Title: A Framework for Automatic and Unsupervised Detection of Multiple Changes in Multitemporal Images

This paper appears in: IEEE Transactions on Geoscience and Remote Sensing

Date of Publication: May 16, 2012

Author(s): Francesca Bovolo, Silvia Marchesi, Lorenzo Bruzzone

Volume: 50, Issue: 6

Page(s): 2196 - 2212

DOI: 10.1109\TGRS.2011.2171493

A FRAMEWORK FOR AUTOMATIC AND UNSUPERVISED DETECTION OF MULTIPLE CHANGES IN MULTITEMPORAL IMAGES

Francesca BOVOLO, Silvia MARCHESI, Lorenzo BRUZZONE

Information Engineering and Computer Science Dept., University of Trento, Via Sommarive, 14, I-38123, Povo, Trento, Italy

E-mail: {francesca.bovolo, silvia.marchesi}@disi.unitn.it, lorenzo.bruzzone@ing.unitn.it

***Abstract* – The detection of multiple changes (*i.e.*, different kinds of change) in multitemporal remote sensing images is a complex problem. When multispectral images having *B* spectral bands are considered, an effective solution to this problem is to exploit all available spectral channels in the framework of supervised or partially supervised approaches. However, in many real applications it is difficult/impossible to collect ground truth information for either multitemporal or single date images. On the opposite, unsupervised methods available in the literature are not effective in handling the full information present in multispectral and multitemporal images. They usually consider a simplified sub-space of the original feature space having small dimensionality and thus characterized by a possible loss of change information. In this paper, we present a framework for the detection of multiple changes in bi-temporal and multispectral remote sensing images that allows one to overcome the limits of standard unsupervised methods. The framework is based on: i) a compressed yet efficient 2-dimensional (2D) representation of the change information; and ii) a 2-step automatic decision strategy. The effectiveness of the proposed approach has been tested on two bi-temporal and multispectral data sets having different properties. Results obtained on both data sets confirm the effectiveness of the proposed approach.**

Key words – **multitemporal images, multiple changes, change detection, change vector analysis, low dimensional representation, thresholding procedure, Bayes decision rule, remote sensing.**

I. Introduction

In the literature, the problem of multiple-change detection (*i.e.*, the detection of different kinds of change) has been usually treated as a problem of explicitly detecting land-cover transitions according to (semi-, partially-) supervised methods [1],[2]. The easiest approach in such cases is Post-Classification Comparison (PCC), where two multitemporal images, acquired over the same area at different times, are independently classified and land-cover transitions are estimated according to a pixel-by-pixel comparison of classification maps [3]. The performance of this approach critically depends on the accuracies of the single classification maps and (under the assumption of independent errors in the maps) it is close to the product of the accuracies yielded at the two times. This method has the drawback that it does not consider temporal correlation between available acquisitions. A possible alternative is given by Direct Multidate Classification (DMC) [3]. In this technique the two images acquired at different dates are simultaneously classified by stacking their feature vectors. Each possible transition is considered as a class, and thus a training set made up of pixels with labels for both available acquisitions should be defined. In real applications this represents a strong constraint. In order to overcome the drawbacks that affect the two aforementioned approaches, recently other methods have been developed in the context of partially supervised or domain adaptation techniques. These methods assume that ground truth information is available for only one acquisition date while it is not given for the second one. Information about class transitions is obtained by jointly exploiting unlabeled patterns from the second acquisition and labeled patterns available for the first one [4]-[6]. However, also these methods require the availability of ground truth information for at least one of the images to be analyzed. When dealing with real applications, the ground truth information collection requires a significant effort from the economical and practical viewpoint. Moreover, in many cases, due to real application constraints, it is almost impossible to retrieve such kind of information. In order to cope with these situations, unsupervised techniques have been developed, which do not require any prior information about land-cover classes. Exhaustive surveys of unsupervised change-detection methods for multispectral images acquired by passive sensors can be found in [3],[7]-[16]. Despite these methods can perform change detection without prior information and with a reduced computational burden, most of them allow only the detection of presence/absence of changes but do not discriminate different kinds of change. In the literature some examples exist of methods that try to distinguish in an unsupervised way between different kinds of change [17]. However, often they require the selection of only 2 (or few) spectral channels among the available ones. This process may lead to a significant loss of information, a degradation of the accuracy of the change-detection process and a failure to identify some kinds of change. Moreover unsupervised methods for the detection of multiple changes at the state-of-art (included [17]) do not address the problem of the change information extraction in an automatic way, neither in the full-dimensional nor in a lower dimensional representation of SCVs. From this analysis it emerges a lack of unsupervised methods being able to properly detect the presence of multiple changes in an unsupervised and automatic way.

In this manuscript we propose a framework for the detection of multiple changes in bi-temporal and multispectral remote sensing images, which allows one to overcome the limits of standard unsupervised methods. The framework is based on: i) a compressed yet efficient 2-dimensional (2D) representation of the change information; and ii) a 2-step automatic decision strategy. First the multidimensional feature space of SCVs is compressed into a 2-dimensional one without neglecting any available spectral band (and thus possible information about changes). This representation allows one to easily display and understand change information in a polar coordinates system. Second, an automatic 2-step method for both separating unchanged from changed patterns and distinguishing different kinds of change is presented.

The rest of the paper is organized into six sections. The next section introduces mathematical notation and summarizes the background behind the proposed framework. Section III introduces the proposed compressed 2D representation and the characterization of the change information in bi-temporal multispectral images. Section IV presents the proposed technique for extracting multiple-change information and computing the final change-detection map. Section V illustrates the experimental results obtained on two real multitemporal datasets acquired by Landsat-5 and Quickbird satellites multispectral sensors. Finally, Section VI draws the conclusion of this work.

II. NOTATION AND BACKGROUND

Let us consider two coregistered multispectral images, $\mathbf{X}_1$ and $\mathbf{X}_2$ of size $I \times J$ acquired over the same geographical area at different times t_1 and t_2, respectively. Let $X_{b,t}$ be the image representing the bth (b=1,…,B) component of the multispectral image $\mathbf{X}_t$ (t=1,2). Unsupervised change-detection methods usually exploit the multispectral difference image $\mathbf{X}_D$ by subtracting the spectral feature vectors associated with each corresponding spatial position in the two considered images $\mathbf{X}_1$ and $\mathbf{X}_2$, *i.e.*,

$$\mathbf{X}_D = \mathbf{X}_2 - \mathbf{X}_1 \tag{1}$$

Let $X_{b,D}$ be the image representing the bth (b=1,…,B) component of $\mathbf{X}_D$. Finally, let $\boldsymbol{\Omega}=\{\omega_n, \boldsymbol{\Omega}_c\}$ be the set of classes to be identified. ω_n represents the class of no-changed pixels and $\boldsymbol{\Omega}_c=\{\omega_{c1}, \omega_{c2}, \ldots, \omega_{cK}\}$ is a meta-class that gathers all the K possible classes (kinds) of change occurred in the considered area.

In the past, many unsupervised approaches have been developed for exploiting the information present in $\mathbf{X}_D$ [3],[8],[9]. The most common and easiest one reduces the BD problem to a 1D problem [18]-[20], by considering only the magnitude ρ of spectral change vectors (SCVs). A simple thresholding of the magnitude variable allows one to obtain a change-detection map that highlights the presence/absence of changes [18]-[20]. However, in this way no information can be retrieved about possible different kinds of change (i.e., multiple changes cannot be distinguished).

The above mentioned drawback can be addressed by the definition of more advanced techniques that try to solve the change-detection problem by including all available features in the decision process.[1] In this case the detection of changes requires the solution of a complex BD problem, where an unsupervised analysis would imply the application of clustering algorithms to BD vectors [17],[21]. However, in real applications, the data complexity and the noise present in the BD feature space (refer to [17] for further details on this issue) affect the performance of clustering procedure, which in many cases result in change-detection accuracies smaller than those provided by a simple thresholding of the 1D magnitude of SCVs [3],[19]. A further drawback of working in a BD space is that this space is difficult or impossible to visualize when the considered dimensions are more than 2. This implies that the process of understanding the change-detection problem structure when semi-automatic interactive solutions are investigated with the support of an expert can become rather complex.

A possible alternative to solve the BD problem with a limited loss of information is to split it in a set of $\binom{B}{2}$ 2D problems by considering all possible pairs of spectral channels. The obvious drawback of this approach is the need of defining an effective strategy for combining the $\binom{B}{2}$

[1] The BD feature space could be either the one of the multispectral difference image or an alternative multidimensional representation of it like the one achieved by Principal Component Analysis (PCA).

solutions in the final decision step (It is worth noting that in the literature it doesn't exist a change-detection method based on this strategy). Therefore the most common practice is to select only one out of all the possible 2D problems (*i.e.*, neglecting *B*-2 spectral bands) and to use this sub-optimal representation as the solution to the initial *B*D change-detection problem [22],[23]. This is for instance done in PCA-based change-detection approaches where only the first 2 (or few) principal components are selected for the solution of the *B*D problem [24]. In practice, the two selected spectral channels of $\mathbf{X}_D$ are commonly used to represent the change-detection problem in 2D polar coordinates (2D-CVA) according to the magnitude and direction variables:

$$\begin{cases} \rho = \sqrt{\mathrm{X}_{m,D}^2 + \mathrm{X}_{n,D}^2} \\ \vartheta = \arctan\left(\dfrac{\mathrm{X}_{m,D}}{\mathrm{X}_{n,D}}\right) \end{cases} \tag{2}$$

where $\mathrm{X}_{n,D}$ and $\mathrm{X}_{m,D}$ represent the considered *n*th and *m*th spectral channels of $\mathbf{X}_D$, respectively.

Independently on the selected channels, the drawback of this strategy is that the change-detection solution is usually affected by a loss of information (except for simple cases) with respect to the original multitemporal and multispectral feature space (or to the *B*D SCVs feature space). To limit this effect, prior knowledge on the specific considered problem (*i.e.*, on the kinds of change occurred on the ground) could be employed to select the 2 most relevant channels [19],[25]. However, in most of the practical applications, prior information is not available and it is not possible to assure that change information is constrained to only two channels (*e.g.*, there are different kinds of change that affect the spectral signatures of the land covers in different bands). Thus the 2D representation can result in poor change-detection performance. Nevertheless, it shows the advantage of being easy to visualize and analyze.

Table I summarizes advantages and disadvantages of the different representations available

TABLE I ADVANTAGES AND DISADVANTAGES OF THE DIFFERENT POSSIBLE REPRESENTATION OF THE *B*D CHANGE-DETECTION PROBLEM GIVEN BY THE MULTISPECTRAL DIFFERENCE IMAGE.

Representation	Unsupervised approach to CD	Advantages	Disadvantages
***B*D (*B*>2)**	• *B*D clustering.	• Information about multiple kinds of change is preserved.	• Complex to manage. • Change information cannot be to visualized. • Clustering techniques are often not effective.
***B*D (sub-optimal) (*B*>2)**	• Solve $\binom{B}{2}$ 2D problems. • Combine $\binom{B}{2}$ solutions.	• Information about multiple kinds of change is preserved.	• Sub-optimal detection of changes. • The combination of 2D solutions for generating the *B*D solution requires an additional step. • Combination strategies are not available in the literature yet.
2D	• Select 2 out of *B* bands. • Threshold magnitude and direction variables.	• Easy and intuitive to visualize. • Different kinds of change can be detected.	• Requires prior information about changes for band selection. • Depending on selected bands some kinds of change can be lost.
1D	• Threshold the magnitude variable.	• Easy to manage and visualize.	• Only information about presence/absence of changes can be extracted.

in the literature. From its analysis and recalling that different kinds of change have different effects on different features (*i.e.*, all spectral channels are potentially useful for solving the change-detection problem and no channel can be neglected a priori), it emerges the need of defining a framework where the information about multiple changes can be easily managed in a 2 dimensional feature space without completely neglecting any spectral channel (and the information about changes in them). Moreover the framework should integrate effective change-detection techniques able to distinguish multiple changes in an unsupervised and automatic way.

III. PROPOSED COMPRESSED REPRESENTATION OF THE CHANGE INFORMATION

In order to preserve the most of the available information present in the *B*D feature space and to obtain a feature space easy to visualize and manage from a user point of view, here we propose a transformation that maps the *B*D feature space into a 2D feature space without the need of any prior information about the specific change-detection problem. The two considered features are: i) the magnitude of spectral change vectors, and ii) a direction variable that models the information about different kinds of change without rejecting any spectral channel. The two features define a space in which the change information can be effectively and intuitively represented and extracted.

A. Magnitude of Spectral Change Vectors

The first of the considered features is the well known and widely used magnitude ρ of multidimensional spectral change vectors in $\mathbf{X}_D$ defined as:

$$\rho = \sqrt{\sum_{b=1}^{B} X_{b,D}^2} = \sqrt{\sum_{b=1}^{B} \left(X_{b,2} - X_{b,1}\right)^2}, \quad \rho \in [0, \rho_{max}]. \tag{3}$$

where ρ_{max} is the maximum value assumed by the magnitude for the considered image pair. Theoretically ρ_{max} could tend to infinity, however in practical applications it is bounded by the digital nature of the data.

As widely known [3],[19], the magnitude carries information about the presence/absence of changes. On this feature changed pixels show a relative high value whereas unchanged pixels assume a relatively low value [3],[17]-[19]. Despite the magnitude does not carry information about different kinds of change, it represents a valuable and robust variable for distinguishing changed from no-changed pixels. In the literature several automatic and unsupervised approaches to change detection that analyze the magnitude variable are available [3],[7],[9]. Among them the most widely used are those based on automatic thresholding techniques [9],[18],[19].

B. Direction of Spectral Change Vectors

As the magnitude of SCVs does not include any information about different kinds of change, a complementary feature is proposed to distinguish multiple changes. A measure α [alternative to ϑ, see eq. (2)] is proposed that effectively compresses the information about different kinds of change to a 1-dimensional variable. α is defined as an angle computed in radians between two multidimensional vectors t and r

$$\alpha = \arccos\left(\sum_{b=1}^{B}(t_b r_b) \Big/ \sqrt{\sum_{b=1}^{B} t_b^2 \sum_{b=1}^{B} r_b^2}\right), \quad \alpha \in [0, \pi] \tag{4}$$

where t_b and r_b are the bth components of BD vectors t and r, respectively [26].

Such kind of measure has been already successfully employed in the context of: i) supervised approaches to geological mapping in high and very high geometrical resolution images [27]-[29]; ii) supervised classification and/or clustering of hyperspectral and multispectral images [30],[31]; iii) supervised change detection [32]; iv) spectral unmixing [33],[34]; v) target detection in hyperspectral images [35]; and vi) pansharpening quality assessment. In these applications the angle defined in (4) is used as a supervised similarity measure between a given spectral signature **X** and reference spectra $\mathbf{X}_{ref}$ (*i.e.*, spectral libraries or end-members stored in a database) and is commonly referred to as Spectral Angle Mapper (SAM) [36]. In such cases, equation (4) can be rewritten as

$$\alpha = \arccos\left(\sum_{b=1}^{B}(\mathrm{X}_b \mathrm{X}_{b,ref}) \Big/ \sqrt{\sum_{b=1}^{B} \mathrm{X}_b^2 \sum_{b=1}^{B} \mathrm{X}_{b,ref}^2}\right), \quad \alpha \in [0, \pi] \tag{5}$$

where X_b and $\mathrm{X}_{b,ref}$ are the bth components of BD vectors **X** and $\mathbf{X}_{ref}$, respectively.

However, as we are dealing with an unsupervised approach to change detection no libraries for $\mathbf{X}_{ref}$ are available. Therefore in this work we propose an alternative way to use the angular distance measure defined in (4). In the BD feature space of the multispectral difference image $\mathbf{X}_\mathrm{D}$, we define the vector t as the spectral change vector associated to each spatial position and r as a BD unit vector u with elements u_b (b=1,…,B) all equal to each other. The latter choice corresponds to use a reference vector in which the same weight is given to all spectral channels in the change analysis procedure. This is due to the absence of prior information about changes occurred on the ground that doesn't make it possible to establish a relative relevance of spectral channels and thus a more specific reference vector. Without loss of generality, in order to define a normalized vector u_b we set elements of u equal to $\sqrt{B}/B$. Therefore the desired angle variable can be written as:

$$\alpha = \arccos\left(\sum_{b=1}^{B} \left(\mathrm{X}_{b,\mathrm{D}} u_b \right) \Big/ \sqrt{\sum_{b=1}^{B} \mathrm{X}_{b,\mathrm{D}}^2 \sum_{b=1}^{B} u_b^2} \right) = \arccos\left[\frac{1}{\sqrt{B}} \left(\sum_{b=1}^{B} \mathrm{X}_{b,\mathrm{D}} \Big/ \sqrt{\sum_{b=1}^{B} \mathrm{X}_{b,\mathrm{D}}^2} \right) \right], \quad \alpha \in [0, \pi] \tag{6}$$

It is worth noting that if any information is available to establish a relative relevance of the spectral channels, different choices can be made for the elements of u.

C. Proposed Compressed Change Vector Analysis

Following an approach similar to the one in [17], the properties of the two features defined according to (3) and (6) can be exploited for defining a compressed polar representation of the change-detection problem represented by the multispectral SCVs in a 2D feature space. We call this feature space as Compressed CVA (C^{2}VA) domain. The C^{2}VA domain is bounded by the ranges of existence of ρ and α, *i.e.*,

$$\mathrm{C}^2\mathrm{VA} = \left\{ \rho \in [0, \rho_{max}] \ \text{ and } \ \alpha \in [0, \pi] \right\} \tag{7}$$

Eq. (7) represents a semi-circle that includes all SCVs of the considered images (see Figure 1). Within this domain, regions of interest can be identified associated to different classes in Ω. Since no-changed pixels are expected to have a magnitude close to zero, whereas changed pixels are expected to show a magnitude far from zero [3],[17],[19], the C^{2}VA domain can be divided into two regions with respect to the magnitude variable. The first region is associated with unchanged pixels, whereas the second one is associated with changed pixels. The two regions can be separated according to the optimal (in the sense of the theoretical Bayesian decision theory) threshold value T that separates pixels belonging to ω_n from pixel belonging to Ω_c (dark and light gray areas in Figure 1, respectively) [17],[19].

The first region is the *semicircle SC_n of no-changed pixels* (dark gray area in Figure 1) located close to the origin of the C^{2}VA domain. This region is defined mathematically as follows

$$SC_n = \{\rho, \alpha : 0 \le \rho < T \text{ and } 0 \le \alpha \le \pi\} \tag{8}$$

The second region is the *semi-annulus of changed pixels SA_c* (light gray area in Figure 1) located far from the origin of the C^{2}VA domain, and is mathematically defined as

$$SA_c = \{\rho, \alpha : T < \rho \le \rho_{\max} \text{ and } 0 \le \alpha \le \pi\} \tag{9}$$

Let us now consider the information carried out by the direction variable α. As it represents the similarity between each considered SCV and a reference vector, it is expected that different kinds of change will be characterized by different values of α. According to this observation, within the semi-annulus of changed pixels different *annular sectors S_k* (k=1,…,K) *of the semi-annulus SA_c* can be detected along α, and defined as

$$S_k = \{\rho, \alpha : \rho \ge T \text{ and } \alpha_{k_1} \le \alpha < \alpha_{k_2},\ 0 \le \alpha_{k_1} < \alpha_{k_2} \le \pi\} \tag{10}$$

where α_{k_1} and α_{k_2} are the two angular thresholds that bounds the sector S_k. Each sector (hatched area in Figure 1) can be associated in principle with a specific kind of change $\omega_{c_k} \in \Omega_c$ occurred on the ground.

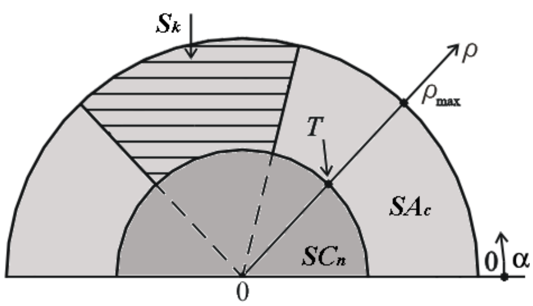


Figure 1. Regions of interest for the compressed 2D representation of the change-detection problem.

D. Discussion

Despite the applied compression considers all spectral channels, some ambiguity rises from the process of information compression which is mainly due to the simplified representation given by the direction variable. The loss of information may result in similar values of α for different kinds of change. In this situation, each detected sector S_k must be associated to more kinds of change in Ω_c instead of only one. However, this is a drawback common to other low dimensional representations (*i.e.*, lower than the original one) usually adopted for unsupervised change detection. Nevertheless, the defined 2D feature space has two valuable advantages: i) the ambiguity does not affect the detection of changes (the magnitude is used like in standard CVA) but just the possible merging (in some specific cases) of different kinds of change; and ii) it considers in the solution of the change-detection problem all available spectral bands thus avoiding the need of prior information about relevant features. This is important because can result in the lost of unexpected kinds of changes having a high value in the operational applications. Moreover the obtained 2D representation makes it easy to visualize the B-dimensional change-detection problem for interaction with the end-user.

In order to better understand the importance of the two mentioned advantages let us compare the proposed representation of change information with the ones that can be obtained by applying either the standard CVA to a pair of spectral channels (2D-CVA) or the PCA-based change-detection methods. In 2D-CVA, 2 out of B spectral channels are selected according to prior information about the spectral behavior of possible kinds of changes occurred in the considered area. Even if good prior information is available, this implies a complete loss of information about those changes (maybe unexpected but also for this reason possibly important) that are visible only in the neglected spectral channels, and a partial loss for the ones visible both in neglected and selected ones. Thus the 2D-CVA allows one to obtain a suitable 2-dimensional representation of the change-detection problem, but can entirely miss some kinds of changes. In PCA-based approaches, the relevant information is compressed in the first principal components (*i.e.*, the ones with the highest eigenvalues). However, when dealing with changes, it is not possible to ensure that the change information is associated only to the first principal components. In other words, we cannot know a priori in which principal components the change information is represented (when changes are a minority in the image, their information can fall also in the last principal components). Furthermore, in the selection of principal components, prior information about the considered data set can not be used like in 2D-CVA, as there is no explicit relationship between each principal component and the physic of the problem. Thus, also for the PCA-based methods the selection of 2 out of B principal components (to obtain a 2D representation of the change-detection problem) results in the possible loss of information about some kinds of changes. This reasoning can be easily generalized to other kinds of transformation techniques that result in more than 2 transformed components.

Concluding, for both the 2D-CVA and the PCA-based change-detection methods, the selection of 2 features out of B can result in losing the information on some specific kinds of change, whereas the proposed representation results, in the worst case, in misclassification among kinds of changes but not in misdetections. Moreover, given a pair of images of the same area. the proposed compressed representation of the change information results in a well defined 2D feature space. On the contrary, for both 2D-CVA and PCA-based change-detection methods the 2D representation of change information significantly changes according to the selected features.

IV. PROPOSED TECHNIQUE FOR THE DETECTION OF MULTIPLE CHANGES

The proposed 2D representation suggests a change-detection approach based on a 2-step procedure: i) identification of the *semicircle SC_n of no-changed pixels* and of the *semi-annulus SA_c of changed pixels* (*i.e.*, separation of changed and unchanged patterns) by the analysis of the distribution of the magnitude variable ρ; and ii) identification of *annular sectors S_k* (k=1,…,K) in the *semi-annulus SA_c* (*i.e.*, detection of different kinds of change within the set of changed patterns identified in the first step) by the analysis of the distribution of the direction variable α. It is worth noting that from a theoretical point of view the identification of the different regions in the C^{2}VA domain should be carried out by jointly analyzing ρ and α. Nonetheless we simplify the process by analyzing separately ρ and α thus implicitly assuming the independence between them.

A. Discrimination between Changed and Unchanged Pixels

In the first step changed and unchanged pixels are distinguished from each other according to a well known and widely used unsupervised technique based on the Expectation-Maximization algorithm [37],[38].

Let $P(\omega_n)$, $P(\Omega_c)$, $p(\rho|\omega_n)$ and $p(\rho|\Omega_c)$ be the prior probabilities and the conditional probability density functions of class ω_n and meta-class Ω_c, respectively. Let us assume that the distribution of the observed magnitude variable can be expressed as a mixture density distribution, *i.e.*:

$$p(\rho) = P(\omega_n)\, p(\rho \mid \omega_n) + P(\Omega_c)\, p(\rho \mid \Omega_c) \tag{11}$$

Under simple assumptions it is possible to prove that in the 2D case the magnitude of changed and unchanged classes can be modeled by a Rayleigh and a Rice density function, respectively (see [17] for greater details). However, in the considered case these hypotheses are not satisfied as: i) more than two spectral channels are involved in the process; ii) Ω_c can include in general more than one kind of change; and iii) no assumption is made on the statistical parameters of changed and unchanged pixels in the domain of the multispectral difference image. According to these considerations, the simplifying assumption that $p(\rho|\omega_n)$ and $p(\rho|\Omega_c)$ follow a Gaussian distribution seems a reasonable and simple approximation. The threshold T that separates class ω_n and meta-class Ω_c can be computed according to the Bayes decision theory after retrieving the class prior probabilities $P(\omega_n)$ and $P(\Omega_c)$ and the class statistical parameters (the mean values $\mu_{c,\rho}$ and $\mu_{n,\rho}$ and variances $\sigma^2_{c,\rho}$ and $\sigma^2_{n,\rho}$ in the magnitude domain ρ of change and no-change classes, respectively). As change detection is approached in an unsupervised way, the well know Expectation-Maximization algorithm [37],[38] can be used for estimating these parameters. After initialization, the following iterative equations that allow us to solve the estimation problem under Gaussian assumption can be applied [19]:

$$P^{s+1}(\omega_n)=\frac{1}{IJ}\sum_{\rho(i,j)\in\rho}\frac{P^s(\omega_n)p^s(\rho(i,j))|\omega_n)}{p^s(\rho(i,j))} \tag{12}$$

$$\mu_{n,\rho}^{s+1}=\frac{\sum_{\rho(i,j)\in\rho}\frac{P^s(\omega_n)p^s(\rho(i,j))|\omega_n)}{p^s(\rho(i,j))}\rho(i,j)}{\sum_{\rho(i,j)\in\rho}\frac{P^s(\omega_n)p^s(\rho(i,j))|\omega_n)}{p^s(\rho(i,j))}} \tag{13}$$

$$(\sigma_{n,\rho}^2)^{s+1}=\frac{\sum_{\rho(i,j)\in\rho}\frac{P^s(\omega_n)p^s(\rho(i,j))|\omega_n)}{p^s(\rho(i,j))}\left[\rho(i,j)-\mu_{n,\rho}^s\right]^2}{\sum_{\rho(i,j)\in\rho}\frac{P^s(\omega_n)p^s(\rho(i,j))|\omega_n)}{p^s(\rho(i,j))}} \tag{14}$$

where $\rho(i,j)$ is the magnitude value of pixel in spatial position (i,j) within the magnitude image. Superscript s indicates the iteration. Initial values for the statistical parameters of both classes are computed as sample mean and variance and relative frequency of pixels within a set of patterns with a high probability to belong to the two classes, respectively. Such sets are built by applying two thresholds for selecting the pixels with very high (meta-class Ω_c) or very low magnitude values (class ω_n) according to the properties of the ρ [19]. The iterative process stops when the likelihood function reaches a local maximum.

Once class statistical parameters are estimated, the Bayes decision rule can be used for pattern labeling, *i.e.*,

$$\omega_h = \arg \max_{\omega_i \in \{\omega_n, \Omega_c\}} \{P(\omega_i \mid \rho(i,j))\} = \arg \max_{\omega_i \in \{\omega_n, \Omega_c\}} \{P(\omega_i) p(\rho(i,j) \mid \omega_i)\} \tag{15}$$

The explicit solution of (15) leads to the definition of a Bayesian decision threshold T [19]. Thus each pixel $x(i,j)$ in spatial position (i,j) is assigned to the class of changed or unchanged pixels according to the following decision rule

$$x(i,j) \in \begin{cases} \omega_n & if \quad \rho(i,j) < T \\ \Omega_c & if \quad \rho(i,j) \geq T \end{cases} \tag{16}$$

B. Identification of Different Kinds of Change

Once changed pixels have been separated from no-changed ones, the attention is focused on the set of changed pixels only (i.e., in the semi-annulus of changed pixels SA_c). The aim of this step is to separate the contributions of possible different kinds of change within the meta-class Ω_c. This can be done by exploiting the direction variable.

Let $P(\omega_{ck} \mid \rho \geq T)$ and $p(\alpha \mid \omega_{ck}, \rho \geq T)$ (k=1, …, K) be the prior probability and the conditional probability density function of the class $\omega_{ck} \in \Omega_c$, k=1,…,K, given that changes occurred (*i.e.*, given that the magnitude variable is higher than the threshold T). Under this hypothesis, the observed direction variable in the semi-annulus of changed pixels can be written as a mixture density distribution:

$$p(\alpha \mid \rho \geq T) = \sum_{k=1}^{K} P(\omega_{c_k} \mid \rho \geq T) p(\alpha \mid \omega_{c_k}, \rho \geq T) \tag{17}$$

The derivation of the analytical expression for the conditional probability density function $p(\alpha|$

ω_{ck} ,$\rho \geq T$) is a complex task [17] and results in distributions that are difficult to be used in the context of automatic techniques. Thus for simplicity, the statistical distribution of each class of change $\omega_{ck} \in \Omega_c$, k=1,…,K is approximated by a Gaussian function. Under this simplified approximation, the generic class of change ω_{ck} can be statistically described with its class prior probability $P(\omega_{ck} | \rho \geq T)$, the mean value ($\mu_{c_k,\alpha}$) and the variance value ($\sigma^2_{c_k,\alpha}$) computed along the direction variable α. Given the statistical parameters of classes, labeling can be performed according to

$$\omega_h = \arg\max_{\omega_i \in \Omega_c} \{ P[\omega_i | \alpha(i,j), \rho \geq T] \} = \arg\max_{\omega_i \in \Omega_c} \{ P(\omega_i | \rho \geq T) p[\alpha(i,j) | \omega_i, \rho \geq T] \} \quad (18)$$

Statistical parameters of each class $\omega_i \in \Omega_c$ can be estimated according to the iterative equations as in (12)-(14). However, unlike the case of the magnitude variable where a reasonable initial assumption can be done on the position of classes, along the direction variable no hypotheses can be formulated on the location of the classes associated to different kinds of change. Therefore in this case a K-mean clustering [39] is applied in order to determine in an unsupervised way reasonable initial seeds for the iterative algorithm. K-mean clustering requires the definition of the number of expected clusters, *i.e.*, the number K of expected kinds of change occurred on the ground. This information can be recovered according to: i) some prior knowledge on the considered problem; ii) interactions with the end-user; iii) a visual analysis of the number of clusters represented in the C^{2}VA domain; or iv) methods for validation of clustering results [40]-[43].

The explicit solution of (18) leads to the definition in the direction domain of a pair of thresholds α_{k_1} and α_{k_2} for each kind of change. Each pixel $x(i,j)$ that falls to the SA_c (i.e., $\rho(i,j)>T$) is assigned to one of the detected kinds of change $\omega_{ck} \in \Omega_c$ according to the following decision rule:

$$x(i,j)\in\begin{cases}\omega_{c1} & if\ \ \alpha_{11}<\alpha(i,j)\leq\alpha_{12}\\ \vdots & \vdots\\ \omega_{cK} & if\ \alpha_{K1}<\alpha(i,j)\leq\alpha_{K2}\end{cases} \tag{19}$$

V. EXPERIMENTAL RESULTS AND DISCUSSION

In order to assess the reliability of both the proposed 2D representation in the C^2VA domain and the effectiveness of the proposed change-detection technique, several experiments were carried out on two multispectral and bitemporal datasets. The first data set is made up of two images acquired by the Thematic Mapper sensor mounted on the Landsat 5 satellite and represents a 6D problem. The second data set includes two very high geometrical resolution images of an area nearby the city of Trento (Italy) acquired by the multispectral sensor mounted on the Quickbird satellite and represents a 4D problem.

The reliability of the proposed C^2VA representation was studied by a comparison with the standard CVA polar framework [17]. We briefly recall here that the CVA framework is defined by the magnitude ρ ($\rho\in[0, \rho_{max}]$) and the direction ϑ ($\vartheta\in[0,2\pi)$) computed by selecting 2 out of B spectral channels according to (2) (it is worth stressing that the selection of 2 out of B spectral channels may result in a significant loss of information). Following [17], the domain of interest is represented by a circle with outer radius $\rho_{\max}$. Within this domain one can identify: i) a circle of no-changed pixels (C_n); and ii) an annulus of changed pixels (A_c) separated by a threshold T. Within A_c, sectors of annuls S_k that correspond to different kinds of change occurred on the ground can be defined bounded by two angular thresholds ϑ_{k1} and ϑ_{k2} (see Figure 2). The magnitude ρ and direction ϑ variables observed in the CVA polar framework can be described as mixture of Gaussian distributed densities [i.e., $p(\rho)$ and $p(\vartheta|\rho\geq T)$, respectively], similarly to what done for ρ and α variables in the C^2VA domain. Thus, thanks to this similarity, the proposed

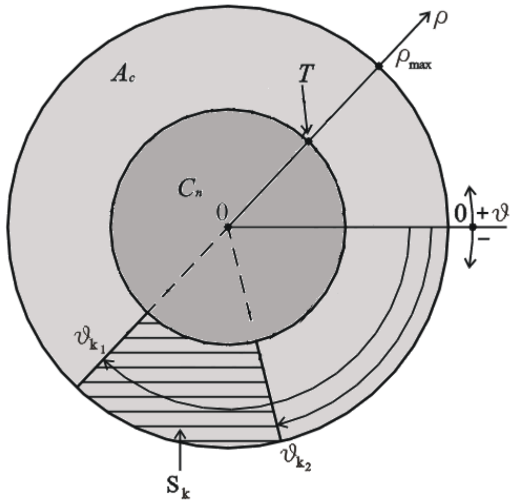


Figure 2. Representation of the regions of interest for the CVA technique in the Polar coordinate system.

automatic technique for the detection of multiple changes can be effectively applied also to the 2D CVA.

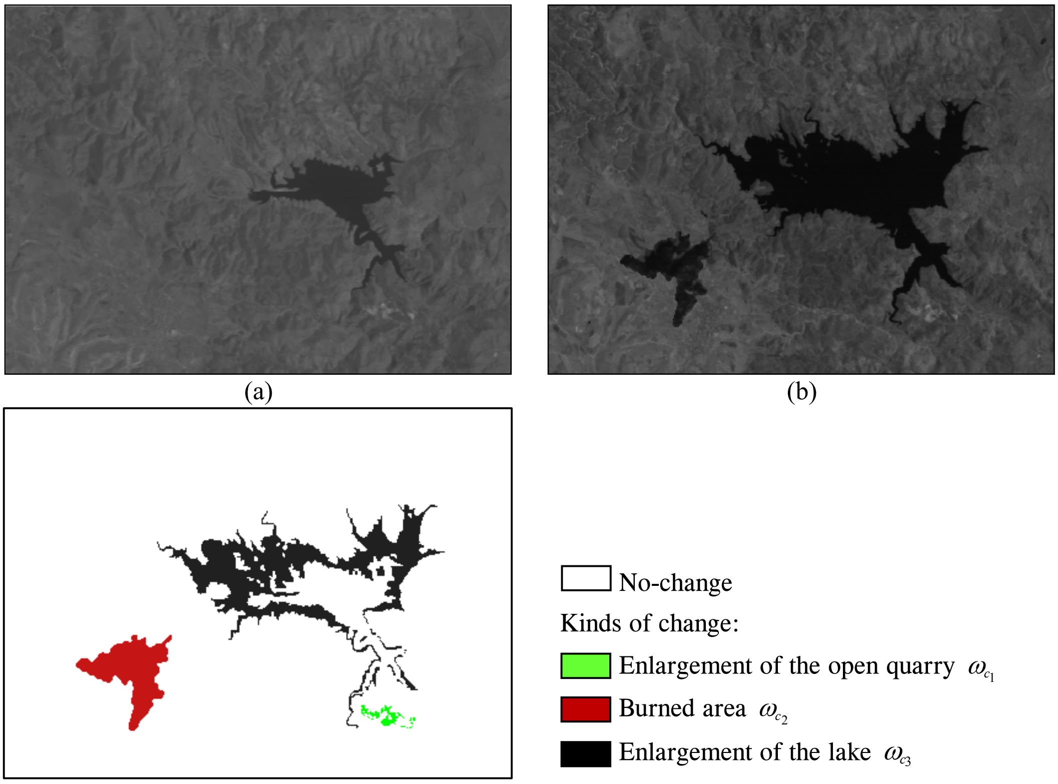
In our experiments, for each data set, CVA in polar coordinates is applied to two different pairs of spectral channels: i) one made up of a couple of bands chosen in a random way (this simulates problems in which no prior information on the kinds of change is available); and ii) one made up of two spectral channels chosen according to some prior knowledge on the considered changes occurred on the ground.

The effectiveness of the proposed framework was evaluated according to: i) a qualitative comparison between the capabilities in representing the change information of the proposed C^2VA and the standard 2D CVA;ii) a quantitative analysis of the performance of the proposed technique for the detection of multiple changes (which was conducted according to an available reference map) applied to both the C^2VA and 2D CVA, and iii) a comparison of the performance obtained with the proposed automatic and unsupervised method with those achieved by an empirical manual trial-and-error procedure (MTEP), *i.e.*, a procedure that selects the threshold values both along magnitude and direction in a supervised way by minimizing the overall error with respect to the available reference map. It is worth noting that, the accuracy obtained by the MTEP can be considered as an upperbound for the one achieved by the proposed automatic method.

A. Dataset 1*: Thematic Mapper Images of Landsat-5*

The first data set is made up of a couple of images acquired on the Sardinia Island (Italy) in September 1995 and July 1996, respectively, by the Thematic Mapper sensor mounted on the Landsat 5 satellite. This data set is characterized by a spatial resolution of 30m×30m. The selected area is a section (412×300 pixels) of two full scenes including Lake Mulargia. In the pre-processing phase the two images were radiometrically corrected and co-registered in order to make them as more comparable as possible. As an example of the images used in this experiment, Figures 3 (a) and (b) show band 4 of the September and July images, respectively. Between the two acquisition dates three kinds of change occurred (K=3): i) an enlargement of an open quarry between the two branches of the lake (bottom right part of the scene, ω_{c1}); ii) a burned area (bottom left part of the scene, ω_{c2}) (this is a simulated change, refer to [17] for further details on how the change has been included in a realistic way); and iii) an enlargement of the lake surface associated to an increase of the water volume of Lake Mulargia (centre of the scene, ω_{c3}). A reference map of the analyzed site was defined according to a detailed visual analysis of the bitemporal images and some prior information. The obtained reference map contains 10180 changed pixels and 113492 unchanged pixels. In greater details, 214 pixels are related to ω_{c1}, 2414 to ω_{c2} and 7480 to ω_{c3} (see Figure 3 (c)).

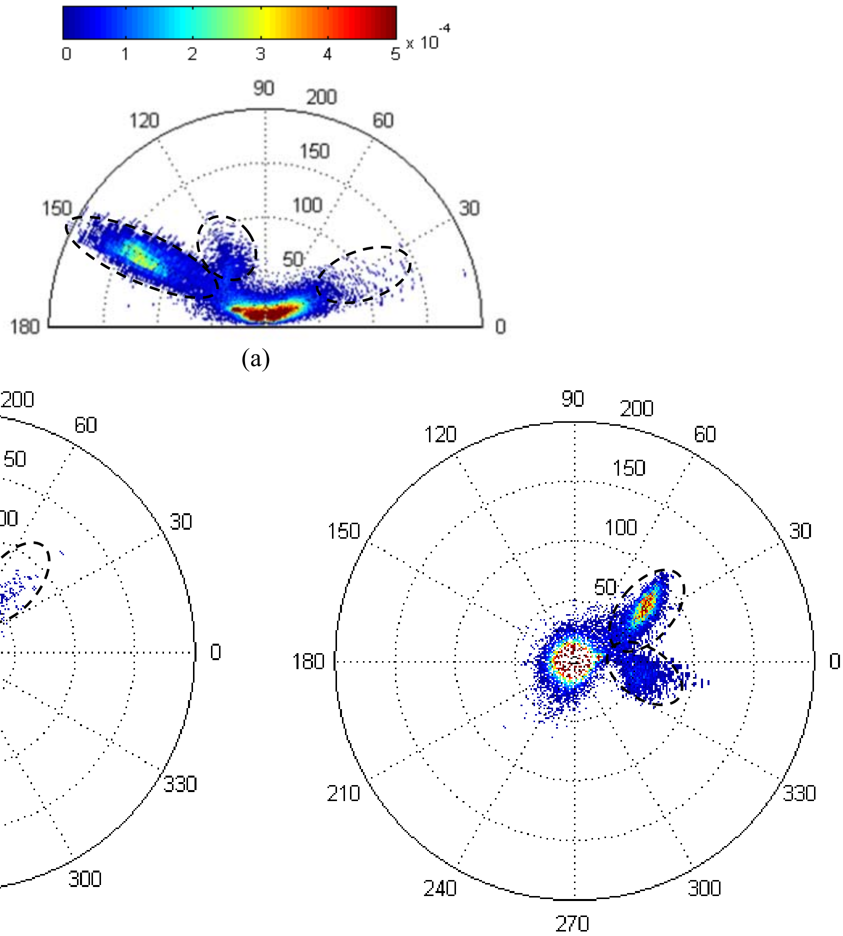
First of all we represented the change information in the Compressed CVA (C^{2}VA) domain. To this purpose, according to the procedure described in Section III, we reduced the dimension of the feature space from 6 (*i.e.*, the number of spectral channels of the TM images, excluding



(c)

Figure 3. Images of Lake Mulargia (Italy) acquired by the Thematic Mapper sensor of the Landsat 5 satellite; (a) channel 4 of the image acquired in September 1995 and (b) channel 4 of the image acquired in July 1996; (c) reference map.

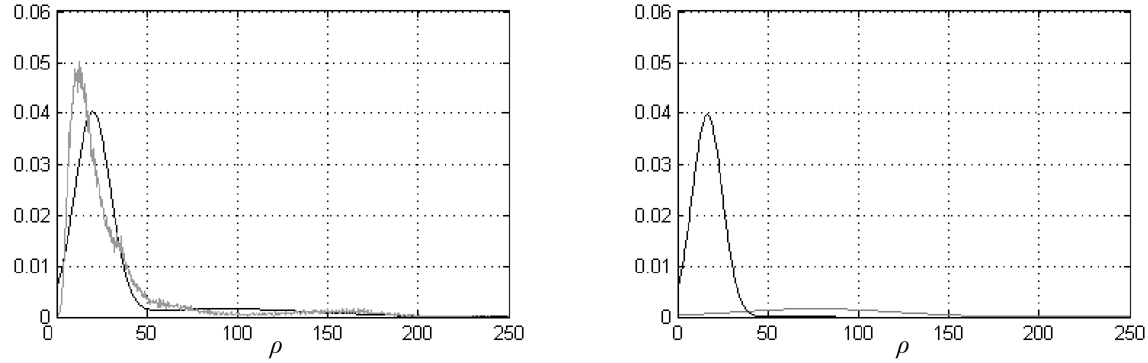
the thermal channel) to 2 computing the magnitude of the multispectral difference image according to (3) and the angle α according to (6). Elements of vector u were all set to $\sqrt{6}/6$. Figure 4 (a) shows the scatterogram in the C^{2}VA domain. Figure 4 (b) and (c) show the scatterogram in the polar domain obtained by applying CVA to two pairs of spectral channels: i) Figure 4 (b) is obtained from the analysis of bands 1 and 3 (which were randomly selected), whereas ii) Figure 4 (c) is obtained from the analysis of bands 4 and 7 (which were selected according to prior knowledge about changes related to water and burned areas). In the scatterogram obtained with the proposed representation (Figure 4 (a)) three main clusters can be easily identified showing a high magnitude and specific preferential values along α. As expected



(a) (b) (c)

Figure 4. Scatterograms obtained by: (a) the proposed 2D C^{2}VA representation; the polar CVA applied to spectral channels (b) 1 and 3, and (c) 4 and 7 (Landsat 5 dataset).

(and confirmed by our experimental analysis) in the other two representations (Figure 4 (b) and (c)), only two clusters can be clearly identified with a high magnitude and a preferred direction and therefore only two types of change can be detected. The proposed approach to multiple change detection estimated a threshold value T (that separates along ρ pixels belonging to ω_n from pixels belonging to Ω_c) equal to 45 when ρ was computed considering all spectral channels (C^{2}VA), whereas it was equal to 35 considering spectral channels 1 and 3, and to 31 considering bands 4 and 7 (2D CVA). As an example, Figure 5 reports the distribution of the SCVs along the magnitude variable. In particular, Figure 5 (a) shows the real distribution derived from the histogram ($h(\rho)$) of the magnitude of SCVs (grey line) and the distribution estimated as a



(a) (b)

Figure 5. (a) $h(\rho)$ (grey line) and $p(\rho)$ (black line) obtained with all spectral channels; and (b) $P(\omega_n)p(\rho|\omega_n)$ (black line) and $P(\Omega_c)p(\rho|\Omega_c)$ (grey line) estimated by the EM algorithm under Gaussian assumption (Landsat 5 dataset).

mixture of Gaussians ($p(\rho)$), while Figure 5 (b) shows separately the distributions $P(\omega_n)p(\rho|\omega_n)$ of the class ω_n (black line) and the distribution $P(\Omega_c)p(\rho|\Omega_c)$ of the class Ω_c (grey line) estimated along the magnitude variable by the EM algorithm (eq. (12)-(14)) for the C^{2}VA. The fitting of the two distributions (the estimated and real ones) confirms the reasonable approximation obtained with the Gaussian distributions [Figure 5 (a)]. Then we derived the distribution $p(\alpha|\rho\geq T)$ of SCVs along the direction variable considering only patterns labeled as changed. According to Sec. III.D, threshold values were identified in order to separate contributions from different kinds of change. Here, we inferred the information about K from a visual analysis of the scatterograms (Figure 4) and of the histograms along the direction variable of SCVs in the semi-annulus (or annulus for CVA) of changed pixels ($h(\alpha|\rho\geq T)$) (Figure 6). In the C^{2}VA domain K was set equal to 3 (three clusters having relatively high magnitude values are present in the scatterograms of Figure 4 (a); thus, the histograms in Figure 6 (a) presents 3 main peaks in positions corresponding to the ones of clusters in the scatterogram). Differently, in the two polar CVA representations the value of K was set equal to 2 (see Figure 4 (b) and (c) and Figure 6 (b) and (c)). As one can see from Figure 6, the distribution estimated with the EM algorithm (black line Figure 6) matches well the real distribution of SCVs (grey line Figure 6), thus confirming the reasonable approximation with a Gaussian distribution. In order to improve the visual quality of Figure 6 (and of similar ones in the following) the level corresponding to zero occurrence was moved from the origin to a semicircle/circle (perimeter of the grey semicircle/circle) slightly shifted from the origin of the plot itself. This choice avoids the bias in the visualized information due to the collapse of probability density functions in the origin due to polar coordinates. The separation of the three different kinds of change is achieved by applying the Bayes decision rule in (19). This operation results in the identification of three threshold values and three annular sectors (S_1, S_2 and S_3) in SA_c corresponding to one of the different kinds of change. The first annular sector is defined as $S_1=\{\rho,\alpha : \rho\geq 45$ and $0°\leq \alpha\leq 70°\}$. All SCVs that fall into S_1 are labeled as ω_{c1} and are associated to the change caused by the quarry enlargement. The second annular sector is defined as $S_2=\{\rho,\alpha : \rho\geq 45$ and$70°\leq \alpha\leq 142°\}$. All SCVs that fall into S_2 are labeled as ω_{c2} (*i.e.*, forest fire). Finally, the third annular sector is defined as $S_3=\{\rho,\alpha : \rho\geq 45$ and $142°\leq \alpha\leq 180°\}$. All SCVs that fall into S_3 are labeled as ω_{c3} and are associated to the change related to the enlargement of the lake surface. Concerning the 2D CVA approach, the analysis of the first pair of channels (1 and 3) led to the identification of pixels belonging to ω_{c1} (SCVs with $\vartheta\in[0°,182°)$, $\rho\geq 35$) and pixels belonging to ω_{c3} ($\vartheta\in[182°,360°)$, $\rho\geq 35$). Considering bands 4 and 7, it is possible to isolate changes due to ω_{c2} (SCVs with $\vartheta\in[323°,360°]\cup[0°,28°)$, $\rho\geq 31$) and to ω_{c3} (SCVs with $\vartheta\in[28°,323°)$, $\rho\geq 31$). A further analysis of all the $\binom{6}{2}$ possible combinations of 2D spectral representations pointed out that it is not possible to identify a pair of spectral channels including information about all mentioned kinds of change (this analysis is not reported for space constraints).

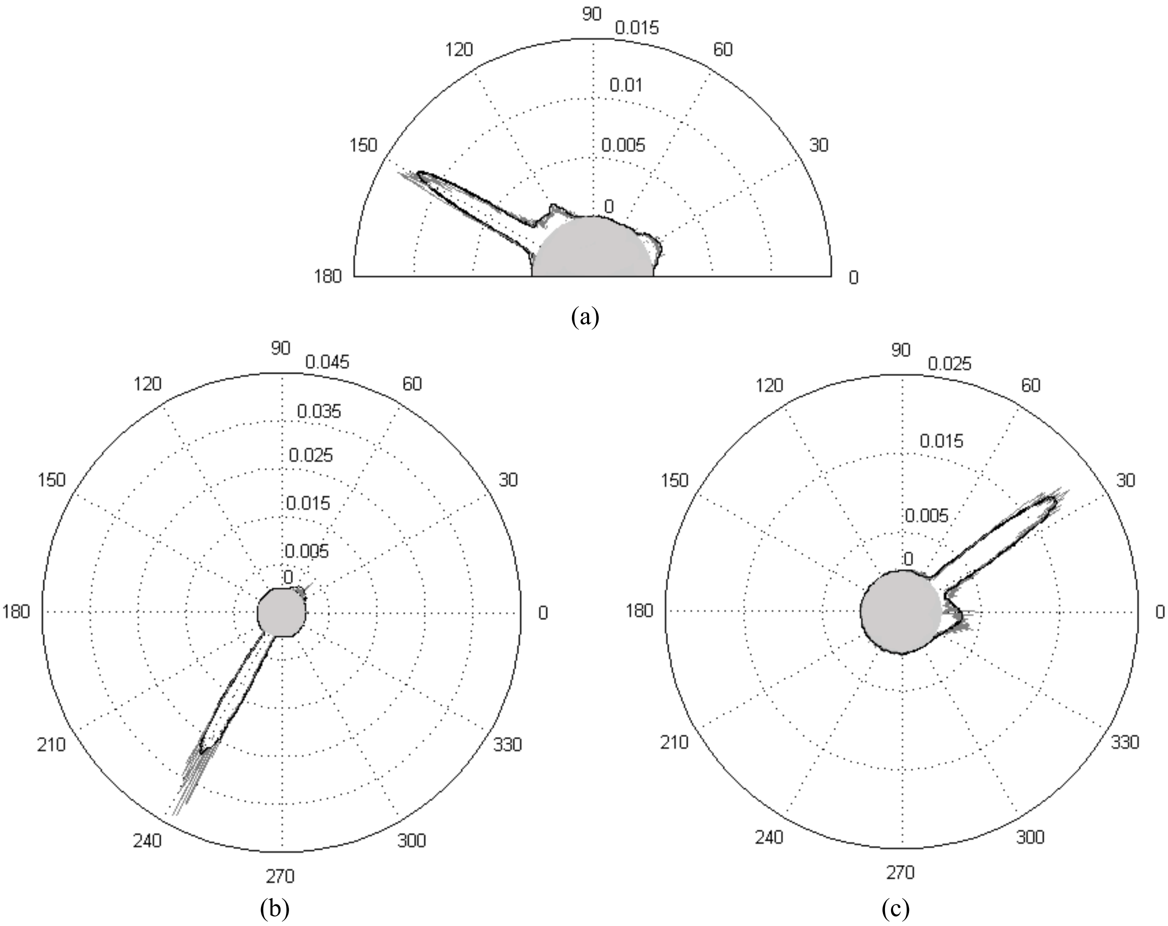


Figure 6. (a) $p(\alpha|\rho\geq T)$ (black line) and $h(\alpha|\rho\geq T)$ (grey line) in SA_c; (b) $p(\vartheta|\rho\geq T)$ (black line) and $h(\vartheta|\rho\geq T)$ (grey line) in A_c when using spectral channels 1 and 3; and (c) $p(\vartheta|\rho\geq T)$ (black line) and $h(\vartheta|\rho\geq T)$ (grey line) in A_c when using spectral channels 4 and 7 (Landsat 5 dataset).

Using the derived threshold values (both in magnitude and direction) a change-detection map is computed for each change-detection problem representation. Figure 7 (a) shows the change-detection map obtained by isolating the three clusters in SA_c according to (19). Each kind of change is clearly identified with a different color (Figure 7). Figure 7 (b) and (c) show the change-detection maps obtained using the two couples of spectral channels with the CVA. As expected, in these maps only two out of three changes appear (ω_{c1} and ω_{c3} considering spectral channels 1 and 3; and ω_{c2} and ω_{c3}, considering bands 4 and 7).

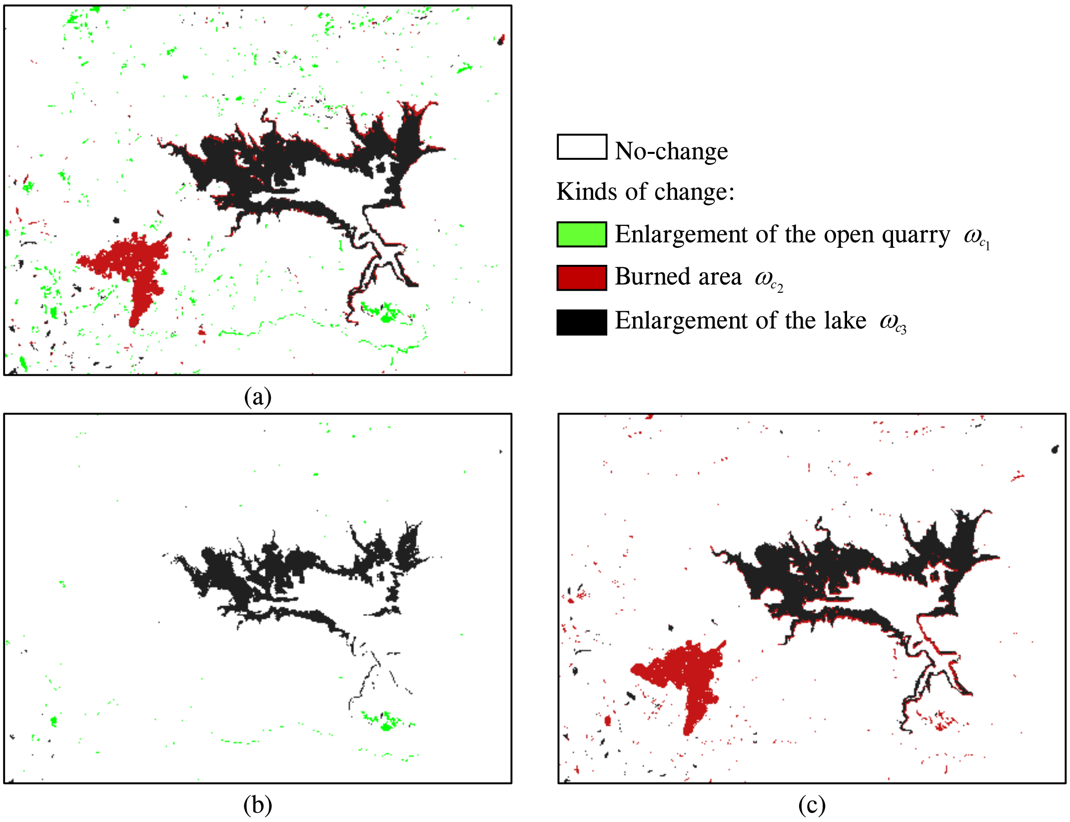


Figure 7. Change-detection maps obtained with the proposed change-detection technique applied to: (a) the proposed C^2VA representation; (b) the polar framework (spectral channels 1 and 3); and (c) the polar framework (spectral channels 4 and 7). (Landsat 5 dataset).

A comparison of these maps with the reference map in Figure 3 (c) allows us a quantitative evaluation of performance. Tables II-IV report the confusion matrices for the three considered cases. As one can see, the overall accuracies computed on the three change-detection maps are very similar to each other, and always higher than 96%. However the proposed representation allowed us to retain the main information related to changes and to distinguish all different kinds of change. This is because C^2VA preserves the most of the information, although it maps a feature space of dimension 6 into one of dimension 2. It is worth stressing that this result is achieved without the need of any prior information about the kinds of change occurred on the ground. On the contrary, the representations obtained considering only couples of spectral channels [17],[19] resulted in total (or partial) loss of change information depending on the considered pair of bands. One can observe that the CD map obtained considering all the spectral channels suffers of a higher impact of noisy components than the other two. This is because the use of all spectral channels not only preserves change information, but also introduces some noise. However, according to the previous considerations, the slightly higher amount of false alarms that affects the C^2VA change-detection map becomes acceptable from an application point of view, where the possible loss of a kind of change could be more critical. The false alarms can then be reduced with the application of proper pre-processing techniques.

TABLE II. CHANGE-DETECTION RESULTS OBTAINED BY THE PROPOSED CHANGE-DETECTION METHOD BASED ON THE C^2VA REPRESENTATION APPLIED TO ALL THE SPECTRAL CHANNELS (LANDSAT 5 DATASET).

		True Class				User Accuracy
		ω_{c_1}	ω_{c_2}	ω_{c_3}	ω_n	
Estimated Class	ω_{c_1}	185	0	0	1487	11.06
	ω_{c_2}	5	2160	445	736	64.55
	ω_{c_3}	19	24	7032	1525	81.77
	ω_n	5	230	3	109744	99.78
Producer Accuracy		86.45	89.48	94.01	96.70	
Kappa Accuracy		0.7966				
Overall Accuracy		96.38				

In order to further assess the effectiveness of the proposed approach its performance are compared with those provided by the MTEP. As demonstrated from Table V, the proposed method and MTEP lead to quite similar threshold values and therefore to very close overall accuracies (96.38% versus 96.73%). This confirms the validity of both the automatic procedure and the selected approximated statistical model for class distributions. An analysis of these results points out that the errors of omission and commission among classes have to be ascribed to the overlapping of classes in the considered problem rather than to the proposed automatic method.[2]

[2] Similar observations hold for the results achieved by the 2D CVA, as well for the ones obtained on the Quickbird data set that are not reported for space constraints.

TABLE III. CHANGE-DETECTION RESULTS OBTAINED BY THE PROPOSED CHANGE-DETECTION METHOD APPLIED TO THE POLAR FRAMEWORK (SPECTRAL CHANNELS 1 AND 3) (LANDSAT 5 DATASET).

		True Class				User Accuracy
		ω_{c_1}	ω_{c_2}	ω_{c_3}	ω_n	
Estimated Class	ω_{c_1}	133	0	0	187	41.56
	ω_{c_2}	0	0	0	0	0.00
	ω_{c_3}	10	0	5390	163	96.89
	ω_n	71	2414	2090	113142	96.11
Producer Accuracy		62.15	0.00	72.06	99.69	
Kappa Accuracy		0.6747				
Overall Accuracy		96.01				

TABLE IV. CHANGE-DETECTION RESULTS OBTAINED BY THE PROPOSED CHANGE-DETECTION METHOD APPLIED TO THE POLAR FRAMEWORK (SPECTRAL CHANNELS 4 AND 7) (LANDSAT 5 DATASET).

		True Class				User Accuracy
		ω_{c_1}	ω_{c_2}	ω_{c_3}	ω_n	
Estimated Class	ω_{c_1}	0	0	0	0	0.00
	ω_{c_2}	0	2354	109	478	80.04
	ω_{c_3}	105	3	7364	1815	79.29
	ω_n	109	57	7	111199	99.84
Producer Accuracy		0.00	97.51	98.45	97.98	
Kappa Accuracy		0.8705				
Overall Accuracy		97.83				

TABLE V. THRESHOLD VALUES OBTAINED BY THE PROPOSED CHANGE-DETECTION METHOD AND THE OPTIMAL SUPERVISED MTEP ON THE C^2VA REPRESENTATION APPLIED TO ALL THE SPECTRAL CHANNELS (LANDSAT 5 DATASET).

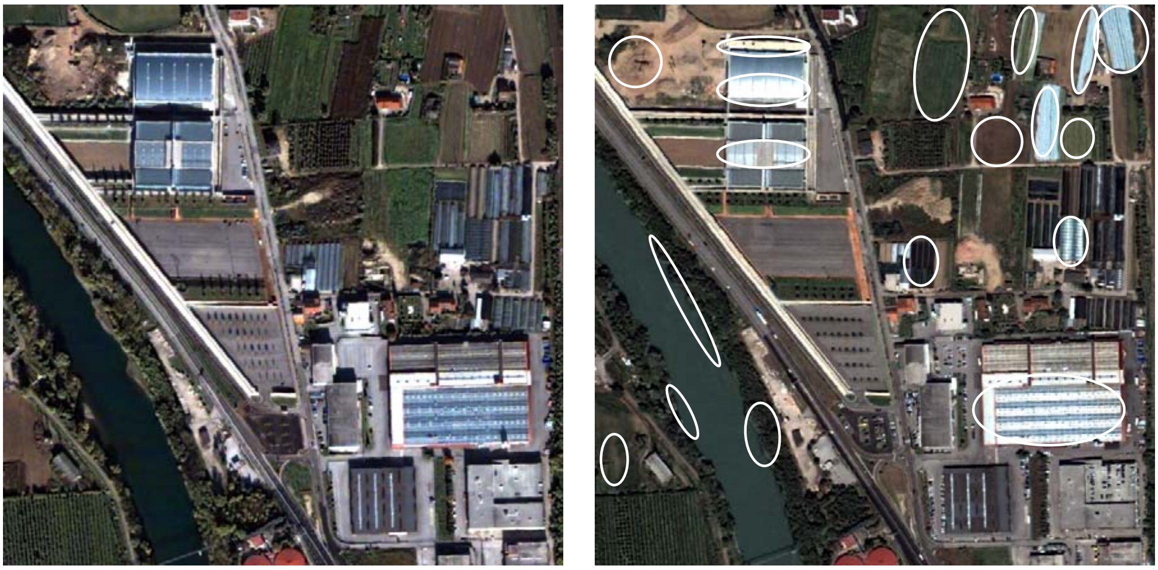
	T	α_{1_1}	$\alpha_{1_2} \equiv \alpha_{2_1}$	$\alpha_{2_2} \equiv \alpha_{3_1}$	α_{3_2}
C^2VA	45	0°	70°	142°	180°
MTEP	50	0°	60°	140°	180°

It is worth noting that, in absence of any prior information about relevant spectral bands with respect to the considered problem (*i.e.*, no spectral bands can be neglected) the standard automatic unsupervised procedures simply threshold the magnitude variable obtained from all spectral channels (*i.e.*, only first step of the proposed change-detection procedure is applied). The resulting change-detection map is as the one in Figure 7 (a) but different kinds of change are not distinguished. It follows that the proposed technique can significantly improve the change information extracted from the considered dataset by allowing the separation of the contributions from different kinds of change in a automatic way.

As a final remark, it is worth noting that the proposed C^2VA representation made it possible to identify a third kind of minority change in the considered data set that was not observed in previous works neither by photointerpreters, nor by automatic techniques based on the exploitation of spectral channels 4 and 7 in the CVA framework. Moreover, despite the possible loss of change information induced by the 2D representations, numerical results allow one to conclude that the proposed automatic technique for the detection of multiple changes is effective when applied to both C^2VA and 2D CVA representations. In all the cases the proposed technique extracted all information about changes available in the considered representation.

B. Dataset 2: Quickbird Images

Experiments similar to the ones conducted on the Thematic Mapper data set were carried out on a pair of very high geometrical resolution images acquired by the Quickbird sensor in October 2005 and July 2006 on the city of Trento (Italy) (Figure 8). In the pre-processing phase the two images were: i) pan-sharpened; ii) radiometrically corrected; and iii) co-registered. In particular, we considered pan-sharpened images as we expect that the pan-sharpening process can improve the results of the change-detection process, as demonstrated in previous work [44]. To this purpose we applied the minimum mean square error (MMSE) pansharpening method [45] to the panchromatic channel and the four bands of the multispectral images. Concerning radiometric corrections, we simply normalized the images by subtracting from each spectral channel of the two considered images its mean value. The registration process was carried out in a simple way by using a polynomial function of order 2 according to 12 ground control points (GCPs), and by applying a nearest neighbor interpolation [46]. The final data set is made up of images of 992x992 pixels with spatial resolution on the ground of 0.7m. Between the two acquisition dates



(a) (b)

Figure 8. True color composition of an area nearby the city of Trento (Italy) acquired by the Quickbird VHR multispectral sensor in (a) July 2005 and (b) October 2006. White circles identify the main areas affected by changes.

some changes related to urban and rural areas occurred on the ground. In particular, three different kinds of change can be observed, *i.e.* K=3 (see circles in Figure 8): i) changes in the a cover of both buildings (*i.e.*, changes in roofs related to saturation problems of the sensor) and crop fields (*i.e.*, new structures built for covering fields) that have the same spectral signature, ω_{c_1} ; ii) seasonal changes in vegetated areas, both in crop fields and wooded zones, ω_{c_2} ; and iii) changes along the river bank due to an increase of the water level, ω_{c_3} . In order to perform quantitative analysis on this data set, we defined a sampled ground truth (based on a spatial random sampling as we do not have a complete knowledge of the changes occurred on the ground) containing: 22652 pixels labeled as ω_{c_1} , 27660 as ω_{c_2} , 6554 as ω_{c_3} and 383396 pixels of no change.

As for the Sardinia data set, we reduced the size of the feature space from 4 (the number of the multispectral channels of the Quickbird images) to 2, computing the magnitude of the multispectral difference image according to (3) and the angle according to (6). In this case $u=[\sqrt{4}/4, \sqrt{4}/4, \sqrt{4}/4, \sqrt{4}/4]$. Figure 9 (a) shows the scatterogramms that represent the considered change-detection problem within the proposed C^{2}VA domain. We compared this plot with the polar scatterograms obtained by applying the CVA technique to: channels 2 and 3 (Figure 9 (b)), which were randomly selected; and channels 3 and 4 (Figure 9 (c)), which were selected according to some prior knowledge about changes occurred on the ground.

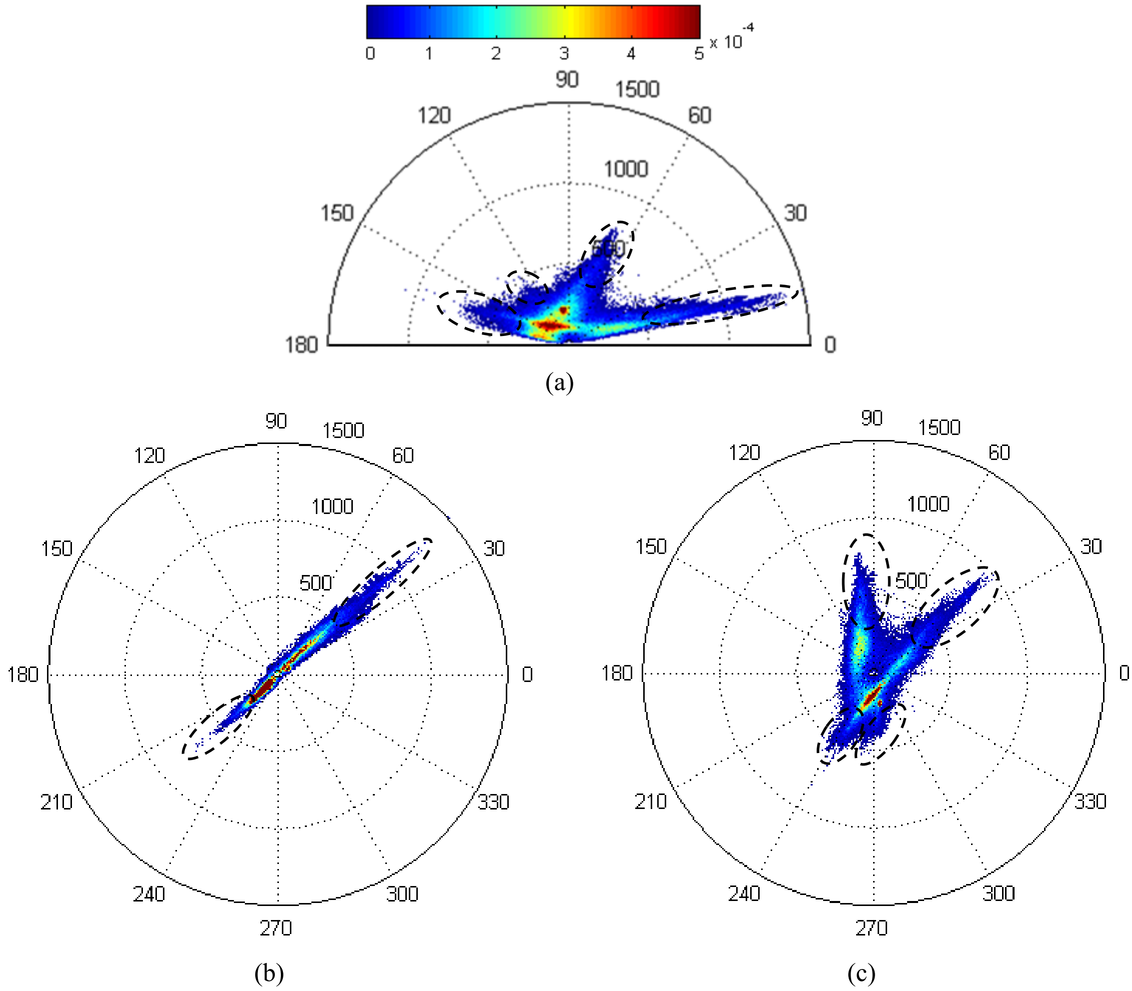


Figure 9. Scatterograms obtained by applying: (a) the proposed 2D representation; (b) the polar CVA to spectral channels 2 and 3; and (c) the polar CVA to spectral channels 3 and 4 (Quickbird Dataset).

The threshold value T for C^{2}VA which separates the SC_n from the SA_c resulted equal to 350. Four main clusters were identified in SA_c in the scatterogram (see dashed circles in Figure 9 (a)) and four modes are present in $h(\alpha|\rho \geq T)$ (see grey line in Figure 10 (a)), therefore K was set equal to 4.[3] The visual analysis of Figure 10 (a) points out that the estimated distribution $p(\alpha|\rho \geq T)$ (black line) fits quite well with the behavior of the real histogram $h(\alpha|\rho \geq T)$ (grey line), thus confirming the reliability of the Gaussian approximation. The second step of the thresholding procedure led us to the definition of the following four annular sectors:

$$\begin{aligned} S_1 &= \{\rho, \alpha : \rho \geq 350 \text{ and } 0° \leq \alpha < 27°\} \\ S_2 &= \{\rho, \alpha : \rho \geq 350 \text{ and } 27° \leq \alpha < 110°\} \\ S_3 &= \{\rho, \alpha : \rho \geq 350 \text{ and } 110° \leq \alpha < 156°\} \\ S_4 &= \{\rho, \alpha : \rho \geq 350 \text{ and } 156° \leq \alpha \leq 180°\} \end{aligned} \tag{20}$$

Analyzing each sector it is possible to observe that S_1, S_2 and S_3 are associated to: ω_{c_1}, ω_{c_2}, and ω_{c_3}, respectively, whereas S_4 is mainly related to the effects of registration noise. This result was expected as registration noise in VHR images significantly affects the change-detection process introducing clusters with a high magnitude and preferred direction that have properties similar to changed pixels [47]. An analysis of this kind of noise within C^2VA domain is out of the purposes of this work. Therefore in the following SCVs that fall in S_4 and that are identified as being of registration noise will be neglected from further analysis and classified as unchanged patterns. The reader is referred to [47],[48] for further details on this challenging problem and on techniques for reducing registration noise impacts on the change-detection process.

With regard to the analysis in the polar domain, as for the analysis in the proposed C^2VA domain, we retrieved the threshold value T (T=350 when considering bands 2 and 3 and T=300 for spectral channels 3 and 4). According to the analysis of both the scatterograms (Figure 9 (b) e (c)) and the histograms (Figure 10 (b) e (c)), the value of K was set to 2 for the case of bands 2

[3] It is worth noting that also in this case a light grey semi-circle/circle is introduced to slightly shift from the origin of the plot the level corresponding to zero occurrences, thus avoiding a bias in the information visualization.

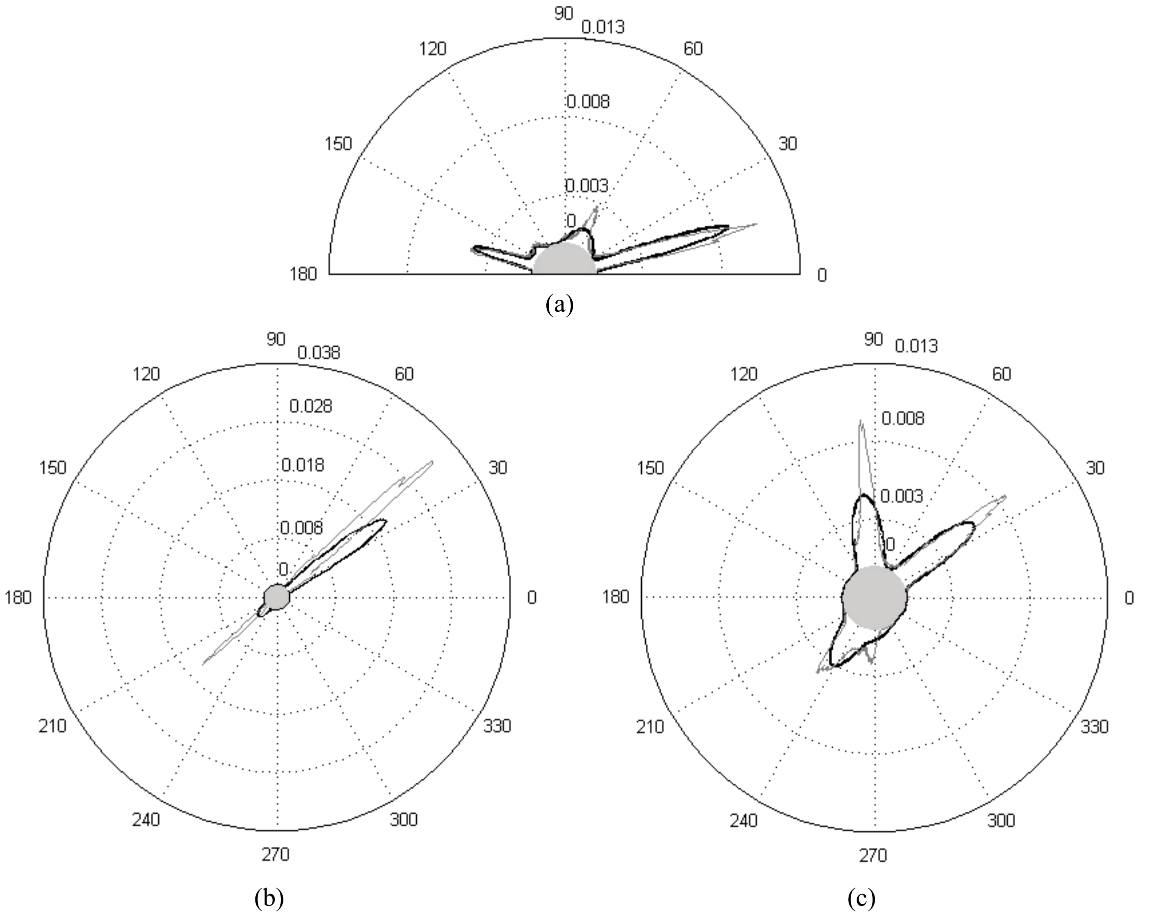


Figure 10. (a) Estimated $p(\alpha|\rho\geq T)$ (black line) and $h(\alpha|\rho\geq T)$ (grey line) in SA_c; (b) estimated $p(\vartheta|\rho\geq T)$ (black line) and $h(\vartheta|\rho\geq T)$ (grey line) in A_c when using spectral channels 2 and 3; and (c) estimated $p(\vartheta|\rho\geq T)$ (black line) and $h(\vartheta|\rho\geq T)$ (grey line) in A_c when using spectral channels 3 and 4 (Quickbird Dataset).

and 3, and to 4 for bands 3 and 4. The second step of the proposed thresholding procedure was applied to distinguish the different contributions to Ω_c. Concerning the first pair of bands (2 and 3) two sectors were identified, S_1 made up of SCVs with $\vartheta\in[0°,103°]$ and S_2 with $\vartheta\in[103°,360°)$. It is possible to show that S_1 is related to changes in building or crop covers (ω_{c_1}), whereas S_2 is related to registration noise effects. This means that considering only spectral channels 2 and 3 it is not possible to extract information about ω_{c_2} and ω_{c_3}. The analysis conducted on the second pair of spectral channels results in the definition of 4 sectors, S_1 made up of SCVs with $\vartheta\in[0°,63°)$, S_2 with $\vartheta\in[63°,168°)$, S_3 with $\vartheta\in[168°,253°)$, and S_4 with

$\vartheta \in [253°, 360°)$. It can be shown that, as for the proposed method, pixels associated to ω_{c_1} fall in S_1, pixels belonging to ω_{c_2} fall in S_2, pixels belonging to ω_{c_3} fall in S_3, and pixels in S_4 are associated to registration noise. It is worth stressing that the considered spectral channels were selected according to some prior information about changes, whereas the proposed method achieves similar results (*i.e.*, it detects all kinds of change present in the multitemporal data set)

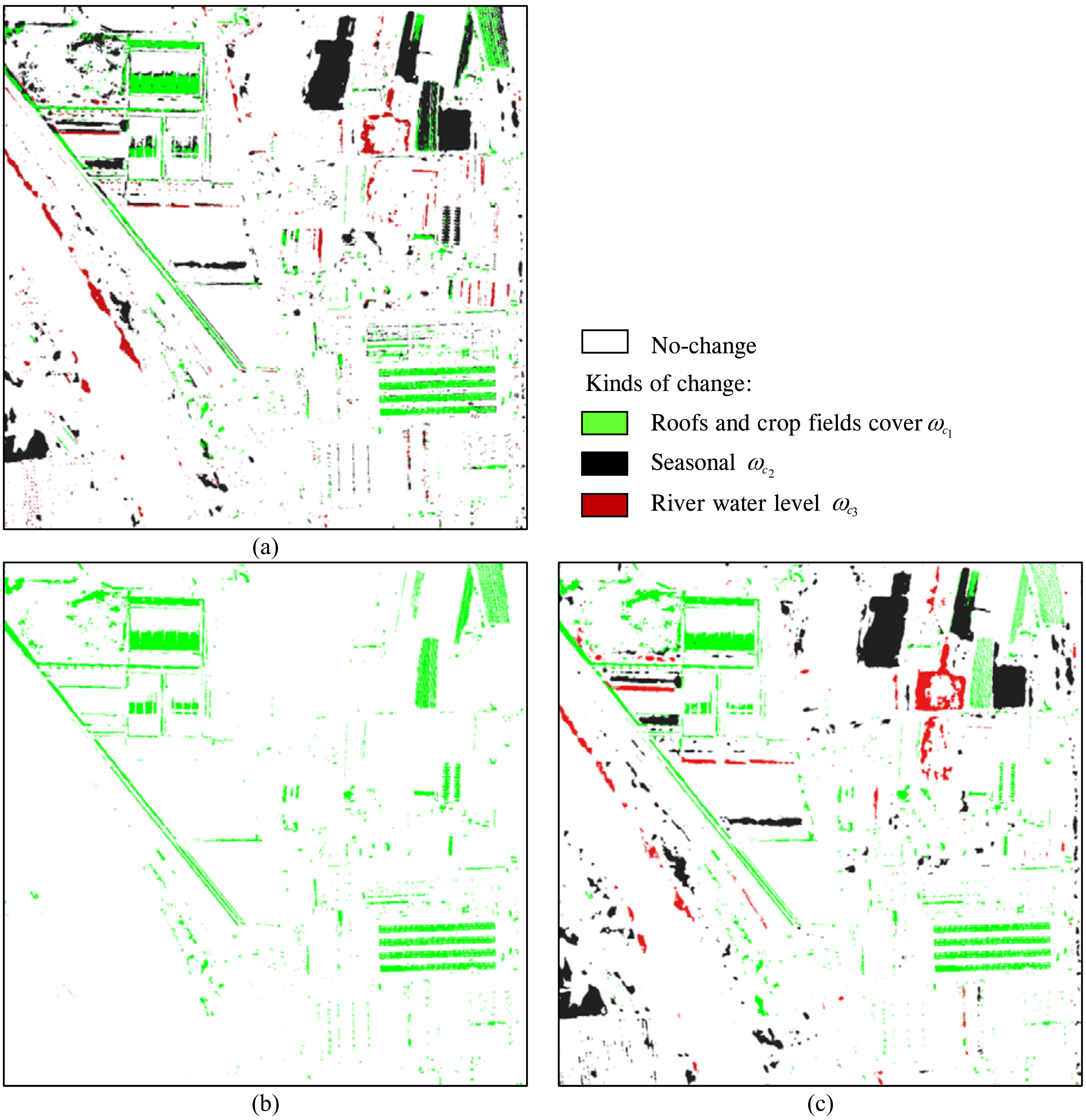


Figure 11. Change-detection map obtained with the proposed change-detection technique applied to: (a) the proposed C^{2}VA domain data representation; (b) the polar framework (spectral channels 2 and 3); and (c) the polar framework (spectral channels 3 and 4). (Quickbird Dataset)

without any prior information.

According to the threshold values estimated with the proposed technique for each of the three representations change-detection maps are generated (see Figure 11). A quantitative analysis of the results achieved on the three considered representations for the reference data set is reported in Tables VI-VIII. These tables confirm the qualitative evaluation. The proposed technique for multiple change detection applied to both the C^2VA and the CVA (spectral channels 3 and 4) achieved similar results (overall accuracy equal to 95.0% and 95.5% respectively). However, CVA requires prior information about possible kinds of change for selecting spectral channels. Moreover, the proposed multiple-change detection technique permits to identify and separate all different kinds of change, showing good accuracies for all of them (higher than 80 % for user accuracy and higher than 70% for the producer accuracy). On the contrary, the standard CVA on randomly selected spectral channels (*i.e.*, bands 2 and 3) allows us to identify only the changes in building and crop covers. All other kinds of change are undetected.

It is worth noting that some registration noise effects are still visible in the change-detection maps affecting significantly the user accuracy which is always smaller than 83% for all the kinds of change in the three analyzed cases. Advanced change-detection techniques developed for VHR images (*i.e.* context sensitive or multiscale techniques [48],[49]) could be employed for

TABLE VI. CHANGE-DETECTION RESULTS OBTAINED BY THE PROPOSED CHANGE-DETECTION METHOD BASED ON THE C^2VA REPRESENTATION APPLIED TO ALL THE SPECTRAL CHANNELS (QUICKBIRD DATASET).

		True Class				User Accuracy
		ω_{c_1}	ω_{c_2}	ω_{c_3}	ω_n	
Estimated Class	ω_{c_1}	18728	823	0	4896	76.61
	ω_{c_2}	3479	24948	139	6706	70.73
	ω_{c_3}	0	12	5691	2473	69.61
	ω_n	445	1858	722	358671	99.16
Producer Accuracy		82.68	90.26	86.86	96.22	
Kappa Accuracy		0.8077				
Overall Accuracy		94.98				

reducing the effects of the residual registration noise in change-detection maps. These strategies can be easily extended to the proposed C^2VA domain. However, this is out of the purpose of this work, for which we just consider the comparison of C^2VA and CVA at pixel level.

As for the Landsat-5 data set a comparison with MTEP results leads to the conclusion that the proposed procedure as well as the assumption of Gaussian distributed classes are effective and reliable. The presence of mislabeled pixels is therefore due to the complexity of the considered problem. According to the analysis of results it is possible to conclude that the proposed representation allows us to preserve the information about all the possible kinds of change, even by reducing dimensionality from 4 to 2 (and thus introducing ambiguity in the process). On the contrary, the representation obtained considering only couples of channels may result in a total (or partial) loss of information related to specific changes. This depends on the selected spectral bands and thus on the available prior information. Furthermore, the proposed automatic technique for the detection of multiple changes demonstrated to be successful when applied to both C^2VA and 2D CVA representations. In all the cases the proposed technique effectively detected all information about changes available in the considered representation.

TABLE VII. CHANGE-DETECTION RESULTS OBTAINED BY THE PROPOSED CHANGE-DETECTION METHOD APPLIED TO THE POLAR FRAMEWORK (SPECTRAL CHANNELS 2 AND 3) (QUICKBIRD DATASET).

		True Class				User Accuracy
		ω_{c_1}	ω_{c_2}	ω_{c_3}	ω_n	
Estimated Class	ω_{c_1}	21872	1782	0	6034	73.67
	ω_{c_2}	0	0	0	0	0.00
	ω_{c_3}	0	0	0	0	0.00
	ω_n	780	25874	6547	372700	91.82
Producer Accuracy		96.56	0.00	0.00	98.41	
Kappa Accuracy		0.4944				
Overall Accuracy		90.58				

TABLE VIII. CHANGE-DETECTION RESULTS OBTAINED BY THE PROPOSED CHANGE-DETECTION METHOD APPLIED TO THE POLAR FRAMEWORK (SPECTRAL CHANNELS 3 AND 4) (QUICKBIRD DATASET).

		True Class				User Accuracy
		ω_{c_1}	ω_{c_2}	ω_{c_3}	ω_n	
Estimated Class	ω_{c_1}	20463	2168	0	5573	72.55
	ω_{c_2}	4	23902	161	4794	82.82
	ω_{c_3}	1	35	4877	2744	63.69
	ω_n	2184	1538	8	360283	98.98
Producer Accuracy		90.34	86.47	96.65	96.49	
Kappa Accuracy		0.8226				
Overall Accuracy		95.52				

VI. CONCLUSION

In this paper an automatic technique for the detection of multiple changes in multitemporal and multispectral remote sensing images has been presented. The proposed method compresses the original BD feature space to be explored for the solution of the change-detection problem (B is the number of spectral channels acquired by the considered sensor) to a 2D space and applies a 2-step decision strategy for detecting changes. The compression is accomplished by computing the magnitude of spectral change vectors, and the angle (direction) between the spectral difference vector and a reference one. In this way we obtain a 2D representation of the change-detection problem that preserves the relevant information present in all available spectral channels. The change information can be represented according to the two proposed variables in a 2D domain, which is defined as Compressed Change Vector Analysis (C^2VA) domain. The proposed transformation leads to a 2D representation of the change-detection problem that can be visualized without the need of selecting a pair of spectral channels as usually done in standard approaches. This represents a valuable advantage as spectral channel selection would require some prior knowledge about possible changes occurred on the ground which often is not available or incomplete. Accordingly, missed alarms associated to possible unexpected kinds of change only visible in non-selected spectral bands are reduced.

Qualitative and quantitative results obtained on both Lansat-5 and Quickbird images confirmed the effectiveness of the proposed automatic technique for the detection of multiple changes when applied to both C^2VA and standard 2D CVA. This also confirms the reliability of the Gaussian approximation for the distribution of the classes (however statistical models different from the Gaussian one could be integrated within the proposed method). Furthermore, results pointed out the better capabilities in representing the change information of the proposed C^2VA representation with respect to the standard 2D CVA. In C^2VA, although the information is projected from a *B*D into a 2D space, it is possible to retrieve the main information related to changes and to distinguish all different kinds of change occurred on the ground. When the C^2VA representation is used the advantage of identifying all kinds of change by using all spectral channels implies an increase of false alarms due to noisy components.

As a final remark, it is worth noting that in complex change-detection problems some ambiguity may rise from the dimension reduction process, mainly due to the simplified representation of the angle variable. This may result in loss of information about the distribution of different kinds of change. Anyway it is preferable to more standard representations based on the use of couple of spectral channels that often implies a significant loss of information about kinds of change.

As future work we plan to exploit the potentialities of the proposed technique in the context of more complex approaches to change detection like those that exploit multiscale/multiresolution information intrinsically present in VHR images and the ones robust to registration noise.

REFERENCES

[1] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. L. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 46, no. 6, pp. 1822-1835, Jun. 2008.

[2] N. Ghoggali and F. Melgani, "Genetic SVM approach to semisupervised multitemporal classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 212-216, Apr. 2008.

[3] A. Singh, "Digital Change Detection Techniques Using Remotely-Sensed Data," *Int. J. Rem. Sens.*, vol. 10, no. 6, pp. 989-1003, 1989.

[4] L. Bruzzone and S. B. Serpico, "An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images," *IEEE Transactions on Geosceince and Remote Sensing*, vol. 35, no. 4, pp. 858-867, 1997.

[5] L. Bruzzone and M. Marconcini, "Domain Adaptation Problems: a DASVM Classification Technique and a Circular Validation Strategy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 770-787, 2010.

[6] L. Bruzzone, R. Cossu, and G. Vernazza, "Detection of land-cover transitions by combining multidate classifiers," *Pattern Recogniton Letters*, vol. 25, no. 13, pp. 1491-1500, 2004.

[7] P. R. Coppin, I. Jonckheere, and K. Nachaerts, "Digital Change Detection Techniques," *Int. J. Rem. Sens.*, vol. 25, no. 9, pp. 1565-1596, 2004.

[8] D. Lu, P. Mausel, E. Brondízio, and E. Moran, "Change Detection Techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, p. 2365–2407, 2004.

[9] A. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image Change Detection Algorithms: A Systematic Survey," *IEEE Trans. on Image Proc.*, vol. 14, no. 3, pp. 294-307, 2005.

[10] F. Pacifici and F. D. Frate, "Automatic change detection in very high resolution images with pulse-coupled neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 1, pp. 58-62, Jan. 2010.

[11] Y. Bazi, F. Melgani, and H. D. Al-Sharari, "Unsupervised change detection in multispectral remotely sensed imagery with level set methods," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 3178-3187, Aug. 2010.

[12] L. Gueguen and M. Datcu, "Mixed Information Measure: Application to Change Detection in Earth Observation," in *Fifth International Wotkshop on the Analysis of Multi-Temporal Remote Sensing Images*, Mystic, Connecticut, USA, 2009.

[13] G. Mercier, G. Moser, and S. B. Serpico, "Conditional copulas for change detection in heterogeneous remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 5, pp. 1428-1441, May 2008.

[14] D. Faur, C. Vaduva, I. Gavat, and M. Datcu, "An information theory based image processing chain for change detection in earth observation," in *15th International Conference on Systems, Signals and Image Processing*, Bratislava, 2008, pp. 129-132.

[15] A. A. Nielsen, "The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data," *IEEE Transaction on Image Processing*, vol. 16, no. 2, pp. 463-478, Feb. 2007.

[16] J. M. Canty and A. A. Nielsen, "Visualization and unsupervised classification of changes in multispectral satellite imagery," *International Journal of Remote Sensing*, vol. 27, no. 18, pp. 3961-3975, Sep. 2006.

[17] F. Bovolo and L. Bruzzone, "A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in Polar Domain," *IEEE Trans. Geosci. and Rem. Sens.*, vol. 45, no. 1, pp. 218-236, 2007.

[18] 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, 2000.

[19] L. Bruzzone and D. Fernández-Prieto, "Automatic Analysis of the Difference Image for Unsupervised Change detection," *IEEE Trans. Geosci. Rem. Sens.*, vol. 38, no. 3, pp. 1170-1182, 2000.

[20] T. Celik and K. K. Ma, "Unsupervised change detection for satellite images using dual-tree complex wavelet transform," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1199-1210, Mar. 2010.

[21] T. Celik, "Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-

Means Clustering," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 4, pp. 772-776, Oct. 2009.

[22] W. A. Malila, "Change vector analysis: An approach for detecting forest," in *Proc. LARS Mach. Process. Remotely Sensed*, W. Lafayette, IN, 1980, p. 326–336.

[23] 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, p. 369–379, 2003.

[24] J. S. Denga, K. Wanga, Y. H. Dengb, and G. J. Qic, "PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data," *International Journal of Remote Sensing*, vol. 29, no. 16, pp. 4823-4838, 2008.

[25] L. Bruzzone and S. B. Serpico, "Detection of changes in remotely-sensed images by the selective use of multi-spectral information," *International Journal of Remote Sensing*, vol. 18, no. 18, pp. 3883-3888, 1997.

[26] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, and K. B. Heidebrecht, "The Spectral Image Processing System (SIPS)—Interactive Visualization and Analysis of Imaging Spectrometer Data," *Remote Sens. Environ.*, vol. 44, no. 2/3, pp. 145-163, 1993.

[27] G. Girouard, A. Bannari, A. Harti, and A. Desrochers, "Validated Spectral Angle Mapper Algorithm for Geological Mapping: Comparative Study Between Quickbird and Landsat-TM," in *Proc. of 20th International Society for Photogrammetry and Remote Sensing Congress*, Istanbul, Turkey, 2004, pp. 599-605.

[28] C. Hecker, M. van der Meijde, H. van der Werff, and F. D. van der Meer, "Assessing the Influence of Reference Spectra on Synthetic SAM Classification Results," *IEEE Trans. Geosci. Rem. Sens.*, vol. 46, no. 12, pp. 4162-4172, 2008.

[29] Y. Amano, N. Takagi, and A. Getz, "A study on the classification of urban region using Hyper-spectrum data at AVIRIS," in *IEEE International Geoscience and Remote Sensing Symposium*, Boston, USA, 2008, pp. IV-687-IV-690.

[30] Y. Sohn and N. S. Rebello, "Supervised and Unsupervised Spectral Angle Classifiers," *Photogram. Eng. Rem. Sens.*, vol. 68, no. 12, pp. 1271-1280, 2002.

[31] N. Keshava, "Distance Metrics and Band Selection in Hyperspectral Processing with Application to Material Identification and Spectral Libraries," *IEEE Trans. Geosci. Rem. Sens.*, vol. 42, no. 7, pp. 1552-1565, 2004.

[32] D. Bash, P. Muela, and Y. M. Villegas, "Assessing land cover changes using standardized principal component and spectral angle mapping techniques," in *Proc. of Pecora 15/Land Satellite Information IV/ISPRS Commission I/FIEOS*, 2002.

[33] F. Van Der Meer, "Spectral unmixing of Landsat Thematic Mapper data," *Int. J. Rem. Sens.*, vol. 16, no. 16, pp. 3189-3194, 1995.

[34] P. E. Dennisona, K. Q. Halliganb, and D.A.Robertsc, "A Comparison of Error Metrics and Constraints for Multiple Endmember Spectral Mixture Analysis and Spectral Angle Mapper," *Rem. Sens. Environ.*, vol. 93, p. 359–367, 2004.

[35] D. Gillis and J. Bowles, "Target Detection in Hyperspectral Imagery Using Demixed Spectral Angles," in *Proc. Conf. Image and Signal Process. Remote Sens. IX, SPIE*, Barcelona, Spain, 2004, pp. 244-254.

[36] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, 2nd ed. San Diego: CA: Academic, 1997.

[37] T. K. Moon, "The expectation-maximization algorithm," *Signal Processing Mag.*, vol. 13, no. 6, p. 47–60, 1996.

[38] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc.*, vol. 39, no. 1, pp. 1-38, 1977.

[39] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," in *Proceedings of 5-th Berkley Symposium on Mathematical Statistics and Probability*, Berkley, University of California, 1967, pp. 281-297.

[40] D. Pelleg and A. W. Moore, "X-means: Extending K-means with Efficient Estimation of the Number of Clusters," in *Proc. 17th International Conference on Machine Learning*, 2000, pp. 727-734.

[41] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 2, pp. 411-423, 2001.

[42] M. Halkidi and M. Vazirgiannis, "Clustering Validity Assessment: Finding the Optimal Partitioning of a Data

Set," in *First IEEE International Conference on Data Mining (ICDM'01)*, 2001, p. 187.

[43] A. K. Jain, "Cluster analysis," in *Handbook of Pattern Recognition and Image Processing*, T. Y. Young and K. S. Fu, Eds. New York: Academic, 1986.

[44] F. Bovolo, et al., "Analysis of the Effects of Pansharpening in Change Detection on VHR Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 1, pp. 53-57, Jan. 2010.

[45] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE Pansharpening of Very High Resolution of Multispectral Images," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 46, no. 1, pp. 228-236, Jan. 2008.

[46] RSI. (2003) ENVI User Manual. [Online]. HYPERLINK "http://www.RSInc.com/envi" http://www.RSInc.com/envi

[47] F. Bovolo, L. Bruzzone, and S. Mrchesi, "Analysis and Adaptive Estimation of Registration Noise in VHR Images," *IEEE Trans. Geosci. Rem. Sens.*, vol. 47, no. 8, pp. 2658-2671, 2009.

[48] S. Marchesi, F. Bovolo, and L. Bruzzone, "A Context-Sensitive Technique Robust to Registration Noise for Change Detection in VHR Multispectral Images," *IEEE Transaction on Image Processing*, p. inress, 2010.

[49] F. Bovolo, "A Multilevel Parcel-Based Approach to Change Detection in Very High Resolution Multitemporal Images," *IEEE Geosci. Rem. Sens. Lett.*, vol. 6, no. 1, pp. 33-37, 2009.

[50] L. Bruzzone and S. B. Serpico, "An Iterative Technique for the Detection of Land-Cover Transitions in Multitemporal Remote-Sensing Images," *IEEE Trans. on Geosci. and Rem. Sens.*, vol. 35, no. 4, pp. 858-867, 1997.

[51] L. Bruzzone and D. Fernández-Prieto, "Unsupervised retraining of a maximum-likelihood classifier for the analysis of multitemporal remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 2, pp. 456-460, 2001.

[52] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, "Cluster validity methods: part I," *ACM SIGMOD Record*, vol. 31, no. 2, pp. 40-45, Jun. 2002.

[53] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, "Clustering validity checking methods: part II," *ACM SIGMOD Record*, vol. 31, no. 3, pp. 19-27, Sep. 2002.