# A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in the Polar Domain

Francesca Bovolo, *Student Member, IEEE*, and Lorenzo Bruzzone, *Senior Member, IEEE*

***Abstract*—This paper addresses unsupervised change detection by proposing a proper framework for a formal definition and a theoretical study of the change vector analysis (CVA) technique. This framework, which is based on the representation of the CVA in polar coordinates, aims at: 1) introducing a set of formal definitions in the polar domain (which are linked to the properties of the data) for a better general description (and thus understanding) of the information present in spectral change vectors; 2) analyzing from a theoretical point of view the distributions of changed and unchanged pixels in the polar domain (also according to possible simplifying assumptions); 3) driving the implementation of proper preprocessing procedures to be applied to multitemporal images on the basis of the results of the theoretical study on the distributions; and 4) defining a solid background for the development of advanced and accurate automatic change-detection algorithms in the polar domain. The findings derived from the theoretical analysis on the statistical models of classes have been validated on real multispectral and multitemporal remote sensing images according to both qualitative and quantitative analyses. The results obtained confirm the interest of the proposed framework and the validity of the related theoretical analysis.**

***Index Terms*—Change detection, change vector analysis (CVA), multitemporal images, polar representation, remote sensing, spherical representation, statistical models, unsupervised techniques.**

## I. Introduction

UNSUPERVISED change detection plays an important role in many application domains related to the exploitation of multitemporal remote sensing images. The availability of images acquired on the same geographical area by satellite sensors at different times makes it possible to identify and label possible changes that have occurred on the ground. In this context, in order to properly exploit the huge amount of data acquired by current remote sensing satellites, it is mandatory to develop effective unsupervised and automatic change-detection techniques.

Several unsupervised change-detection methodologies have been proposed in the literature [1]–[3]. Among them, a widely used technique is change vector analysis (CVA). CVA is typically applied to multispectral images acquired by passive sensors, by considering more than one spectral channel in order to exploit all the available information about the considered event of change. However, usually CVA is used in an empirical way without referring to a specific theoretical framework capable to properly and formally represent all the information contained in the spectral change vectors (SCVs) obtained by subtracting corresponding spectral bands of two images acquired at different dates. In addition, in most of the applications, only the magnitude of the SCVs is exploited in order to identify changed pixels. Only in a few applications, also the direction of the vector is empirically used for deriving information on the kind of change that occurred on the ground [4]–[13]. This lack of a formal framework and of a proper analysis of the statistics of data results in suboptimal applications of the CVA or, in some cases, in a noncomplete understanding of the richness of the information present in SCVs. This involves incomplete exploitation of all the available information and/or the definition of change-detection algorithms that are not based on a solid theoretical background and on proper analysis procedures.

In order to address the aforementioned problems, in this paper, we present a consistent theoretical framework for proper representation, modeling, and exploitation of the information present in the SCVs computed according to the CVA technique. The proposed framework and the related analysis are developed in the context of a polar representation of the CVA. In particular, the proposed novel contributions of this paper consist in: 1) the introduction of formal definitions for the characterization of the information present in SCVs; 2) a theoretical analysis on the distributions of changed and unchanged pixels in the polar domain under both general conditions and proper simplifying assumptions; 3) the introduction of proper guidelines for defining effective preprocessing strategies based on the expected properties of the theoretical distributions of changed and unchanged pixels; and 4) the definition of a solid background for the development of advanced and accurate automatic change-detection algorithms in the polar domain. A validation of the theoretical analysis, carried out on two multispectral and multitemporal data sets, is also reported. The first data set is made up of two real multitemporal remotely sensed images including only one kind of change, while the second data set is obtained from the first one by properly simulating a second kind of change.

This paper is organized into seven sections. The next section introduces the change-detection problem and the formulation of the CVA technique in both Cartesian and hyperspherical domains. Section III gives some basics on the models for joint conditional class distributions in both Cartesian and polar coordinate systems. Section IV presents the proposed theoretical analysis on the models of marginal conditional distributions

Manuscript received January 27, 2006; revised July 19, 2006.

The authors are with the Department of Information and Communication Technologies, University of Trento, 38050 Trento, Italy (e-mail: lorenzo.bruzzone@ing.unitn.it).

Digital Object Identifier 10.1109/TGRS.2006.885408

of magnitude and direction; furthermore, it proposes a critical analysis on the importance and effects of image-preprocessing procedures (e.g., radiometric corrections, coregistration, etc.) on the data distributions. The validation of the proposed theoretical analysis carried out on single-change and double-change data sets is reported in Sections V and VI, respectively. Finally, Section VII discusses the obtained results and draws the conclusions of this paper.

## II. Proposed Polar-Representation Framework for CVA

### A. Background and CVA Formulation

Let us consider two coregistered multispectral images, $\mathbf{X}_1$ and $\mathbf{X}_2$, of size $I \cdot J$ acquired over the same area at different times $t_1$ and $t_2$.[1] Let $X_1$ and $X_2$ be two multidimensional random variables that represent the statistical distributions of pixels in images $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively. Let $X_{b,t}$ be the random variable representing the $b$th component of the multispectral image $\mathbf{X}_t (t = 1, 2)$ in the considered feature space. Let $\Omega = \{\omega_{\rm n}, \Omega_{\rm c}\}$ be the set of classes of unchanged and changed pixels to be identified. In greater detail, $\omega_{\rm n}$ represents the class of unchanged pixels and $\Omega_{\rm c} = \{\omega_{c_1}, \ldots, \omega_{c_K}\}$ is the set of the $K$ possible classes (kinds) of change occurred in the considered area.

The first step of the most widely used change-detection techniques presented in the literature performs comparison between the two considered images according to a proper operator [1]. When dealing with multispectral images, the comparison operator is usually the vector difference, which is applied to a $n$-dimensional feature space in order to give as input to the change-detection process all the relevant spectral information. This technique is known as CVA [1], [5] and has been successfully used in many different application domains [4]–[13].[2] CVA first computes a multispectral difference image $(\mathbf{X}_{\rm D})$ subtracting the spectral feature vectors associated with each corresponding spatial position in the two considered images $\mathbf{X}_1$ and $\mathbf{X}_2$. Let $X_{\rm D}$ be the multidimensional random variable representing the SCVs in the difference image obtained as follows [1]:

$$\mathbf{X}_{\rm D} = \mathbf{X}_2 - \mathbf{X}_1. \tag{1}$$

Each SCV is usually implicitly represented in polar coordinates with its magnitude and direction. Although the direction of the SCVs is rich of information (e.g., on the kind of changes occurred on the ground and on the distribution of registration noise), in the most of the applications it is not considered. Among the few studies reported in the literature where magnitude and direction expressed as cosine functions are considered together for change detection, we recall [5]–[13]. In 1980, Malila [5] first formulated the concept of change vector and, then, used both magnitude and direction in a two-dimensional (2-D) space for identifying changes due to plants' clearcut and regrowth in the northern Idaho (U.S.) forest. In [5]–[8], the direction variable was subdivided into a fixed number of sectors, each of them corresponding to positive or negative changes in one of the $B$ considered features (i.e., spectral channels or linear combinations of them, like tasseled-cap transformation). This kind of quantization leads to the definition of a maximum of $2^B$ sectors and, hence, of types of changes. The major drawback of this approach is that different kinds of changes could assume the same sector code. In [9], the CVA sector-coding approach was extended to the solution of multivariate, full-dimensional, and also multi-interval problems (i.e., applications involving more than two acquisition dates). In [10], Allen and Kupfer introduced in the CVA technique the use of direction cosines for the description of SCV directions. They applied a hierarchical linear-discriminant analysis for testing predictive power of magnitude and vector angles in solving change-detection problems. Direction cosines were used also in [11] and [12]. In these works, the authors first identified changed pixels on the basis of magnitude values, then image-classification algorithms were applied to direction cosines for discriminating the different kinds of change. In [13], Nackarets *et al.* defined a modified CVA technique, where polar coordinates are transformed back into a Cartesian coordinate system to overcome discontinuity between 0 and $2\pi$ and different kinds of changes are then detected by using either supervised or unsupervised-clustering algorithms. A different approach to the use of the direction information has been presented in [4], where Bruzzone and Cossu proposed a method for estimating and reducing the effects of the registration noise. The method is based on a joint exploitation of the magnitude and direction components.

However, most of the analyses reported in the literature have been carried out in an empirical way, without a rigorous characterization of the statistical models of information classes and without referring to a proper theoretical framework for complete understanding and proper processing of the information present in SCVs. In this paper, in order to fill the aforementioned gaps, we propose a rigorous framework for CVA in the polar domain.

### B. General Framework for Hyperspherical Representation

Given a $B$-dimensional feature space, the SCV associated with a pixel of the analyzed scene can be described with its magnitude value and $B - 1$ directions. In this paper, we propose to represent the properties of the SCVs, instead of using a Cartesian coordinate system, by plotting SCVs in a $B$-dimensional hyperspherical coordinate system.[3] Thus, the multidimensional random variable $X_{\rm D}$ can be represented with a random variable $\rho \in [0, +\infty)$ that models the statistical distribution of the change vector magnitude, and $B - 1$ random variables $[\vartheta, \varphi_1, \varphi_2, \ldots, \varphi_{B-2}]$ that represent the distribution of the change vector angular coordinates ($\vartheta \in [0, 2\pi)$ and $\varphi_k \in [0, \pi], k = 1, \ldots, B - 2$). It is worth noting that changing the coordinate system has a dramatic impact on the statistical distributions of the considered classes. This aspect will be analyzed in the next section.

[1] In this paper, only the case of pairs of images is discussed. It is worth noting that the proposed framework can be applied to a multitemporal sequence made up of more than two images by analyzing separately couples of images.

[2] The particular case of working with a single feature reduces the CVA to the univariate image-differencing technique [1].

[3] In the particular case of $B = 2$ the hyperspherical coordinate system is said polar-coordinate system.

Let $X_{1,\mathrm{D}},\ldots,X_{B,\mathrm{D}}$ be the random variables representing the distributions of SCVs along the $B$ dimensions (spectral channels) of the considered Cartesian coordinate system; then, the relations between the random variables modeling SCVs in the Cartesian and in the hyperspherical coordinates are the following:

$$\begin{aligned} X_{b,\mathrm{D}} &= \rho\left(\prod_{k=1}^{b-1}\sin\varphi_k\right)\cos\varphi_b, \qquad 1\le b\le B-2 \\ X_{B-1,\mathrm{D}} &= \rho\left(\prod_{k=1}^{B-2}\sin\varphi_k\right)\cos\vartheta \\ X_{B,\mathrm{D}} &= \rho\left(\prod_{k=1}^{B-2}\sin\varphi_k\right)\sin\vartheta \end{aligned} \tag{2}$$

where $\rho\in[0,+\infty)$, $\varphi_k\in[0,\pi], b=1,\ldots,B-2$, and $\vartheta\in[0,2\pi)$.

In the following, for simplicity, we will assume that the CVA technique is applied only to two spectral channels of the considered multitemporal images, i.e., that a 2-D coordinate system is sufficient to completely describe SCVs. However, the analysis can be generalized to the case of more spectral channels by considering more direction contributions for describing each SCV (see Appendix). It is worth noting that, although the aforementioned generalization is possible, the assumption of working with a couple of spectral channels is reasonable in many change-detection problems [14]–[16]. This choice is often due to the need of isolating the most informative features with respect to the specific considered problem without including noisy and misleading spectral channels in the analysis. In the above assumption, in the Cartesian coordinate system, only random variables $X_{1,\mathrm{D}}$ and $X_{2,\mathrm{D}}$ are necessary to describe SCVs, whereas in the polar-coordinate-system random variables representing the magnitude $\rho$ and the direction $\vartheta$ are required for each SCV. The relation between Cartesian and polar representation of the difference image is as follows:

$$\begin{cases} \rho = \sqrt{(X_{1,\mathrm{D}})^2+(X_{2,\mathrm{D}})^2} \\ \vartheta = \arctan\left(\frac{X_{2,\mathrm{D}}}{X_{1,\mathrm{D}}}\right). \end{cases} \tag{3}$$

### C. Proposed Polar Framework for the CVA Technique: Definitions

In this section, we propose a rigorous characterization of the polar framework for the CVA technique. First of all, observe that in the polar representation, all the change vectors of a given scene are included in a magnitude-direction domain MD defined as (see Fig. 1)

$$\mathrm{MD} = \{\rho\in[0,\rho_{\max}] \text{ and } \vartheta\in[0,2\pi)\} \tag{4}$$

where $\rho_{\max}$ is the maximum value assumed by the magnitude on the considered image, i.e.,

$$\rho_{\max} = \max\left\{\sqrt{(X_{1,\mathrm{D}})^2+(X_{2,\mathrm{D}})^2}\right\}. \tag{5}$$

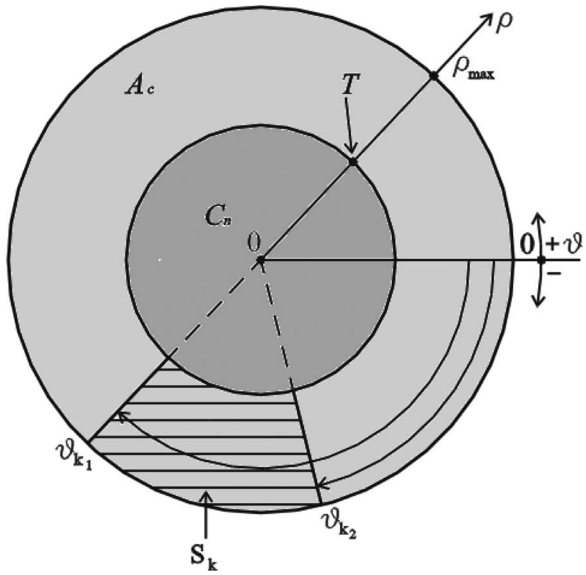



Fig. 1. Representation of the regions of interest for the CVA technique in the polar-coordinate system.

According to the given definition of the magnitude-direction domain MD, in order to establish a clear framework for CVA, we propose to identify different regions in MD for pointing out the information present in SCVs. From the definition in (4) and following indications in [16], we expect that unchanged pixels have magnitude close to zero (often not exactly zero due to the presence of noise components), while changed pixels have magnitude far from zero. Consequently, it is possible to identify two different regions associated with: 1) unchanged and 2) changed pixels. Thus, the polar domain can be split into two parts: 1) circle $C_{\mathrm{n}}$ of no-changed pixels and 2) annulus $A_{\mathrm{c}}$ of changed pixels. This can be done according to the optimal (in the sense of the theoretical Bayesian decision theory) threshold value $T$ that separates pixels belonging to $\omega_{\mathrm{n}}$ from pixel belonging to $\Omega_{\mathrm{c}}$ (dark gray and light gray areas in Fig. 1, respectively).

*Definition 1:* The circle of no-changed pixels $C_{\mathrm{n}}$ is defined as

$$C_{\mathrm{n}} = \{\rho,\vartheta : 0\le\rho<T \text{ and } 0\le\vartheta<2\pi\}. \tag{6}$$

$C_{\mathrm{n}}$ can be represented in the polar domain as a circle with radius $T$. From this definition, we can state that for the generic pixel (spatial position) $(i,j)$, it holds that

$$(i,j)\in\omega_{\mathrm{n}} \quad\Leftrightarrow\quad (i,j)\in C_{\mathrm{n}} \tag{7}$$

or in other words

$$(i,j)\in\omega_{\mathrm{n}} \quad\Leftrightarrow\quad \rho(i,j)<T. \tag{8}$$

This means that the unchanged pixels satisfy (7) [or equivalently (8)] and are included in $C_{\mathrm{n}}$.

*Definition 2:* The annulus of changed pixels $A_{\mathrm{c}}$ is defined as

$$A_{\mathrm{c}} = \{\rho, \vartheta : T \leq \rho \leq \rho_{\max} \text{ and } 0 \leq \vartheta < 2\pi\}. \quad (9)$$

$A_{\mathrm{c}}$ can be represented in the polar domain as a ring with inner radius $T$ and outer radius $\rho_{\max}$. From this definition, we can state that for the generic spatial position $(i, j)$, it holds that

$$(i,j) \in \Omega_{\mathrm{c}} \quad \Leftrightarrow \quad (i,j) \in A_{\mathrm{c}} \quad (10)$$

or in other words

$$(i,j) \in \Omega_{\mathrm{c}} \quad \Leftrightarrow \quad T \leq \rho(i,j) \leq \rho_{\max}. \quad (11)$$

This means that all the changed pixels satisfy (10) [or equivalently (11)] and are included in $A_{\mathrm{c}}$.

According to the above definitions, the polar domain can be described as the union of $A_{\mathrm{c}}$ and $C_{\mathrm{n}}$, i.e.,

$$\mathrm{MD} = A_{\mathrm{c}} \cup C_{\mathrm{n}}. \quad (12)$$

The previous definitions have been based on the values of the magnitude, independently on the direction variable. A further important definition is related to sectors in the polar domain, which are mainly related to the direction of the change vectors and, therefore, to the kinds of change occurred on the ground.

*Definition 3:* The annular sector $S_k$ of change $\omega_{c_k} \in \Omega_{\mathrm{c}}$ is defined as

$$S_k = \{\rho, \vartheta : \rho \geq T \text{ and } \vartheta_{k_1} \leq \vartheta < \vartheta_{k_2}, 0 \leq \vartheta_{k_1} < \vartheta_{k_2} < 2\pi\}. \quad (13)$$

$S_k$ can be represented in the polar domain as a sector of change within the annulus of changed pixels [see (9)] and bounded from two angular thresholds $\vartheta_{k_1}$ and $\vartheta_{k_2}$ (see Fig. 1). We expect that pixels that belong to the same kind of change are included in the same annular sector. In the polar-coordinate system, two angular coordinates identify two sectors: 1) a convex sector (i.e., the smallest one) and 2) a concave sector (i.e., the largest one). As it is reasonable to expect that pixels belonging to the same change class have a low variance, generally, the sector we are interested to is the convex one. It is worth noting that this condition is no longer satisfied if the convex sector covers the discontinuity between 0 and $2\pi$. In this case, the variance of the changed pixels is high and the relation between the two angular coordinates is inverted, i.e., $\vartheta_{k_1} > \vartheta_{k_2}$; thus, the definition of $S_k$ becomes

$$S_k = \{\rho, \vartheta : \rho \geq T \text{ and } \vartheta_{k_1} \leq \vartheta < 2\pi \cup 0 \leq \vartheta < \vartheta_{k_2}, \; 0 \leq \vartheta_{k_1} < \vartheta_{k_2} < 2\pi\}. \quad (14)$$

Fig. 1 depicts an example of annular sector as a hatched sector of annulus that overlaps region $A_{\mathrm{c}}$ between angular coordinates $\vartheta_{k_1}$ and $\vartheta_{k_2}$.

In real applications, often the pixels with magnitude close to the optimal (in the sense of the theoretical Bayesian decision theory) threshold value $T$ and with direction close to the two angular threshold values $\vartheta_{k_1}$ and $\vartheta_{k_2}$ cannot be accurately labeled according to a simple thresholding procedure due to the intrinsic uncertainty present in the data. In these cases, by taking into account that the spatial autocorrelation function of the images is not impulsive (i.e., pixels are spatially correlated),[4] it is possible to increase the reliability of the decision process according to a context-sensitive analysis of the investigated pixel [16]. This analysis is aimed at exploiting the spatial correlation as an additional information source in the decision process. In order to model and represent this uncertainty in the proposed framework, we can define two additional regions: 1) the annulus $A_u$ of uncertain pixels with respect to the magnitude, and 2) the annular sector $S_{\mathrm{ku}_r}$ for uncertain pixels with respect to change $\omega_{c_k} \in \Omega_{\mathrm{c}}$ and, thus, directions $\vartheta_{k_r}$ (where $r = 1, 2$).

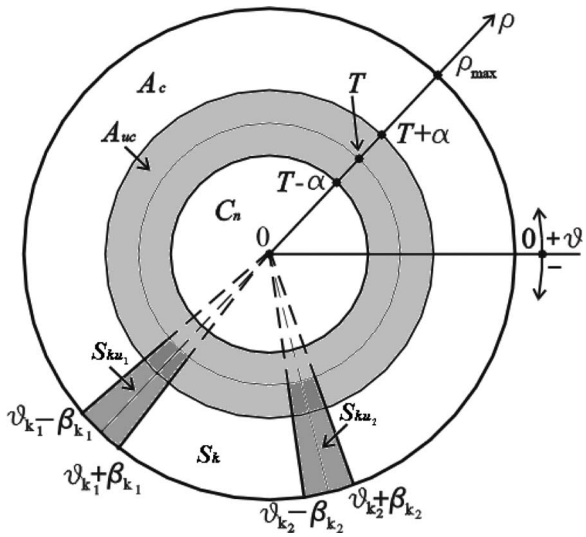



Fig. 2. Representation of the regions of uncertainty for the CVA technique in the polar-coordinate system.

*Definition 4:* The annulus $A_u$ of uncertain pixels is defined as

$$A_u = \{\rho, \vartheta : T - \alpha \leq \rho \leq T + \alpha \text{ and } 0 \leq \vartheta < 2\pi\} \quad (15)$$

where $\alpha$ is a parameter that defines the margin around $T$, in which pixels cannot be easily identified as either changed or unchanged. $A_u$ can be represented in the polar domain as a ring with inner radius $T - \alpha$ and outer radius $T + \alpha$ (light gray annulus in Fig. 2). From this definition, we can state that

$$(i,j) \in A_u \quad \Leftrightarrow \quad T - \alpha \leq \rho(i,j) \leq T + \alpha. \quad (16)$$

This means that all the uncertain pixels satisfying (16) are included in $A_u$. The use of this definition depends on the specific data-analysis strategy (if no contextual information is considered, it is assumed that $\alpha = 0$ and, consequently, $A_u = \varnothing$).

[4]This is true under the reasonable assumption that the geometrical resolution of the considered multispectral sensor is proper for the analyzed scene.

*Definition 5:* The annular sector $S_{\mathrm{ku}_r}$ of uncertain pixels of change $\omega_{c_k} \in \Omega_{\mathrm{c}}$ close to $\vartheta_{k_r}$ is defined as

$$S_{\mathrm{ku}_r} = \{\rho, \vartheta : \rho \geq T \text{ and } \vartheta_{k_r} - \beta_{k_r} \leq \vartheta < \vartheta_{k_r} + \beta_{k_r}, \quad 0 \leq \vartheta_{k_r} < 2\pi\} \qquad r = 1 \text{ or } 2 \quad (17)$$

where $\beta_{k_r}$ is a parameter that defines the margin around $\vartheta_{k_r}$, in which pixels cannot be easily identified as belonging to $\omega_{c_k}$ (dark gray annular sectors in Fig. 2). We expect that in general $\beta_{k_1} = \beta_{k_2}$ (as for the annulus $A_u$ of uncertain pixels, if no contextual information is considered, it is assumed that $\beta_{k_r} = 0$ and, consequently, $S_{\mathrm{ku}_r} = \varnothing$).

In the following of the paper, for simplicity, we will assume the absence of uncertainty, i.e, we will assume $A_u = \varnothing$ and $S_{\mathrm{ku}_r} = \varnothing (r = 1, 2)$.

## III. Analysis of the Joint Conditional Distributions of Classes

The definition of the different regions of interest in the polar domain allows a better representation of the change-detection problem and drives to the analysis of another important problem that concerns the expected distribution of classes of interest in the change-detection problem.

### A. Class Distributions in the Cartesian Domain

As known from the literature [17], the statistical distribution of natural classes in images acquired by multispectral passive sensors can be considered approximately Gaussian. Thus, both multidimensional random variables $X_1$ and $X_2$ can be modeled as a mixture of multidimensional Gaussian distributions in the Cartesian domain. As $X_{\mathrm{D}}$ is obtained subtracting $X_1$ from $X_2$, its distribution can be also reasonably represented as a mixture of multidimensional Gaussian distributions, each of them associated with a class $\omega_i$, $\omega_i \in \Omega = \{\omega_{\mathrm{n}}, \Omega_{\mathrm{c}}\} = \{\omega_{\mathrm{n}}, \omega_{c_1}, \ldots, \omega_{c_K}\}$:

$$\begin{aligned} p(X_{\mathrm{D}}) &= P(\omega_{\mathrm{n}}) p(X_{\mathrm{D}}|\omega_{\mathrm{n}}) + P(\Omega_{\mathrm{c}}) p(X_{\mathrm{D}}|\Omega_{\mathrm{c}}) \\ &= P(\omega_{\mathrm{n}}) p(X_{\mathrm{D}}|\omega_{\mathrm{n}}) + \sum_{k=1}^{K} P(\omega_{\mathrm{c_k}}) p(X_{\mathrm{D}}|\omega_{c_k}) \end{aligned} \quad (18)$$

where $p(X_{\mathrm{D}}|\omega_i)$ is a normal conditional density that models the distribution of the class $\omega_i$ in the multivariate-difference image. As classes in $X_{\mathrm{D}}$ can be approximated as jointly Gaussian distributed, it is possible to show that all the components $X_{b,\mathrm{D}}$, obtained subtracting corresponding spectral bands $(b = 1, 2)$, are also a mixture of normally distributed random variables. This consideration and the assumption in (18) are the starting point for analyzing the statistical distributions of the no-change class and of the $K$ classes of change in the polar domain.

### B. Class Distributions in the Polar Domain

From (18), it is shown that the analytical expression of class distributions in the polar domain can be obtained by computing the joint conditional-density functions of the magnitude and direction of SCVs [see (3)] of the 2-D random variable $X_{\mathrm{D}}$. A simplifying hypothesis is to consider features $X_{1,\mathrm{D}}$ and $X_{2,\mathrm{D}}$ as independent (see Section IV-C for a detailed discussion on this hypothesis). Under this assumption, the distribution of the class $\omega_i (\omega_i \in \Omega)$ in a Cartesian coordinate system can be written as the product of the two marginal densities $p(X_{b,\mathrm{D}}|\omega_i)$ of the class $\omega_i (b = 1, 2)$, i.e.,

$$p(X_{\mathrm{D}}|\omega_i) = \frac{1}{2\pi\sigma_{1,i}\sigma_{2,i}} \times \exp\left[-\frac{(X_{1,\mathrm{D}} - \mu_{1,i})^2}{2\sigma_{1,i}^2} - \frac{(X_{2,\mathrm{D}} - \mu_{2,i})^2}{2\sigma_{2,i}^2}\right] \quad (19)$$

where $\mu_{b,i}$ and $\sigma_{b,i}$ are the mean value and the standard deviation, respectively, of the Gaussian distributed marginal density of class $\omega_i$ over the $b$th considered feature $(b = 1, 2)$. Applying the transformation from Cartesian to polar-coordinate system, the joint conditional distribution can be written as

$$p(\rho, \vartheta|\omega_i) = \frac{1}{2\pi\sigma_{1,i}\sigma_{2,i}} \times \exp\left[-\frac{(\rho\cos\vartheta - \mu_{1,i})^2}{2\sigma_{1,i}^2} - \frac{(\rho\sin\vartheta - \mu_{2,i})^2}{2\sigma_{2,i}^2}\right]. \quad (20)$$

According to this general equation (which is the starting point for the statistical analysis reported in the next section), we can define the quantities typically used for evaluating the performance of change-detection algorithms (i.e., false alarms and missed alarms) in the context of the proposed framework. Let us define the following decision regions: 1) The region of changed pixels $(R_{\mathrm{c}})$ that corresponds to the union of all identified nonoverlapping annular sectors $S_k (k = 1, \ldots, K)$ (i.e. $R_{\mathrm{c}} = \bigcup_{k=1}^{K} S_k$), and 2) the region of no-changed pixels $(R_{\mathrm{n}})$ that corresponds to the union of circle of no-change $C_{\mathrm{n}}$ and the region $A_{\mathrm{c}} - \bigcup_{k=1}^{K} S_k$ complementary to the region of changed pixels with respect to $A_{\mathrm{c}}$, (i.e., $R_{\mathrm{n}} = C_{\mathrm{n}} \bigcup \{A_{\mathrm{c}} - \bigcup_{k=1}^{K} S_k\}$). In ideal situations, no pixels will fall in region $A_{\mathrm{c}} - \bigcup_{k=1}^{K} S_k$; whereas in real situations (when noise affects the data), we may have patterns in the annulus of changed pixels that do not belong to any of the sectors of changed pixels (these patterns are usually few and relatively isolated).

Given the aforementioned decision regions and the joint conditional distributions for the classes of change and no-change in (20), it is possible to analytically define the false and missed alarms.

False alarms occur when unchanged pixels are identified as changed. The probability of this kind of error $P_f$ can be written as the integral of the joint conditional density function given the class of no-change over the region of changed pixels, i.e.,

$$P_f = \int_{R_{\mathrm{c}}} p(\rho, \vartheta|\omega_{\mathrm{n}}) d\rho d\vartheta. \quad (21)$$

Missed alarms occur when changed pixels are identified as unchanged. The probability of this kind of error $P_m$ can be written as the sum of the $K$ integrals (one for each class of change $\omega_{c_k}$) of the joint conditional-density function given the

class $\omega_{c_k}$ of change over the region associated to unchanged pixels, i.e.,

$$P_m = \sum_{k=1}^{K} \int_{R_n} p(\rho, \vartheta|\omega_{c_k}) d\rho d\vartheta. \quad (22)$$

## IV. Analysis of the Marginal Conditional Distributions of Magnitude and Direction

The joint conditional density in (20) is too general to be efficiently used in solving change-detection problems. A more suitable way to approach the problem is to compute the marginal conditional densities of the magnitude $\rho$ and the direction $\vartheta$. Starting from (20), these two densities can be computed for each class $\omega_i$ by integrating (20) over the range of $\vartheta$ and $\rho$, respectively.

Let us first consider the marginal conditional density of the magnitude $p(\rho|\omega_i)$. Integrating (20) over the range of $\vartheta$ leads to the following equation:

$$p(\rho|\omega_i) = \frac{\rho}{2\pi\sigma_{1,i}\sigma_{2,i}} \times \int_0^{2\pi} \exp\left[-\frac{(\rho\cos\vartheta - \mu_{1,i})^2}{2\sigma_{1,i}^2} - \frac{(\rho\sin\vartheta - \mu_{2,i})^2}{2\sigma_{2,i}^2}\right] d\vartheta. \quad (23)$$

This integral cannot be expressed in a closed form, but it can be reduced to an infinite series of Bessel functions. By following [18], it is possible to show that $p(\rho|\omega_i)$ can be written as

$$p(\rho|\omega_i) = \frac{\rho}{\sigma_{1,i}\sigma_{2,i}} \exp(-V) \sum_{p=0}^{\infty} (-1)^p \varepsilon_p I_p(P) I_{2p}\left(\sqrt{Q^2+R^2}\right) \times \cos\left(2p \arctan\frac{R}{Q}\right), \qquad \rho \geq 0 \quad (24)$$

where $I_p(z)$ is the $p$th-order modified-Bessel function of the first kind defined as

$$I_p(z) = \frac{1}{2\pi} \int_C^{C+2\pi} \exp\left(-z\cos(u) + jpu\right) du \quad (25)$$

where $C$ is a constant, $j$ is the imaginary unit, and $D$, $P$, $Q$, $R$ and $\varepsilon_p$ are defined as follows:

$$V = \frac{\mu_{1,i}^2}{2\sigma_{1,i}^2} + \frac{\mu_{2,i}^2 + \rho^2}{2\sigma_{2,i}^2} + \frac{\sigma_{2,i}^2 - \sigma_{1,i}^2}{4\sigma_{1,i}^2\sigma_{2,i}^2}\rho^2$$

$$P = \frac{\sigma_{2,i}^2 - \sigma_{1,i}^2}{4\sigma_{1,i}^2\sigma_{2,i}^2}\rho^2$$

$$Q = \frac{\mu_{1,i}\rho}{\sigma_{1,i}^2}$$

$$R = \frac{\mu_{2,i}\rho}{\sigma_{2,i}^2}$$

$$\varepsilon_p = \begin{cases} 1, & \text{for } p = 0 \\ 2, & \text{for } p \neq 0. \end{cases} \quad (26)$$

The marginal conditional density of the direction $p(\vartheta|\omega_i)$ can be obtained by integrating (20) over the range of $\rho$. The integral can be written in a closed form. It is possible to prove that after some handling we obtain

$$p(\vartheta|\omega_i) = \frac{\left(1+\tan^2(\vartheta)\right)}{4\pi\sqrt{\left(\sigma_{1,i}^2\tan^2(\vartheta)+\sigma_{2,i}^2\right)^3}} \times \left\{\exp\left(-\frac{\mu_{2,i}^2\sigma_{1,i}^2+\mu_{1,i}^2\sigma_{2,i}^2}{2\sigma_{1,i}^2\sigma_{2,i}^2}\right) \times \left[\sqrt{2\pi}\exp\left(-\frac{\left(\mu_{2,i}\sigma_{1,i}^2\tan(\vartheta)+\mu_{1,i}\sigma_{2,i}^2\right)^2}{2\sigma_{1,i}^2\sigma_{2,i}^2\left(\sigma_{1,i}^2\tan^2(\vartheta)+\sigma_{2,i}^2\right)}\right) \times \left(\mu_{2,i}\sigma_{1,i}^2\tan(\vartheta)+\mu_{1,i}\sigma_{2,i}^2\right) \times \left[1+\operatorname{erf}\left(\frac{\mu_{2,i}\sigma_{1,i}^2\tan(\vartheta)+\mu_{1,i}\sigma_{2,i}^2}{\sigma_{1,i}\sigma_{2,i}\sqrt{2\left(\sigma_{1,i}^2\tan^2(\vartheta)+\sigma_{2,i}^2\right)}}\right)\right] + 2\sigma_{1,i}\sigma_{2,i}\sqrt{\left(\sigma_{1,i}^2\tan^2(\vartheta)+\sigma_{2,i}^2\right)}\right]\right\}, \quad \vartheta \in [0, 2\pi). \quad (27)$$

Figs. 3 and 4 show examples of the behaviors of the magnitude and direction marginal conditional densities versus the mean values $(\mu_{1,i}, \mu_{2,i})$ and the standard deviations $(\sigma_{1,i}, \sigma_{2,i})$, respectively, of the class $\omega_i$ characterized by a Gaussian distribution in the Cartesian coordinate system. It is worth noting that the periodicity of the direction distribution depends on the tangent function; in real applications, the proper maximum should be selected according to the data distribution.

Equations (24) and (27) represent two complex mathematical expressions. In real applications, usually additional hypotheses can be made in order to simplify the analytical expressions of the probability density functions. In the change-detection problem, different assumptions can be formulated for the classes of changed and unchanged pixels. In the following, the cases related to the two classes of interest will be addressed separately and in greater detail.

### A. Statistical Models for the Class of Unchanged Pixels

As stated in Section III-B, we assume that images $\mathbf{X}_1$ and $\mathbf{X}_2$ have been coregistered [19], [20] and, that possible differences in the light and atmospheric conditions at the two times have been corrected [21].[5] Under these hypotheses, we can reasonably assume that in unchanged areas, natural classes do not significantly change their distributions between the two acquisition dates. This simplifies the computation of distributions in the polar domain as we can write

$$\mu_{1,\omega_n} \cong \mu_{2,\omega_n} \approx 0 \quad (28)$$

$$\sigma_{1,\omega_n} \cong \sigma_{2,\omega_n} = \sigma_{\omega_n}. \quad (29)$$

[5]This assumption will be discussed in Section IV-C.

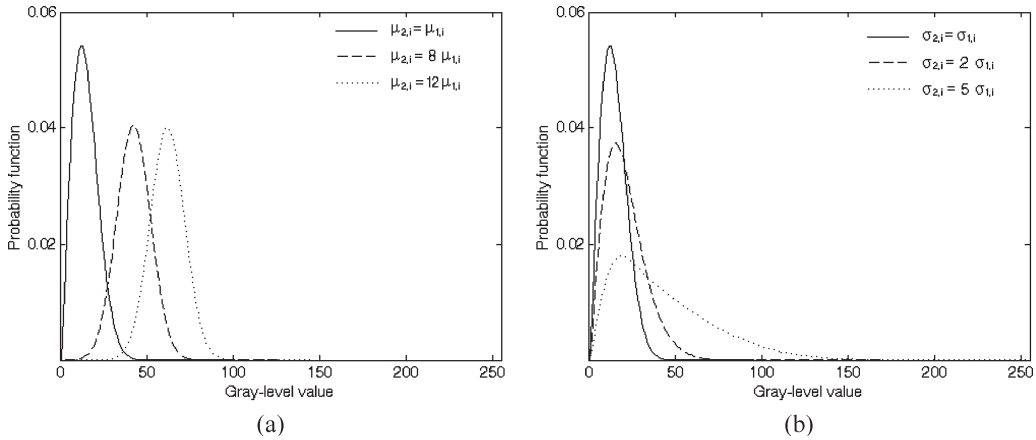



Fig. 3. Examples of conditional distributions of the magnitude (a) with respect to different values of $\mu_{2,i}$ ($\mu_{1,i} = 5$ and $\sigma_{1,i} = \sigma_{2,i} = 10$) and (b) with respect to the different values of $\sigma_{2,i}$ ($\mu_{1,i} = \mu_{2,i} = 5$ and $\sigma_{1,i} = 10$).

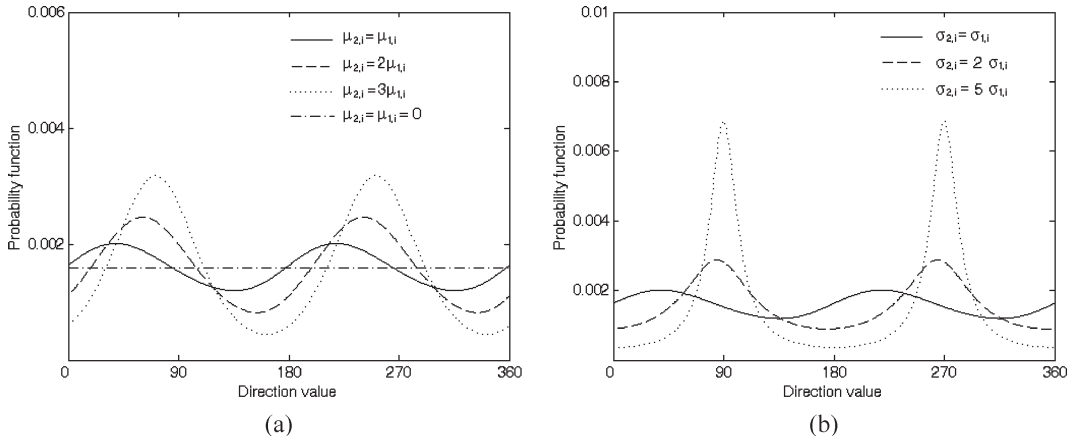



Fig. 4. Examples of conditional distributions of the direction (a) with respect to different values of $\mu_{2,i}$ ($\mu_{1,i} = 5$ and $\sigma_{1,i} = \sigma_{2,i} = 10$) and in the case $\mu_{1,i} = \mu_{2,i} = 0$ (which leads to the uniform distribution) and (b) with respect to different values of $\sigma_{2,i}$ ($\mu_{1,i} = \mu_{2,i} = 5$ and $\sigma_{1,i} = 10$).

Substituting both expressions (28) and (29) into (23) and solving the integral, we get for the magnitude random variable the following probability density function:

$$p(\rho|\omega_{\mathrm{n}}) = \frac{\rho}{\sigma^2_{\omega_{\mathrm{n}}}} \exp\left(-\frac{\rho^2}{2\sigma^2_{\omega_{\mathrm{n}}}}\right), \qquad \rho \geq 0 \tag{30}$$

which is commonly known as the Rayleigh distribution.

Concerning the statistical distribution of the direction variable for the class of unchanged pixels, it can be obtained substituting (28) and (29) into (27), i.e.,

$$p(\vartheta|\omega_{\mathrm{n}}) = \frac{1}{2\pi}, \qquad \vartheta \in [0, 2\pi). \tag{31}$$

This means that the statistical distribution of the direction is uniform within $[0, 2\pi)$.

### B. Statistical Models for the Classes of Changed Pixels

The analytical study of the distribution of the generic class $\omega_{c_k} \in \Omega_{\mathrm{c}}$ (for simplicity of notation in the following $\omega_{c_k}$ will be indicated as $\omega_k$) of changed pixels is more complex than the one carried out for the class of unchanged pixels. In this case, assumption (28) is no further valid, as changes in land-cover types modify the mean values of the natural classes in different ways in different spectral channels (this depends on the kind of change). This leads to the following condition:

$$\mu_{1,\omega_k} \neq \mu_{2,\omega_k} \neq 0. \tag{32}$$

It is worth noting that if $\mu_{1,\omega_k} = \mu_{2,\omega_k}$, the analysis of the distribution is simplified. In order to further simplify the computation of the magnitude and direction statistical distributions, we can assume that

$$\sigma_{1,\omega_k} \approx \sigma_{2,\omega_k} = \sigma_{\omega_k}. \tag{33}$$

In some applications, this assumption is reasonable, but its validity should be verified for any specific case considered.[6] Thus, rewriting (23) according to (32) and (33) and solving the integral, it is possible to show that the random variable

[6]If the assumption is not verified, the general (24) and (27) should be used for modeling the statistical distributions of magnitude and direction, respectively, or a proper preprocessing should be applied to the data before using the simplified model (see Section IV-C).

TABLE I
SUMMARY OF THE THEORETICAL MARGINAL CONDITIONAL DISTRIBUTIONS OF MAGNITUDE AND DIRECTION FOR THE CHANGE AND NO-CHANGE CLASSES UNDER SIMPLIFYING ASSUMPTIONS

| Class | Conditional distribution | |
|---|---|---|
| | Magnitude ($\rho$) | Direction ($\vartheta$) |
| Unchanged pixels ($\omega_n$) | *Rayleigh* | *Uniform* |
| Changed pixels ($\omega_k$) | *Rice* | *Non-Uniform* |

representing the magnitude is Ricean distributed, with conditional density function given by

$$p(\rho|\omega_k) = \frac{\rho}{\sigma_{\omega_k}^2} \exp\left(-\frac{\rho^2 + M_{\omega_k}^2}{2\sigma_{\omega_k}^2}\right) I_0\left(\frac{\rho M_{\omega_k}}{\sigma_{\omega_k}^2}\right), \qquad \rho \geq 0 \tag{34}$$

where $I_0(\cdot)$ is the modified zeroth-order Bessel function of the first kind [see (24)] and $M_{\omega_k}$ is the noncentrality parameter of the class of change $\omega_k$:

$$M_{\omega_k} = \sqrt{\mu_{1,\omega_k}^2 + \mu_{2,\omega_k}^2}. \tag{35}$$

It is worth noting that as $M_{\omega_k}$ becomes much larger than the standard deviation $\sigma_{\omega_k}$, then the Ricean distribution tends to become Gaussian.

Concerning the density of the direction of the class of changed pixels $\omega_k$, it is possible to prove that in the aforementioned assumptions, (27) can be simplified leading to the following nonuniform distribution:

$$\begin{aligned} p(\vartheta|\omega_k) = \exp&\left(-\frac{\mu_{2,\omega_k}^2 + \mu_{1,\omega_k}^2}{2\sigma_{\omega_k}^2}\right) \\ &\times \left\{ \frac{1}{2\pi} + \frac{1}{2\sigma_{\omega_k}\sqrt{2\pi}} \frac{\mu_{2,\omega_k}^2 \tan(\vartheta) + \mu_{1,\omega_k}^2}{\sqrt{(1+\tan^2(\vartheta))}} \right. \\ &\times \exp\left(-\frac{(\mu_{2,\omega_k}\tan(\vartheta) + \mu_{1,\omega_k})^2}{2\sigma_{\omega_k}(1+\tan^2(\vartheta))}\right) \\ &\left. \times \left[1 + \operatorname{erf}\left(-\frac{\mu_{2,\omega_k}\tan(\vartheta) + \mu_{1,\omega_k}}{\sigma_{\omega_k}\sqrt{2(1+\tan^2(\vartheta))}}\right)\right] \right\}, \\ &\vartheta \in [0, 2\pi). \end{aligned} \tag{36}$$

### C. Discussion

In the previous subsections, we analyzed the statistical models more suitable to represent class distributions in the polar domain in the general case and in some simplifying assumptions. Since the use of simplified models is of great importance for the development of effective and adequately complex automatic change-detection techniques, here, we report a critical discussion on the assumptions considered for modeling the distributions of the classes of changed and unchanged pixels. In addition, we analyze practical implications of the theoretical analysis, in order to suggest criteria for driving the definition of proper preprocessing techniques for an effective data representation. Table I reports a summary of the theoretical statistical distributions derived (under simplifying assumptions) in Sections IV-A and IV-B for magnitude and direction of change and no-change classes.

First of all, it is important to point out that a hypothesis at the basis of the theoretical analysis reported in the previous subsections consists in assuming independence among features describing SCVs in the Cartesian domain. The validity of this assumption depends on the considered images and applications, as well as on the subset of investigated spectral channels. Significant deviations from this assumption affect the precision of the analytical distributions derived for describing the behaviors of changed and unchanged pixels. Nonetheless, we expect that the obtained distributions are more precise than simple empirical models used in the literature. In addition, if for a generic data set the aforementioned assumption is not reasonable, it is possible to transform data from the original feature space to a transformed domain, in which features can be approximately modeled as uncorrelated. This can be obtained by applying a principal component transformation (PCT) to the features characterizing the SCVs [22]. In this way, at the cost of an additional transformation applied to the data, it is possible to properly adopt the analytical models described in Sections IV-A and IV-B in the development of change-detection techniques.[7]

The aforementioned assumption is at the basis of the presented theoretical analysis. All the other assumptions (discussed in the following) allow only to simplifying the statistical distributions with respect to the general models in (24) and (27), which can be included in automatic techniques for operational change-detection algorithms but are rather complex. For this reason, in the following, we analyze the simplifying assumptions in greater detail.

An important hypothesis that deserves to be discussed concerns the assumption that different features in the Cartesian SCV domain have similar standard deviations. The validity of the assumption in the original feature space depends on the considered images and applications, as well as on the investigated spectral channels. However, as for the assumption of the independence, if this approximation is not acceptable for the considered data set, it is possible to transform the original feature space according to a procedure of diagonalization and whitening [22] and to apply change-detection algorithms to the transformed space.[8] (As the application of this procedure modifies the relation between the covariance terms of the spectral channels, it should be applied with particular attention).

A further relevant assumption to be analyzed for the class of unchanged pixels consists in the hypothesis that the mean-vector components of the SCVs are equal to zero. This assumption is verified, if the images are radiometrically corrected, so that the mean vectors are the same at the two dates (this condition can be always satisfied according to proper preprocessing strategies). Under this assumption, we obtain that the magnitude has a Rayleigh distribution and the direction has a uniform distribution. However, in some practical cases,

[7]It is worth noting that the PCT guarantees independence of features on the basis of the global distributions of patterns in the feature space. This means that after transformation, the feature independence on the classes of changed and unchanged pixels can be assumed only in an approximate way.

[8]As stated for PCT also, whitening produces effects on the basis of the global distributions of patterns in the feature space. This means that after transformation the standard deviation of the classes can be considered similar only in an approximate way

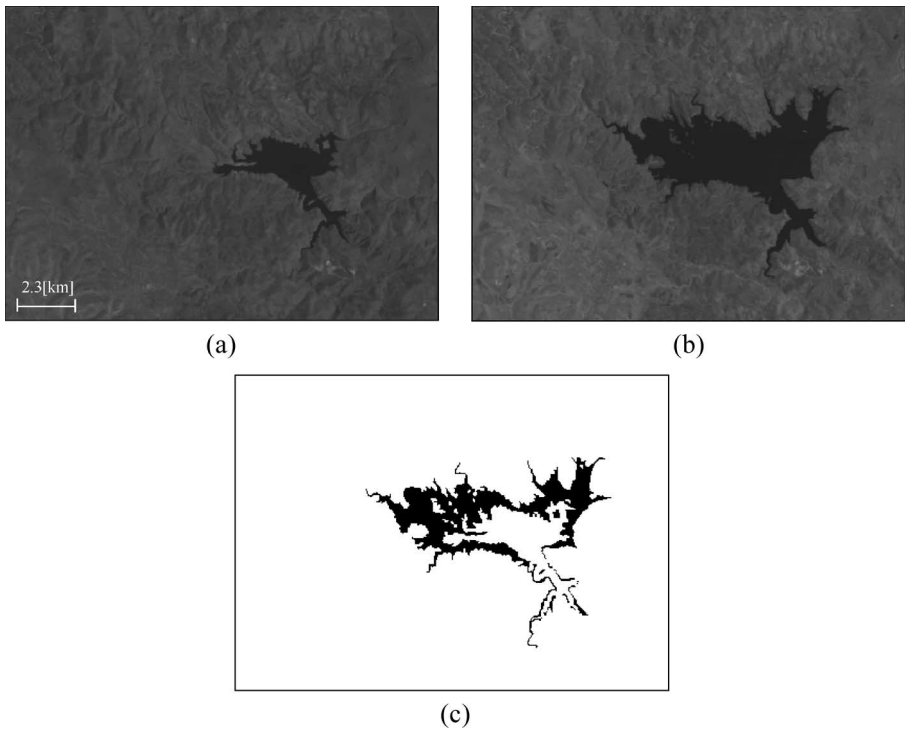



Fig. 5. Images of the Lake Mulargia (Italy), used in the experiments. (a) Channel 4 of the Landsat-5 TM image acquired in September 1995. (b) Channel 4 of the Landsat-5 TM image acquired in July 1996. (c) Available reference map of changed areas.

images are not radiometrically corrected and preprocessing procedures for matching the light conditions are neglected. According to the presented theoretical analysis, this may result in two main very critical effects: 1) a possible increase of the overlapping of the classes of unchanged and changed pixels in the magnitude domain and 2) a strong deviation of the conditional distribution of the direction of unchanged pixels from the expected uniform model. The first effect is due to the fact that, although differences in light conditions result in a bias common to all classes in the Cartesian domain, when the nonlinear magnitude operator is applied, the bias may result in an increase of overlapping between classes (this behavior will be shown in the experimental analysis reported in Section V). The second effect results from the observation that if the mean-value components of the SCVs are different from zero, the direction distribution of the class of unchanged pixels is no longer uniform but assumes a completely different behavior, which should be modeled with (36) (see Fig. 4). This has a dramatic impact on the data-processing strategy, as it completely changes the distribution of the direction with respect to what expected in the ideal case. These observations confirm the importance of the radiometric-correction step in the CVA technique.

Finally, another important implication derived from the theoretical analysis concerns the behavior of the distributions of the direction for the unchanged and changed classes, which are uniform and nonuniform, respectively. This means that it is possible to exploit the direction information (and in particular, the modes associated to distribution for each changed class) for reducing the effects of the residual sources of noise present in the preprocessed multitemporal images (e.g., the registration noise that appears outside the change modes can be easily removed as it will be shown in the experimental analysis reported in Section V-A). This confirms from a theoretical point of view the analysis carried out in [4], where the direction information was used for identifying, modeling, and reducing registration noise.

In order to illustrate the use of the proposed CVA framework and to assess its effectiveness, in the next two sections, we consider two different real multitemporal data sets: 1) a single-change data set and 2) a double-change data set.

## V. Experimental Results: Single-Change Case

### A. Data-Set Description and Experiment Design

The first data set is made up of two multispectral images acquired by the Thematic Mapper (TM) multispectral sensor of the Landsat 5 satellite on the island of Sardinia, Italy, in September 1995 $(t_1)$ and July 1996 $(t_2)$. Both images have a spatial resolution of 30 $\times$ 30 m. The area selected for the experiments is a section (412 $\times$ 300 pixels) of the two scenes including Lake Mulargia. As an example of the images used in the experiments, Fig. 5(a) and (b) shows Channel 4 of the September and July images, respectively.

Between the two acquisition dates, only one kind of change occurred in the investigated area, which is related to the extension of the lake surface (the water volume of the Lake Mulargia increased producing an enlargement of the lake surface). The multitemporal images were coregistered and radiometrically corrected. A reference map of the analyzed site was defined according to a detailed visual analysis of both the available multitemporal images and the difference image. The obtained reference map contains 7480 changed pixels and 116 120 unchanged pixels [see Fig. 5(c)]. The extracted information was used for both computing the parameters of the statistical distributions for the classes of interest and evaluating the performances (in terms

of false and missed alarms) of the change-detection process carried out by using the proposed statistical models.

The considered change-detection problem is relatively simple, and thus, it is suitable for a proper understanding of the properties and potentialities of the proposed framework. In this particular single-change problem, we define $\Omega_{\rm c} = \{\omega_{\rm c_1}\} = \{\omega_{\rm c}\}$. In this paper, we considered only the two spectral channels 4 and 7 of the TM, i.e., the near and the middle infrared, as they are the most reliable for detecting changed areas. For simplicity of notation, in the following, these channels will be referred with subscripts 1 and 2, respectively.

In order to assess the effectiveness of the proposed framework, three different experiments have been carried out.

In the first experiment, a qualitative analysis of the true distributions of data in the polar domain versus different preprocessing applied to the images is carried out. This experiment is aimed at pointing out the effects of the preprocessing procedures on both the data distributions and the precision of the models introduced in Section IV for data representation.

The second experiment is aimed at validating the accuracy of the theoretical models of distributions presented in Section IV in fitting the true data distributions for both the magnitude and the direction, under the simplifying assumptions introduced in Sections IV-A and IV-B. Furthermore, the goodness-to-fit of the Rayleigh and Rice distributions adopted for the magnitude of the change and no-change classes, respectively, is compared with the goodness-to-fit of the widely used Gaussian model. Here, the well-known Kolmogorov–Smirnov (KS) test is used for establishing whether a statistical model fits or not the true distribution [23]. The KS statistical test determines if two sets of data are drawn from the same statistical distribution. The test is based on the comparison between the cumulative distribution functions of the true data $S_n(x)$[9] (or empirical distribution function) and the expected one $F(x)$ (i.e., the cumulative distribution of the density function adopted for modeling the data) [23]. The KS test compares the cumulative distributions by a difference operator computing the so-called KS-statistic $D_n$

$$D_n = \sup_x \{|F(x) - S_n(x)|\}. \tag{37}$$

It is worth noting that $D_n$ is a random variable, whose distribution does not depend upon $F(x)$, i.e., the KS test is nonparametric and distribution free. The output of the KS test consists in the acceptance of the assumption that the true data distribution follows the selected model, if $D_n \leq D_n^\alpha$ with a high probability $P_{\rm KS}$; otherwise, the hypothesis is rejected and the two distributions are considered different. $D_n^\alpha$ is the critical value that depends on both the desired confidence level $\alpha$, and the number of samples $n$ used for estimating the empirical distribution function (numerical values of $D_n^\alpha$ for different combination of $\alpha$ and $n$ are well-known tabulated values [23]).

The third and last experiment is aimed at establishing the possible improvements that can be obtained on the accuracy of the change-detection process (in terms of false and missed alarms, as well as total errors) by adopting the derived theoretical statistical models (for approximating the magnitude distributions of change and no-change pixels) rather than the widely used Gaussian model. In addition, an analysis on the impact of a poor preprocessing phase on the change-detection accuracy is also reported.

[9] $x$ is the set of values for which both the cumulative densities are known.

### *B. Qualitative Analysis of the Class Distributions in the Polar Domain*

The aim of the first experiment is to qualitatively show the effects of an inaccurate preprocessing phase (in terms of radiometric differences and/or misregistration noise) on the statistical distributions of magnitude and direction in the polar-coordinate system. In order to accomplish this analysis, we study the statistical distributions of SCVs obtained by applying the CVA technique to multitemporal images in three different cases: 1) radiometrically corrected and coregistered images; 2) coregistered images without radiometric corrections; and 3) radiometrically corrected images with a poor coregistration (a residual shift of about two pixels on ground control points was accepted).

As expected from the theoretical analysis, in all cases, it is possible to identify two clusters in the polar domain. In the case of corrected images [Fig. 6(a)], the first cluster is centered in the origin of the polar plot and shows high occurrences (red color) close to zero and a uniform distribution with respect to the direction domain. This cluster is associated to the unchanged SCVs. The second cluster shows a preferred direction and is located relatively far from the origin. This cluster is related to the SCVs associated with changed pixels. In this case, it is quite easy from a qualitative viewpoint to identify the decision boundary (threshold value on the magnitude) between the circle $C_{\rm n}$ of no-changed pixels and the annulus $A_{\rm c}$ of changed pixels. Furthermore, also the sector $S$ of the changed pixels is clearly visible [see Fig. 6(a)].

The situation is significantly different in the second case, i.e., if no radiometric corrections are applied to the original images. Radiometric differences between the two acquisitions [see Fig. 6(b)] have a dramatic impact on the distribution of the no-change class. As one can see, the cluster of unchanged pixels is no longer centered in zero; thus, the direction distribution is no further uniform but assumes values in a subset of the domain, which is defined by the difference of the mean values of unchanged pixels at the two dates in the two considered spectral bands. In greater detail, in this condition, the no-change-class distribution with nonzero mean can be approximated with the model described in Section IV-B for the class of changed pixels. This behavior points out that the use of the uniform model for the approximation of the distribution of the direction of the class of unchanged pixels in the data-analysis phase when images are not radiometrically corrected is not acceptable. Furthermore, by analyzing Fig. 6(b), it is possible to observe that the mean value of the magnitude of the unchanged pixels increases (with respect to the case of radiometrically corrected images), while the mean value of the magnitude of the changed pixels decreases. This means that if only the magnitude is used for the change detection (like in many real applications), the classes result more overlapped. This effect, which is due to the nonlinearity of the magnitude operator, involves a higher change-detection error with respect to the case of radiometrically corrected data. In other words, the absence of radiometric corrections does not result in a bias contribution

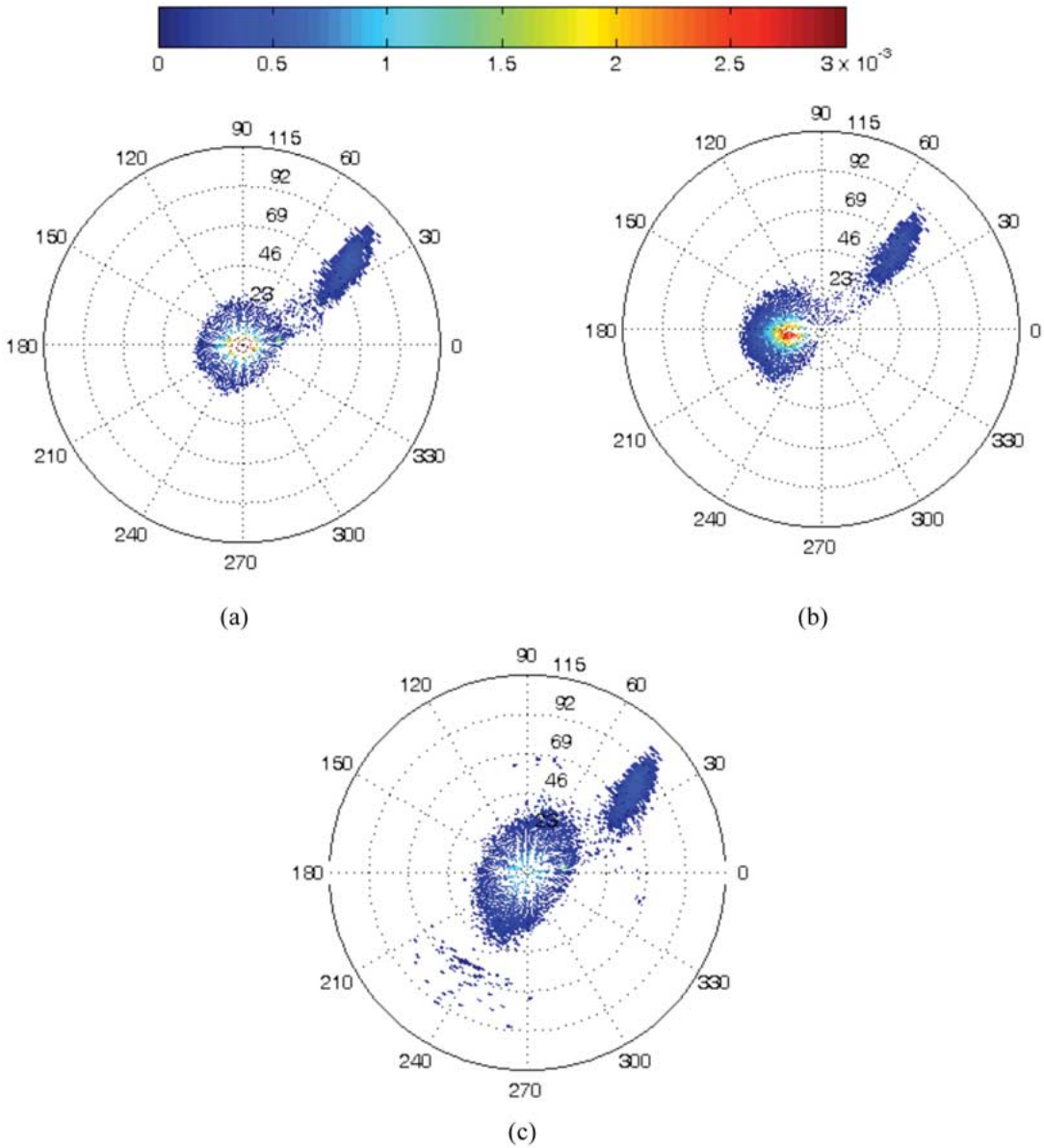



Fig. 6. Histograms in the polar-coordinate system obtained after applying CVA to (a) multitemporal radiometrically corrected and coregistered images, (b) coregistered multitemporal images without radiometric corrections, and (c) multitemporal radiometrically corrected images with a significant residual registration noise (single-change case).

common to both classes, but may decrease significantly the separability between them in the magnitude domain.

In the third case, image misregistration generates in the histogram plotted in the polar domain: 1) more spread SCV distributions and 2) the presence of unchanged SCVs that are out of $C_{\mathrm{n}}$ and assume values in the entire direction domain. The spread increment is related to the nonperfect correspondence between multitemporal pixels, which leads to an increase of the variances of classes. The presence of pixels outside $C_{\mathrm{n}}$ and $S$ is mainly due to the effects induced from border regions and details, which lead to the comparison of pixels belonging to completely different classes. These pixels have high-magnitude values but have direction that may differ from those of true changed pixels [Fig. 6(c)]. This behavior points out a very important guideline for practical applications, i.e., in situations where the residual misregistration between images cannot be neglected, the use of the direction variable in addition to the magnitude can reduce false alarms due to registration noise.

On the basis of the aforementioned analysis, it is clear that the polar representation results in a useful qualitative tool for easily understanding the effectiveness of the preprocessing applied to the data.

### C. Quantitative Analysis of the Accuracy of the Statistical Models of Class Distributions in the Polar Domain

This experiment aims at a quantitative validation of the analytical models defined for approximating the statistical distributions of the magnitude and direction of the classes of changed and no-changed pixels. The validation is carried out according to the KS test. In these trials, only the radiometrically corrected and coregistered images were considered.

TABLE II
MEAN VALUES AND STANDARD DEVIATIONS FOR THE CLASS OF UNCHANGED AND CHANGED PIXELS IN THE CARTESIAN COORDINATE SYSTEM (SINGLE-CHANGE CASE)

| $b$ | $\mu_{b,\omega_n}$ | $\sigma_{b,\omega_n}$ | $\mu_{b,\omega_c}$ | $\sigma_{b,\omega_c}$ |
|---|---|---|---|---|
| **1 (TM4)** | 0 | 10.26 | 58.94 | 8.90 |
| **2 (TM7)** | 0 | 8.73 | 43.47 | 10.67 |

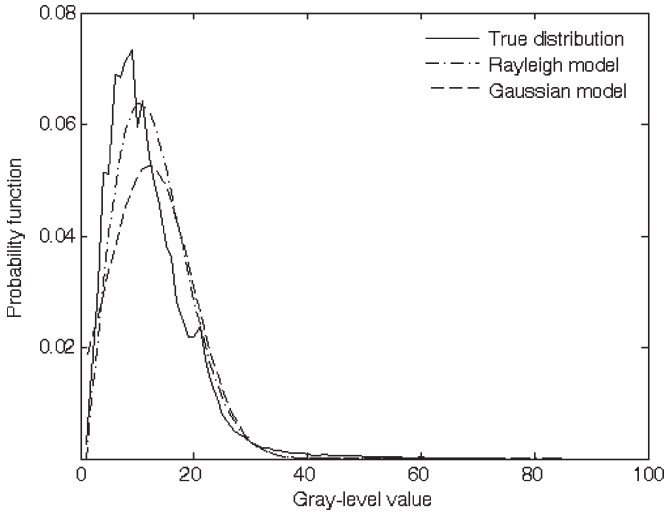



Fig. 7. Comparison between the behavior of the distribution of the magnitude of the unchanged pixels and its approximations obtained with Gaussian and Rayleigh models (single-change case).

In order to perform the validation of the derived statistical models, the true mean values and standard deviations of the change and no-change distributions were computed from $X_{\mathrm{D}}$ in the Cartesian coordinate system on the basis of the available reference map. The obtained values are summarized in Table II. In addition, for the magnitude of the change and no-change classes a comparison between the goodness-to-fit of Rice and Rayleigh models, respectively, and the commonly used Gaussian model is performed. In the following, the analysis of the results obtained on the two classes are considered separately.

*1) Statistical Models for the Class of No-Changed Pixels:* In order to adopt statistical models in (30) and (31) for the magnitude and the direction of unchanged SCVs, respectively, it should be verified if the hypotheses in (28) and (29) hold. By observing numerical values of standard deviations $\sigma_{b,\omega_{\mathrm{n}}}$ in Table II, it is reasonable to conclude that they are similar to each other and can be approximated to the mean of the standard deviations (which is 9.49), thus satisfying (29). For the mean values, no approximations should be introduced as (28) is verified (owing to the use of the radiometrically corrected images).

Let us first consider the magnitude variable. In Fig. 7, it is possible to see that the Rayleigh model approximates with good accuracy the distribution of the unchanged pixels (extracted from the reference map). In greater detail, this model fits better the data than the Gaussian model. This is confirmed by the KS test that results in a significantly higher $P_{\mathrm{KS}}$ value for the Rayleigh model than for the Gaussian one (0.6396 versus $2 \cdot 10^{-4}$).

Let us now consider the behaviors of the distributions in the direction dimension. The KS test states that the SCV directions are uniformly distributed with a $P_{\mathrm{KS}}$ value equal to 0.9939.

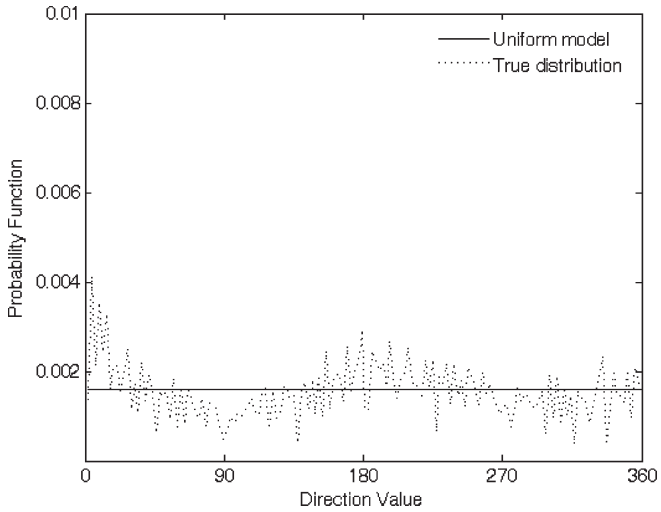



Fig. 8. Comparison between the behavior of the distribution of the direction of the unchanged pixels and its approximation obtained with the uniform model (single-change case).

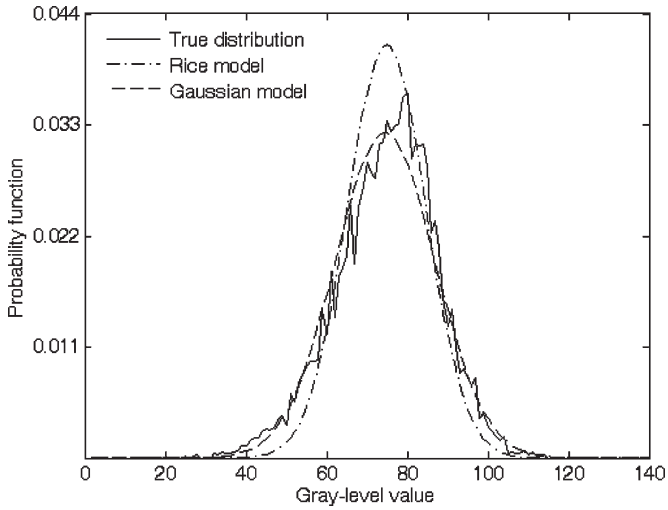



Fig. 9. Comparison between the behavior of the distribution of the magnitude of the changed pixels and its approximations obtained with Gaussian and Rice models (single-change case).

This reasonable result is confirmed also by a qualitative visual comparison between the true data distribution and the uniform distribution (Fig. 8).

*2) Statistical Models for the Class of Changed Pixels:* In order to adopt statistical models in (34) and (36) for the magnitude and the direction of changed SCVs, respectively, it should be verified if the hypothesis in (33) hold. Similarly to the no-change class, from numerical values in Table II, it is possible to observe that it is reasonable to approximate the standard-deviation values $\sigma_{b,\omega_{\mathrm{c}}}$ to the mean of the standard deviations, i.e., 9.77. This condition satisfies (33).

From Fig. 9, it is possible to see that the Rice model fits well the data in general, and only slightly better than the Gaussian model. This is confirmed by the KS test that results in a slightly higher $P_{\mathrm{KS}}$ value for the Rice model than for the Gaussian one (0.9993 versus 0.9961). The small difference in the two statistical models for this data set is that the noncentrality parameter (35) is much larger than the standard deviation, thus the Ricean distribution tends to become Gaussian.

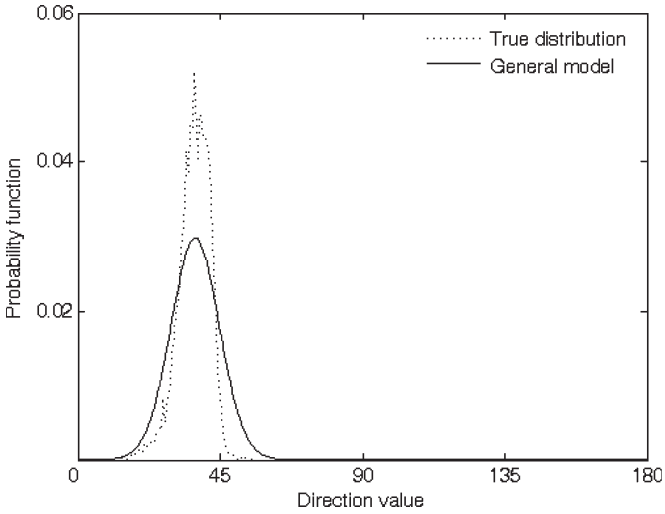



Fig. 10. Comparison between the behavior of the distribution of the direction of the changed pixels and its approximation obtained with the nonuniform model (single-change case).

The KS test states that the SCV direction is distributed according to (36) with a $P_{\mathrm{KS}}$ value equal to 0.3306. The relatively small value of $P_{\mathrm{KS}}$ is due to the presence of SCVs whose direction differs from the expected one (see Fig. 10), as SCV direction is highly sensitive to noise components. Such outliers can be related to the presence of residual misregistration noise.

### D. Analysis of the Effectiveness of the Proposed Framework for Solving Change-Detection Problems

This experimental part has two goals: 1) the first is to evaluate the impact of radiometric corrections on the performances of CVA; and 2) the second is to assess the improvement of the change-detection accuracy obtained by adopting the proposed statistical models rather than the Gaussian one for the change and no-change classes in the magnitude domain.

In order to evaluate the impact of the radiometric corrections on the performances of the CVA technique, we have compared the accuracies yielded by thresholding the magnitude variable in the case of: 1) radiometrically corrected and coregistered images and 2) coregistered images without radiometric corrections. To this end, the threshold values were defined according to a supervised manual trial-and-error procedure (MTEP), i.e., the minimum-error threshold was derived by performing a nonautomatic evaluation of the overall change-detection errors versus all the possible values of the decision threshold and, then, the threshold value that yielded the minimum overall error was chosen. From the qualitative analysis carried out in Section V-A, we expect that the change-detection accuracy is lower when no radiometric corrections are applied. As shown from Table III, the MTEP procedure applied to the magnitude of the original data set resulted in 1803 errors, while we obtained only 704 errors when thresholding was applied to the magnitude after a very simple radiometric-correction procedure, which adjusted the mean value of the images. The overall error is more than halved. In greater detail, after rediometrically correcting the multitemporal images, both missed and false alarms decreased significantly from 529 to 369 pixels and from 1274 to 335 pixels, respectively.

TABLE III
OVERALL ERROR, FALSE ALARMS, AND MISSED ALARMS (IN NUMBER OF PIXELS) AND THRESHOLD VALUE RESULTING FROM THE MTEP APPLIED TO THE COREGISTERED IMAGES WITH AND WITHOUT RADIOMETRIC CORRECTIONS (SINGLE-CHANGE CASE)

| Radiometric corrections | False alarms | Missed alarms | Overall errors | Threshold value ($T$) |
|---|---|---|---|---|
| None | 529 | 1274 | 1803 | 47 |
| Mean adjustment | 369 | 335 | 704 | 51 |

TABLE IV
OVERALL ERROR, FALSE ALARMS, AND MISSED ALARMS RESULTING FROM THE SELECTION OF THE DECISION THRESHOLD VALUES CARRIED OUT BY USING MTEP AND BDR WITH THE PROPOSED STATISTICAL MODELS AND THE STANDARD GAUSSIAN STATISTICAL MODEL (SINGLE-CHANGE CASE)

| | False alarms | Missed alarms | Overall errors | Threshold value ($T$) |
|---|---|---|---|---|
| MTEP | 369 | 335 | 704 | 51 |
| BDR proposed models | 855 | 101 | 956 | 43 |
| BDR *Gaussian* model | 1086 | 57 | 1143 | 40 |

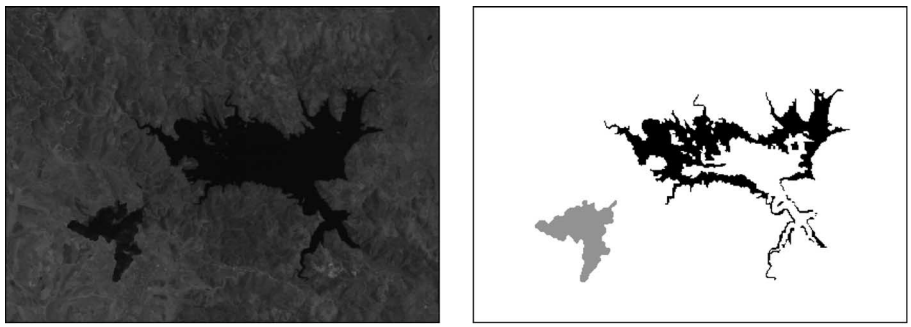
In order to asses the effectiveness of the proposed statistical models in solving change-detection problems, we considered only the coregistered and radiometrically corrected multitemporal data set. Here, the MTEP results have been compared with the performances obtained by solving the change-detection problem according to the Bayes decision rule (BDR) [22] under two different assumptions on statistical distributions: 1) the proposed statistical models (i.e., Rayleigh model for the class of unchanged pixels and Rice model for the class of changed pixels) and 2) the widely used Guassian model for both classes (Table IV).

As expected, owing to the capability of the Rayleigh and Rice density functions to better model the true data distributions (see Section V-B), the proposed models allow to obtain a lower amount of total errors with respect to the model based on the Gaussian distribution (956 versus 1143). This is that using the proposed model, the obtained threshold value (i.e., 43) is much closer to the optimal one (i.e., 51) than the threshold computed with the Gaussian model (i.e., 40). It is worth noting that although significant, the difference of threshold values between the Gaussian and the proposed models is relatively small. This depends on the fact that on the considered data set the Rice distribution is close to the Gaussian one, as the noncentrality parameter is much larger then the standard deviation. We expect that higher improvements can be obtained in more general cases.

## VI. EXPERIMENTAL RESULTS: DOUBLE-CHANGE DATA SET

### A. Data-Set Description and Experiment Design

The second data set is made up of the same two multispectral images acquired on the island of Sardinia, Italy, in September 1995 $(t_1)$ and July 1996 $(t_2)$ described in Section V-A, in which a second kind of change was introduced. In order to simulate the presence of the new change, a data set made up of two multispectral images acquired by the Landsat-5 TM multispectral sensor on the island of Elba, Italy, in August 1994 $(t_1)$ and September 1994 $(t_2)$ was considered [16]. Between these two acquisitions, a wildfire destroyed a large part of the



(a) (b)

Fig. 11. Simulated double-change data set. (a) Channel 4 of the simulated image at $t_2$. (b) Reference map of changed areas (burned area in light gray color, lake enlargement in black color).

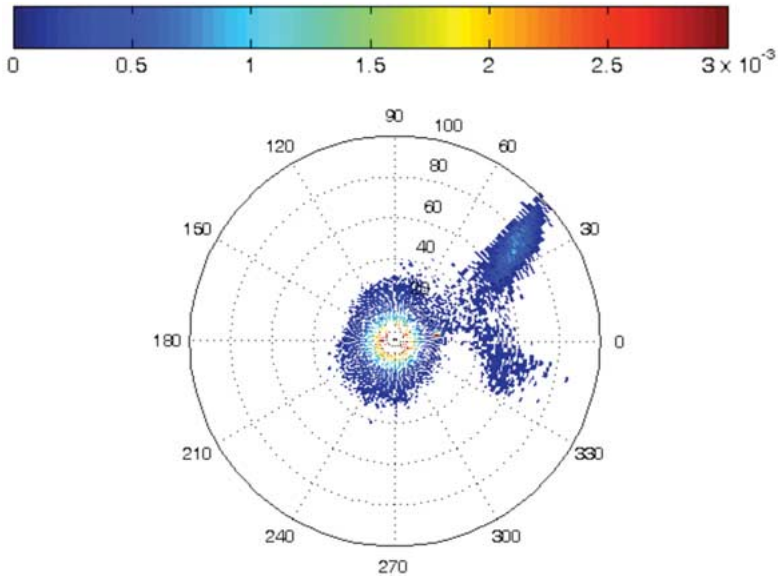



Fig. 12. Histograms in the polar-coordinate system obtained after applying CVA to multitemporal radiometrically corrected and coregistered images (double-change case).

vegetation in the northwest part of the island. On the basis of the available ground truth, the area affected from the change was isolated on spectral channels 4 and 7 of the TM images of both the August and September acquisitions. The selected area was inserted in the lower left part of the spectral channels 4 and 7 of the September and July images of the Sardinia Island. It is worth noting that in order to obtain a realistic representation of the statistics of the change, both radiances at $t_1$ and $t_2$ of the Elba data set were inserted in the Sardinia data set. In this way, we obtained a data set that properly represents the properties of the different kinds of changes considered. The reference map of this data set was built by modifying the reference map of the Sardinia data set accordingly to the introduced change. As an example of the resulting data set, Fig. 11(a) and (b) show the channel 4 of the simulated image at $t_2$ and the double-change reference map, respectively. The reference map has 7480 pixels belonging to the changed class $\omega_{c_1}$ (which is associated with the change in water level in the lake, black color), 2414 pixels belonging to the changed class $\omega_{c_2}$ (which is associated with the burned area, light gray color), and 113 706 unchanged pixels (white color).

The experiments for this data set were carried out only on the coregistered and radiometrically corrected pair of images in order to evaluate the effectiveness of the proposed theoretical framework for analyzing and extracting SCV information about different kinds of change. Three different experiments were carried out: 1) a qualitative analysis of the distribution of the classes in the polar domain; 2) a validation of the accuracy of the theoretical models in fitting the true data distributions using the KS test for the change class $\omega_{c_2}$ (we refer the reader to Section V for the analysis of magnitude and phase distributions of the classes of unchanged pixels and of changed pixels associated with the water level in the lake); and 3) an evaluation of the accuracy of the change-detection process (in terms of confusion matrix and Kappa coefficient) obtained by adopting the derived theoretical statistical models for both magnitude and direction distributions.

### *B. Qualitative and Quantitative Analysis of the Class Distributions in the Polar Domain*

As expected from the theoretical analysis, the presence of a second kind of change resulted in the appearance of an additional cluster in the polar domain (Fig. 12). Thus, the number of clusters is now three. As in the data set with a single change, the cluster associated with the unchanged SCVs is centered in

TABLE V
MEAN VALUES AND STANDARD DEVIATIONS FOR THE CLASS OF CHANGED PIXELS $\omega_{c_2}$ IN THE CARTESIAN COORDINATE SYSTEM

| $b$ | $\mu_{b,\omega_{c_2}}$ | $\sigma_{b,\omega_{c_2}}$ |
|---|---|---|
| **1 (TM4)** | 55.34 | 11.87 |
| **2 (TM7)** | -11.77 | 9.79 |

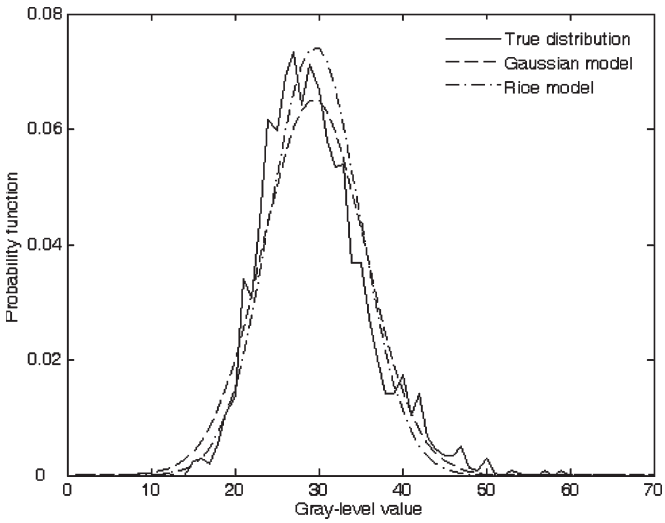



Fig. 13. Comparison between the behavior of the distribution of the magnitude of the changed pixels of class $\omega_{c_2}$ and its approximations obtained with Gaussian and Rice models (single-change case).

the origin of the polar plot and shows high occurrences (red color) close to zero and a uniform distribution with respect to the direction variable. Both the clusters related to the SCVs associated with changed pixels are located relatively far from the origin and show a quite similar magnitude but different preferred directions. These differences are due to the different radiometric variations induced from the two kinds of changes on the considered spectral channels.

As in Section V-B, hypothesis in (33) should hold for adopting statistical models in (34) and (36) for the magnitude and the direction of changed SCVs, respectively. From numerical values in Table V, it is possible to observe that also for the new kind of change it is reasonable to approximate the standard-deviation values $\sigma_{b,\omega_{c2}}$ to the mean of the standard deviations, i.e., 10.83. This condition satisfies (33).

The effectiveness of the theoretical model in fitting the true data distribution is easy to understand in a qualitative way by observing Fig. 13, where it is possible to see that the Rice model fits well the data in general and only slightly better than the Gaussian model. This is confirmed also by the KS test that results in a slightly higher $P_{\mathrm{KS}}$ value for the Rice model than for the Gaussian one (i.e., 0.9903 versus 0.9461). As for the class of change associated with the enlargement of the lake surface, the reason of the small difference for this class is that the noncentrality parameter (35) is much larger than the standard deviation and, thus, the Ricean distribution tends to become Gaussian.

Analogously, from Fig. 14, it is possible to conclude that the true direction distribution of SCVs of class $\omega_{c_2}$ is very close to the general model in (35). The KS test confirmed this observation resulting in a $P_{\mathrm{KS}}$ value equal to 0.9683.

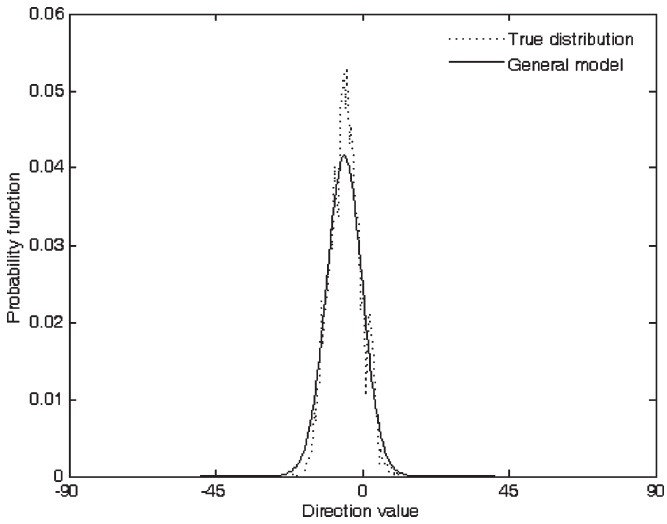



Fig. 14. Comparison between the behavior of the distribution of the direction of the changed pixels of class $\omega_{c_2}$ and its approximation obtained with the nonuniform model (single-change case).

### C. Analysis of the Effectiveness of the Proposed Framework for Solving Change-Detection Problems

Finally, in order to assess the effectiveness of the proposed framework for solving change-detection problems, we compared the results yielded by the proposed statistical models for both the magnitude and the phase distributions with the results obtained applying a supervised MTEP. In order to distinguish the classes of unchanged pixels and the two classes of changed pixels, it is necessary to identify: 1) the threshold value $T$ that separates the circle $C_{\mathrm{n}}$ of no-changed pixels from the annulus $A_{\mathrm{c}}$ of changed pixels ($T$ is identified by considering the two kinds of changes as a single class)[10]; 2) the two angular threshold values $\vartheta_{1_1}$ and $\vartheta_{1_2}$ that bound the annular sector $S_1$ of changed pixels associated with the enlargement of the lake $(\omega_{c_1})$; and 3) the two angular threshold values $\vartheta_{2_1}$ and $\vartheta_{2_2}$ that bound the annular sector $S_2$ of changed pixels associated with the burned area $(\omega_{c_2})$.

In this case, the MTEP was applied in two sequential steps. The first step considers only the magnitude and identifies, among all possible values, the threshold value $T$ that separates the class of unchanged pixels from the two classes of changed pixels (considered as a single class in the magnitude domain) with the minimum overall error. The second step considers also the direction and identifies, given the value of $T$, the four angular thresholds that maximize the Kappa coefficient of the final change-detection map (in this case, the two kinds of changes are separated from the no-change class and also from each other).

Also, the solution of the change-detection problem by adopting the proposed statistical models for the magnitude and the direction distributions was obtained with a two-step procedure.

[10]It is worth noting that it is possible to optimize the change-detection results by selecting a different threshold value in the magnitude domain for each annular sector (i.e., for each kind of change present in the considered scene).

TABLE VI
THRESHOLD VALUE $T$ AND ANGULAR THRESHOLD VALUES $\vartheta_{1_1}$, $\vartheta_{1_2}$, $\vartheta_{2_1}$, AND $\vartheta_{2_2}$ OBTAINED BY ADOPTING THE MTEP AND THE BDR WITH THE PROPOSED STATISTICAL MODELS (DOUBLE-CHANGE CASE)

| Change-detection procedure | $T$ | $\vartheta_{1_1}$ | $\vartheta_{1_2}$ | $\vartheta_{2_1}$ | $\vartheta_{2_2}$ | K |
|---|---|---|---|---|---|---|
| MTEP | 45 | 13.17° | 45.84° | 323.87° | 12.02° | 0.9272 |
| BDR | 40 | 13.17° | 57.30° | 323.30° | 13.17° | 0.9270 |

TABLE VII
CONFUSION MATRICES FOR THE CHANGE-DETECTION MAPS OBTAINED WITH (a) THE MTEP AND (b) THE BDR WITH THE PROPOSED STATISTICAL MODELS (DOUBLE-CHANGE CASE)

| | | True Class | | |
|---|---|---|---|---|
| | | $\omega_n$ | $\omega_{c_1}$ | $\omega_{c_2}$ |
| Estimated Class | $\omega_n$ | 113172 | 294 | 338 |
| | $\omega_{c_1}$ | 499 | 7020 | 8 |
| | $\omega_{c_2}$ | 35 | 166 | 2068 |

(a)

| | | True Class | | |
|---|---|---|---|---|
| | | $\omega_n$ | $\omega_{c_1}$ | $\omega_{c_2}$ |
| Estimated Class | $\omega_n$ | 112807 | 114 | 120 |
| | $\omega_{c_1}$ | 798 | 7120 | 13 |
| | $\omega_{c_2}$ | 101 | 246 | 2281 |

(b)

The first step identifies the threshold value $T$ applying the BDR under the hypothesis that the class of unchanged pixels is Rayleigh distributed and the class of changed pixels is given by the sum of two Rice distributions. The second step detects the angular threshold values $\vartheta_{1_1}$, $\vartheta_{1_2}$, $\vartheta_{2_1}$, and $\vartheta_{2_2}$ by applying the BDR under the assumption that the direction distribution of unchanged pixels is uniform in $[0, 2\pi)$ and, that, both direction distributions of classes $\omega_{c_1}$ and $\omega_{c_2}$ follow (36).

As shown in Table VI, also in this case, the threshold value $T$ selected on the basis of the proposed statistical models (i.e., 40) is similar to the threshold value obtained with the MTEP procedure (i.e., 45). Furthermore, also in the direction domain, the threshold-selection procedure based on the proposed statistical models resulted in a pair of threshold values for each annular sector (i.e., $\vartheta_{1_1} = 13.17°$ and $\vartheta_{1_2} = 57.30°$ for $S_1$, and $\vartheta_{2_1} = 323.30°$ and $\vartheta_{2_2} = 13.17°$ for $S_2$) very similar to the optimal ones (i.e., $\vartheta_{1_1} = 13.17°$ and $\vartheta_{1_2} = 45.84°$ for $S_1$, and $\vartheta_{2_1} = 323.87°$ and $\vartheta_{2_2} = 12.02°$ for $S_2$). The similarity of the threshold values resulted in almost the same value of the Kappa coefficient of accuracy for both the change-detection procedures (i.e., 0.9372 for the MTEP and 0.9370 for the BDR). These good results are further confirmed by the analysis of the confusion matrices (see Table VII).

## VII. DISCUSSION AND CONCLUSION

In this paper, a polar theoretical framework for unsupervised change detection based on of the CVA technique has been presented. The main motivation of this paper relies on the observation that the CVA is a widely used technique for unsupervised change detection in multispectral and multitemporal remote sensing images, but a precise theoretical framework concerning its definition and use has not been proposed in the literature (in many applications, CVA is used without a complete understanding of the implications of the representation of the change information in the magnitude-direction domain). In this work, we aimed at filling this gap by introducing: 1) a proper polar framework for the representation and the analysis of multitemporal data in the context of the CVA technique; 2) a set of formal definitions (which are linked to the properties of the data) related to pattern representation in the polar domain; 3) a theoretical analysis of the distributions of changed and unchanged pixels in the polar domain; 4) a critical analysis of the theoretical study of distributions aimed at driving a proper exploitation of the information present in the polar representation; and 5) two examples of use of the proposed framework in change-detection problems.

In the light of the aforementioned contributions, we expect that the main impact of this work in the remote sensing community can be focused on the following issues:

1) possibility to use in all practical applications of the CVA technique (irrespectively of the specific change-detection problem considered) a uniform polar representation with proper formal definitions of the different regions of interest based on the proposed framework;
2) better understanding of the statistical properties of SCVs in the polar domain and of the impact of the simplifying assumptions usually considered in the literature in the development of automatic data-analysis algorithms;
3) presentation of a solid background for the development of advanced and accurate automatic algorithms for change detection, which properly takes into account the statistical properties of data in the polar domain;
4) better understanding of the fundamental role played by a proper preprocessing step (e.g., radiometric corrections, coregistration, etc.) for driving a correct design and use of efficient automatic data-processing algorithms.

The presented analysis points out that some of the simplifying assumptions usually adopted for representing data distributions in the polar domain can become critical if a precise modeling of the change-detection problem is desired. Among the other properties discussed in the Section IV-C, we stress the following four observations.

1) The theoretical and experimental analyses confirm that it may be critical solving the change-detection problems by representing the magnitude of classes of unchanged and changed pixels with Gaussian distributions, rather than using the more accurate models described in this paper (i.e., the Rayleigh distribution for unchanged pixels and the Rice distribution for changed pixels).
2) Radiometric corrections play a fundamental role in unsupervised change detection based on CVA: a) for increasing the separability between the classes of changed and unchanged pixels (by increasing the distance between the mean values of the two classes in the magnitude domain) and b) for properly exploiting the direction information in the data-processing phase. (If the radiometric corrections

are neglected, the direction distribution of the class of unchanged pixels is completely different from the expected uniform model, resulting in an important source of errors in the design of automatic data-processing techniques).

3) The use of the direction information in the change-detection algorithms can be very important for reducing the false alarms induced from registration noise.
4) The use of the direction information in the change-detection algorithms is very important for distinguishing different kind of changes.

The effectiveness of the proposed polar framework and of the related statistical analysis, as well as the importance of their implications, have been verified on two change-detection problems by analyzing qualitatively and quantitatively the reliability of the simplifying assumptions considered in the theoretical analysis of data distributions and their impact on the precision of the models. The results obtained confirm that the theoretical models presented in this paper are suitable for a proper representation of the considered data sets when the CVA technique is adopted.

As a final remark, it is worth noting that in the proposed framework, according to the majority of the literature, we analyze separately the magnitude and direction variables in both the statistical-modeling phase and the threshold-selection process. Although, this is a simplification from the viewpoint of the Bayesian decision theory (which seems reasonable for studying with a limited complexity the statistics of classes), it results in a better understanding of both the physical meaning of these variables and the different roles they play in the change-detection problem. An alternative strategy for a joint analysis of the magnitude and direction variables could be the use of clustering procedures. However, it is difficult to properly model the cluster joint statistical distributions in the MD space.

As future developments of this work, we are: 1) studying the reformulation of some threshold-selection algorithms developed in the literature [15], [16] according to the distributions derived from the theoretical analysis reported in this paper; 2) considering the properties of the direction information for: a) devising effective unsupervised change-detection algorithms capable of automatically identifying different kinds of change in a generic multitemporal data set; b) better reducing the effects of the registration noise in the polar domain; and c) optimizing the framework and the procedure for threshold selection in the magnitude domain by considering different threshold values for different kinds of change.

## APPENDIX

In this Appendix, we provide the guidelines for extending the proposed framework to the $B$-dimensional case ($B > 2$).

Let us consider the general case of CVA applied to $B$ spectral channels of the considered multitemporal images. As stated in Section III-A for the 2-D case, also in the $B$-dimensional case, it is reasonable to represent $X_{\mathrm{D}}$ as a mixture of multidimensional Gaussian distributions as it is obtained by subtracting the multidimensional random variables $X_1$ and $X_2$, both modeled with a mixture of multidimensional Gaussian distributions. Furthermore, also in this case, we assume that features $X_{b,\mathrm{D}} (b = 1, \ldots, B)$ are statistically independent.[11] Under this assumption, the distribution of the class $\omega_i (\omega_i \in \Omega)$ in a Cartesian coordinate system can be written as the product of the $B$ marginal densities $p(X_{b,\mathrm{D}}|\omega_i)$ related to the class $\omega_i (b = 1, \ldots, B)$, i.e.,

$$p(X_{\mathrm{D}}|\omega_i) = \frac{1}{(2\pi)^{\frac{B}{2}} \prod_{b=1}^{B} \sigma_{b,i}} \exp\left[ -\sum_{b=1}^{B} \frac{(X_{b,\mathrm{D}} - \mu_{b,i})^2}{2\sigma_{b,i}^2} \right] \quad (38)$$

where $\mu_{b,i}$ and $\sigma_{b,i}$ are the mean values and the standard deviations, respectively, of the Gaussian distributed marginal density of class $\omega_i$ over the $b$th considered feature ($b = 1, \ldots, B$). Equation (38) can be written in hyperspherical coordinates using (2), as given in (39), shown at the top of the next page.

The marginal conditional densities of the random variables $\rho$, $\vartheta$, and $\varphi_1, \varphi_2, \ldots, \varphi_{B-2}$ for each class $\omega_i$ can be computed by proper integrations of (39). The general formulation of the integrals to be evaluated is given in the following:

$$p(\rho|\omega_i) = \int_0^{2\pi} \left[ \int_0^{\pi} \cdots \int_0^{\pi} p(\rho, \vartheta, \varphi_1, \ldots, \varphi_{B-2}|\omega_i) \times d\varphi_1 \ldots d\varphi_{B-2} \right] d\vartheta \quad (40)$$

$$p(\vartheta|\omega_i) = \int_0^{+\infty} \left[ \int_0^{\pi} \cdots \int_0^{\pi} p(\rho, \vartheta, \varphi_1, \ldots, \varphi_{B-2}|\omega_i) \times d\varphi_1 \ldots d\varphi_{B-2} \right] d\rho \quad (41)$$

$$p(\varphi_h|\omega_i) = \int_0^{+\infty} \int_0^{2\pi} \left[ \int_0^{\pi} \cdots \int_0^{\pi} p(\rho, \vartheta, \varphi_1, \ldots, \varphi_{B-2}|\omega_i) \times d\varphi_1 \ldots d\varphi_{B-2} \right] d\rho d\vartheta \quad (42)$$

where $d\varphi_h$ is not included in the integration variables.

In the general case, the models to be adopted should be based on the solution of the integral (40)–(42). These solutions can be simplified in some particular cases. In the following, we report the simplified equations for a pair of important situations analogous to those shown in Sections IV-A and IV-B.

As stated in Section IV-A, it is reasonable to expect that the mean along each of the $B$ spectral channels for the class of unchanged pixels is equal to zero (i.e., $\mu_{b,\omega_{\mathrm{n}}} \cong 0, b = 1, \ldots, B$).

[11] Discussion on this hypothesis is reported in Section IV-C.

$$p(\rho,\vartheta,\varphi_1,\ldots,\varphi_{B-2}|\omega_i)=\frac{1}{(2\pi)^{\frac{B}{2}}\prod_{b=1}^{B}\sigma_{b,i}}\exp\left[-\frac{\left(\rho\left(\prod_{k=1}^{B-2}\sin\varphi_k\right)\sin\vartheta-\mu_{B,i}\right)^2}{2\sigma_{B,i}^2}\right.$$
$$\left.-\frac{\left(\rho\left(\prod_{k=1}^{B-2}\sin\varphi_k\right)\cos\vartheta-\mu_{B-1,i}\right)^2}{2\sigma_{B-1,i}^2}-\sum_{b=1}^{B-2}\frac{\left(\rho\left(\prod_{k=1}^{b-1}\sin\varphi_k\right)\cos\varphi_b-\mu_{b,i}\right)^2}{2\sigma_{b,i}^2}\right] \quad (39)$$

A significantly simplified equation can be obtained in the approximation that all the standard deviations are similar to each other (i.e., $\sigma_{b,\omega_n}\cong\sigma_{\omega_n}, b=1,\ldots,B$). Under this assumption, we can rewrite (40) and (41) as

$$p(\rho|\omega_k)=\frac{2\rho^{B-1}}{\left(2\sigma_{\omega_k}^2\right)^{\frac{1}{2}B}\Gamma\left(\frac{1}{2}B\right)}\exp\left(-\frac{\rho^2}{2\sigma_{\omega_k}^2}\right),\qquad \rho\geq 0$$
$$p(\vartheta|\omega_n)=\frac{1}{2\pi},\qquad \vartheta\in[0,2\pi). \quad (43)$$

Considering changed pixels, it is possible to derive a significantly simplified equation for the distribution of the magnitude in the case in which the mean along each of the $B$ dimensions is not equal to zero (i.e., $\mu_{p,\omega_k}\neq\mu_{q,\omega_k}\neq 0,\ p,\ q=1,\ldots,B$ with $p\neq q$) and under the assumption that all the standard deviations are similar to each other (i.e., $\sigma_{b,\omega_k}\cong\sigma_{\omega_k},\ b=1,\ldots,B$). Under these hypotheses, (40) can be rewritten as

$$p(\rho|\omega_k)=\frac{M_{\omega_k}}{\sigma_{\omega_k}^2}\left(\frac{\rho}{M_{\omega_k}}\right)^{\frac{1}{2}B}\exp\left(-\frac{\rho^2+M_{\omega_k}^2}{2\sigma_{\omega_k}^2}\right)$$
$$\times\,\mathrm{I}_{\frac{1}{2}(B-2)}\left(\frac{\rho M_{\omega_k}}{\sigma_{\omega_k}^2}\right),\qquad \rho\geq 0. \quad (44)$$

## ACKNOWLEDGMENT

The authors would like to thank the anonymous referees for their constructive criticism.

## REFERENCES

[1] A. Singh, "Digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, 1989.
[2] P. R. Coppin, I. Jonckheere, and K. Nachaerts, "Digital change detection in ecosystem monitoring: A review," *Int. J. Remote Sens.*, vol. 25, no. 9, pp. 1565–1596, May 2004.
[3] D. Lu, P. Mausel, E. Brondízio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2407, Jun. 2004.
[4] L. Bruzzone and R. Cossu, "Adaptive approach to reducing registration noise effects," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2455–2465, Nov. 2003.
[5] W. A. Malila, "Change vector analysis: An approach for detecting forest changes with Landsat," in *Proc. LARS Mach. Process. Remotely Sensed Data Symp.*, W. Lafayette, IN, 1980, pp. 326–336.
[6] J. L. Michalek, T. W. Wagner, J. J. Luczkovich, and R. W. Stoffle, "Multispectral change vector analysis for monitoring coastal marine environments," *Photogramm. Eng. Remote Sens.*, vol. 59, no. 3, pp. 381–384, 1993.
[7] L. A. Virag and J. E. Colwell, "An improved procedure for analysis of change in Thematic Mapper image-pairs," in *Proc. 21st Int. Symp. Remote Sens. Environ.*, Ann Arbor, MI, Oct. 26–30, 1987, pp. 1101–1110.
[8] J. R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.
[9] R. D. Johnson and E. S. Kasischke, "Change vector analysis: A technique for multispectral monitoring of land cover and condition," *Int. J. Remote Sens.*, vol. 19, no. 3, pp. 411–426, 1998.
[10] T. R. Allen and J. A. Kupfer, "Application of spherical statistics to change vector analysis of Landsat data: Southern Appalachian spruce-fir forests," *Remote Sens. Environ.*, vol. 74, no. 3, pp. 482–493, Dec. 2000.
[11] J. Chen, P. Gong, C. He, R. Pu, and P. Shi, "Land-use/land-cover change detection using improved change-vector analysis," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 4, pp. 369–379, 2003.
[12] T. Warner, "Hyperspherical direction cosine change vector analysis," *Int. J. Remote Sens.*, vol. 26, no. 6, pp. 1201–1215, 2005.
[13] K. Nackarets, K. Vaesen, B. Muys, and P. Coppin, "Comparative performance of a modified change vector analysis in forest change detection," *Int. J. Remote Sens.*, vol. 26, no. 5, pp. 839–852, 2005.
[14] E. M. Pereira and A. W. Stezer, "Spectral characteristic of fire scars in Landsat-5 TM images of Amazonia," *Int. J. Remote Sens.*, vol. 14, no. 11, pp. 2061–2078, 1993.
[15] L. Bruzzone and D. Fernández Prieto, "A minimum-cost thresholding technique for unsupervised change detection," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3539–3544, Dec. 2000.
[16] ——, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1170–1182, May 2000.
[17] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. New York: Springer-Verlag, 1999.
[18] P. Beckmann, *Probability in Communication Engineering*. New York: Harcourt, Brace & World, 1967.
[19] J. R. G. Townshend, C. O. Justice, and C. Gurney, "The impact of misregistration on change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 5, pp. 1054–1060, Sep. 1992.
[20] L. Bruzzone and S. B. Serpico, "Detection of changes in remotely-sensed images by the selective use of multi-spectral information," *Int. J. Remote Sens.*, vol. 18, no. 18, pp. 3883–3888, 1997.
[21] P. S. Chavez, Jr., "Radiometric calibration of Landsat Thematic Mapper multispectral images," *Photogramm. Eng. Remote Sens.*, vol. 55, no. 9, pp. 1285–1294, 1989.
[22] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. Boston, MA: Academic, 1990.
[23] P. G. Hoel, S. C. Port, and C. J. Stone, *Introduction to Statistical Theory*. Atlanta, GA: Houghton Mifflin, 1971.

**Francesca Bovolo** (S'05) received the Laurea (B.S.) and the Laurea Specialistica (M.S.) degrees in telecommunication engineering (*summa cum laude*) from the University of Trento, Italy, in 2001 and 2003, respectively, where since 2003, she has been working toward the Ph.D. degree in communication and information technologies.

She is currently with the Remote Sensing Laboratory, Department of Telecommunication and Information Technologies, University of Trento. Her main research activity is in the area of remote sensing image processing; in particular, her interests are related to change detection in multispectral and SAR images. She conducts research on these topics within the frameworks of several national and international projects.

Ms. Bovolo is the recipient of the first-place award in the Student Prize Paper Competition of the 2006 IEEE International Geoscience and Remote Sensing Symposium in Denver, in August 2006. She has served on the Scientific Committee of the International Society for Optical Engineers International Conference on "Image and Signal Processing for Remote Sensing XI," Stockholm, Sweden, in September 2006. She is a Referee for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the *International Journal of Remote Sensing* and *Remote Sensing of Environment*.

**Lorenzo Bruzzone** (S'95–M'98–SM'03) received the Laurea (M.S.) degree in electronic engineering (*summa cum laude*) and the Ph.D. degree in telecommunications from the University of Genoa, Genoa, Italy, in 1993 and 1998, respectively.

From 1998 to 2000, he was a Postdoctoral Researcher at the University of Genoa. From 2000 to 2001, he was an Assistant Professor at the University of Trento, Italy, where from 2001 to 2005, he was an Associate Professor, and since March 2005, he has been a Full Professor of telecommunications, teaching remote sensing, pattern recognition, and electrical communications. He is concurrently the Head of the Remote Sensing Laboratory in the Department of Information and Communication Technology, University of Trento. His current research interests are in the area of remote sensing image processing and recognition (analysis of multitemporal data, feature selection, classification, regression, data fusion, and machine learning). He conducts and supervises research on these topics within the frameworks of several national and international projects. He is the author (or coauthor) of more than 150 scientific publications, including journals, book chapters, and conference proceedings. He is a Referee for many international journals and has served on the Scientific Committees of several international conferences.

Dr. Bruzzone is the recipient of the first-place award in the Student Prize Paper Competition of the 1998 IEEE International Geoscience and Remote Sensing Symposium, Seattle, in July 1998. He was a recipient of the Recognition of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Best Reviewers in 1999 and was a Guest Editor of a Special Issue of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING on the subject of the analysis of multitemporal remote sensing images, in November 2003. He was the General Chair and Cochair of the first and second IEEE International Workshop on the Analysis of Multi-temporal Remote Sensing Images. Since 2003, he has been the Chair of the International Society for Optical Engineers Conference on Image and Signal Processing for Remote Sensing. He is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He is a member of the Scientific Committee of the India–Italy Center for Advanced Research. He is also a member of the International Association for Pattern Recognition and of the Italian Association for Remote Sensing.