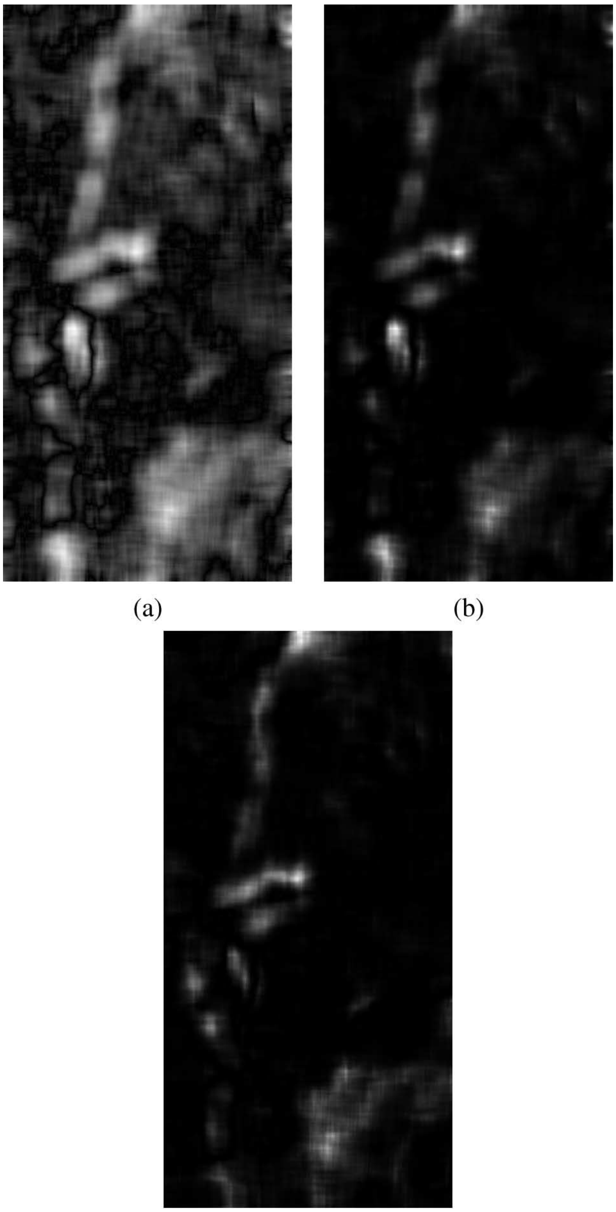
This section shows an example of applications of these algorithms to a real case. A pair of Radarsat images, acquired before and after the eruption of the Nyiragongo volcano (Democratic Republic of the Congo), which occurred in January 2002, were used. Fig. 12 shows the two images to be compared and a change map produced using the ground measures. The images have a ground resolution of 10 m and cover an area of 4×8 km. The images were orthorectified by IGN-F, the French National Geographic Institute, to a UTM35S projection, which was the same as the one used for the reference map. No filtering or calibration was applied to the data. The 16-bit to18-bit conversion was performed using a 3σ thresholding followed by a linear intensity rescaling. It is worth noting that the image resampling applied in the orthoregistration step modifies the local statistics of the image. Indeed, the image resampling implies the local-image interpolation, which is based on approximate interpolators. A bicubic interpolation was used in this case. This type of filter has a smoothing effect, which depends on the local shift [31]. Because of these radiometric artifacts introduced during the geometric preprocessing, the theoretical models for SAR statistics may not hold locally. The area at the bottom right-hand corner of the ground-truth mask corresponds to an area where a severe misregistration exists due to the lack of a proper digital terrain model. Finally, one has to take into account the fact that the ground truth is not perfect and that all results presented in this section should be analyzed rather in a relative manner—one detector with respect to another—rather than in an absolute one—absolute value of detection probabilities.

A. Change Indicator

The comparisons between the result coming from the classical image intensity ratio and the method proposed in this paper are shown in Fig. 13. Fig. 14 gives the ROC plots using the ground truth of Fig. 12(c). It shows that the use of KL approximation by the Edgeworth series outperforms any other methods such as model-based (Gaussian-based or Pearson-based) KL distance, or the ratio measure. As stated in the Introduction, a misdetection behavior of this detector can be observed, because the detector uses the mean pixel values only. It is interesting to underline the fact that the ratio criterion is not always worse than the pdf-based criteria. In fact, a density model has to fit the data in order to yield pertinent results.

For a detection probability below 0.3, it is more interesting to use the ratio criterion instead of a model-based one (by using Gaussian or Pearson assumption) in this example, even if a better change detection could have been expected by using Gaussian or Gamma laws coming from the local analysis of the two Radarsat images.

This point confirms that it is more interesting, for operational use, to consider a more flexible pdf approximation by using



(a) (b) (c)

Fig. 13. Change detection. Comparison between the different change indicators using the same window size (35 ×35 pixels). (a) Intensity ratio. (b) Pearson KL. (c) Cumulant-based KL.

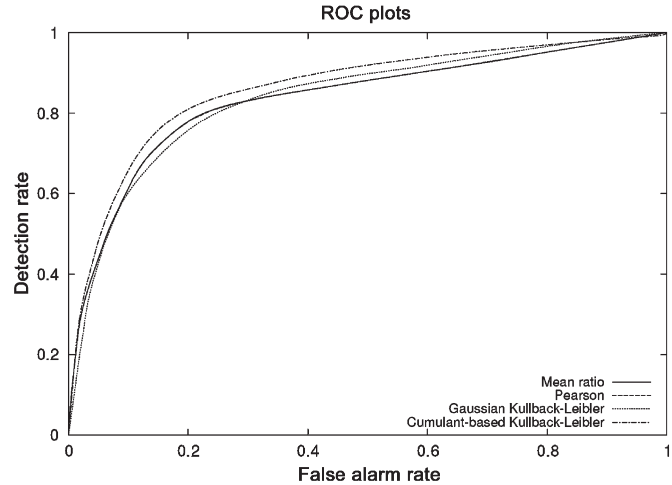


Fig. 14. ROC plots for the different detectors. The CKLD outperforms all other detectors. The Pearson-based detector gives results identical to the classical mean ratio. The Gaussian-based detector shows the worse behavior.

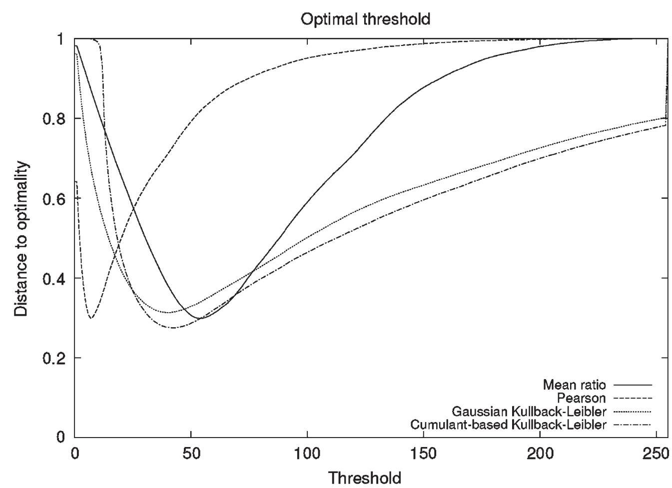


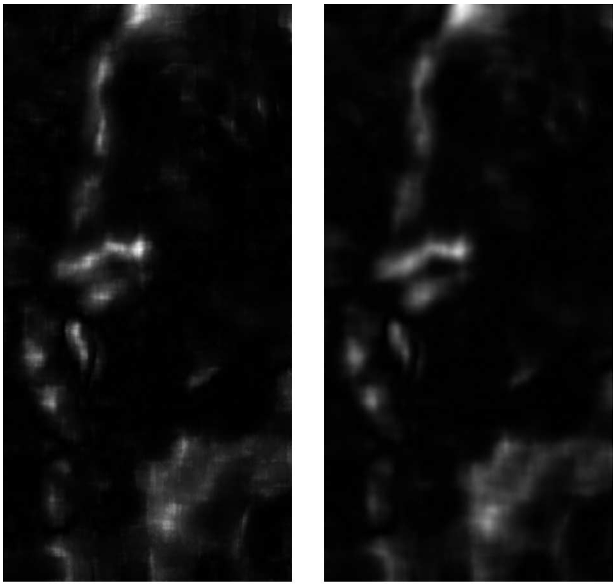
Fig. 15. Distance of ROC curves to the point $(P_d = 1, P_{\mathrm{fa}} = 0)$ for the different detectors. The Pearson detector allows trivial thresholding but is very sensitive. The cumulant-based threshold is less sensitive to threshold variations.

the Edgeworth series instead of a pdf parameterization. The cumulant-based approximation may give equivalent results to the Pearson-based approximation if the estimated cumulants correspond to a pdf belonging to the Pearson system of distributions, even though it may be less appropriate in the case of heavy-tailed distributions (single-look data). If cumulants of orders three and four vanish, the Edgeworth series is equivalent to a Gaussian model. If the variance of X and Y are equivalent, the Edgeworth series yields the same behavior as the ratio measure. However, when the local observations X and Y to be compared do not fit an *a priori* model, the Edgeworth series becomes a more suitable tool.

Fig. 15 draws the minimum distance of ROC curves to the point $(P_d = 1, P_{\mathrm{fa}} = 0)$. It is an interesting point of view to evaluate the threshold to be applied to obtain the best tradeoff between detection and false alarms. The best value of the threshold is to be found at the minimum of the curves.

Fig. 15 shows that this minimum is lower—and therefore, more interesting—for the Edgeworth series than for the Pearson measure or the ratio detector.

When no ground truth is available, the end-user has no *a priori* knowledge to set the value of the threshold. In this case, the Pearson measure seems to be better since a trivial value of zero could be used (i.e., pixels with values greater to zero may be considered as a change). Unfortunately, simulations and comparisons with other sets of images have shown that this trivial threshold is very sensitive to noise and fluctuations. The same observations about the sensitivity hold for the ratio measure. On the contrary, the cumulant-based measure takes its minimum for a wider range of values. Therefore, a threshold chosen *a priori* from the interval [40, 50] gives an almost optimal change map for all cases.



(a) (b)

Fig. 16. Change detection results obtained with the MCP. (a) Maximum of the MCP. (b) First PC of the MCP.

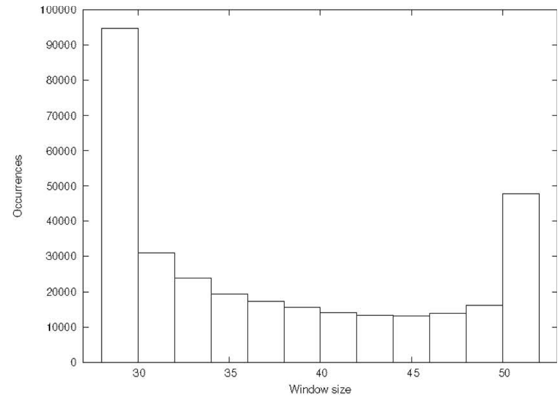
B. Multiscale Change Indicators

As stated in Section IV-B, our goal here is not to find the optimal way of exploiting the MCP but to show only the interest of the concept with simple examples. The results presented here use an MCP with window sizes ranging from 29 $\times$ 29 to 51 $\times$ 51.

In order to select the appropriate analysis window for each pixel in the image, we will choose the maximum of the MCP. The resulting change image is shown in Fig. 16(a). Fig. 17(a) presents the histogram of the sizes of the selected analysis windows when using the maximum of the MCP. It is interesting to observe that there is a high variability of window sizes, meaning that no trivial choice exists, like for instance choosing the largest window in order to increase the number of samples. Nevertheless, two peaks may be observed in the histogram. The first maximum gives the limit of the resolution of the detector and corresponds to areas near the borders of the change and no-change classes. The second one corresponds to homogeneous areas where the window size could continue increasing. Fig. 17(b) shows the map of the selected scales. The histogram bounds of Fig. 17(a) are linearly mapped to the minimum and maximum values of the image. It is interesting to note that large windows are used inside the change and no-change areas and that small window sizes are selected near the boundaries of these areas.

The ROC plots of Fig. 18 show that this simple strategy improves the results with respect to the case where the 35 $\times$ 35 window was used.

As an approach to multiscale fusion, we propose here to use the first principal component of the stack of multiscale detection images. The obtained change image is presented in Fig. 16(b). The ROC plot of Fig. 18 shows that this approach also provides better performance than the monoscale detector.



(a)

(b)

Fig. 17. Analysis of the selected scales using the maximum of the MCP. (a) Histogram of the window sizes. (b) Map of selected scales.

VII. Discussion and Conclusion

In this paper, a new similarity measure between images has been introduced in the context of multitemporal SAR image change detection. This measure is based on the use of the cumulant-based series expansion of the local-image statistics combined with the KL divergence. The concept of MCP has been developed, and a fast and efficient implementation has been proposed. Finally, two simple approaches for the production of the change images containing multiscale information have been presented. The first one is based on the selection of the scale which gives the highest change indicator, and the second one uses the first principal component of the multiscale change image stack. The proposed similarity measure has been compared to the classical ratio of local means and also to other KLDs, which use parametric models (Gaussian- or Pearson-based). The experiments have been carried out on simulated and real data for which a reference change map was available.

The proposed original cumulant-based detector has been shown to have a more robust behavior than other detectors in terms of ROCs. The two simple yet useful schemes for the

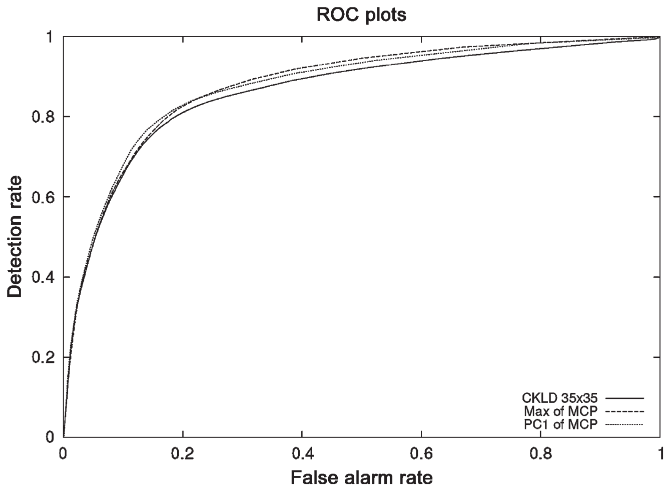


Fig. 18. ROC plots for two possibilities of MCP exploitation: The maximum and the first principal component. They outperform the CKLD for a fixed window size of 35 × 35.

exploitation of the MCP provide better performance than the monoscale detector.

The main advantages of the proposed approach are the following: Our detector needs only the computation of the first four statistical moments and can deal with a great variety of pdfs and the MCP provides change information over a wide range of scales at very low computation cost.

Some improvements could be done in order to use this approach with single-look images, where the heavy-tailed distributions may need other statistical models. The use of Gamma distributions instead of Gaussian for the series expansion seems to be a good starting point.

Some questions still remain open about the use of MCPs. Indeed, it would be interesting to analyze if we could establish a classification of the profiles and thereby derive useful information, not only about the scale of the change but also about its type. This task could be carried out by visual inspection, but automatic clustering techniques, like for instance the self-organizing map [32], could be used. The parametric modeling of the profiles by projection on an orthogonal basis could be envisaged.

Another issue remaining is the automatic thresholding of the change images. Whether it is for the case of a single scale or for the case of a multiscale analysis, the statistics of the change indicators could be used in order to propose adaptive Bayesian thresholding techniques, as done in [13].

Finally, direct classification of multiscale profiles by using support vector machines seems an appropriate choice for the production of binary-change maps in the case of supervised analyses. This approach has successfully been applied to the classification of hyperspectral images [33].

All these aspects will be studied in future work.

References

[1] L. Bruzzone, M. Marconcini, U. Wegmuller, and A. Wiesmann, "An advanced system for the automatic classification of multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1321–1334, Jun. 2004.

[2] P. Lombardo and T. Pellizzeri, "Maximum likelihood signal processing techniques to detect a step pattern of change in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 4, pp. 853–870, Apr. 2002.

[3] 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, Jul. 1993.

[4] 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.

[5] F. Bujor, E. Trouvé, E. Valet, J.-M. Nicolas, and J. Rudant, "Application of log-cumulants to the detection of spatiotemporal discontinuities in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2073–2084, Oct. 2004.

[6] N. Lam and D. Quattrochi, "On the issues of scale, resolution, and fractal analysis in the mapping sciences," *Prof. Geogr.*, vol. 44, no. 1, pp. 88–98, 1992.

[7] D. Marceau and G. Hay, "Remote sensing contributions to the scale issue," *Can. J. Remote Sens.*, vol. 25, no. 4, pp. 357–366, 1999.

[8] L. M. T. Carvalho, L. M. G. Fonseca, F. Murtagh, and J. Clevers, "Digital change detection with the aid of multiresolution wavelet analysis," *Int. J. Remote Sens.*, vol. 22, no. 18, pp. 3871–3876, Dec. 2001.

[9] G. Mercier, S. Derrode, W. Pieczynski, and J. L. Caillec, "Multiscale oil slick segmentation with Markov chain model," in *Proc. IGARSS*, Toulouse, France, Jul. 21–25, 2003, pp. 3501–3503.

[10] 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. 1171–1182, May 2000.

[11] ——, "An adaptive parcel-based technique for unsupervised change detection," *Int. J. Remote Sens.*, vol. 21, no. 4, pp. 817–822, 2000.

[12] L. Bruzzone and S. B. Serpico, "An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 858–867, Jul. 1997.

[13] 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.

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

[15] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 1, pp. 165–168, Mar. 1980.

[16] 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.

[17] D. Kuan, A. A. Sawchuk, T. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-7, no. 2, pp. 165–177, Mar. 1985.

[18] K. Ward, "Compound representation of high resolution sea clutter," *Electron. Lett.*, vol. 17, no. 16, pp. 561–565, 1981.

[19] Y. Delignon, R. Garello, and A. Hillion, "Statistical modelling of ocean SAR images," *Proc. Inst. Electr. Eng.—Radar Sonar Navig.*, vol. 144, no. 6, pp. 348–354, Dec. 1997.

[20] N. Johnson and S. Kotz, *Distributions in Statistics: Continuous Univariate Distributions*. New York: Wiley Interscience, 1969.

[21] J. Inglada, "Change detection on SAR images by using a parametric estimation of the Kullback-Leibler divergence," in *Proc. IGARSS*, Toulouse, France, Jul. 21–25, 2003, pp. 4104–4106.

[22] J.-M. Nicolas, "Introduction aux statistiques de deuxième espèce: Applications des log-moments et des log-cumulants à l'analyse des lois d'images radar," *Trait. Signal*, vol. 19, no. 3, pp. 139–167, 2002.

[23] J.-M. Nicolas, C. Tison, and F. Tupin, "Comparison of Beta1-Beta2 diagram and Logcumulant2-Logcumulant3 diagram for SAR image texture classification," in *Proc. EUSAR*, May 2004, pp. 133–136.

[24] C. Tison, J.-M. Nicolas, and F. Tupin, "Accuracy of Fisher distributions and log-cumulant estimation in a Markovian segmentation of high resolution SAR images over dense urban areas," in *Proc. IGARSS*, Jul. 2003, vol. 3, pp. 1999–2001.

[25] C. Tison, J. Nicolas, F. Tupin, and H. Maître, "A new statistical model of urban areas in high resolution SAR images for Markovian segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2046–2057, Oct. 2004.

[26] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework*. Englewoods Cliff, NJ: Prentice-Hall, 1993.

[27] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, 5th ed. London, U.K.: Arnold, 1991.

[28] P. McCullagh, *Tensor Methods in Statistics*. London, U.K.: Chapman & Hall, 1987.

[29] J. Lin, N. Saito, and R. Levine. (1999). "Edgeworth approximation of the Kullback-Leibler distance towards problems in image analysis," Univ. California, Davis. Tech. Rep. [Online]. Available: http://www.math.ucdavis.edu/˜saito

[30] F. Bovolo and L. Bruzzone, "A wavelet-based change-detection technique for multitemporal SAR images," in *Proc. Int. Workshop Anal. Multi-Temporal Remote Sens. Images*, May 16–18, 2005, pp. 85–89.

[31] J. Inglada, V. Muron, D. Pichard, and T. Feuvrier, "Analysis of artifacts in subpixel remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 254–264, Jan. 2007.

[32] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, no. 9, pp. 1464–1480, Sep. 1990.

[33] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote-sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

Jordi Inglada received the degree in telecommunications engineering from the Universitat Politècnica de Catalunya, Barcelona, Spain, and the École Nationale Supérieure des Télécommunications de Bretagne, Brest, France, in 1997 and the Ph.D. degree in signal processing and telecommunications from the Université de Rennes 1, Rennes, France, in 2000.

He is currently with the Centre National d'Études Spatiales, Toulouse, France, working in the field of remote-sensing image processing. He is in charge of the development of image-processing algorithms for the operational exploitation of Earth Observation images, mainly in the fields of image registration, change detection, and object recognition.

Grégoire Mercier (M'01) was born in France in the 1971. He received the degree in engineering from the Institut National des Télécommunications, Evry, France in 1993 and the Ph.D. degree from the University of Rennes, Rennes, France, in 1999.

Since 1999, he has been with the École Nationale Supérieure des Télécommunications de Bretagne, Brest, France, where he is currently an Associate Professor in the Image and Information Processing Department. His research interests are in remote-sensing image compression and segmentation, especially in hyperspectral and synthetic aperture radar. Specifically, his research is dedicated to change detection and combating pollution.