Unsupervised Change Detection From Multichannel SAR Data by Markovian Data Fusion

Gabriele Moser, *Member, IEEE*, and Sebastiano B. Serpico, *Fellow, IEEE*

***Abstract*—In applications related to environmental monitoring and disaster management, multichannel synthetic aperture radar (SAR) data present a great potential, owing both to their insensitivity to atmospheric and Sun-illumination conditions and to the improved discrimination capability they may provide as compared with single-channel SAR. However, exploiting this potential requires accurate and automatic techniques to generate change maps from (multichannel) SAR images acquired over the same geographic region in different polarizations or at different frequencies at different times. In this paper, a contextual unsupervised change-detection technique (based on a data-fusion approach) is proposed for two-date multichannel SAR images. Each SAR channel is modeled as a distinct information source, and a Markovian approach to data fusion is adopted. A Markov random field model is introduced that combines together the information conveyed by each SAR channel and the spatial contextual information concerning the correlation among neighboring pixels and formulated by using "energy functions." In order to address the task of the estimation of the model parameters, the expectation–maximization algorithm is combined with the recently proposed "method of log-cumulants." The proposed technique was experimentally validated with semisimulated multipolarization and multifrequency data and with real SIR-C/XSAR images.**

***Index Terms*—Change detection, data fusion, expectation–maximization (EM), Markov random fields (MRFs), multichannel SAR, synthetic aperture radar (SAR).**

I. Introduction

TECHNIQUES allowing one to detect the changes that occurred in a given area between different acquisition dates can represent efficient data-analysis tools in application-oriented contexts such as environmental monitoring and disaster management. Changed and unchanged areas can often be effectively discriminated by operating with optical multispectral data in the related multidimensional feature space [32], [60]. However, such data are affected by atmospheric and Sun-illumination conditions. On the other hand, synthetic aperture radar (SAR) [25], [26], [39], [49] is less sensitive to such issues, and multitemporal SAR imagery can be expected to play an important role, for instance, in ecological applications [29] or in disaster assessment and prevention [20], [59]. Unfortunately, SAR often allows only one amplitude/intensity observation, thus possibly resulting in poor discrimination between changed and unchanged areas. In this respect, multichannel (i.e., multipolarization and/or multifrequency) SAR presents a great potential: as compared with single-channel SAR, it is expected to provide an increased discrimination capability while maintaining the insensitivity to atmospheric and Sun-illumination conditions. The availability, granted by current and future missions (e.g., COSMO/SkyMed and TerraSAR), of multichannel SAR images obtained at (very) short revisit times further confirms this potential [8].

Due to the multiplicative nature of speckle, unsupervised change detection in SAR images is often addressed by image ratioing, i.e., by computing a ratio image obtained by dividing the pixel intensities in the image acquired at the first observation date by the ones in the image acquired at the second observation date (or vice versa) [49]. The ratio image is then analyzed in order to quantitatively distinguish between changed and unchanged areas. Interactive thresholding procedures are often applied toward this end [49].

Some efforts have recently been aimed at the automation of this thresholding process in the case of single-channel SAR. The unsupervised change-detection method proposed in [37] for optical multispectral data and based on the Kittler and Illingworth thresholding algorithm (K&I) is extended in [3] by introducing a generalized Gaussian model for a logarithmic ratio image and by integrating it with K&I. Similar approaches are developed in [2] and [4] by combining generalized Gaussian distributions with the expectation–maximization (EM) technique for parametric probability density function (pdf) estimation [53] and with a multiple-threshold version of K&I, respectively. In [43], three SAR-specific parametric models for the pdfs of SAR amplitude ratios are described, and an unsupervised change-detection method is proposed by combining image ratioing with K&I and with the recently proposed "method of log-cumulants" (MoLC) for parameter estimation [48], [63].

Other approaches have also been proposed in the literature to deal with single-channel SAR change detection, including multitemporal coherence analysis [54], integration of segmentation with multilayer perceptron and Kohonen neural networks [67], fuzzy-rule-based image analysis [33], multisource (optical and SAR) and multitemporal data fusion [11], spatio-temporal contextual classification [36], [38], [58], likelihood ratio tests [15], [35], Kullback–Leibler divergence and cumulant-based expansions [26], information-theoretic multitemporal feature extraction [1], multiscale wavelet analysis [9], split-based adaptive analysis of large multitemporal images [10], joint detection of spatial and temporal discontinuities by log-cumulants and fuzzy reasoning [12], fuzzy hidden Markov chains [13], fusion of feature-level and pixel-level analysis [22], signal subspace processing [52], and spatial chaotic modeling [66].

Manuscript received December 29, 2007; revised October 15, 2008. First published April 28, 2009; current version published June 19, 2009.

The authors are with the Department of Biophysical and Electronic Engineering, University of Genoa, 16145 Genova, Italy (e-mail: sebastiano.serpico@unige.it).

Digital Object Identifier 10.1109/TGRS.2009.2012407

However, when considering multichannel SAR images, change detection is still mostly an open issue, often addressed by using time-expensive "trial-and-error" procedures. In [18], image ratioing is applied to C- and L-band airborne SAR intensity images and compared with the use of the correlation coefficient and the phase difference between copolarized channels, with special attention on the detection of changes involving small objects and linear image features; thresholds to be applied to the ratio channels are selected according to a Neyman–Pearson-type approach [65] by manually choosing an acceptable value for the false-alarm probability. In [15], the problem of change detection in multilook fully polarimetric (complex) SAR images is formalized as a hypothesis testing problem, and case-specific likelihood ratio tests are derived according to the well-known Wishart distribution for polarimetric covariance matrices [49]. Polarimetric basis transformations are used in [51] to optimize the similarity of the scattering mechanisms characterizing the images acquired at different observation dates and to define quantitative measures of change in multitemporal polarimetric SAR images to be interactively analyzed to distinguish between changed and unchanged areas. "Contrast ratio" and "ellipticity" statistics are proposed in [30], as quantitative measures of change, computed from an eigenvalue analysis of the polarimetric covariance matrices of a single pixel at different observation times. Thresholds must be selected by trial and error for these measures of change as well as for the change statistics proposed in [15] and [51]. An experimental comparative analysis between these measures can be found in [42]. A change-detection method is proposed for polarimetric SAR in [41], using a content-based image retrieval system that is based on self-organizing maps and on polarimetric feature extraction. In [45], the EM algorithm with semilabeled samples developed by Landgrebe and Jackson (LJ-EM for short) [28] is combined with Markov random fields (MRFs) and with a SAR-specific version of the Fisher transform [21] to iteratively compute a scalar transformation of the multichannel ratio image that optimally discriminates between "change" and "no-change." The method is automatic, but its main drawback is that, even though a convergent behavior was experimentally observed, no theoretical proof of convergence is available yet.

This paper proposes an unsupervised automatic contextual change-detection method for multichannel amplitude SAR images, based on image ratioing and MRFs. In particular, the task of change detection is formalized in terms of an unsupervised classification problem. Each channel of the SAR amplitude ratio image is considered as a separate "information source," and an additional source is derived from the spatial context; the multisource data-fusion task is addressed together with the classification task using an MRF-based approach [62]. This formalization allows the joint exploitation of information conveyed by the ratios that correspond to all available SAR amplitude (or intensity) channels while avoiding the difficult task of modeling and accurately estimating the joint pdf of such ratio variables. Information about correlations among the ratio channels is disregarded in the proposed method since the adopted MRF model only involves the related marginal pdfs, whose unsupervised estimation can be carried out in an easier and more robust way. The resulting decision strategy is equivalent to the well-known "logarithmic opinion pool" (LOGP) consensus theory for multisource data fusion [5]–[7]. In order to estimate model parameters, a case-specific novel formulation of LJ-EM is developed and combined with MoLC. This choice also overcomes the convergence drawback of the approach proposed in [45] because of the robust analytical properties of EM-based estimation procedures. The primary novelties of this paper lie in formalizing the SAR multichannel change-detection problem as a Markovian data-fusion problem and in developing the EM-based algorithm to estimate the MRF model parameters.

This paper is organized as follows. The proposed technique is described from the methodological viewpoint in Section II. Experimental results on semisimulated data and on real SIR-C/XSAR images are then presented in Section III, and conclusions are drawn in Section IV. Appendix provides all analytical details and proofs concerning parameter-estimation issues involved in the proposed algorithm.

II. Proposed Method

A. Proposed MRF Model

Let us consider the detection of changes in a pair of multichannel amplitude images; if a pair of complex-valued polarimetric SAR images is available, the amplitude images may easily be computed. The change-detection problem is formulated as a binary hypothesis testing problem [65] by denoting the "change" and "no-change" hypotheses as H_0 and H_1, respectively. We denote by $\mathcal{I}_0$ and $\mathcal{I}_1$ two SAR images, composed of n SAR amplitude channels each, acquired over the same geographical area at times t_0 and t_1, respectively $(t_1 > t_0)$, and coregistered. The well-known image-ratioing approach is adopted, which generates a ratio image $\mathcal{R}$ by dividing, pixel by pixel and channel by channel, the amplitudes in $\mathcal{I}_0$ by the amplitudes in $\mathcal{I}_1$ [18]. The resulting ratio image $\mathcal{R}$ is an n-channel image as well. Specifically, $\mathcal{R}$ is modeled as a set $\{u_1, u_2, \ldots, u_N\}$ of n-variate identically distributed random vectors (N being the number of pixels). This definition for $\mathcal{R}$ is aimed at detecting decreases in the n SAR amplitudes due to ground changes during the time interval $[t_0, t_1]$. If the ground changes yield increases in the SAR amplitudes, $\mathcal{R}$ could be redefined by dividing (pixel by pixel and channel by channel) the amplitudes in $\mathcal{I}_1$ by the amplitudes in $\mathcal{I}_0$. The proposed method could then be applied as described in the following paragraphs with no modifications. If there are both increases and decreases in the SAR amplitudes, the method could be separately applied to the two possible formulations of the multichannel ratio image [43]. Alternatively, the two formulations could be combined into a single "normalized ratio image" by computing, pixel by pixel and band by band, the minimum of the two aforementioned ratios [12], [64]; this choice would require modifying the parameter-estimation procedure described hereinafter. Note also that the image-ratioing approach implicitly refers to the situation in which $\mathcal{I}_0$ and $\mathcal{I}_1$ are acquired by the same sensor in the same configuration, so that differences between pixel values may not be ascribed to differences in the acquisition processes at the two observation times.

The key ideas of the proposed method are to model each channel of $\mathcal{R}$ as a distinct information source and to formalize the change-detection problem as a multichannel data-fusion problem that can be addressed by using MRFs. Such fields represent a wide family of contextual image models, defined according to neighborhood systems. Owing to the well-known Hammersley–Clifford theorem [23], they allow the (intractable) global "maximum *a posteriori*" decision rule (i.e., the maximization of the joint probability distribution of all pixel labels, conditioned to all pixel intensities) to be expressed as the (tractable) minimization of an energy function, defined locally according to the neighborhood system [19]. Let $\ell_k \in \{H_0, H_1\}$ be the hypothesis label of the kth pixel $(k = 1, 2, \ldots, N)$; the label configuration $\{\ell_1, \ell_2, \ldots, \ell_N\}$ is assumed to be an MRF [19]. Accordingly, the following (local) posterior probability distribution is introduced $(k = 1, 2, \ldots, N; i = 0, 1)$ [19]:

$$P\{\ell_k = H_i | u_k, \mathcal{C}_k\} = \frac{\exp\left[-U(H_i|u_k, \mathcal{C}_k)\right]}{\sum_{j=0}^{1} \exp\left[-U(H_j|u_k, \mathcal{C}_k)\right]} \quad (1)$$

where $\mathcal{C}_k$ is the set of labels of the neighbors (i.e., the "spatial context") of the kth pixel and $U(H_i|u_k, \mathcal{C}_k)$ is the "energy function" of the MRF model $(i = 0, 1; k = 1, 2, \ldots, N)$. In accordance with the MRF-based approach to data fusion, this energy function is expressed as a linear combination of energy contributions, each related either to the label configuration (contextual information) or to the information conveyed by each channel of the ratio image [62]. The former information is formulated as a second-order isotropic Potts model [19], and the latter is related directly to the single-variate pdf of the corresponding SAR ratio channel conditioned to "change" and "no-change." Hence, the proposed MRF model is defined by the following energy function[1] $(k = 1, 2, \ldots, N; i = 0, 1)$:

$$U(H_i|u_k, \mathcal{C}_k, \theta) = \sum_{r=1}^{n} \alpha_r \left[-\ln p_{ir}(u_{kr}|\xi_{ir})\right] - \beta m_{ik} \quad (2)$$

where $p_{ir}(\cdot|\xi_{ir})$ is a parametric model for the pdf of the rth amplitude ratio u_{kr} conditioned to H_i, ξ_{ir} is the vector of the parameters of this pdf, m_{ik} is the number of labels equal to H_i in the neighborhood of the kth pixel, α_r $(r = 1, 2, \ldots, n)$ and β are positive parameters, and $\theta = (\beta, \alpha_1, \alpha_2, \ldots, \alpha_n, \xi_{01}, \xi_{11}, \xi_{02}, \xi_{12}, \ldots, \xi_{0n}, \xi_{1n})$ is a vector collecting all the model parameters to be estimated (the dependence of the energy on θ has been included in the related notation). Owing to the aforementioned minimum-energy MRF decision rule, the spatial energy contribution $(-m_{ik})$ favors the generation of homogeneous regions made up of pixels assigned to the same hypothesis, thus reducing the impact of speckle on the classification results (further details on this spatial contribution can be found in [19] and [27]). Note that the MRF model in (1) and (2) is intrinsically stationary; this property is usually violated by real images but is often accepted as a simplifying analytical assumption [27].

[1]Given $v \in \mathbb{R}^m$, we denote by v_j the jth component of v $(j = 1, 2, \ldots, m)$.

The pdf of a SAR amplitude ratio is known to be strongly non-Gaussian [49]. Several parametric models can be used to describe the marginal distribution of each feature u_{kr} conditioned to H_i $(k = 1, 2, \ldots, N; r = 1, 2, \ldots, n; i = 0, 1)$, such as the log-normal (LN), Nakagami-ratio (NR), and Weibull-ratio (WR) models [43]. These distributions are analytically derived (under suitable assumptions)[2] from the well-known Nagakami-Gamma, Weibull, and LN distributions for the SAR amplitudes acquired at the two observation dates (further details on these derivations and on the related assumptions can be found in [43]) and are given by the following $(k = 1, 2, \ldots, N; r = 1, 2, \ldots, n; i = 0, 1)$.

1) LN distribution for u_{kr}, conditioned to H_i $(\xi_{ir} = (\mu_{ir}, \sigma_{ir}), \mu_{ir} \in \mathbb{R}, \sigma_{ir} > 0)$

$$p_{ir}(u_{kr}|\xi_{ir}) = \frac{1}{\sigma_{ir} u_{kr} \sqrt{2\pi}} \exp\left[-\frac{(\ln u_{kr} - \mu_{ir})^2}{2\sigma_{ir}^2}\right], \quad u_{kr} > 0. \quad (3)$$

2) NR distribution for u_{kr}, conditioned to H_i $(\xi_{ir} = (L_{ir}, \gamma_{ir}), L_{ir}, \gamma_{ir} > 0)$

$$p_{ir}(u_{kr}|\xi_{ir}) = \frac{2\Gamma(2L_{ir})}{\Gamma^2(L_{ir})} \frac{\gamma_{ir}^{L_{ir}} u_{kr}^{2L_{ir}-1}}{(\gamma_{ir} + u_{kr}^2)^{2L_{ir}}}, \qquad u_{kr} > 0 \quad (4)$$

where $\Gamma(\cdot)$ is the usual Gamma function [55].

3) WR distribution for u_{kr}, conditioned to H_i $(\xi_{ir} = (\eta_{ir}, \lambda_{ir}), \eta_{ir}, \lambda_{ir} > 0)$

$$p_{ir}(u_{kr}|\xi_{ir}) = \eta_{ir} \lambda_{ir}^{\eta_{ir}} \frac{u_{kr}^{\eta_{ir}-1}}{\left(\lambda_{ir}^{\eta_{ir}} + u_{kr}^{\eta_{ir}}\right)^2}, \qquad u_{kr} > 0. \quad (5)$$

Note that such models for the pdf of u_{kr}, given H_1 $(k = 1, 2, \ldots, N; r = 1, 2, \ldots, n)$, are monomodal. Hence, we implicitly focus on the case in which a single typology of change is present. It is also worth remarking that, in the case of the method proposed in [45], the choice of an LN distribution was almost mandatory due to the specific analytical formulation of the technique. For the present approach, this constraint does not hold because only the marginal pdfs of the amplitude ratio channels are involved in the MRF model. In particular, this method does not require the joint n-variate pdf of the multichannel ratio vector u_k conditioned to H_i to be modeled because the MRF formulation in (2) is equivalent to replacing the joint pdf by the product $p_{i1}(u_{k1}|\xi_{i1})^{\alpha_1} p_{i2}(u_{k2}|\xi_{i2})^{\alpha_2} \ldots p_{in}(u_{kn}|\xi_{in})^{\alpha_n}$ $(k = 1, 2, \ldots, N; i = 0, 1)$ of the related n marginal pdfs, exponentially weighted by the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$. This represents a more tractable simplified model. The method is not equivalent to assuming that $u_{k1}, u_{k2}, \ldots, u_{kn}$ are independent when conditioned to H_i $(i = 0, 1; k = 1, 2, \ldots, N)$, since the parameters α_r's are

[2]If two SAR images acquired at different observation dates have zero complex correlation and if the related SAR amplitudes are Nakagami-distributed with the same number of looks, then the resulting amplitude ratio is NR-distributed. Similarly, if the two amplitudes are independent and Weibull-distributed with the same value of the shape parameter of the Weibull model, then the amplitude ratio follows a WR model. If the two amplitudes have LN pdfs, also the related amplitude ratio is LN-distributed (no further assumptions are needed in this case) [43].

generally not unitary $(r = 1, 2, \ldots, n)$ and, consequently, $p_{i1}(u_{k1}|\xi_{i1})^{\alpha_1} p_{i2}(u_{k2}|\xi_{i2})^{\alpha_2} \ldots p_{in}(u_{kn}|\xi_{in})^{\alpha_n}$ is not a n-variate pdf (its integral is generally not unitary if $\alpha_r \neq 1$ for some $r \in \{1, 2, \ldots, n\}$). More precisely, unlike an MRF model involving the joint H_i-conditional $(i = 0, 1)$ pdfs of the ratio channels, the MRF defined by (2) does not model the possible correlations between distinct ratio channels. Instead, it allows one to quantify the reliability associated with each channel through the parameters $\alpha_1, \alpha_2, \ldots, \alpha_n$. This is formally similar to the consensus-theoretic LOGP approach to multisource fusion adopted in [5] and [7] for supervised remote-sensing image classification. Plugging information about interchannel correlations into the MRF model would require accurate parametric models for the joint pdfs of the amplitude ratios $u_{k1}, u_{k2}, \ldots, u_{kn}$ conditioned to H_i $(i = 0, 1; k = 1, 2, \ldots, N)$. Well-known models such as the Wishart distribution [15] are available for the joint pdf of complex-valued polarimetric SAR features; however, the problem of parametrically modeling the joint pdf of n amplitude or intensity ratio features remains an open issue. A simple solution may be to adopt the LN distribution for related marginal pdfs, since this distribution allows a natural multivariate generalization (the so-called "multivariate LN"). This possible alternative model will be considered again in Section III.

The parameters β, $\alpha_1, \alpha_2, \ldots, \alpha_n$ tune the relative importance of the energy terms. In particular, α_r plays the role of a "reliability factor" assigned to the information source represented by the rth channel $(r = 1, 2, \ldots, n)$ and is usually assumed to take on values in [0, 1]. This constraint can be analytically expressed as follows[3]:

$$\|2\alpha - \mathbf{1}\|_\infty \leq 1 \tag{6}$$

where $\mathbf{1}$ is an n-dimensional vector with n unitary components and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$. As detailed in the Appendix, a differentiable constraint is desirable for analytical convenience. Hence, we replace the nondifferentiable constraint in (6) by the following one:

$$\|2\alpha - \mathbf{1}\|_q \leq 1 \tag{7}$$

where $q \geq 2$ is an even integer number. For sufficiently large values of q (e.g., $q = 10$), (7) is basically a differentiable approximation of (6). Furthermore, it also allows undesired degenerate parameter values to be avoided [e.g., the undesired configuration $\alpha = 0$, which removes from the MRF model all the energy contributions related to the ratio channels, is permitted by (6) but forbidden by (7)].

B. EM-Based Parameter Estimation and Classification

The proposed change-detection method postulates the MRF model of (1) and (2) to hold and aims at generating a change map that minimizes the corresponding energy function. Hence, suitable parameter-estimation and classification strategies need to be defined. Toward this end, a case-specific variant of the LJ-EM method is developed by integrating this approach with MoLC. LJ-EM is an iterative parameter-estimation technique developed for the computation of maximum-likelihood estimates in problems characterized by data incompleteness [28].

The application of LJ-EM to the MRF model defined by (1) and (2) gives rise to two difficulties. First, a large computational burden is typically involved by the combination of EM-type algorithms and contextual image models, due to the need to run a Gibbs sampler or a Monte Carlo technique at each EM iteration [14]. In order to overcome this drawback, the "mode-field" approach is adopted that is based on an approximation in the log-likelihood function: the approximation allows the execution time to be sharply reduced and usually represents a good tradeoff between computational burden and classification accuracy (a general description of the "mode-field" approach can be found in [14]; similar techniques have been applied to remote-sensing data in [16], [27], and [45]). A further critical point lies in the fact that the maximization problems involved by LJ-EM have no closed-form solutions, when one uses the NR and WR models [43], [46]. In order to address this issue, we propose to combine LJ-EM with MoLC, a recently developed parameter-estimation method that exploits the Mellin transform theory [61] to state a set of equations relating the unknown parameters to logarithmic moments or logarithmic cumulants of the gray-level distribution, thus expressing the parameter-estimation problem as the solution of a set of (typically nonlinear) equations [48], [63].

MoLC has been developed in the context of SAR image analysis and has been proven to be numerically feasible and fast for many SAR-specific models and to exhibit good theoretical estimation properties [47], [48]. Moreover, it has turned out to be feasible for all the LN, NR, and WR distributions [43]. Denoting by $\kappa_{1ir} = E\{\ln u_{kr}|\ell_k = H_i\}$ and $\kappa_{2ir} = \mathrm{Var}\{\ln u_{kr}|\ell_k = H_i\}$ the logarithmic mean and variance of u_{kr}, given H_i $(k = 1, 2, \ldots, N; r = 1, 2, \ldots, n; i = 0, 1)$, MoLC allows the following equations to be obtained for the LN, NR, and WR distributions [43]:

$$\begin{aligned}
\text{LN}: &\begin{cases} \kappa_{1ir} = \mu_{ir} \\ \kappa_{2ir} = \sigma_{ir}^2 \end{cases} \\
\text{NR}: &\begin{cases} 2\kappa_{1ir} = \ln \gamma_{ir} \\ 2\kappa_{2ir} = \Psi(1, L_{ir}) \end{cases} \\
\text{WR}: &\begin{cases} \kappa_{1ir} = \ln \lambda_{ir} \\ \kappa_{2ir} = 2\Psi(1,1)\eta_{ir}^{-2} \end{cases}
\end{aligned} \tag{8}$$

where $\Psi(1, \cdot)$ is the first-order polygamma function (i.e., the second logarithmic derivative of $\Gamma(\cdot)$ [61]). The MoLC equations for LN simply state the identity between the distribution parameters and the logarithmic mean and variance. On the other hand, the MoLC equations for NR and WR represent a set of nontrivial equations relating the distribution parameters to κ_{1ir} and κ_{2ir} $(r = 1, 2, \ldots, n; i = 0, 1)$.

LJ-EM is iterative and initialized with the single-channel strategy described in Section II-C. As detailed in the Appendix, when applied according to an LN model, it computes, at each

[3]Given $v \in \mathbb{R}^m$ and $q \geq 1$, we denote by $\|v\|_\infty$ and $\|v\|_q$ the Tchebitchev and the qth-order Minkowski norms (∞-norm and q-norm, for short) of v, i.e., $\|v\|_\infty = \max\{|v_1|, |v_2|, \ldots, |v_m|\}$ and $\|v\|_q = (|v_1|^q + |v_2|^q + \cdots + |v_m|^q)^{1/q}$ [56].

iteration, the estimates of the conditional logarithmic means and variances. In order to integrate LJ-EM with MoLC in the cases of WR and NR, we propose to update, at each iteration, estimates of the conditional logarithmic means and variances as in the LN version of LJ-EM. Then, we derive the estimates of the NR and WR distribution parameters (i.e., (L_{ir}, γ_{ir}) and $(\eta_{ir}, \lambda_{ir})$, respectively) by solving the related MoLC equations. A similar hybrid MoLC-EM approach was applied in [46] for pdf modeling purposes.

Accordingly, the following operations are performed at the tth iteration of the proposed algorithm $(t = 0, 1, 2, \ldots)$.

1) Compute, for each kth pixel, the current estimate of $P\{\ell_k = H_i | u_k, \mathcal{C}_k^t, \theta^t\}$ $(k = 1, 2, \ldots, N; i = 0, 1)$ by using (1) and (2).
2) Update the label of each kth pixel according to the MRF minimum-energy rule [19] by setting ℓ_k^{t+1} as the label H_i that corresponds to the smallest value of $U(H_i | u_k, \mathcal{C}_k^t, \theta^t)$ $(k = 1, 2, \ldots, N; i = 0, 1)$.
3) Update the estimates of the spatial parameter β and of the logarithmic means and variances as follows $(r = 1, 2, \ldots, n; i = 0, 1)$:

$$\begin{cases} \kappa_{1ir}^{t+1} = \dfrac{\sum_{k=1}^{N} w_{ik}^t \ln u_{kr}}{\sum_{k=1}^{N} w_{ik}^t} \\ \kappa_{2ir}^{t+1} = \dfrac{\sum_{k=1}^{N} w_{ik}^t \left(\ln u_{kr} - \kappa_{1ir}^{t+1}\right)^2}{\sum_{k=1}^{N} w_{ik}^t} \\ \beta^{t+1} = \arg\max\limits_{\beta > 0} \sum\limits_{k=1}^{N} \left[\beta \sum\limits_{i=0}^{1} w_{ik}^t m_{ik}^t - \ln \sum\limits_{i=0}^{1} \exp\left(\beta m_{ik}^t\right) \right] \end{cases} \tag{9}$$

   where $w_{ik}^t = P\{\ell_k = H_i | u_k, \mathcal{C}_k^t, \theta^t\}$ if the kth pixel has been assigned to H_i in step 2, and $w_{ik}^t = 0$ otherwise $(k = 1, 2, \ldots, N; i = 0, 1)$.
4) Compute, for each hypothesis H_i and each rth channel, the updated estimate ξ_{ir}^{t+1} of the parameter vector ξ_{ir} of the adopted pdf model (i.e., NR, WR, or LN) by plugging the logarithmic mean and variance estimates κ_{1ir}^{t+1} and κ_{2ir}^{t+1} computed in step 3 in the corresponding MoLC equations [see (8)] and by solving for the unknown parameters $(r = 1, 2, \ldots, n; i = 0, 1)$.
5) Update the estimate of the reliability factor α_r of each rth channel as follows $(r = 1, 2, \ldots, n$; see Appendix):

$$\alpha_r^{t+1} = \frac{1}{2} + \frac{1}{2} \sqrt[q-1]{\frac{c_r^t}{\|c^t\|_{q'}}} \tag{10}$$

   where

$$q' = \frac{q}{q-1} \qquad c_r^t = \sum_{k=1}^{N} \sum_{i=0}^{1} w_{ik}^t \ln p_{ir}\left(u_{kr} | \xi_{ir}^{t+1}\right) \tag{11}$$

   and $\|c^t\|_{q'}$ is the q'-norm of the vector $c^t = (c_1^t, c_2^t, \ldots, c_n^t)$.

The formulations of the updated estimates of the spatial parameter (step 3) and of the reliability factors (step 5) are proven in the Appendix. The maximization problem to compute β^{t+1} [see (9)] is numerically addressed by the Newton–Raphson algorithm [50]. Step 4 requires solving the MoLC equations for the adopted pdf model. No actual solution process is required for LN, since the pdf parameters are exactly equal to the logarithmic mean and variance. The solution process for WR can be analytically carried out in closed form. A numerical procedure is needed for NR in order to solve the equation involving κ_{2ir} and L_{ir} $(r = 1, 2, \ldots, n; i = 0, 1)$. However, owing to the strict monotonicity of the involved polygamma function, the simple bisection method [50] can be effectively applied. The updates of the estimates of the reliability factors (step 5) are performed by a closed-form calculation.

The procedure is iterated until the differences between the parameter estimates computed at the successive iterations are below a threshold (equal to 0.001 in our experiments). Since the iterative procedure performed by the proposed change-detection method is actually derived from EM (see Appendix), a convergent behavior of the method can be theoretically expected [28], [53], [68]. We shall denote the proposed technique by "data-fusion based multichannel change detection" (DF-MCD for short).

C. Initialization With the GKIT Algorithm

The single-channel "generalized Kittler and Illingworth" technique (GKIT) [43] is used to compute the initial change map. GKIT can be based on an LN, NR, or WR conditional model for SAR amplitude ratios and automatically computes the optimal threshold to be applied to a single-channel amplitude ratio image (in order to distinguish between changed and unchanged areas) by minimizing a "criterion function" related to the probability of error of a Bayesian binary classifier [43].

Here, the version of GKIT based on the adopted parametric distribution (either LN, NR, or WR) is applied. Denoting by $J_r(\cdot)$ the criterion function computed by GKIT when applied to the rth channel in $\mathcal{R}$, the corresponding optimal threshold is $\tau_r^* = \arg\min_\tau J_r(\tau)$ $(r = 1, 2, \ldots, n)$. By thresholding the rth channel in $\mathcal{R}$ by this value, a change map $\mathcal{M}_r^*$ is obtained $(r = 1, 2, \ldots, n)$ [43]. The initial change map for the proposed method is defined as the map $\mathcal{M}_r^*$ corresponding to the smallest optimal value $J_r(\tau_r^*)$ of the criterion function $(r = 1, 2, \ldots, n)$. Owing to the aforementioned relationship between the criterion function and the probability of error [31], [43], this map is expected to be at least a suboptimal choice from among the single-channel maps.

For initialization purposes, the initial map is used to define the initial context $\mathcal{C}_k^0$ of each kth pixel $(k = 1, 2, \ldots, N)$ and the initial estimates of the pdf parameters. A sample mean estimate κ_{1ir}^0 of κ_{1ir} and a sample-variance estimate κ_{2ir}^0 of κ_{2ir} are computed according to the set of samples assigned to H_i in this map $(i = 0, 1; r = 1, 2, \ldots, n)$.

III. Experimental Results

A. Data Set for Experiments

The proposed method was experimentally validated with both semisimulated and real data sets. The latter was composed of two 700 × 280 pixel-sized coregistered multipolarization and multifrequency SAR images, taken over an agricultural

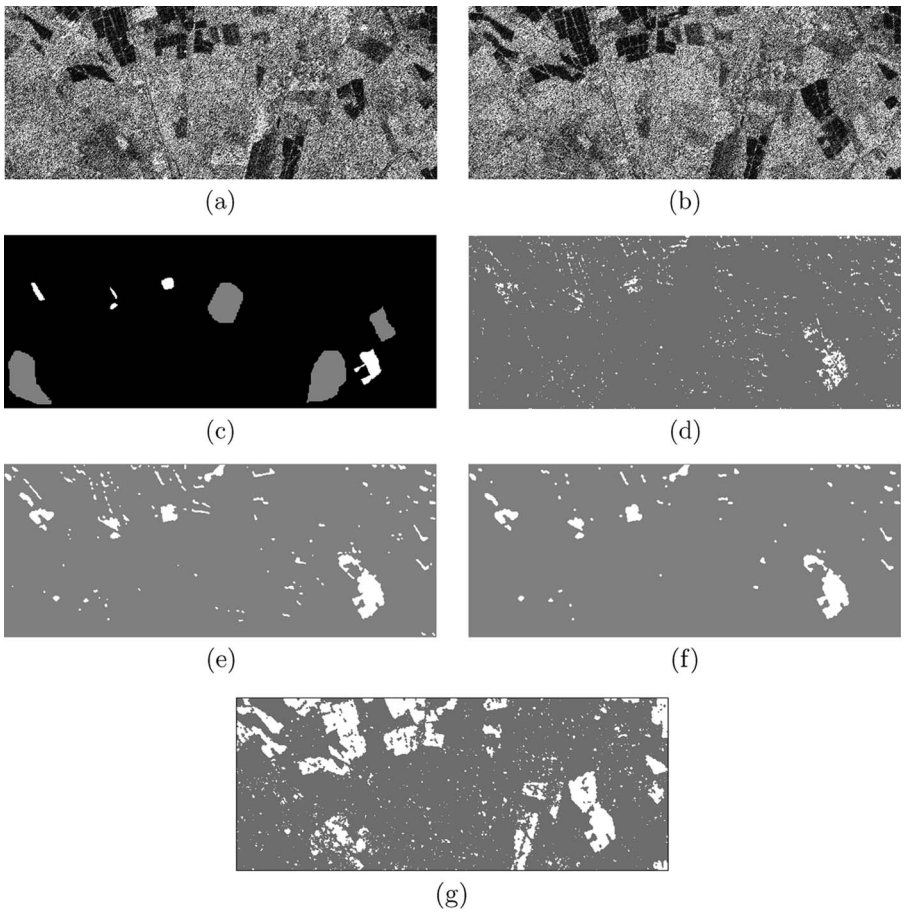

Fig. 1. Real data set. XSAR channels acquired on (a) April 16, 1994 and (b) April 18, 1994 (after histogram equalization). (c) Test map. Change maps generated (d) by the initialization GKIT procedure, by DF-MCD, applied with (e) $q = 2$ and (f) $q = 10$, and (g) by the method based on the energy function in (12) with a multivariate LN model. All maps (d)–(f) were obtained using the LN model. Legend to the change maps: White = "change" and gray = "no-change."

area near the city of Pavia (Italy) on April 16 and 18, 1994. At each date, four channels were available, i.e., a four-look XSAR image [X-band, VV polarization; Fig. 1(a) and (b)] and three 4-look SIR-C channels (C-band, HH and HV polarizations, and TP (total power) channels). Ground changes were present because several rice fields were artificially flooded for cultivation purposes. A test map, including 11 287 "no-change" test pixels and 1870 "change" test pixels, was available for this data set [Fig. 1(c)]. As the method is unsupervised, the test map was used only to estimate the accuracies of the resulting change maps.

The semisimulated data set [Fig. 2(a)–(c)] was constructed from a nine-channel single-look SAR amplitude image acquired in August 1989 over the agricultural region of Feltwell (U.K.) by a fully polarimetric PLC-band NASA/JPL airborne sensor. The image included HH, HV, and VV polarizations acquired at C-, L-, and P-bands [57]. Changes were simulated by moving a region to a different position in the image area and by introducing multiplicative noise to both the original image and the image with the translated region. In order to simulate the speckle behavior in the amplitude images, a unitary mean Nakagami-distributed noise was introduced with an equivalent number of looks (ENL) equal to five. Because the changed area was simulated, the test map for this data set was exhaustive [Fig. 2(c)].

B. Experimental Results by DF-MCD

The proposed method can be run in three different versions that correspond to the three pdf models considered. In order to validate the method, all three of the versions were applied to both data sets. In all the cases, the α_r and β parameters were initialized with unitary values.

The initialization GKIT procedure was applied to the real and semisimulated data sets (using either the LN, NR, or WR models) and selected the change maps generated by the SIR-C-HV and C-HH channels, respectively. As theoretically expected, the proposed method reached convergence in all cases (fewer than 50 iterations were sufficient in all the experiments). Table I shows the detection accuracies (i.e., the percentages of correctly labeled "change" test pixels), false-alarm rates (i.e., the percentages of erroneously labeled "no-change" test pixels), and overall error rates (i.e., the percentages of erroneously labeled test pixels) of the resulting change maps, which were computed according to the test maps in Figs. 1(c) and 2(c).

Focusing first on the results obtained with $q = 2$ [i.e., the Euclidean norm in the constraint of (7)], accurate change maps were generated by all versions of the method with both data sets, with very low false-alarm rates (never above 1% and often below 0.2%) and high detection accuracies (above 92% and 88% for all experiments with the real and semisimulated data sets, respectively), even though the initialization maps

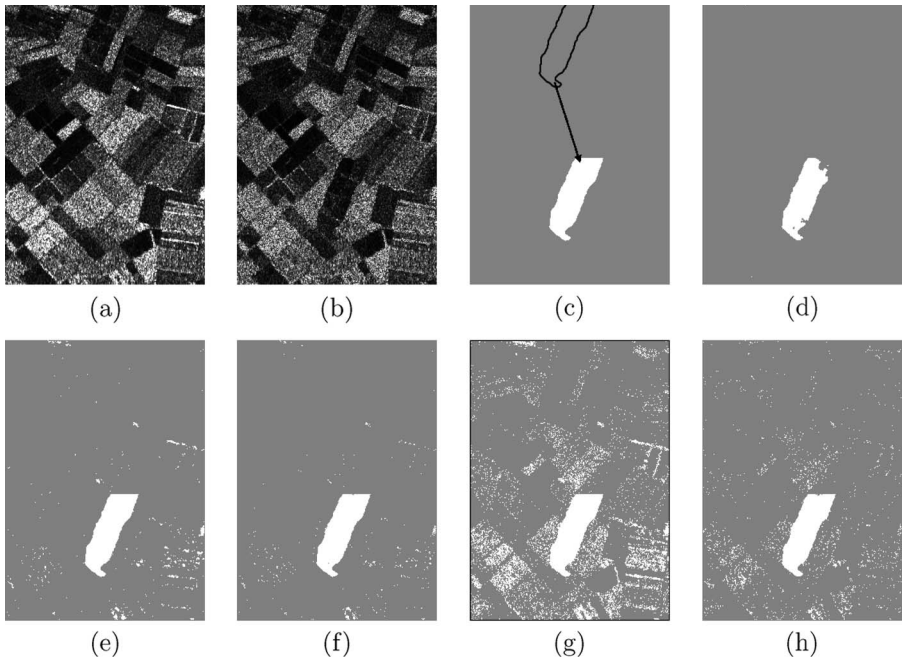

Fig. 2. Semisimulated data set. L-HH channels of the (a) first and (b) second date image. (c) Test map. Change maps generated by DF-MCD when applied with $q = 2$ and (d) LN, (e) NR, and (f) WR. Change maps generated (g) by the method based on the energy function in (12) with a multivariate LN model and (h) by the method in [45]. Legend to the change maps: White = "change" and gray = "no-change." The black line in the test map highlights the region that has been moved in order to simulate changes.

TABLE I
CHANGE-DETECTION PERFORMANCES OF THE PROPOSED DF-MCD METHOD, OF THE INITIALIZATION GKIT ALGORITHM, OF THE PREVIOUS TECHNIQUE DEVELOPED IN [45], AND OF THE METHOD BASED ON THE ENERGY FUNCTION IN (12) WITH A MULTIVARIATE LN MODEL

Method	PDF model	Real data set			Semisimulated data set		
		False-alarm rate	Detection accuracy	Overall error rate	False-alarm rate	Detection accuracy	Overall error rate
initialization (GKIT)	LN	0.33%	49.09%	7.52%	0.15%	2.24%	3.79%
	NR	0%	15.67%	11.99%	0.14%	2.05%	3.78%
	WR	0.02%	45.03%	7.83%	0.10%	0.67%	3.80%
DF-MCD (2-norm)	LN	0.16%	95.08%	0.84%	0.00%	88.57%	0.43%
	NR	0.11%	92.24%	1.20%	1.00%	99.97%	0.96%
	WR	0.62%	97.05%	0.95%	0.50%	99.82%	0.49%
DF-MCD (10-norm)	LN	0%	94.70%	0.75%	0.00%	84.43%	0.58%
	NR	0%	91.01%	1.27%	0.01%	93.99%	0.23%
	WR	0.62%	97.05%	0.95%	0.50%	99.82%	0.49%
Multivariate LN		2.11%	92.24%	2.81%	9.77%	100.00%	9.40%
Method in [45]		0%	93.10%	0.98%	2.47%	99.54%	2.40%

were very poor (see Table I). This illustrates the capability of DF-MCD to jointly exploit information conveyed by the ratio channels and the spatial information related to the correlations among neighboring pixels by Markovian data fusion. When applied to the real data set, the three pdf models yielded quite similar results. Specifically, the lowest overall error rate (0.84%) was obtained by LN, while the highest detection accuracy (97.05%) was achieved by WR, and a slightly poorer result was obtained by NR. A visual analysis of the change maps (see, for instance, Fig. 1(e), which refers to the case of LN) further confirms both the limited impact of speckle and the correct detection of the changed areas. Several false alarms can also be seen in Fig. 1(e), particularly at the borders of the agricultural fields in the image. These false alarms are mainly due to small misregistration errors. When using the semisimulated data set, the lowest overall error rate (0.43%) was achieved by LN, albeit with a detection accuracy of just 88.57%; the highest detection accuracy (99.97%) was achieved by NR, but with a 1% false-alarm rate. WR yielded an error rate and a detection accuracy (i.e., 0.49% and 99.82%) very similar to those of LN and NR,

TABLE II
COMPUTATION TIMES OF DF-MCD (APPLIED TO BOTH THE REAL AND THE SEMISIMULATED DATA SETS WITH THE LN, NR, OR WR MODELS AND WITH $q = 2$ OR $q = 10$) ON A PERSONAL COMPUTER WITH 2.51-GHz CPU AND 2-GB RAM. THE RELATED IMAGE SIZES ARE ALSO REPORTED

Method	PDF model	Real data set	Semisimulated data set
DF-MCD (2-norm)	LN	14 s	7 s
	NR	3 min 5 s	3 min 11 s
	WR	43 s	25 s
DF-MCD (10-norm)	LN	22 s	6 s
	NR	3 min 23 s	3 min 10 s
	WR	43 s	25 s
Image size [pixel]		700×280	250×350

respectively, thus representing a good tradeoff between false- and missed-alarm errors for this data set. Fig. 2(d)–(f) visually confirms this observation.

The results were shown to be remarkably stable when the Minkowski-norm index q was varied in the range [2, 10] (the results of $q = 10$ are given in Table I; those of $q = 4, 6, 8$ are similar, so they are omitted for the sake of brevity). When using WR, the results with both data sets were actually unaffected by the variation of $q \in [2, 10]$; when either LN or NR was employed, only small variations in numerical accuracies were observed (Table I). A visual analysis of the related change maps [see, for example, Fig. 1(f)] reveals that such variations are actually related to a stronger spatial regularization effect obtained with the 10-norm, as compared with the 2-norm. Many false alarms were therefore seen, due to misregistration errors, in the maps obtained with the real data set when $q = 2$; such false alarms were not present when $q = 10$ [Fig. 1(f)]. On the other hand, a slightly stronger smoothing can be noticed in the borders between changed and unchanged areas in the change map generated by using the 10-norm than in the map obtained by the 2-norm. This can be interpreted as a consequence of the different convergence values reached with the estimated reliability factors. Focusing, for instance, on the case of LN and the real data set, $\alpha = (0.224, 0.238, 0.281, 0.260)$ and $\alpha = (0.060, 0.062, 0.071, 0.066)$ were reached at convergence in the 2- and 10-norm cases, respectively. In other words, much smaller values were obtained by $q = 10$ than by $q = 2$. On the other hand, the method provided almost unitary values for the estimate of β in both cases. Therefore, in the 10-norm case, the spatial energy contribution was given a comparatively larger relative weight than those of the energy terms related to the ratio channels, as compared with the 2-norm case; thus, a stronger spatial regularization effect was obtained.

DF-MCD required short computation times with both $q = 2$ and $q = 10$ cases and both data sets (see Table II), always less than 30 s with LN, less than a minute with WR, and around 3–4 min with NR on a personal computer with 2.51-GHz CPU and 2-GB RAM. These slight differences in computation time are expected based on the different solution methods. Using NR requires numerically solving the related MoLC equations at each iteration of the method, whereas a closed-form analytical solution is available for the MoLC equations of WR and no actual solution process is needed with LN.

C. Experimental Comparisons

Several experimental comparisons were carried out in order to better investigate the properties of DF-MCD. The first focus was on the proposed MRF model and on the related choice to formalize each ratio channel as a separate information source. A comparison was made between the proposed MRF model and one based on a multivariate LN model for the joint pdf of all amplitude ratios conditioned to "change" and "no-change." Following the notations in Section II-A, we denote the pdf of the vector u_k of the amplitude ratios in all n channels, conditioned to H_i, by $p_i(u_k|\kappa_{1i}, K_{2i})$ $(k = 1, 2, \ldots, N; i = 0, 1)$, and we model this pdf as a multivariate LN [34]. Therefore, it is parameterized by the n-dimensional vector κ_{1i} whose rth component is the H_i-conditional log-mean κ_{1ir} and by the $n \times n$ matrix K_{2i} whose (r, s)th entry is the conditional log-covariance $E\{(\ln u_{kr} - \kappa_{1ir})(\ln u_{ks} - \kappa_{1is})|\ell_k = H_i\}$ $(i = 0, 1; r, s = 1, 2, \ldots, n)$. Assuming, as with DF-MCD, a Potts model for the joint statistics of neighboring pixel labels, a classical MRF-based approach to binary unsupervised classification is characterized by the following energy function $(i = 0, 1; k = 1, 2, \ldots, N)$:

$$U(H_i|u_k, \mathcal{C}_k, \theta) = -\ln p_i(u_k|\kappa_{1i}, K_{2i}) - \beta m_{ik} \qquad (12)$$

where $\theta = (\beta, \kappa_{10}, \kappa_{11}, K_{20}, K_{21})$. The results of DF-MCD were compared with those obtained with this MRF model. Specifically, the parameter vector θ was estimated by LJ-EM with the mode-field approximation; the resulting iterative equations were a straightforward adaptation of the well-known iterative equations for EM-like parameter estimation for a Gaussian mixture [14], [28], [53]. Only the LN pdf was considered for this experiment, since no multivariate extensions have been proposed so far for the NR and WR distributions.

When applied to the two data sets, this approach generated the change maps in Figs. 1(g) and 2(g), in which most (or all) changed pixels were detected, but many false alarms were also generated. Indeed, high detection accuracies were obtained for both data sets, but with much higher error rates than DF-MCD (see Table I). The diminished accuracy from the multivariate modeling-based approach is interpreted as a consequence of the more limited accuracy of an estimated LN model for the joint amplitude ratio pdf than for the univariate distribution of each single ratio channel. Focusing on the real data set, a visual analysis of the map in Fig. 1(g) specifically identifies most false alarms as concentrated in "no-change" areas occupied by wet soil in the April 18th image. This result may be a consequence of a possible multimodal distribution of "no-change" in the 4-D feature space of the amplitude ratios. Since the multivariate LN is a monomodal distribution, this multimodality may severely affect the accuracy of the unsupervised parameter estimates, causing one or more modes of the "no-change" distribution to be erroneously merged with "change." On the contrary, the parameter-estimation process used by DF-MCD, which only

involves the marginal distributions of the ratios, proves to be more robust in handling the possible multimodality of "no-change." Parallel conclusions apply for the semisimulated data set as well. Note that similar conclusions were also drawn in [37] with regard to multivariate Gaussian models in change detection with multispectral data and in [45] with regard to noncontextual change detection with multichannel SAR.

In order to thoroughly investigate the data-fusion capabilities of DF-MCD, the results were compared by separately applying the method to each channel.[4] When applied to the X-band of the real data set with any of the three pdf models, DF-MCD did not provide significant results. In this case, the initialization GKIT procedure could not effectively discriminate "change" and "no-change" due to the strong presence of speckle in the X-band of the multitemporal image pair (a more detailed discussion can be found in [43]). Due to this failure during initialization, DF-MCD could not be effectively applied to this channel. The same conclusion holds for the L-HH, L-HV, P-HH, and P-HV channels of the semisimulated data set.

When separately applied to the C-HH, C-HV, and C-TP channels of the real data set, DF-MCD generated change maps with detection accuracies and overall error rates ranges of [32.67%, 90.86%] and [1.30%, 9.57%], respectively (the false-alarm rates were always below 0.05%). The most accurate result was obtained using the C-TP channel and the LN model. Thus, the change-detection performances of DF-MCD, when applied to single channels, were always worse than those achieved by jointly exploiting all available channels. Similarly, when applying DF-MCD to the single channels of the semisimulated data set, the detection accuracies fell in the range of [0%, 97.95%], and the overall error rates were [0.79%, 3.73%]. The most accurate result (with a 97.52% detection accuracy and a 0.79% error rate) was obtained with the C-VV channel and with WR. Compared to this most accurate method, the joint use of all nine channels by DF-MCD with the same pdf model yielded a 2.30% increase in the detection accuracy and a 0.29% decrease in the overall error rate (see Table I). This confirms the capability of the proposed MRF-based fusion approach to jointly exploit information conveyed by different channels.

A comparison was also made with the results given by the previous method proposed in [45] (see Table I). On average, the results obtained by applying this approach to the real data set were quite similar to those generated via DF-MCD, with slightly higher accuracy (in terms of overall error rate) given by DF-MCD with LN and WR than by the technique in [45] and with slightly lower accuracy given by NR. In particular, the detection accuracy of DF-MCD when applied with WR was almost 4% higher than that provided by the technique in [45]. When considering the semisimulated data set, the performance difference between the DF-MCD and the method in [45] was even greater; compared to DF-MCD, the latter obtained a very high detection accuracy (99.54%) but also a 2.40% overall error rate due to a much larger number of false alarms. A visual analysis of the change map in Fig. 2(h) confirms this conclusion. This result indicates, at least for the adopted data sets, that the data-fusion approach is more effective than the multichannel Fisher transformation approach for solving the multichannel SAR change-detection problem.

In order to further investigate this aspect, another experiment was conducted by focusing on the real data set and corrupting all channels of the images acquired at each observation date with multiplicative noise of increasing variance. Specifically, a multiplicative unitary-mean Nakagami-distributed noise was employed, with an ENL range of $\{4; 3.5; 3; \ldots; 1\}$. Since the technique in [45] is based on the LN model, DF-MCD was also applied with LN to ensure a more homogeneous comparison between the two approaches. The experiment was independently run ten times; the detection accuracies and overall error rates of the two methods, averaged over the ten runs, are shown in Fig. 3, along with the related standard deviations (the false-alarm rates were below 0.06% for all runs). DF-MCD provided higher average detection accuracies and lower average error rates than the algorithm in [45] for all seven tested noise ENL values. The performances were essentially stable when the noise ENL was ≥ 3 but worsened with lower values of ENL. The higher noise variance was shown to affect the technique from [45] more severely than DF-MCD; the differences between average detection accuracies for ENL $= 4$ and ENL $= 1$ were 3.17% for DF-MCD and 4.37% for [45].

Finally, the DF-MCD results were compared with those generated by the method described in [15], which is based on a likelihood-ratio test for the Wishart distribution of the complex-valued covariance matrices for multilook polarimetric SAR data on different observation dates. Particular cases of the method are also introduced in [15] to account for covariance matrices that are either diagonal or exhibit azimuthal symmetry. The diagonal case was considered in order to compare this approach with DF-MCD fairly, since it only involves the SAR intensities acquired on the two dates. The method in [15] does not directly produce a binary change map but rather extracts a quantitative measure of change from a multitemporal image pair; the choice of suitable threshold values for this extracted feature then yields the change map. Fig. 4 shows the (empirical) ROC curves obtained by computing this measure of change for both data sets and then plotting the false-alarm rate and detection accuracy values as a function of varied threshold. Since the technique in [15] is noncontextual while DF-MCD is contextual, we performed a fair comparison by plotting the ROC curves obtained by applying this method not only to the original multichannel intensity images but also to speckle-filtered intensities (a 7×7 Gamma-MAP filter [49] was used). The points corresponding to the DF-MCD performances (applied with all three possible pdf models and with 10-norm) are also shown in the 2-D space of false-alarm rate and detection accuracy. For the real data set, a much more accurate result is obtained by DF-MCD than by the technique in [15], even when the latter was also applied after despeckling. A smaller difference in performance is seen in the case of the semisimulated data set; the result of DF-MCD with LN was almost directly on the ROC curve of [15] with despeckling, while the results

[4]Note that in each single-channel experiment, only one weight parameter α_r was present in the MRF model, thus making the method insensitive to the value of q.

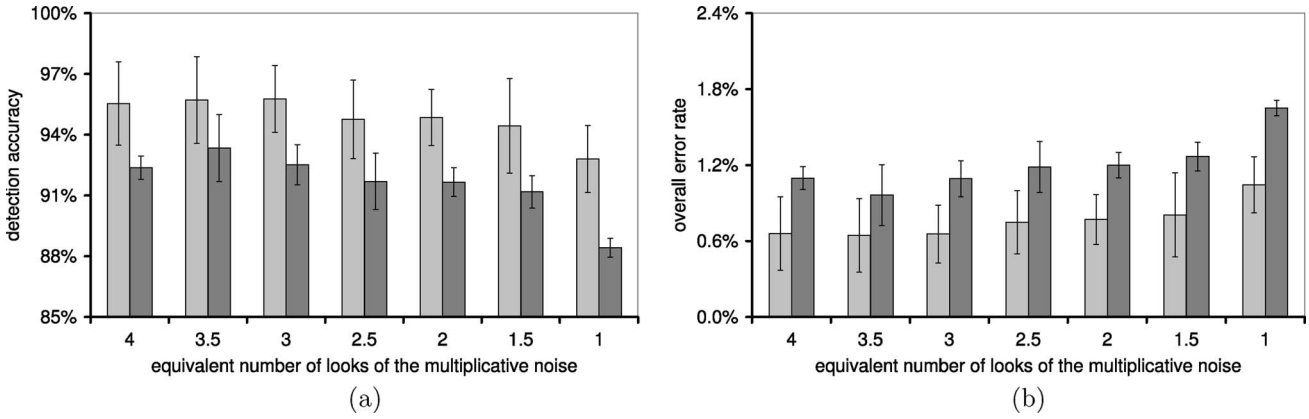

Fig. 3. Real data set. Behavior of the (a) average detection accuracy and (b) overall error rate of DF-MCD (applied with the LN model) and of the method in [45] as a function of the noise ENL (unitary-mean multiplicative Nakagami-distributed noise is introduced). The results were obtained by averaging over ten independent runs of the experiment, and the error bars denote related standard deviations. Legend: Light gray = DF-MCD and dark gray = method in [45].

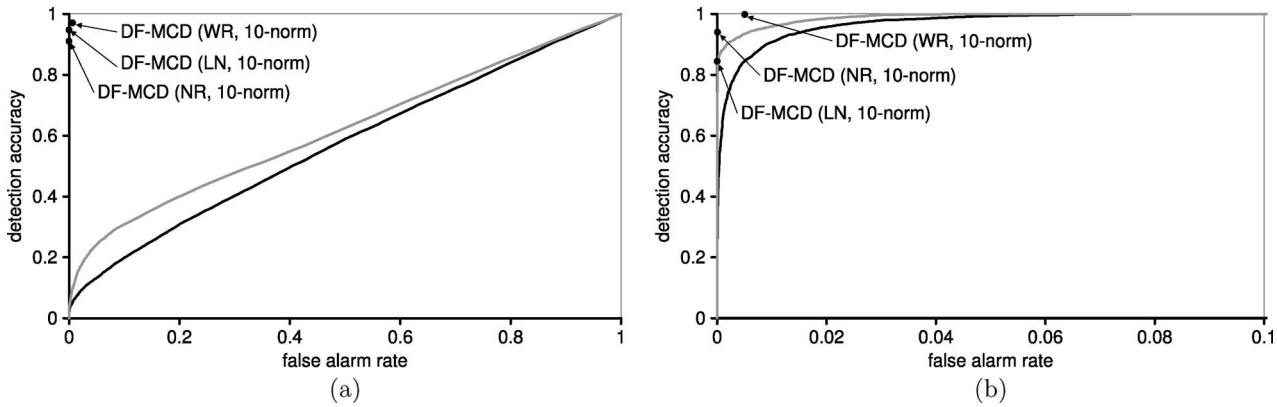

Fig. 4. ROC curve of the method in [15], applied with a diagonal covariance matrix to both (a) the real and (b) the semisimulated data sets with (gray line) and without (black line) despeckling. The points corresponding to DF-MCD performance (applied with all three possible pdf models and with the 10-norm) are also shown, in the 2-D space of false-alarm rate and detection accuracy.

with NR and WR fell above both ROC curves. These results suggest that the proposed MRF-based data-fusion approach to SAR change detection is more effective than the approach in [15], at least with respect to its formulation based on diagonal covariance matrices.

IV. Conclusion

An unsupervised contextual technique is proposed for the detection of changes in multichannel amplitude SAR images. The method combines the Markovian approach, which is used to fuse information derived from the different SAR channels with one another and with the spatial contextual information, with the LJ-EM algorithm, which is used for parameter estimation. In particular, the proposed method utilizes the single ratio channels as different information sources and fuses them within an energy function in accordance with the typical MRF approach. The resulting model disregards information regarding interchannel correlation, which would actually be captured intrinsically by the joint multivariate pdf of ratio images derived from the multichannel data. On the other hand, our method requires that the statistics of single ratio channels be modeled and estimated by their marginal pdfs alone; this proves a much easier task for most statistics typically considered for SAR signals and a more robust process with respect to the modeling and estimation of the joint multivariate pdfs. Although the interchannel correlation is disregarded, the proposed method is not equivalent to the simple statistical model based on the independent channel assumption, because the adopted MRF model utilizes an adaptive parameter for each channel to "modulate" its relevance (or reliability) in the model.

Accurate change maps were provided by the proposed DF-MCD method when applied to real and semisimulated multipolarization and multifrequency SAR data. Three versions of the method were applied according to three different parametric pdf families (i.e., LN, WR, and NR) utilized to model the statistics of the ratio images. In the case of the real data set, the three methods provided quite similar results with slightly lower overall error rates for the LN distribution and slightly higher detection accuracies for the WR family. In the case of the semisimulated data set, the most accurate results were achieved with WR. Each adopted model for the ratio statistics was derived from a corresponding model for the pdf of the SAR amplitude acquired at each date. Such amplitude pdfs are known from the SAR literature to correctly model the gray-level statistics of given land covers [49]. Therefore, a general suggestion on how to apply the proposed method can be to use the amplitude-ratio parametric family corresponding to the

amplitude pdf that optimally models the land covers in the SAR images acquired at the two observation dates.

An experimental comparison with the polarimetric SAR change-detection technique in [15] showed that the proposed approach yielded a more accurate result. This higher accuracy is credited to the effectiveness of combining contextual MRF-based processing and EM- and MoLC-based parameter estimation. Furthermore, the approach in [15] requires the user to interactively select a threshold value, whereas the proposed method does not require this interaction. On the other hand, the method in [15] was developed for (complex-valued) polarimetric SAR covariance matrices, whereas it could be run only in the particular case of diagonal matrices for the available data sets in this study. DF-MCD is not directly applicable to multichannel complex-valued SAR images. In order to extend the method to this data typology, parametric models for complex-valued SAR should be integrated into the proposed MRF-EM-MoLC framework. Integrating the Wishart-based approach in [15] into DF-MCD would be an interesting extension worth investigating, particularly since it would allow the additional exploitation of the SAR phase (and interchannel phase difference) information within the change-detection process.

No preliminary despeckling was required before the application of DF-MCD, owing to the MRF capability of exploiting the spatial contextual information in the classification process. Experiments confirmed that even when the initialization map was highly inaccurate, DF-MCD was able to converge toward an accurate final change map. However, when the initialization map does not separate at all "change" and "no-change" (which may happen in cases where speckle is very strong), DF-MCD cannot be initialized. In such cases, the use of local mean ratios [64] and/or despeckling [43] would be necessary (at least, for the initialization step).

The proposed technique is iterative, and the experiments showed good convergent behaviors in small numbers of iterations, although the method was initialized with (quite) low-accuracy change maps. Slightly longer execution times were required by DF-MCD with NR, due to the need to numerically solve nonlinear equations at each iteration. Anyway, the computation time was small in all the cases, and no more than a few minutes were needed to reach convergence in any run of the method. The convergence of DF-MCD is a theoretically expected result, owing to the well-known asymptotic properties of EM-like procedures. This represents a significant theoretical advantage of DF-MCD over the previous automatic change-detection method proposed in [45]. In addition, an experimental comparison of the two approaches illustrates a slight improvement with DF-MCD over the other approach in the case of the real data set (particularly with WR and in terms of detection accuracy) and shows a more significant improvement in the case of the semisimulated data set. This conclusion was also confirmed by a further experiment in which multiplicative noise with decreasing ENL was introduced, and both DF-MCD (in the LN version) and the technique in [45] were applied. Even when corrupting the input image pair by single-look multiplicative noise, the average detection accuracy and overall error rate of DF-MCD remained around 93% and 1%, respectively. This may suggest the robustness of the method when applied to the detection of changes in very noisy multitemporal SAR data.

DF-MCD presents an internal parameter (i.e., the index q of the Minkowski norm involved in the related MRF model) whose value has to be preliminarily chosen. Anyway, a remarkable stability of the change-detection results as functions of q was pointed out by the experiments. This suggests that a preliminary choice of a proper value for q is not critical.

The impact of the specific EM algorithm adopted within the proposed method has also been investigated. Different variants of EM were studied and experimented with using the real data set. In particular, a further version of DF-MCD, based on the standard EM technique, has also been developed. Slightly worse results were given by this version, as compared with LJ-EM (details can be found in [44]).

According to the monomodal model for the "change" statistics, only one typology of change (either an increase or a decrease in the backscattering coefficient) is assumed to be present in the image. This assumption was satisfied for the data sets used for experiments since a single type of change was included in the semisimulated data set and the ground changes in the real data set were essentially due only to the process of flooding rice fields (the time gap between the two images in this data set was two days, which made phenological changes in vegetation negligible). When both an increase and a decrease in the backscattering coefficient occur, the proposed method can be separately applied to the two possible ratio images [43], [49] (preliminary experiments, not reported here for brevity, have confirmed the effectiveness of this procedure). As an alternative, the two possible ratio images could be combined in a single normalized ratio image [12], [49], [64]. With this approach, DF-MCD could be generalized by integrating the parametric pdf developed in [64] into the adopted MRF framework to model the statistics of the normalized amplitude ratio. However, when many typologies of change are present, the adopted two-component model may lead to an incorrect merging of some "change" and "no-change" subclasses. A multiclass extension of the method should be developed to allow multiple types of change to be differentiated from one another (e.g., to distinguish changes caused by a flood from those related to phenological development in vegetation in a damage-assessment application). As an alternative, a supervised extension could be developed that would use training data about the land covers from each observation date to associate each of these multiple typologies of change to a given land-cover temporal transition [11].

Image ratioing implicitly assumes that the images acquired on the two observation dates are from the same sensor, thus avoiding potential differences in preprocessing, calibration, and normalization between distinct sensors that may affect the ratio image (such differences would likely result in an increased false-alarm rate). This is an intrinsic limitation of DF-MCD and other change-detection approaches involving ratioing or log-ratioing. To mitigate this limitation, the technique in [40] could be used to generate a transformed image, one that emphasizes the distinction between changed and unchanged areas, from a multitemporal pair of heterogeneous images (e.g., images collected by different sensors or with different

acquisition modalities). In order to integrate DF-MCD with this approach, suitable parametric models should be developed for the statistics of the pixel intensities in this transformed image.

Future extensions of this work could aim at integrating more sophisticated MRF models to further reduce the error rates. For example, image nonstationarity could be accounted for [17]; edge information could be included [23]; the proposed method could be combined with object-based approaches [24] to better adapt them for application to very high-resolution images; DF-MCD could be integrated with adaptive split-based change-detection approaches for application to very large images [10] or with multispectral data and/or other (e.g., ancillary) information sources (by extending the related MRF-based data-fusion approach); and the initialization of DF-MCD could be generalized by performing a decision fusion of the change maps obtained by separately processing the single ratio channels. It would also be worth experimentally investigating the possible correlation (suggested by the similarities between the proposed approach and the consensus-theoretic framework) between the values of the reliability factors obtained at the convergence of DF-MCD and the capabilities of the single channels to discriminate changed and unchanged areas.

APPENDIX
PROOFS RELATED TO DF-MCD

The mode-field EM algorithm expresses the parameter estimation problem as the iterative maximization of a modified log-likelihood function [14]. Adopting the notations in Section II, the "mode-field" EM algorithm for the MRF model related to DF-MCD computes the updated parameter vector θ^{t+1} at the tth iteration by maximizing, with respect to θ, the following pseudolikelihood function $(t = 0, 1, 2, \ldots)$ [14]:

$$Q(\theta|\theta^t) = \sum_{k=1}^{N}\sum_{i=0}^{1} w_{ik}^t \times \left[\ln P\left\{\ell_k = H_i|\mathcal{C}_k^t, \beta\right\} + \sum_{r=1}^{n} \alpha_r \ln p_{ir}(u_{kr}|\xi_{ir})\right] \quad (13)$$

where $w_{ik}^t = P\{\ell_k = H_i|u_k, \mathcal{C}_k^t, \theta^t\}$. Owing to the Markovianity of the label configuration, the local prior probability distribution $P\{\ell_k = H_i|\mathcal{C}_k^t, \beta\}$ of the kth pixel label ℓ_k conditioned to its context is proved to be $(k = 1, 2, \ldots, N)$ [19]

$$P\left\{\ell_k = H_i|\mathcal{C}_k^t, \beta\right\} = \frac{\exp\left(\beta m_{ik}^t\right)}{\sum_{j=0}^{1}\exp\left(\beta m_{jk}^t\right)}. \quad (14)$$

Hence, we have

$$Q(\theta|\theta^t) = \Phi^t(\beta) + \sum_{r=1}^{n}\alpha_r\sum_{i=0}^{1}\Phi_{ir}^t(\xi_{ir}) \quad (15)$$

where

$$\begin{cases} \Phi^t(\beta) = \sum_{k=1}^{N}\left[\beta\sum_{i=0}^{1} w_{ik}^t m_{ik}^t - \ln\sum_{i=0}^{1}\exp\left(\beta m_{ik}^t\right)\right] \\ \Phi_{ir}^t(\xi_{ir}) = \sum_{k=1}^{N} w_{ik}^t \ln p_{ir}(u_{kr}|\xi_{ir}). \end{cases} \quad (16)$$

The updated parameter vector θ^{t+1} is obtained by maximizing the function in (15) subject to the constraint in (7) on $\alpha_1, \alpha_2, \ldots, \alpha_n$, to a positivity constraint on β, and to possible further constraints on each parameter vector ξ_{ir} (e.g., in the case of NR, both components of $\xi_{ir} = (L_{ir}, \gamma_{ir})$ are constrained to be positive). Since only the first term $\Phi^t(\beta)$ depends on β, the expression in (9) for the updated spatial parameter β^{t+1} immediately follows. Concerning the pdf parameters, since ξ_{ir} is included only in the term $\Phi_{ir}^t(\xi_{ir})$, which is weighted in (15) by the positive factor α_r, the mode-field EM algorithm updates ξ_{ir} as follows $(r = 1, 2, \ldots, n; i = 0, 1)$:

$$\xi_{ir}^{t+1} = \arg\max_{\xi_{ir}} \Phi_{ir}^t(\xi_{ir}). \quad (17)$$

If the LN model is used, then $\xi_{ir} = (\mu_{ir}, \sigma_{ir})$; this maximization problem is fairly a standard one in the EM literature and yields $(r = 1, 2, \ldots, n; i = 0, 1)$ [53]

$$\begin{cases} \mu_{1ir}^{t+1} = \frac{\sum_{k=1}^{N} w_{ik}^t \ln u_{kr}}{\sum_{k=1}^{N} w_{ik}^t} \\ \sigma_{2ir}^{t+1} = \left[\frac{\sum_{k=1}^{N} w_{ik}^t \left(\ln u_{kr} - \mu_{1ir}^{t+1}\right)^2}{\sum_{k=1}^{N} w_{ik}^t}\right]^{1/2}. \end{cases} \quad (18)$$

Since the parameters of LN are exactly the logarithmic mean and variance of the distribution, this actually proves (9) in the LN case. On the other hand, if either NR or WR is adopted, the maximization process in (17) does not yield a closed-form solution [43]. On the contrary, the MoLC approach is feasible for such parametric families; hence, we plug this approach in the EM iterative process, as described in Section II-B.

The vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ of the reliability factors is involved only in the double summation term in (15). Hence, plugging the updated pdf parameter estimates ξ_{ir}^{t+1} $(r = 1, 2, \ldots, n; i = 0, 1)$ in (15) and taking into account the definition of the c_r^t coefficients $[r = 1, 2, \ldots, n$; see (11)], the vector α^{t+1} is obtained by maximizing the following function:

$$F^t(\alpha) = \sum_{r=1}^{n}\alpha_r\sum_{i=0}^{1}\Phi_{ir}^t\left(\xi_{ir}^{t+1}\right) = \sum_{r=1}^{n} c_r^t\alpha_r \quad (19)$$

subject to the following constraint:

$$G(\alpha) = \|2\alpha - \mathbf{1}\|_q^q - 1 = \sum_{r=1}^{n}(2\alpha_r - 1)^q - 1 \le 0. \quad (20)$$

The domain $\Omega = \{\alpha \in \mathbb{R}^n : G(\alpha) \le 0\}$ defined by the constraint is a compact subset of $\mathbb{R}^n$, and F^t is continuous on Ω: Hence, a global constrained maximum point of F^t exists, owing to the Weierstrass theorem [55]. Both F^t and G are differentiable, and F^t has a constant and (in general) nonzero gradient. Hence, the constrained maximum point will lie at the boundary of Ω, i.e., it will satisfy the equality condition $G(\alpha) = 0$. According to the Lagrange multiplier theorem, a necessary condition for α to be a solution of this constrained maximization problem is expressed as

$$\frac{\partial F^t(\alpha)}{\partial\alpha_r} + \lambda\frac{\partial G(\alpha)}{\partial\alpha_r} = 0, \qquad r = 1, 2, \ldots, n \quad (21)$$

where λ is a Lagrange multiplier. Therefore

$$c_r^t + \lambda \cdot 2q(2\alpha_r - 1)^{q-1} = 0, \qquad r = 1, 2, \ldots, n \tag{22}$$

and since $q \geq 2$ is an even integer number (and so, $(q-1)$ is odd), we obtain

$$2\alpha_r - 1 = -\sqrt[q-1]{\frac{c_r^t}{2\lambda q}}, \qquad r = 1, 2, \ldots, n. \tag{23}$$

Plugging this result in the equality constraint, we have

$$\begin{aligned} G(\alpha) &= \frac{1}{|2\lambda q|^{q/(q-1)}} \sum_{r=1}^{n} |c_r^t|^{q/(q-1)} - 1 \\ &= \left(\frac{\|c^t\|_{q'}}{|2\lambda q|} \right)^{q'} - 1 = 0 \Longrightarrow |2\lambda q| = \|c^t\|_{q'}. \end{aligned} \tag{24}$$

This yields two solutions, α^+ and α^-, i.e.,

$$\alpha_r^{\pm} = \frac{1}{2} \pm \frac{1}{2} \sqrt[q-1]{\frac{c_r^t}{\|c^t\|_{q'}}}, \qquad r = 1, 2, \ldots, n. \tag{25}$$

They correspond to local extrema of F^t subject to the constraint defined by G. In order to choose the maximum point, we substitute (25) into (19), which yields, after standard algebraic calculations

$$F^t(\alpha^{\pm}) = \frac{1}{2} \left(\sum_{r=1}^{n} c_r^t \pm \|c^t\|_{q'} \right). \tag{26}$$

Hence, α^+ is the solution corresponding to the maximum constrained value of F^t, and the reliability factors are updated as in (10).

The use of LJ-EM instead of the standard EM algorithm does not directly affect this analytical formulation. The parameter-update equations of LJ-EM are obtained by modifying the EM weights w_{ik}^t as defined in Section II-B (i.e., by setting $w_{ik}^t = 0$ if the kth pixel is not assigned to H_i at the tth iteration) and by letting the EM equations be unchanged [28]. This modification is aimed at reducing the overlapping between the two hypotheses and gives rise to the final formulation in (9).

Note that the aforementioned calculations were feasible, owing to the differentiability of both the objective function F^t and the constraint G. The latter holds after the replacement of the ∞-norm with the q-norm. Furthermore, the ∞-norm constraint in (6) means that $0 \leq \alpha_r \leq 1$ for $r = 1, 2, \ldots, n$, i.e., α belongs to the n-dimensional unitary hypercube $[0, 1]^n$. Since F^t is linear, if the ∞-norm constraint were used, a linear programming problem would result, and the corresponding solution would be, in general, one of the vertices of the hypercube. Hence, binary values of the reliability factors would be obtained. This is not formally forbidden but is intuitively quite undesired from a data-fusion viewpoint, because this solution does not actually fuse the information conveyed by the n ratio channels but simply selects a subset of channels while discarding the remaining ones. A specific experiment (not reported for brevity) was performed to confirm this drawback of using the ∞-norm constraint (details can be found in [44]).

ACKNOWLEDGMENT

The authors would like to thank Dr. P. Gamba from the University of Pavia (Italy) for providing the SIR-C/XSAR data set used in the experiments, A. Sciarrone for his help with the implementation, and the anonymous reviewers for their constructive comments and criticism.

REFERENCES

[1] L. Alparone, B. Aiazzi, S. Baronti, F. Nencini, and A. Garzelli, "Robust change analysis of SAR data through information-theoretic multitemporal features," in *Proc. IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, pp. 3883–3886.

[2] Y. Bazi, L. Bruzzone, and F. Melgani, "Change detection in multitemporal SAR images based on generalized Gaussian distribution and EM algorithm," in *Proc. SPIE Conf. Image Signal Process. Remote Sens. X*, Gran Canaria, Spain, 2004, pp. 364–375.

[3] 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.

[4] Y. Bazi, L. Bruzzone, and F. Melgani, "Automatic identification of the number and values of decision thresholds in the log-ratio image for change detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 3, pp. 349–353, Jul. 2006.

[5] J. Benediktsson and I. Kanellopoulos, "Classification of multisource and hyperspectral data based on decision fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1367–1377, May 1999.

[6] J. Benediktsson, J. R. Sveinsson, and P. H. Swain, "Hybrid consensus theoretic classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 833–843, Jul. 1997.

[7] J. Benediktsson and P. H. Swain, "Consensus theoretic classification method," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, no. 4, pp. 688–704, Jul./Aug. 1992.

[8] G. Boni, F. Castelli, L. Ferraris, N. Pierdicca, S. B. Serpico, and F. Siccardi, "High resolution COSMO/SkyMed SAR data analysis for civil protection from flooding events," in *Proc. IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, pp. 6–9.

[9] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2963–2972, Dec. 2005.

[10] F. Bovolo and L. Bruzzone, "A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1658–1670, Jun. 2007.

[11] L. Bruzzone, D. F. Prieto, and S. B. Serpico, "A neural-statistical approach to multitemporal and multisource remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1350–1359, May 1999.

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

[13] C. Carincotte, S. Derrode, and S. Bourennane, "Unsupervised change detection on SAR images using fuzzy hidden Markov chains," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 432–441, Feb. 2006.

[14] G. Celeux, F. Forbes, and N. Peyrand, "EM procedures using mean field-like approximations for Markov model-based image segmentation," *Pattern Recognit.*, vol. 36, no. 1, pp. 131–144, Jan. 2003.

[15] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 4–19, Jan. 2003.

[16] R. Cossu, S. Chaudhuri, and L. Bruzzone, "A context-sensitive Bayesian technique for the partially supervised classification of multitemporal images," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 3, pp. 352–356, Jul. 2005.

[17] X. Descombes, M. Sigelle, and F. Preteux, "Estimating Gaussian Markov random field parameters in a nonstationary framework: Application to remote sensing imaging," *IEEE Trans. Image Process.*, vol. 8, no. 4, pp. 490–503, Apr. 1999.

[18] W. Dierking and H. Skriver, "Change detection for thematic mapping by means of airborne multitemporal polarimetric SAR imagery,"

IEEE Trans. Geosci. Remote Sens., vol. 40, no. 3, pp. 618–636, Mar. 2002.

[19] R. C. Dubes and A. K. Jain, "Random field models in image analysis," *J. Appl. Stat.*, vol. 16, no. 2, pp. 131–163, 1989.

[20] J. E. S. Fransson, F. Walter, K. Blennow, A. Gustavsson, and L. M. H. Ulander, "Detection of storm-damaged forested areas using airborne CARABAS-II VHF SAR image data," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2170–2175, Oct. 2002.

[21] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. New York: Academic, 1990.

[22] P. Gamba, F. Dell'Acqua, and G. Lisini, "Change detection of multitemporal SAR data in urban areas combining feature-based and pixel-based techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2820–2827, Oct. 2006.

[23] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.

[24] G. G. Hazel, "Object-level change detection in spectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 553–561, Mar. 2001.

[25] 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, vol. 6, pp. 4104–4106.

[26] 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.

[27] Q. Jackson and D. Landgrebe, "Adaptive Bayesian contextual classification based on Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2454–2463, Nov. 2002.

[28] Q. Jackson and D. A. Landgrebe, "An adaptive classifier design for high-dimensional data analysis with a limited training data set," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 12, pp. 2664–2679, Dec. 2001.

[29] E. S. Kasischke, J. M. Melack, and M. C. Dobson, "The use of imaging radars for ecological applications—A review," *Remote Sens. Environ.*, vol. 59, no. 2, pp. 141–156, Feb. 1997.

[30] P. R. Kersten, J. S. Lee, and T. L. Ainsworth, "A comparison of change detection statistics in POLSAR images," in *Proc. IGARSS*, Seoul, Korea, Jul. 25–29, 2005, pp. 4836–4839.

[31] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, no. 1, pp. 41–47, Jan./Feb. 1986.

[32] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. Hoboken, NJ: Wiley-InterScience, 2003.

[33] X. Li and A. G. Yeh, "Multitemporal SAR images for monitoring cultivation systems using case-based reasoning," *Remote Sens. Environ.*, vol. 90, no. 4, pp. 524–534, Apr. 2004.

[34] P. Lombardo, C. J. Oliver, T. Macrì Pellizzeri, and M. Meloni, "A new maximum-likelihood joint segmentation technique for multitemporal SAR and multiband optical images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2500–2518, Nov. 2003.

[35] P. Lombardo and T. Macrì 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.

[36] F. Melgani, "Classification of multitemporal remote-sensing images by a fuzzy fusion of spectral and spatio-temporal contextual information," *Int. J. Pattern Recogn. Artif. Intell.*, vol. 18, no. 2, pp. 143–156, 2002.

[37] F. Melgani, G. Moser, and S. B. Serpico, "Unsupervised change detection methods for remote sensing images," *Opt. Eng.*, vol. 41, no. 12, pp. 3288–3297, Dec. 2002.

[38] F. Melgani and S. B. Serpico, "A statistical approach to the fusion of the spectral and spatio-temporal contextual information for the classification of remote sensing images," *Pattern Recognit. Lett.*, vol. 23, no. 9, pp. 1053–1061, Jul. 2002.

[39] G. Mercier and S. Derrode, "SAR image change detection using distance between distributions of classes," in *Proc. IGARSS*, Anchorage, AK, 2004, vol. 6, pp. 3872–3875.

[40] G. Mercier, G. Moser, and S. B. Serpico, "Conditional copula for change detection on heterogeneous data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1428–1441, May 2008.

[41] M. Molinier, J. Laaksonen, Y. Rauste, and T. Hame, "Detecting changes in polarimetric SAR data with content-based image retrieval," in *Proc. IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, pp. 2390–2393.

[42] M. Molinier and Y. Rauste, "Comparison and evaluation of polarimetric change detection techniques in aerial SAR data," in *Proc. IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, pp. 2386–2389.

[43] G. Moser and S. B. Serpico, "Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2972–2982, Oct. 2006.

[44] G. Moser and S. B. Serpico, "Unsupervised change detection by multichannel SAR data fusion," in *Proc. IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, pp. 4854–4857.

[45] G. Moser, S. B. Serpico, and G. Vernazza, "Unsupervised change detection from multichannel SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 278–282, Apr. 2007.

[46] G. Moser, J. Zerubia, and S. B. Serpico, "Dictionary-based stochastic expectation-maximization for SAR amplitude probability density function estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 1, pp. 188–200, Jan. 2006.

[47] G. Moser, J. Zerubia, and S. B. Serpico, "SAR amplitude probability density function estimation based on a generalized Gaussian model," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1429–1442, Jun. 2006.

[48] J.-M. Nicolas, "Introduction aux statistiques de deuxième espéce: Applications des log-moments et des log-cumulants a l'analyse des lois d'images radar," *Trait. Signal*, vol. 19, pp. 139–167, 2002.

[49] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. Norwood, MA: Artech House, 1998.

[50] W. H. Press, S. A. Teukolsky, W. T. Wetterling, and B. P. Flannery, *Numerical Recipes in C*. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[51] M. Qong, "Polarization state conformation and its application to change detection in polarimetric SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 304–308, Oct. 2004.

[52] K. I. Ranney and M. Soumekh, "Signal subspace change detection in averaged multilook SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 1, pp. 201–213, Jan. 2006.

[53] R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood, and the EM algorithm," *SIAM Rev.*, vol. 26, no. 2, pp. 195–239, Apr. 1984.

[54] 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.

[55] W. Rudin, *Principles of Mathematical Analysis*, 2nd ed. New York: McGraw-Hill, 1976.

[56] W. Rudin, *Functional Analysis*, 2nd ed. New York: McGraw-Hill, 1990.

[57] S. B. Serpico, L. Bruzzone, and F. Roli, "An experimental comparison of neural and statistical non-parametric algorithms for supervised classification of remote sensing images," *Pattern Recognit. Lett.*, vol. 17, no. 13, pp. 1331–1341, Nov. 1996.

[58] S. B. Serpico and G. Moser, "MRF-based remote-sensing image classification with automatic model-parameter estimation," in *Signal and Image Processing for Remote Sensing*, C. H. Chen, Ed. New York: Taylor & Francis, 2006, pp. 305–326.

[59] F. Siegert and A. A. Hoffmann, "The 1998 forest fires in East Kalimantan (Indonesia): A quantitative evaluation using high resolution ERS-2 SAR images and NOAA AVHRR Hotspot data," *Remote Sens. Environ.*, vol. 72, no. 1, pp. 64–77, Apr. 2000.

[60] A. Singh, "Digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, 1989.

[61] I. Sneddon, *The Use of Integral Transforms*. New York: McGraw-Hill, 1972.

[62] A. H. S. Solberg, T. Taxt, and A. K. Jain, "A Markov random field model for classification of multisource satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 1, pp. 100–113, Jan. 1996.

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

[64] R. Touzi, A. Lopez, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 6, pp. 764–773, Nov. 1988.

[65] H. L. V. Trees, *Detection, Estimation and Modulation Theory*, vol. 1. New York: Wiley, 1968.

[66] Y. C. Tzeng, S. H. Chiu, D. Chen, and K. S. Chen, "Change detections from SAR images for damage estimation based on a spatial chaotic model," in *Proc. IGARSS*, Barcelona, Spain, Jul. 23–27, 2007, pp. 1926–1930.

[67] R. G. White, "Change detection in SAR imagery," *Int. J. Remote Sens.*, vol. 12, no. 2, pp. 339–360, 1990.

[68] C. F. J. Wu, "On the convergence properties of the EM algorithm," *Ann. Stat.*, vol. 11, no. 1, pp. 95–103, 1983.

Gabriele Moser (S'03–M'05) received the Laurea (M.S.) degree (*summa cum laude*) in telecommunications engineering and the Ph.D. degree in space sciences and engineering from the University of Genoa, Genoa, Italy, in 2001 and 2005, respectively.

Since 2001, he has been a Research Fellow with the Signal Processing and Telecommunications Research Group, Department of Biophysical and Electronic Engineering, University of Genoa, working in the field of remote-sensing image analysis. From January to March 2004, he was a Visiting Student with the Institut National de Recherche en Informatique et en Automatique, Sophia Antipolis, France, working with the "Ariana" research group on the problem of SAR data modeling. His research activity is focused on image-processing and image-analysis methodologies for remote-sensing data interpretation. In particular, his current research interests include SAR data analysis, multitemporal image classification, hyperspectral image analysis, contextual classification, and geo/biophysical parameter estimation. He has been a reviewer for several international journals.

Dr. Moser has been an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS since 2008.

Sebastiano B. Serpico (M'87–SM'00–F'08) received the Laurea degree in electronic engineering and the Ph.D. degree in telecommunications from the University of Genoa, Genoa, Italy, in 1982 and 1989, respectively.

Since 1982, he has been with the Department of Biophysical and Electronic Engineering (DIBE), University of Genoa, working in the field of image processing and recognition. He was an Assistant Professor, from 1990 to 1998, and an Associate Professor of telecommunications, from 1998 to 2004, with the Faculty of Engineering, University of Genoa, where he taught signal theory, pattern recognition, telecommunication systems, and electrical communications. He is currently a Full Professor of telecommunications. From 1995 to 1998, he was the Head of the Signal Processing and Telecommunications Research Group (SP&T), DIBE, and currently, he is the Head of the SP&T Laboratory. He is currently the Chairman of the Institute of Advanced Studies in Information and Communication Technologies. His current research interests include the application of pattern recognition (feature selection, classification, change detection, and data fusion) to remotely sensed images. He is the author or coauthor of more than 150 scientific publications, including journals and conference proceedings.

Dr. Serpico is a member of the International Association for Pattern Recognition Society. Since 2001, he has been an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He coedited a Special Issue of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING on the subject of the analysis of hyperspectral image data (July 2001) and a Special Issue on advances in techniques for analysis of remotely sensed data (March 2005). He was the recipient of the Recognition of TGARS Best Reviewers from the IEEE Geoscience and Remote Sensing Society in 1998.