Title: The Time Variable in Data Fusion: A Change Detection Perspective

This paper appears in: IEEE Geoscience and Remote Sensing Magazine

Date of Publication: September, 2015

Author(s): Francesca Bovolo, Lorenzo Bruzzone

Volume: 3, Issue: 3

Page(s): 8 - 26

DOI: 10.1109/MGRS.2015.2443494

THE TIME VARIABLE IN DATA FUSION: A CHANGE DETECTION PERSPECTIVE

Francesca Bovolo[1], *Lorenzo Bruzzone*[2]

1 Center for Information and Communication Technologies, Fondazione Bruno Kessler

Via Sommarive, 18 I-38123, Povo, Trento, Italy

2 Department of Information Engineering and Computer Science, University of Trento

Via Sommarive, 5 I-38123, Povo, Trento, Italy

e-mail: bovolo@fbk.eu

Abstract-This paper presents an overview on the image fusion concept in the context of multitemporal remote sensing image processing. In the remote sensing literature, multitemporal image analysis mainly deals with the detection of changes and land-cover transitions. Thus the paper presents and analyses the most relevant literature contributions on these topics. From the perspective of change detection and detection of land-cover transitions, multitemporal image analysis techniques can be divided into two main groups: i) those based on the fusion of the multitemporal information at feature level, and ii) those based on the fusion of the multitemporal information at decision level. The former mainly exploit multitemporal image comparison techniques, which aim at highlighting the presence/absence of changes by generating change indices. These indices are then analysed by unsupervised algorithms for extracting the change information. The latter rely mainly on classification and include both supervised and semi/partially-supervised/unsupervised methods. The paper focuses the attention on both standard (and largely used) methods and techniques proposed in the recent literature. The

analysis is conducted by considering images acquired by optical and SAR systems at medium, high and very high spatial resolution.

***Keywords* –** Multitemporal data fusion, Fusion at feature level, Fusion at decision level, Multitemporal images, Time series, Change detection, Image comparison, Land-cover transitions, Multitemporal image classification, Unsupervised methods, Supervised methods, Semi-supervised methods.

1 INTRODUCTION

In the last years a strong interest has been devoted to the development of novel methodologies for multitemporal information extraction and analysis. This is demonstrated by the sharp increase in the number of papers published on the major remote sensing journals, the increased number of sessions in international conferences and the increased number of projects related to multitemporal images and data.

The main reasons for this are: i) the increased number of satellites with higher revisit period that allow the acquisition of either long time series or frequent bi-temporal images, ii) the new policy for data distribution of archive data that makes it possible a retrospective analysis on large scale (e.g., the Landsat Thematic Mapper archive), and iii) the new policies for the distribution of new satellites data (e.g., ESA Sentinel).

Multitemporal information extraction methodologies differ on the basis of both the specific investigated application and the kind of data available. Accordingly, different kinds of multitemporal products are more suitable to be considered in certain applications than others. The most widely addressed applications are related to products obtained through change-detection analysis, multitemporal classification and trend analyses of temporal series of data (for change identification or forecasting/prediction).

According to an information theory perspective, the information in multitemporal data is associated with the dynamic of the variables that are measured, which is linked with the changes occurred between successive acquisitions. Thus the most interesting applications are related to the fusion/integration of multitemporal data/image for the detection of changes. We can distinguish among abrupt changes that occur in a short time (e.g., the ones caused by forest fires, floods and earthquakes) or medium/long term changes, which can be appreciated only by comparing long time series of images (e.g., desertification, urban growth). The above-mentioned applications can be addressed by using images acquired at different times by: i) the same sensor; ii) different sensors with similar properties; iii) different sensors with different properties. Accordingly, multitemporal data fusion can be integrated with multisensor fusion (when multitemporal images are acquired by different sensors either at the same time or at different times) or multisource fusion (when ancillary data are used for representing the information at given times of the considered temporal acquisition). However, in this paper we focus on the time variable only, assuming that the temporal images are taken from the same sensor at different times and analyse in detail the problem of change detection.

The main methodological approaches proposed in the literature to the automatic analysis of changes in multi-temporal remote-sensing images can be categorized in relation to the different levels at which fusion in the time domain can be conducted. Two main categories of algorithms can be defined:

1. **Fusion at feature level**: includes algorithms where multitemporal information is extracted by means of fusion of multitemporal features/images. Multitemporal information is associated with differences in the spectral signatures (or the backscattering coefficient) of the land-covers. After integration/fusion the separation between changed and unchanged areas (i.e., each pixel is associated with one of two possible classes: the class of changed patterns or the class of unchanged patterns) is performed mainly by unsupervised decision approaches. Sometimes land-cover transitions can be distinguished but without explicit labelling.
2. **Fusion at decision level**: includes algorithms that elaborate the multitemporal signature performing fusion at the level of decision. Approaches in this category are mainly supervised or semi/partially-supervised/unsupervised. They explicitly identify land-cover classes in each considered time instant and thus land-cover transitions (these methods can also be used when there are no changes between images for generating land-cover maps).

Within each category methods can be divided according to the strategy for extracting multitemporal information.

Figure **1** shows the tree of multitemporal data fusion approaches that will be described in the following. It is worth noting that this is a possible categorization of the methods presented in the literature. However, alternative categorizations could be considered.

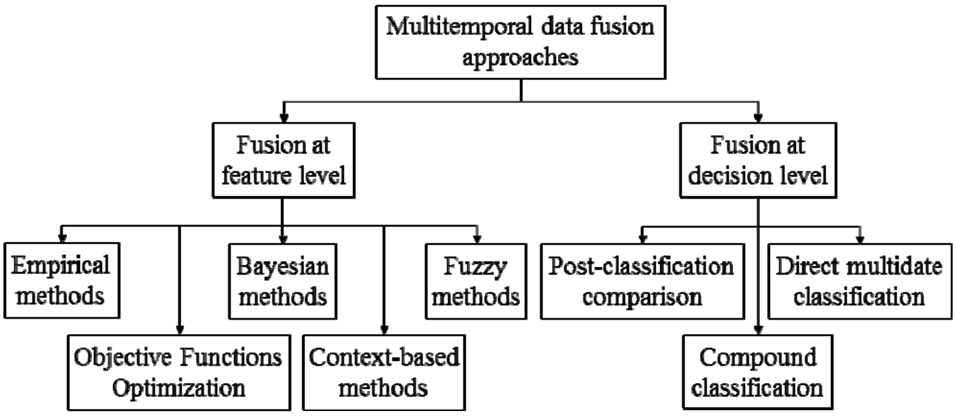


Figure 1 Tree of multitemporal data fusion approaches.

Methods in the tree should be implemented taking into account the characteristics of the considered kind of data. However in general the overall block scheme of multitemporal fusion is the one shown in Figure 2.

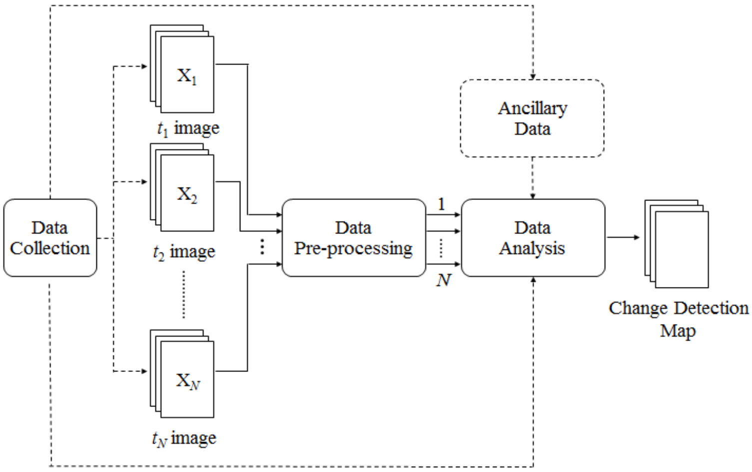


Figure 2 Overall block scheme of multitemporal information fusion.

2 FUSION/INTEGRATION AT FEATURE LEVEL: MULTITEMPORAL IMAGE COMPARISION

The aim of fusion at feature level of multitemporal images is to generate new features that highlight multitemporal information. These features are often referred to as change indices since they are employed to highlight changes occurred in bi-temporal image pairs (X_1 and X_2 acquired over the same area at different times t_1 and t_2). Change indexes are the main input to unsupervised change detection procedures that generate maps without the use of ground reference information. Figure 3 summarizes the basic processing chain for multitemporal information fusion and extraction performed at feature level. Due to their unsupervised nature this kind of approaches are widely employed since at an operational level ground reference information is often not available (e.g., the user is interested to investigate a phenomenon occurred in the past for which no information was collected), costly (i.e., it requires in situ surveys by experts with proper equipment) or impossible (i.e., ground truth is required over a very large or arduous area) to be collected. Unsupervised change detection approaches mainly distinguish between changed and unchanged pixels. Some techniques allow identifying different kinds of changes as well. However they do not give any explicit label to land-cover transitions. In the final map each pixel is associated with one among the following classes: no-change (ω_n) or change (Ω_c). In the case that land-cover transition can be distinguished, the latter class can be further detailed in K kinds of change as $\Omega_c = \{\omega_{c_1}, \omega_{c_2}, ..., \omega_{c_K}\}$ [1]. Fusion at feature level commonly assumes that bi-temporal images are accurately pre-processed in order to mitigate differences that do not depend on real changes occurred on the ground. In other words pre-processing aims at making multitemporal images as similar as possible to each other. Pre-processing usually includes radiometric corrections (relative or absolute), geometric corrections (co-registration, ortho-rectification, geo-referencing), de-noising, etc. These steps have to be conducted in different ways depending on the kind of considered data (either active SAR or passive optical). A literature survey on the pre-processing phase is out of the aim of this paper however a large amount of material can be found in the literature.

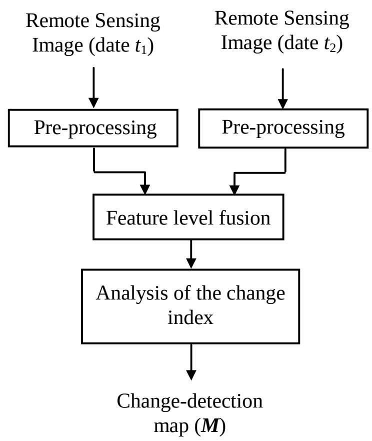


Figure 3 Block scheme of a standard change detection approach based on fusion at feature level.

Several mathematical operators can be applied to extract a change index. The choice of the specific mathematical operator gives rise to different kinds of techniques [1],[3]-[7]. Change indices by bi-temporal image fusion highlight information associated with changes in the spectral signature or the backscattering coefficient depending on whether optical or SAR images are considered, respectively. Given the technological differences in the acquisition processes of the two mentioned kinds of data they need to be treated separately. In order to extract the change information after fusion, a proper unsupervised image analysis technique should be adopted. In the literature approaches have been developed including pixel-based [1],[8],[46], context-based [23],[46], single-scale [1]-[18] and multi-scale [15],[20]-[32] approaches. Among pixel-based techniques, the most widely used is based on the selection of a decision threshold that aims at separating changed from unchanged pixels. The decision threshold can be selected either with a manual trial-and-error procedure (according to the desired trade-off between false and missed alarms) or with automatic techniques (*e.g.*, by analysing the statistical distribution of the image obtained after comparison, by fixing the desired false alarm probability [8],[9], following a Bayesian minimum-error/cost decision rule [2],[46], using methods based on fuzzy theory [17],[18], etc.). Among context-based techniques there are the ones based on fixed size sliding windows [15],[46], and the ones based on adaptive segmentation [14],[97]. Among multi-scale techniques, three main strategies can be identified: adaptive multiscale techniques for SAR images [20], multilevel parcel-based technique suitable for very high resolution images [29],[97], and approaches based on the use of similarity measures [15],[34].

A Fusion of Multitemporal Optical Images

As mentioned above, fusion of multitemporal data can be performed by image comparison. Several techniques can be employed to this end. When dealing with optical data acquired by passive sensor these techniques mainly rely on the difference operator (see Table 1). This is because the noise model in optical images is additive and the natural classes have a Gaussian distribution. Thus the difference operator results to be the most effective one.

The simplest way to apply the difference operator is to consider the same spectral band for $\mathbf{X}_1$ and $\mathbf{X}_2$ and perform subtraction pixel-by-pixel. This technique is referred to as *Univariate Image Differencing* [1],[5]-[7]. The follow up of this approach leads to the use of multiple spectral bands [1]. This technique takes the name of *Change Vector Analysis* because the results of differencing are a multidimensional Spectral Change Vector (SCV) [35],[36],[41],[44]-[55]. Under the assumption of Gaussian distributed natural classes and being the difference a linear operator, classes of change and no-change in the SCV feature space result to be Gaussian distributed as well [41]. However when non-linear features are extracted from SCVs, the analysis becomes more complex. In fact, in order to better characterize the properties of changes it is common to compute the magnitude and the directions of SCVs by applying Cartesian to Spherical coordinates transformation [41],[44]-[55]. The magnitude image is such that pixels associated with land-cover changes present values significantly higher than those of pixels associated to unchanged areas [1],[47]. Both change and no-change classes are often assumed to follow a Gaussian [55] or nearly Gaussian [36] statistical distribution. However, in [47] it has been demonstrated that, under some reasonable assumptions, they are Rayleigh and Rice distributed, respectively. Direction variables carry less information about unchanged samples since they result to be uniformly distributed [47]. They become highly relevant when analysing the classes of change instead, since they characterize different kinds of change. Changes assume preferred directions depending on the kind of change. Examples can be found in the literature where the direction information is used in the change detection process [1],[47]-[55]. Figure 4 gives an example of a change detection problem in multispectral optical images and of fusion at feature level conducted by CVA. A detailed analytical derivation of class statistical distribution in the magnitude-direction domain can be found in [47], which was developed under the assumption that multitemporal images are well co-registered and radiometrically corrected. If the assumptions on image pre-processing are not satisfied, the statistical distribution of spectral change vectors becomes more complex and the change-detection process rather difficult and less effective. This points out the importance of a proper pre-processing [47]. Cartesian to Spherical coordinates transformation preserves the dimensionality problem. Sometimes this can be a drawback since it hampers the visualization of the feature space when the dimensionality becomes higher than 3. A possible alternative is to use Compressed Change Vector Analysis (C^2VA) [50]. C^2VA compresses the information present in SCVs by computing the direction as the angular distance between the multispectral difference vector and a reference vector. If combined with the

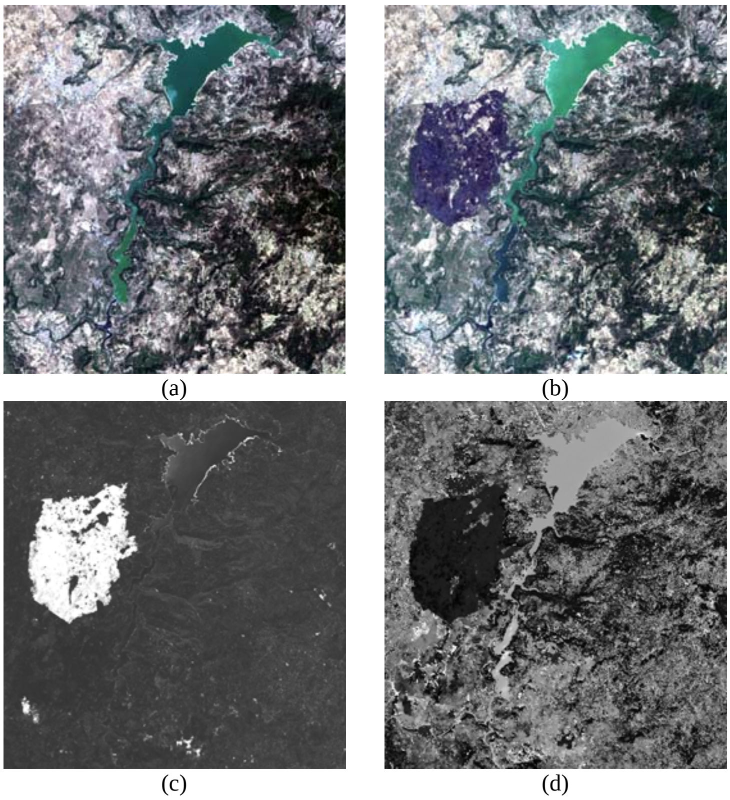
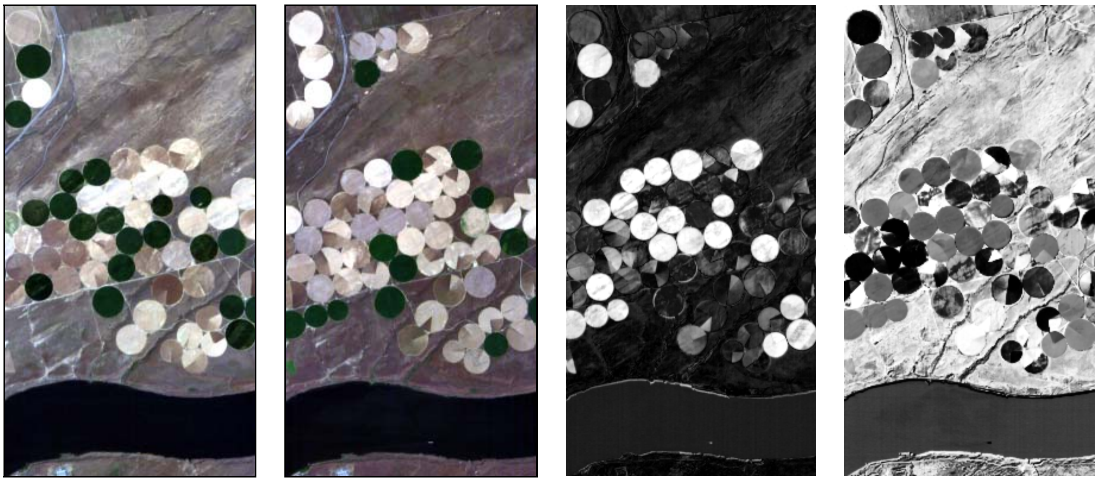


Figure 4 Example of fusion at feature level in multitemporal optical images. RGB true colour composition of Landsat-8 images acquired in: a) July 2013, and b) August 2013. c) Magnitude image and d) direction image computed according CVA. The area of interest is located close to the Lake Omodeo in Sardinia Island (Italy). Changes occurred between acquisition dates are associate to a forest fire (left-top) and the increase of the lake surface (center-top).



(a) (b) (c) (d)

Figure 5 Example of fusion at feature level in multitemporal optical images. RGB true colour composition of Hyperion EO-1 images acquired in: a) 1st May 2004, and b) 1st May 2007 (images downloaded from Geological Survey (USGS) website http://earthexplorer.usgs.gov/). c) Magnitude image and d) direction image computed according C²VA. The area of interest is located close to Hermiston city in Umatilla County, U.S. The study area is an agricultural land and changes are mainly associated to crops.

magnitude we obtain a 2-dimensional feature space that can be easily visualized and where no information is neglected. Both characteristics become highly interesting and are successfully applied when multitemporal hyperspectral images are considered [51]-[53]. A limitation of this approach is that the lossy compression of the direction information may result in recognizing some classes of change as a single class. Figure 5 gives an example of a change detection problem in hyperspectral optical images and of fusion at feature level conducted by C^2VA

CVA and difference operator have been mainly applied to the original image feature space. However examples can be found where they are applied to the posterior probability space [54] as well as to vegetation indices (*Vegetation Index Differencing*) [1],[3] or other linear (*e.g.*, *Tasselled Cap Transformation* [3], Multivariate Alteration Detection [56]-[58],[108]) or non-linear combinations of spectral bands. Transformation-based techniques like Multivariate Alteration Detection (MAD) [108] have been widely investigated resulting in several subsequent amelioration like the Iteratively Reweighted (IR)-MAD [56],[57] combined with Maximum Autocorrelation Factor (MAF) Transformation to find maximum change areas and its kernel version [58]. In the transformed feature space and after differencing, similar to CVA, unchanged and changed areas will show significantly different values. MNF (Maximum Noise Fraction)/MAD [58] has been employed for change detection in multisensor multitemporal images as well. The main characteristics that makes the MNF/MAD suitable for multisensor change detection are: i) random variables associated to the acquisitions before and after the event of change should not necessarily have same dimension (i.e., multitemporal images are not required to have the same number of spectral channels); ii) the MNF/MAD method is invariant to linear transformations, which implies that the impact of missing radiometric normalization and rectification is lower than for other change detection approaches. An alternative approach based on transformations is to use *Principal Component Analysis* (PCA). PCA can be applied separately to the feature space at single time image [1],[4],[7] or jointly to the stacked image features [103],[104]. In the first case, comparison should be performed in the transformed feature space before performing change detection; in the second case the minor components of the transformed feature space contain change information. Other linear transformations have been used such as tasselled cap, and Gram-Schmidt orthogonalization [106].

More recently the multiscale/resolution concept has been introduced in the multitemporal image fusion literature. The first works were devoted to SAR data because of their complexity. However their use resulted to be effective in optical images as well. As an example the Wavelet decomposition was used in [21],[23]-[25], and the Contourlet transform was used in [30]. Such transformations have been applied either before or after applying fusion at feature level. Multiresolution profiles for multitemporal images have been elaborated by using features extracted from multiresolution segmentation [29],[32], morphological profiles and their improvements [35],[36], and methods based on scale-invariant feature transform (SIFT) [37]. More sophisticated approaches to the representation of multiresolution information have been developed when very high spatial resolution (VHR) images should be analysed (Figure 6 shows an example of a change detection problem in VHR optical images). Such approaches aim at effectively model the high level semantic information available in VHR images [41]. Reasoning at a higher level of abstraction makes such approaches intrinsically suitable for multisensor analysis [42]-[44]. Sometimes more than one change index is jointly used in the detection process [45]. Table 1 gives an overview of the most widely used comparison operators for multitemporal fusion at feature level when passive sensor images are considered.

(a) (b)

Figure 6 Example of change detection problem in VHR optical images. True colour composition of QuickBird pansharpened images acquired in: a) October 2005, and b) July 2006. The area of interest is a sub-urban area located close to Trento (Italy). Changes are mainly associated to buildings (see white circles in the right image).

TABLE 1 SUMMARY OF THE MOST WIDELY USED COMPARISON OPERATORS. f_k IS THE CONSIDERED FEATURE AT TIME t_k THAT CAN BE: I) A SINGLE SPECTRAL BAND X_k^b; II) A VECTOR OF M SPECTRAL BANDS $[X_k^1, ..., X_k^m]$; III) A VEGETATION INDEX V_K; IV) A VECTOR OF FEATURES $[P_k^1, ..., P_k^m]$ OBTAINED AFTER TRANSFORMATION; X_D IS THE IMAGE AFTER COMPARISON.

Technique	Feature vector f_k at the time t_k	Comparison operator
Univariate image differencing	$f_k = X_k^b$	$X_D = f_2 - f_1$
Vegetation index differencing	$f_k = V_k$	$X_D = f_2 - f_1$
Regression	$f_k = X_k^b$	$X_D = f_2 - f_1$
Change vector analysis	$f_k = [X_k^1, ..., X_k^m]$	$X_D = f_2 - f_1$
Transformation based (PCA, etc.)	$f_k = [P_k^1, ..., P_k^m]$	$X_D = f_2 - f_1$
MAD	$f_k = [X_k^1, ..., X_k^m]$	$X_D = a^T f_2 - b^T f_1$

B Fusion of Multitemporal SAR images

When dealing with SAR images the additive noise model is no longer valid. For active sensor images the commonly adopted noise model is multiplicative. The direct consequence of the noise model is that the difference operator becomes poorly effective. Let us consider two multilook intensity SAR images. It is possible to show that after subtraction the statistical distribution of the resulting image depends on both the relative change between the intensity values in the two images and a reference intensity value (i.e., the intensity at t_1 or t_2). This leads to a higher change-detection error for changes occurred in high-intensity regions of the image than in low-intensity regions. Although in the past the difference operator was used with SAR data [60], the aforementioned behaviour is an undesired effect that renders the difference operator intrinsically not suited to the statistics of SAR images. To overcome this problem the ratio operator (*Image Rationing*) [1] was introduced in the SAR multitemporal image comparison at feature level literature. The ratio operator demonstrated to be more effective [2],[59],[60] because its distribution depends only on the relative change in the average intensity between the two dates and not on a reference intensity level [2],[59]. Moreover it is possible to prove that the distribution of the ratio image depends on the true change in the radar cross section. Thus changes are detected in the same manner both in high- and low-intensity regions. Moreover rationing allows to reduce common multiplicative error components (which are due to both multiplicative sensor calibration errors and to the multiplicative effects of the interaction of the coherent signal with the terrain geometry [2],[61]), as far as these components are the same for images acquired with the same geometry. In the literature, the ratio image is usually expressed in a logarithmic scale. With this operation the distribution of the classes of interest in the ratio image becomes more symmetrical and the residual multiplicative speckle noise can be transformed in an additive noise component [2],[10],[20],[60]-[62]. Thus the log-ratio operator is typically preferred when dealing with SAR images [2],[10],[20],[60]-[62]. Changes in SAR images can be associated to both increase and decrease of backscattering. The two contributions locate on the left and right side of the no-change class distribution, respectively. Sometimes there is no interest in distinguishing between the two contributions. To model them as a single change class, the normalized log-ratio can be computed.

Another set of comparison operators widely used when performing fusion at feature level in SAR images, is the one based on the use of information theoretical similarity measures [63]. Despite they have been mainly employed for SAR data, they can be successfully applied to optical data as well [34]. In [15],[34] the Kullback–Leibler (KL) divergence was used as change index. The divergence is a function of two probability densities characterizing a random variable that describes the image behaviours in the local neighbourhood of the analysed pixel. If the probability densities are similar (no change), the Kullback-Leibler divergence has a small value otherwise the value is high. The Kullback-Leibler divergence can be applied to the backscattering values or to higher order statistics (e.g., co-occurrence probability texture features). The KL divergence is not symmetric as it stands, but a symmetric version (i.e., the KL distance) may be defined by summing up the two non-symmetric divergences. In order to estimate the KL divergence/distance, the statistical distributions of the random variables associated to the multitemporal images have to be known. In the case of Gaussian distributed variables, the KL distance assumes a closed form. If no assumption on the distribution can be done, a non-parametric estimation needs to be conducted which is computationally demanding. As an example [15] provides an approach to reduce the computational burden, where the shape of the local probability density functions is estimated by the Edgeworth expansion, which is based on the cumulants of the random variables. In the similarity measures family we find the Normalized Information Distance (NID) [115]. It represents the dominant similarity aspect between every two objects and it is defined as the length of the shortest binary program that is needed to transform the two objects into each other. This distance can be interpreted also as being proportional to the minimal amount of energy required to do the transformation from one image to the other. Since the similarity metric is based on the non-computable notion of Kolmogorov complexity, a practical analogue of the NID is based on real-world compressors: the Normalized Compression Distance (NCD) [115]. The NCD is a nonnegative number representing how different the two images are. Smaller NCD values represent no changes and vice versa. This complexity based measure captures full non–linear dependencies. The correlation coefficient can be used as similarity measure as well. The correlation captures the linear dependencies between images, both in the pixel based intensity as well as spatially. Several other statistical similarity measures have been employed for fusion at feature level such as: Mutual Information [64],[65] which quantifies the common information or the statistical independence between a couple of random variables (and its coherent extension to the use in polarimetric SAR data [66]); Variational Information introduced in [67], which quantifies the different information between the couple of random variables; and Mixed Information [68], which unifies the Mutual and Variational Information. In the latter case, the trade-off between the Mutual and Variational Information is reached according to a parameter α in the range [0,1]. Mixed information can be seen as a simple linear combination of the Mutual Information and the joint entropy as well [68]. The Mixed information results in better performance than the Mutual and Variational considered independently. Furthermore, it exhibits good results when fusing multi-temporal or multi-sensor images which suffer from different radiometric conditions. Another example of joint use of more than one comparison operator can be found in [26]. The mentioned similarity measures rely on a sliding window. Larger windows reduce the time consumption but decrease the sensibility to details. Thus a trade-off is required.

More recently the multiscale/resolution concept has been introduced in the multitemporal image analysis. This need emerged because of the complexity of SAR data and because of the intrinsic multiresolution information available in the images acquired by the new generation high spatial resolution sensors. To properly model multiscale/resolution information different approaches have been used either before or after applying fusion at feature level according to the operators listed above. Among the others we recall the Wavelet decomposition [20]-[22],[24]-[28], the Contourlet transform [30],[31], and the local similarity measures computed on varying windows size [15] or multiscale segments [33]. Figure 7 shows an example of a change detection problem in high resolution SAR images and of possible change indices. The multiresolution analysis had a step forward when very high spatial resolution images become available leading to the development of methods based on objects,

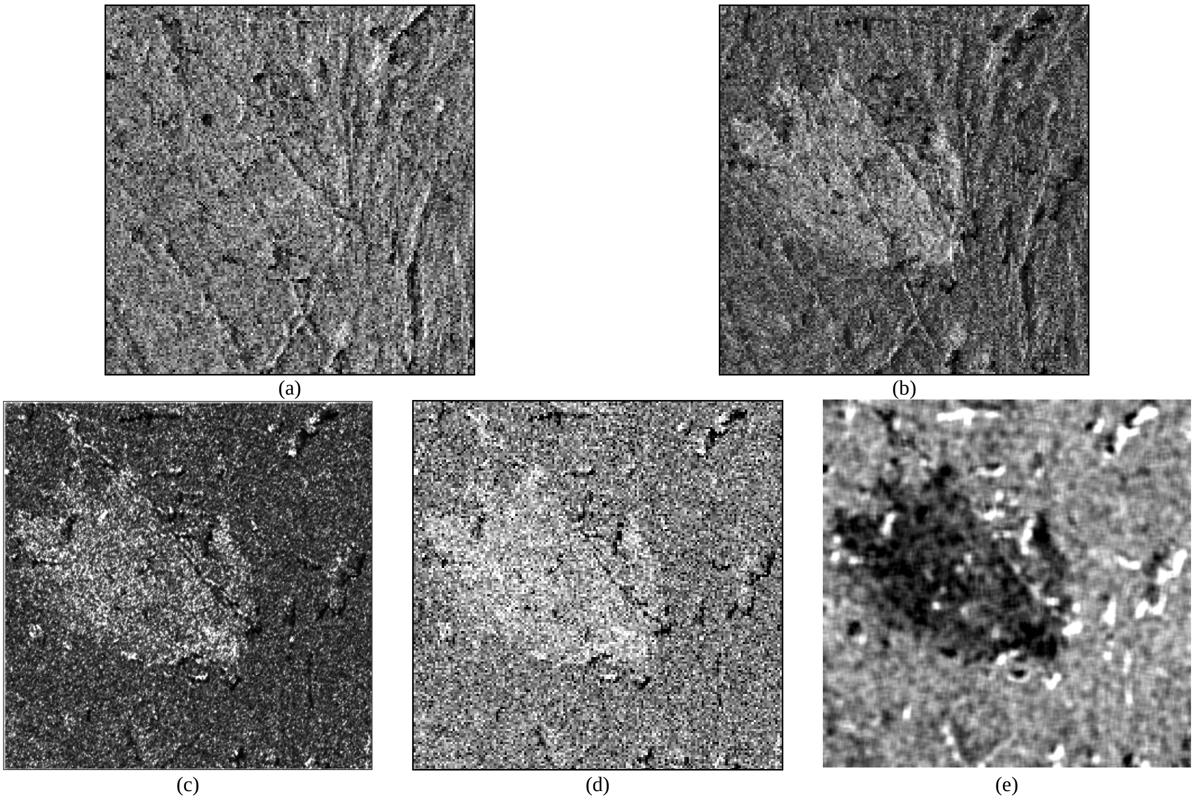


Figure 7 Example of fusion at feature level in multitemporal SAR images. Images acquired from the ERS-1 SAR sensor in (a) July 1995 and (b) October 1995. (c) ratio image, (d) log-ratio image, (e) KL distance computed on a 3x3 sliding window. The area of interest is located in Saskatchewan province, Canada. Changes occurred between acquisition dates are associated to a forest fire.

primitives extraction and modelling [38]-[40] suitable for modelling the high level semantic information available in VHR images. As already mentioned in sec II.A, reasoning at a higher level of abstraction makes such approaches intrinsically suitable for multisensor analysis [42]-[44]. Figure 8 gives an example of a change detection problem in VHR SAR images and of possible change indices. Table 2 gives an overview of the most widely used operators when performing fusion at feature level with SAR multitemporal images.

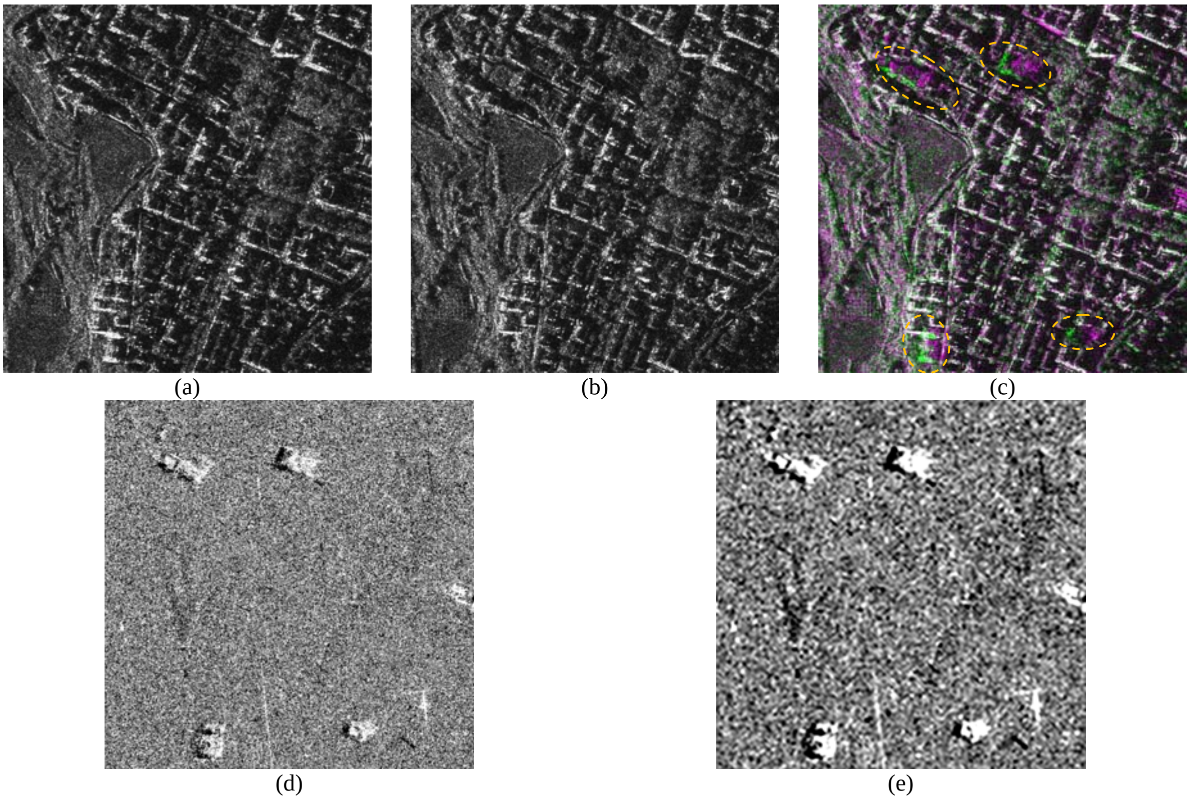


Figure 8 Example of change detection problem in multitemporal VHR SAR images. Images acquired from the COSMOSkyMed constellation in (a) April 2009 and (b) September 2009. (c) Multitemporal false colour composition. (d) and (e) are the log-ratio images obtained at wavelet decomposition level 1 and 3, respectively (for further details on how wavelet decomposition was computed refer to [20]). The area of interest is located in L'Aquila, Italy. Changes occurred between acquisition dates are associated to buildings destroyed by the earthquake occurred in April 2009 (yellow circles in the false colour image). COSMO-SkyMed Product – ©ASI – Agenzia Spaziale Italiana – (2009). All Rights Reserved

Up to now, comparison/fusion techniques for single polarimetric SAR images have been presented. However polarimetric SAR data can be considered for multitemporal fusion as well. When dealing with polarimetric SAR the comparison operators listed above have been used after extracting specific features like: i) the backscattering coefficient [70], ii) Cloude decomposition (or H-α decomposition) [71], iii) Polarimetric signatures (i.e., a by-product of polarimetry synthesis) [72] or the polar azimuthal polarimetric signature [73]. Due to the richer information of polarimetric data with respect to non polarimetric ones, additional specific comparison operators can be considered, e.g., difference of scattering matrix element products or correlation coefficients; polarimetric change indices based on the covariance matrix Σ, Contrast Ratio (or Rayleigh quotient) and Ellipticity indices [75], the test of equality for two complex Wishart matrixes [76], the Bartlett test [75]. Among them the Wishart test in diagonal case appears to be the fastest, while returning satisfying results compared to the full case of the Wishart test. It has low sensitivity to noise, sensitivity to a broader range of changes and it is able to characterize the kind of changes.

TABLE 2 SUMMARY OF THE MOST WIDELY USED COMPARISON OPERATORS. f_k IS THE CONSIDERED FEATURE AND X_k IS THE SAR IMAGE AT TIME t_k, $p(.)$ IS THE PROBABILITY DENSITY.

Technique		Feature vector f_k at the time t_k	Comparison operator
Image differencing		$f_k = X_k$	$X_{LR} = f_2 - f_1$
Image rationing		$f_k = X_k$	$X_R = f_2 / f_1$
Log-ratio		$f_k = X_k$	$X_{LR} = \log f_2 - \log f_1$
Similarity measures	Kullback-Leibler distance	$f_k = X_k$	$KL(f_2 \mid f_1) = \int \log\left[\frac{p(f_1)}{p(f_2)}\right] p(f_1)$
	Mutual information	$f_k = X_k$	$I(f_2, f_1) = \sum_{X_2, X_1} p(f_2, f_1) \log\left[\frac{p(f_1, f_2)}{p(f_1) p(f_2)}\right]$
	Variational information	$f_k = X_k$	$VI(f_1, f_2) = -\sum_{f_2, f_1} p(f_1, f_2) \log\left[\frac{p(f_1, f_2)^2}{p(f_1) p(f_2)}\right]$

C Multitemporal Information Extraction

Fusion at feature level results in features where change information is highlighted. In order to effectively extract such information further steps are required. For sake of completeness we provide a brief summary of methods for multitemporal information extraction after fusion. Depending on the technique used to fuse multitemporal images at feature level, changes can be identified in different positions of the corresponding change index feature space. Methods for information extraction available in the literature can be classified into: i) empirical methods; ii) methods based on the Bayesian decision theory; iii) methods based on the optimization of an objective function; iv) methods based on fuzzy theory; v) methods based on the use of spatial-context information.

In empirical methods, changes are identified by thresholding a change index. The threshold identification can be performed according to empirical strategies [77] that often employ manual trial-and-error procedures, which significantly affect the reliability and accuracy of the final change-detection map. The basic assumption when applying empirical strategies is that changed pixels are few and show values significantly different from the unchanged ones. Thus changed pixels are those far from the mode of the density function associated to the change index. A simple strategy consists in fixing the decision threshold as $n\mu + \sigma$, being μ and σ the mode and the standard deviation of the considered change index, respectively, and n is a real number derived by a trial-and-error procedure. In this context, the selection of the parameter n strongly depends on the end-user's subjective criteria, which may lead to unreliable change-detection results. In addition, such a selection usually requires several trials and hence a non-negligible computation time. An alternative strategy, which is typically adopted in SAR image processing, is to label as changed pixels the ones that modified their backscattering more than $\pm x\,\mathrm{dB}$, where x is a real number depending on the considered scene. The value of x is fixed according to the kind of change and the expected variation in order to obtain a desired probability of correct detection (which is the probability to be over the threshold if a change occurred) or false alarm (which is the probability to be over the threshold if no change occurred). It has been shown that the value of x can be analytically defined as a function of the true change in the radar backscattering and of the equivalent number of looks [59],[60] once detection or false alarm probabilities are fixed. A similar approach is presented [61]; it identifies the decision threshold on the basis of predefined values on the cumulative histogram of the change index. It is worth noting that these approaches are not fully automatic and objective from an application point of view, as they depend on the user sensibility in constraint definition with respect to the considered kind of change. These properties may represent a critical limitation.

An interesting alternative consists in formulating the change-detection problem in the framework of the Bayesian decision theory in order to optimize the separation between changed and unchanged pixels in an unsupervised way. The main problem to be solved for the application of the Bayes decision theory consists in the estimation of the statistical terms associated to the classes of change and no-change (i.e., their prior probabilities and probability density functions) [2],[46] without any ground-truth information (i.e., without any training set). The starting point of methodologies based on the Bayesian decision theory is the hypothesis that the statistical distributions of pixels in the change index can be modelled as a mixture of densities. Mixture components are associated to changed and unchanged pixels. In the literature, explicit estimation of class statistical parameters has been addressed with the Expectation-Maximization (EM) algorithm which is an iterative approach to maximum-likelihood (ML) estimation for incomplete data problems [78]. The iterative equations that characterize the EM algorithm are different according to the statistical model adopted for the distributions of the classes. The most suitable statistical model varies according to the kind of data. If optical passive sensor data are considered the most common statistical models are: i) Gaussian [7],[79]-[81]; ii) mixture of Gaussians [55]; iii) Rayleigh (for the magnitude of unchanged samples computed according to CVA) [47]; vi) Rice (for the magnitude of changed samples computed according to CVA) [47]; v) Uniform (for the direction of unchanged samples computed according to CVA) [47]; vi) Non-uniform (for the direction of changed samples computed according to CVA) [47]. If SAR images are considered, it has been shown that the Generalized Gaussian [2],[82], Weibull or Nakagami-Gamma [83],[84] distributions allow to better handle the complexity of the class distributions. The iterative equations needed for performing EM parameter optimization under the Gaussian, mixture of Gaussian and Generalized Gaussian class models can be found in [2],[55] and [36], respectively, whereas more details on the validity of the Rayleigh and Rice models can be found in [47].

Once the statistical parameters are computed, pixel-based or context-based decision rules from the pattern recognition literature can be applied. The review of such methods is out of the scope of the present manuscript. However, we recall the most widely used approaches in the context of change detection. Concerning pixel-based methods, we can mention: i) Bayes rule for minimum error [2],[36],[47],[80]; ii) Bayes rule for minimum cost [80]; iii) Neyman-Pearson criterion [80]. The Bayes rule for minimum cost and the Neyman-Pearson criterion allow considering the costs of false and/or missed alarms in the decision process. Bayesian decision theory can be used also in multisensor change detection [109]. Here fusion is carried out according to the consensus theory by integrating the estimates of statistical terms over different sensors. In the fusion step a weight is associated to each source according to its expected reliability. Within the Bayesian decision theory framework different techniques for reducing the effects of the residual registration noise between multitemporal images have been integrated [48],[49],[85].

Another set of methods is based on the optimization of objective (cost) functions. The fact that, generally, the change index is one-dimensional makes this process easy. The choice of the cost function plays a fundamental role in the accuracy of the results. In the change detection literature, several objective functions have been employed based on: i) discriminant analysis and inter- and intra-class measures [86]; ii) Bayes decision rule for minimum error [87]; iii) distribution free fuzzy entropy measure [88]. The optimization of objective functions leads to an implicit estimation of the class statistical parameters [87],[89]-[96]. According to the kind of data, different assumptions on the statistical distribution of classes can be made. As an example, the Kittler and Illingworth criterion has been used under both the Gaussian [89] and the Generalized Gaussian [2],[82] assumption for the statistical distributions of classes. Also methods based on Machine Learning and clustering that minimize a cost function can be listed in this category. In the literature examples can be found based on Support Vector Machine [90],[91], clustering and kernel-based clustering [65],[92],[93], neural networks [94],[98].

The use of fuzzy theory is another possibility. These kinds of techniques rely on the assumption that some ambiguity exists that arises from the overlapping nature of classes or image properties [17],[18]. The ambiguity of an image can be expressed in terms of radiometry (e.g., fuzzy entropy, hybrid entropy, correlation, etc.) or geometry (e.g., compactness, high and width, length and breadth, index of area coverage, degree of adjacency, etc.). The decision threshold is selected as the value where the membership function shows a global minimum or maximum depending on the selected ambiguity measure [17],[18]. Fuzzy clustering approaches belongs to this group as well [26].

Some approaches involve spatial-context information in the decision process. This is justified by the reasonable assumption that changes are large if compared with the spatial resolution of the sensor. Thus a pixel is likely to be surrounded by pixels of the same class. The use of interpixel dependence may yield more reliable and accurate change-detection results. A fully automatic approach to the unsupervised analysis of the change index, which exploits the spatial contextual information to reduce the effect of noise in the detection procedure, has been proposed in [2],[36]. The solution is developed in the context of Bayesian decision theory, where the spatial context of each pixel is modelled by the use of Markov Random Fields. Another effective technique capable to consider the spatial-contextual information is based on adaptive parcels, i.e., small homogeneous regions shared by both original images [29],[97]. The adaptive nature of parcels allows spatial-contextual information to be exploited so that noise may be reduced without damaging the boundaries of the changed areas. Spatial correlation between neighbouring pixels has been modelled by using Hopfield neural network [94],[98] as well. This solution is fully automatic and distribution free. Spatial-context features have been used in [49], [99]-[102] with the explicit objective of compensating for co-registration problems. Table 3 summarizes the main land-cover transitions detection approaches based on fusion at decision level.

TABLE 3 SUMMARY OF THE MAIN TECHNIQUES FOR LAND-COVER TRANSITIONS DETECTION PERFORMED AT DECISION LEVEL (THE TABLE IS NOT EXHAUSTIVE).

Change Detection Technique	**Detection Algorithm**	**Reference**	**Kind of Data**
Change vector Analysis (CVA) Compressed CVA (C^2VA) Image Differencing (ID) Vegetation Index Differencing (VID) Principal Component Analysis (PCA)	Empirical thresholding	Fung (1987) Singh (1989) Fung *et al.* (1990) Muchoney (1994) Townshend *et al.* (1995)	Multispectral
	Empirical cost function minimization	Kittler *et al.* (1986) Bruzzone *et al.* (2000, 2002) Melgani *et al.* (2002) Bovolo *et al.* (2008, 2012) Celik (2009) Muñoz-Marí *et al.* (2010) Chen *et al.* (2011) Volpi *et al.* (2012)	
	Fuzzy thresholding	Pal *et al.* (2000, 2001) Di Zenzo (1998)	
	Context-based approaches	Solberg *et al.* (1996) Bruzzone *et al.* (2000) Ghosh *et al.* (2007, 2013) Huo *et al.* (2014) Hao *et al.* (2014)	
	Multiscale/Hierarchical	Bovolo *et al.* (2005) Inglada *et al.* (2007) Dalla Mura *et al.* (2008) Bovolo (2009) Bazi *et al.* (2010) Moser *et al.* (2011) Falco *et al.* (2013) Bruzzone *et al.* (2013) Liu *et al.* (2014, 2015a, 2015b)	
	Reduction registration noise	Bruzzone *et al.* (1997, 2003) Marchesi *et al.* (2010)	
Image (Log-)Rationing (IR)	Empirical thresholding	Singh (1989) Rignot *et al.* (1993) Cihlar *et al.* (1992)	SAR

Change Detection Technique	Detection Algorithm	Reference	Kind of Data
	Thresholding explicitly or implicitly based on the Bayes decision theory	Bazi *et al.* (2004, 2005, 2006) Moser *et al*. (2006)	
	Fuzzy	Pal *et al.* (2000) Gong *et al.* (2012)	
	Wavelet/Contourlet-based multiresolution approach	Bovolo *et al.*(2005) Celik et al. (2009, 2011) Celik (2010) Li *et al.* (2012) Ma *et al.* (2012) Cui *et al.* (2012)	
	Context-based approaches	Bazi *et al.* (2005) Bovolo *et al.* (2008) Moser *et al*. (2009)	
	Multiscale	Gamba *et al*. (2006) Dell'Acqua *et al.* (2006) Bovolo *et al.* (2005, 2013) Inglada *et al.* (2007) Bazi *et al.* (2010)	
	Multivariate Alteration Detection	Nielsen *et al.* (1998)	
Kullback Leibler distance (KLD) Normalized information distance (NID) Mutual Information (I) Variational Information (VI) Mixed Information (MI)	Single scale	Meila (2003) Cebrian *et al.* (2007) Chatelain *et al.* (2007) Gueguen *et al.* (2009, 2011) Erten *et al.* (2012)	Multispectral and SAR
	Multiscale	Inglada *et al.* (2007) Celik (2010)	
	Multimodal	Datcu *et al.*	
Correlation coefficient Contrast Ratio Ellipticity	Empirical thresholding	Dierking et al (2000, 2002) Kersten *et al.* (2005)	Polarimetric SAR data
	Test Statistics	Conradsen *et al.* (2003)	
	Context based	Molinier *et al.* (2007)	
Feature and area based techniques	Thresholding and refinement	Dell'Acqua *et al*. (2004, 2006) Della Ventura *et al.* (1990)	Multispectral and SAR
Multivariate Alteration Detection (MAD) and modifications		Nielsen *et al.* (1998) Nielsen (2001, 2007) Liao *et al.* (2005) Marpu *et al.* (2011)	Multispectral, SAR and multisensor

3 FUSION AT DECISION LEVEL: MULTITEMPORAL IMAGE CLASSIFICATION

As opposed to feature-based multitemporal information fusion for change detection, a set of multitemporal image fusion techniques can be found that aims at multitemporal analysis by fusion at decision level. Methods belonging to this category mainly rely on supervised or semi/partially-supervised/unsupervised classification. The terms partially supervised classification and partially unsupervised classification have been used in the literature for defining change detection problems with bi-temporal images where reference data (and thus a training set) are available only for one of the two images considered. The two terms refer to the same concept considered from the initial perspective either of supervised or unsupervised classification. It is worth noting that most recently similar problems in the context of multitemporal classification have been defined as domain adaptation in the framework of transfer learning. The term semi-supervised refers in general to the use of labelled and unlabelled data in the learning phase of a classifier. Partially supervised and unsupervised methods (as well as domain adaptation methods) exploit semi-supervised techniques implemented across two domains (associated with the two images). After this clarification, for avoiding confusion, in this paper we refer to all these approaches with the terms which they have been presented in the literature. Such kind of approaches do not only highlight the changes, they explicitly identify the pair of classes (i.e., land-cover transition) associated with each detected change. Note that they can be successfully applied to bitemporal images both when there are changes and when there are not. However in the following we concentrate our attention to their use in the context of change detection applications. The (semi)supervised nature of these kind of approaches reduces the sensibility to radiometric differences. The use of fully or partially supervised methods depends on the availability of ground truth information. On the one hand, if multitemporal ground truth information is available supervised techniques can be applied. This information is used in the learning phase of supervised data classification for modelling the kind of land-cover transitions. On the other hand, if ground truth is available for one or some of the images in the multitemporal sequence, partially-supervised techniques should be considered. If no ground truth is available, unsupervised clustering techniques should be used.

Three main general approaches to fusion at decision level can be found in the literature: *Post-Classification Comparison* [1], *Supervised Direct Multidate Classification* [1],[110] and *Compound Classification* [111]-[114]. In the literature many different classifiers have been used in the context of the analysis of temporal series of remote sensing images. Among the others, we recall the Maximum Likelihood classifier [112], Neural Networks [116], Fuzzy Classifiers [116], and Support Vector Machines [117],[118], which are either the most widely used or the most effective ones. The reader is referred to the literature for more details on the behaviour and mathematical details of each single classifier.

The use of supervised classification is in general more accurate than unsupervised approaches. Nevertheless, it is less appealing in operational applications. This is due to the difficulties in collecting proper ground-truth information (necessary for supervised techniques), which is a complex, time consuming and expensive process (in many cases this is not consistent with the application constraints). It becomes even more complex since there is a need of multitemporal information. Semi-supervised approaches represent a trade-off between the two above-mentioned conditions. It is worth noting that all the techniques based on classifiers cited in this section are intrinsically suitable to be used with different kind of data and also with multisensor, multisource and multiresolution information.

The post-classification comparison (PCC) (also referred to as delta classification [1]) is the simplest technique among fusion at decision level approaches. It performs change detection by comparing the classification maps obtained by classifying independently the two images considered. For each change the land cover transition is obtained in an explicit way. The main advantage of delta classification lies in the fact that multitemporal images are classified independently, thereby minimizing the problem of radiometric calibration. Although PCC has been extensively used in several applications, its performance strongly depends on the classification accuracies of the classifier applied to each single image. After multitemporal fusion the accuracy is close to the product of the accuracies yielded by the independent classifiers [1],[113], making the approach often unsatisfactory [119]. This is a direct result of the fact that PCC does not take into account the temporal correlation. However it has been widely employed in the literature both at pixel and region level [127]-[130]. Attempts to increase PCC accuracy have been done by using more than two images in the fusion step [131].

Supervised direct multidate classification (DMC) [1],[113],[132], unlike PCC, takes into account the dependence existing between two images of the same area. The main idea of such a technique is to characterize pixels by a vector obtained by stacking the feature vectors related to the images acquired at the two different times. Then the identification of the land-cover transitions is carried out by considering each transition as a single class and by training a classifier to recognize such transitions. It is worth noting that a complex constraint to satisfy for using this technique is to have a training set composed of training pixels related to the same points on the ground at the two times. In addition, training pixels should represent accurately the proportions of all the transitions in the whole area of interest. This represents a serious drawback as, in real applications, it is difficult to obtain training sets with such characteristics. In [121] Schowengerdt remarked that, since spectral and temporal features have equal status in the combined data set, they cannot be easily separated in the pattern recognition process. As a consequence, class labelling may be difficult if relatively simple classification algorithms are considered. In [132] DMC has been adapted to the use in VHR images.

A more realistic approach to fusion at decision level is compound classification (CC) [113]. Similarly to the DMC, also in this case the objective is to perform the classification of pixels of the two images according to the maximization of the posterior joint probability of classes. Conditional probabilities of classes can be estimated according to different techniques and under different assumption on their statistical distribution. On the one hand, with respect to the PCC, the CC technique allows the temporal correlation between images to be considered in the change-detection process. On the other hand, with respect to the DMC method, the CC technique allows the constraints related to the training sets to be relaxed [113]. In particular, training pixels should not necessarily be related to the same area on the ground [113]-[114].

In real problems it may happen that, given a series of multi-temporal images, a ground truth is not available for all the items of the series. In such realistic cases, the aforementioned supervised approaches cannot be employed. In [122], an ensemble of non-parametric multitemporal partially-supervised classifiers was defined and integrated in the context of a multiple classifier system. Each multitemporal classifier was developed in the framework of the compound classification decision rule. In [114], a partially supervised methodology was proposed able to update the parameters of an already trained parametric maximum-likelihood (ML) classifier whenever a new image lacking the corresponding ground truth has to be analysed. The updating is performed by means of the EM algorithm [78] that allows tuning the parameters of the trained ML classifier on the basis of the distribution of the new image. In this way, it is possible to classify multi-temporal data of a given area (and hence to derive land-covers transition maps) without relying on a multi-temporal ground truth. These methods have been recently referred to as Domain Adaptation (DA) methods. In [123], a partially unsupervised technique based on Markov Random Fields is proposed for the identification of the only land cover transitions of interest for the end-user, by exploiting training samples belonging exclusively to the land covers involved in the specific kind of changes to be mapped. In [124], an advanced context-sensitive classification technique that exploits a temporal series of remote sensing images for a regular updating of land-cover maps is proposed. The authors introduced a classifier which is based on an iterative partially supervised algorithm that jointly estimates the class-conditional densities and the prior model for the class labels on the image to be classified by taking into account spatial-context information.

All the aforementioned supervised and partially (semi) supervised methods based on classification are intrinsically suitable to process also multisensor/multisource data. In fact, if proper distribution-free non-parametric classifiers are used for the analysis of the images, data acquired from different sensors (e.g., multispectral images and SAR images) can be processed to produce the map of land-cover transitions. Under the assumption that the considered images are re-sampled at the same geometrical resolution, it is also possible to employ (semi)supervised approaches with different sensors at the two dates: in fact, the comparison process is carried out at a classification-map level. This property is very important as it allows: i) to produce land-cover transition maps related to large temporal differences (large temporal differences involve the availability of data acquired by different sensors of different generation); ii) to obtain land-cover transitions maps also when data acquired from a specific sensor at the first date are not available at the second date (e.g., multispectral images at the second date might be not available depending on atmospheric conditions. In these cases, SAR images could be compared with multispectral images).

In [126] the authors describe both neuro-fuzzy and statistical approaches to the exploitation of the contextual information and the classification, and different schemes for the multisensor fusion. The presented technique appropriately fuses the information of active and passive sensors and results in a good change detection precision and in the best possible classification accuracy.

Due to the complexity in constructing ground truth information for the training of classifiers, fusion at decision level methods have been less used. However recently a novel interest from the scientific community was devoted to these methods. This is because of the methodological developments in the context of domain adaptation (DA) and active learning (AL) context. As mentioned above, DA approaches allow to take advantage of the ground truth information available for one acquisition (i.e., the source domain) and to adapt this information to images for which ground truth is not available (i.e., the target domain). The adaptation mechanism can be significantly improved by the use of AL approaches that guarantee a minimum amount of new labelled data for the target domain. Thus they handle in a better way the possible significant differences between statistical distributions of the source and target domains. Examples of such approaches to multitemporal image fusion can be found in [133]-[135]

TABLE 4 SUMMARY OF THE MAIN TECHNIQUES FOR MULTICLASS CHANGE DETECTION (THE TABLE IS NOT EXHAUSTIVE).

<table>
<tr><th>Multiclass Change Detection Technique</th><th>Reference</th><th>Kind of Data</th></tr>
<tr><td>Supervised Post-classification Comparison</td><td>Howarth et al. (1981)
Singh(1989)
Hall et al. (1991)
Bruzzone et al. (1997)
Smits et al. (1999)
Serra et al. (2003)
Pacifici et al. (2007)
Ling et al. (2011)
Kempeneers et al. (2012)</td><td rowspan="4">Multispectral
SAR
Multisensor</td></tr>
<tr><td>Supervised Direct-Multidate Classification</td><td>Schowengerdt (1983)
Singh (1989)
Hall et al. (1991)
Bruzzone et al. (1997)
Volpi et al. (2013)</td></tr>
<tr><td>Compound Classification and Domain adaptation</td><td>Bruzzone et al. (1997)
Bruzzone et al. (1999)
Fernández Prieto et al. (2001)
Cossu et al. (2005)
Serpico et al.
Demir et al. (2012, 2013)</td></tr>
<tr><td>Unsupervised approaches</td><td>Byrne et al. (1980)
Häme et al. (1998)
Bovolo et al. (2007)</td></tr>
</table>

4 CONCLUSION

This paper has addressed the data fusion problem in the context of multitemporal remote sensing images. As the topic is wide and in an extremely rapid evolution [136]-[145], it has not been possible to present an exhaustive analysis. Accordingly, we focused the multitemporal data fusion from a change detection perspective. We provided an overview of the main methods and approaches available in the literature, including the standard (and widely used) methods and the most recent developments. From the fusion perspective, multitemporal data analysis techniques can be divided into two main categories: i) fusion at feature level; and ii) fusion at decision level. The former includes mostly unsupervised methods that perform change detection after the computation of a change index. Change indices result from a comparison/fusion at feature level of multitemporal images and aim at highlighting the presence/absence of changes. Several comparison/fusion techniques and analysis algorithms exist depending on the kind of images considered (e.g., SAR or optical images). The latter mainly includes methods based on multitemporal classification. Approaches in this group are intrinsically suitable for both SAR and optical images, and can implicitly process also multisensor images. Among them we can identify methods that require exhaustive multitemporal ground truth information and methods that can work with limited/partial reference data.

As a final remark, we point out that it is not possible to identify a technique better than the others. All the different approaches, as discussed in the paper, have advantages and disadvantages. Thus the choice of the method to use at operational level strongly depends on the considered application and the end-users requirements.

5 REFERENCES

[1] A. Singh, "Digital change detection techniques using remotely-sensed data," *International Journal of Remote Sensing*, Vol. 10, pp. 989-1003, 1989.

[2] Y. Bazi, L. Bruzzone and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, pp. 874-887, 2005.

[3] J.R.G. Townshend and C.O. Justice, "Spatial variability of images and the monitoring of changes in the Normalized Difference Vegetation Index," *International Journal of Remote Sensing*, Vol. 16, pp. 2187–2195, 1995.

[4] T. Fung and E. LeDrew, "Application of principal component analysis to change detection," *Photogrammetric Engineering and Remote Sensing*, 53, 1649–1658, 1987.

[5] P.S. Jr. Chavez and D.J. MacKinnon, "Automatic detection of vegetation changes in the south-western united states using remotely sensed images," *Photogrammetric Engineering and Remote Sensing*, Vol. 60, pp. 571–583, 1994.

[6] T. Fung, "An assessment of TM imagery for land-cover change detection," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, pp. 681–684, 1990.

[7] D.M. Muchoney and B.N. Haack, "Change detection for monitoring forest defoliation," *Photogrammetric Engineering and Remote Sensing*, Vol. 60, pp. 1243–1251, 1994.

[8] S. Quegan, T. Le Toan, J.J. Yu, F. Ribbes and N. Floury,, "Multitemporal ERS SAR analysis applied to forest mapping," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, pp. 741-753, 2000.

[9] W. Dierking and H. Skriver, "Change detection for thematic mapping by means of airborne multitemporal polarimetric SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 40, pp. 618-636, 2002.

[10] K. Grover, S. Quegan and C. da Costa Freitas, "Quantitative estimation of tropical forest cover by SAR," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, pp. 479-490, 1999.

[11] J.G. Liu *et al.*, "Land surface change detection in a desert area in Algeria using multitemporal ERS SAR coherence images," *International Journal of Remote Sensing*, Vol. 2, pp. 2463-2477, 2001.

[12] P. Lombardo and T. Macrì Pellizzeri, "Maximum likelihood signal processing techniques to detect a step pattern of change in multitemporal SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 40, pp. 853-870, 2002.

[13] R.G. White and C.J. Oliver, "Change detection in SAR imagery," *Proceedings of the IEEE International Radar Conference*, pp. 217–222, 1990.

[14] F. Bovolo, L. Bruzzone, "An Adaptive Technique Based on Similarity Measures for Change Detection in Very High Resolution SAR Images", *Proceedings IEEE 2008 International Geoscience and Remote Sensing Symposium*, (IGARSS '08), Vol. 3, pp. 158 – 161, 2008.

[15] J. Inglada, and G. Mercier, "A New Statistical Similarity Measure for Change Detection in Multitemporal SAR Images and Its Extension to Multiscale Change Analysis," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, pp. 1432-1445, 2007.

[16] S.K. Pal, "Fuzzy Image Processing and Recognition," *International Journal on Image and Graphics*, Vol. 1, pp. 169-195, 2001.

[17] S.K. Pal, A. Ghosh, and B.U. Shankar, "Segmentation of Remotely Sensed Images with Fuzzy Thresholding, and Quantitative Evaluation," *International Journal of Remote Sensing*, Vol. 21, pp 2269-2300, 2000.

[18] C. Carincotte, S. Derrode, and S. Bourennane, "Unsupervised Change Detection on SAR Images Using Fuzzy Hidden Markov Chains," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 44, pp. 432-440, 2006.

[19] S. Di Zenzo, L. Cinque, S. Levialdi, "Image Thresholding Using Fuzzy Entropies," *IEEE Transactions on Systems, Man and Cybernetics-part B: Cybernetics*, Vol. 28, pp. 15-23, 1998.

[20] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, pp. 2963-2972, 2005.

[21] T. Celik, "Multiscale Change Detection in Multitemporal Satellite Images," *IEEE Geoscience and Remote Sensing Letters*, Vol. 6, pp. 820-824, 2009.
[22] T. Celik, "A Bayesian approach to unsupervised multiscale change detection in synthetic aperture radar images," *Signal Processing*, Vol. 90, pp. 1471–1485, 2010.
[23] G. Moser, E. Angiati, S.B. Serpico, "Multiscale Unsupervised Change Detection on Optical Images by Markov Random Fields and Wavelets," *IEEE Geoscience and Remote Sensing Letters*, Vol. 8, pp. 725 – 729, 2011.
[24] T. Celik, K.-K. Ma, "Unsupervised Change Detection for Satellite Images Using Dual-Tree Complex Wavelet Transform," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, pp. 1199 – 1210, 2009.
[25] T. Celik, K.-K. Ma, "Multitemporal Image Change Detection Using Undecimated Discrete Wavelet Transform and Active Contours," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 49, pp. 706 – 716, 2011.
[26] M. Gong, Z. Zhou and J. Ma, "Change Detection in Synthetic Aperture Radar Images based on Image Fusion and Fuzzy Clustering," *IEEE Transactions on Image Processing*, Vol. 21, pp. 2141-2152, 2012.
[27] J. Ma, M. Gong and Z. Zhou, "Wavelet Fusion on Ratio Images for Change Detection in SAR Images," *IEEE Geoscience and Remote Sensing Letters*, Vol. 9, pp. 1122-1126, 2012.
[28] S. Cui and M. Datcu, "Statistical Wavelet Subband Modeling for Multi-Temporal SAR Change Detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 5, 1095-1108, 2012.
[29] F. Bovolo, "A Multilevel Parcel-Based Approach to Change Detection in Very High Resolution Multitemporal Images," *IEEE Geoscience and Remote Sensing Letters*, Vol. 6, pp. 33-37, 2009.
[30] S. Li, L. Fang and H. Yin, "Multitemporal Image Change Detection Using a Detail-Enhancing Approach With Nonsubsampled Contourlet Transform," *IEEE Geoscience and Remote Sensing Letters*, Vol. 9, pp. 836-840, 2012.
[31] B. Hou, Q. Wei, Y. Zheng and S. Wang, "Unsupervised Change Detection in SAR Image Based on Gauss-Log Ratio Image Fusion and Compressed Projection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 7, pp. 3297-3317, 2014.
[32] Y. Bazi, F. Melgani, H.D. Al-Sharari, "Unsupervised Change Detection in Multispectral Remotely Sensed Imagery With Level Set Methods," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, pp. 3178 – 3187, 2010.
[33] F. Bovolo, L. Bruzzone, "An Adaptive Technique Based on Similarity Measures for Change Detection in Very High Resolution SAR Images", in *Proceedings of the IEEE* 2008 *International Geoscience and Remote Sensing Symposium* (IGARSS'08), pp. III - 158 - III - 161, 2008.
[34] G. Mercier, G. Moser, S.B. Serpico, "Conditional Copulas for Change Detection in Heterogeneous Remote Sensing Images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, pp. 1428 – 1441, 2008.
[35] N. Falco, M. Dalla Mura, F. Bovolo, J.A. Benediktsson, L. Bruzzone, "Change Detection in VHR Images Based on Morphological Attribute Profiles," *IEEE Geoscience and Remote Sensing Letters*, Vol. 10, pp. 636-640, 2013.
[36] M. Dalla Mura, J.A. Benediktsson, F. Bovolo, L. Bruzzone, "An Unsupervised Technique Based on Morphological Filters for Change Detection in Very High Resolution Images," *IEEE Geoscience and Remote Sensing Letters*, Vol. 5, pp. 433-437, 2008.
[37] L. Huo, X. Feng, C. Huo, Z. Zhou and C. Pan, "Change Field: A New Change Measure for VHR Images," *IEEE Geoscience and Remote Sensing Letters*, Vol. 11, pp. 1812-1816, 2014.
[38] C. Marin, F. Bovolo, L. Bruzzone, "Building Change Detection in Multitemporal Very High Resolution SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 53, pp. 2664-2682, 2015.
[39] P. Gamba, F. Dell'Acqua, and G. Lisini, "Change Detection of Multitemporal SAR Data in Urban Areas Combining Feature-Based and Pixel-Based Techniques," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 44, pp. 2820-2827, 2006.
[40] F. Bovolo, C. Marin, L. Bruzzone, "A Hierarchical Approach to Change Detection in Very High Resolution SAR Images for Surveillance Applications," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 51, pp. 2042-2054, 2013.
[41] L. Bruzzone, F. Bovolo, "A Novel Framework for the Design of Change-Detection Systems for Very-High-Resolution Remote Sensing Images," *Proceedings of the IEEE*, Vol. 101, pp. 609-630, 2013.
[42] D. Brunner, G. Lemoine, L. Bruzzone, "Earthquake damage assessment of buildings using VHR optical and SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, pp. 2403-2420, 2010.
[43] V. Poulain, J. Inglada, M. Spigai, J.-Y. Tourneret and P. Marthon, "High-Resolution Optical and SAR Image Fusion for Building Database Updating," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 49, pp. 2900-2910, 2011.
[44] M.N. Klaric, B.C. Claywell, G.J. Scott, N.J. Hudson, O. Sjahputera, Y. Li, S.T. Barratt, J.M. Keller and C.H. Davis, "GeoCDX: An Automated Change Detection and Exploitation System for High-Resolution Satellite Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 51, pp. 2067-2086, 2013.
[45] P. Du, S. Liu, P. Gamba, K. Tan and J. Xia, "Fusion of Difference Images for Change Detection Over Urban Areas," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 5, pp. 1076-1086, 2012.
[46] L. Bruzzone and D. Fernández Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, pp. 1170-1182, 2000.
[47] F. Bovolo and L. Bruzzone, "A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in Polar Domain", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, pp. 218-236, 2007.
[48] L. Bruzzone and R. Cossu, "Adaptive approach to reducing registration noise effects," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, pp. 2455-2465, Nov. 2003.
[49] S. Marchesi, F. Bovolo, L. Bruzzone, "A Context-Sensitive Technique Robust to Registration Noise for Change Detection in VHR Multispectral Images," *IEEE Transactions on Image Processing*, Vol. 19, pp. 1877 - 1889, 2010.

[50] F. Bovolo, S. Marchesi, L. Bruzzone, “A Framework for Automatic and Unsupervised Detection of Multiple Changes in Multitemporal Images,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 50, pp. 2196-2212, 2012.
[51] S. Liu, L. Bruzzone, F. Bovolo, P. Du, “A Novel Sequential Spectral Change Vector Analysis for Detecting Multiple Changes in Hyperspectral Images,” *IEEE 2014 International Geoscience and Remote Sensing Symposium*, (IGARSS '14), pp. 4656-4659, 2014.
[52] S. Liu, L. Bruzzone, F. Bovolo, P. Du, “Hierarchical Unsupervised Change Detection in Multitemporal Hyperspectral Images,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 53, pp. 244 – 260, 2015.
[53] S. Liu, L. Bruzzone, F. Bovolo, M. Zanetti P. Du, “Sequential Spectal Change Vector Analysis for Change Detection in Multitemporal Hyperspectral Images,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 53, pp. 4363-4378, 2015.
[54] J. Chen, X. Chen, X. Cui and J. Chen, “Change Vector Analysis in Posterior Probability Space: A New Method for Land Cover Change Detection,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 8, pp. 317-321, 2011.
[55] L. Bruzzone and D. Fernández Prieto, “An adaptive semi-parametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images,” *IEEE Transactions on Image Processing*, Vol. 11, pp. 452-446, 2002.
[56] A. Nielsen, "The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data," *IEEE Transactions on Image Processing*, vol. 16, pp. 463-478, 2007.
[57] P.R. Marpu, P. Gamba and M.J. Canty, “Improving Change Detection Results of IR-MAD by Eliminating Strong Changes,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 8, pp. 799-803, 2011.
[58] A.A. Nielsen, “Kernel Maximum Autocorrelation Factor and Minimum Noise Fraction Transformations,” *IEEE Transactions on Image Processing*, Vol. 20, pp. 612-624, 2011.
[59] C.J. Oliver and S. Quegan, Understanding Synthetic Aperture Radar Images, Norwood, MA: Artech House, 1998.
[60] E.J.M. Rignot and J.J. van Zyl, “Change detection techniques for ERS-1 SAR data,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 31, pp. 896-906, 1993.
[61] J. Cihlar, T.J. Pultz, and A.L. Gray, “Change detection with synthetic aperture radar,” *International Journal of Remote Sensing*, Vol. 13, pp. 401-414, 1992.
[62] R.J. Dekker, “Speckle filtering in satellite SAR change detection imagery,” *International Journal of Remote Sensing*, Vol. 19, pp. 1133-1146, 1998.
[63] L. Gueguen, P. Soille, M. Pesaresi, “Change Detection Based on Information Measure,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 49, pp. 4503 – 4515, 2011.
[64] F. Chatelain, J.-Y. Tourneret, J. Inglada and A. Ferrari, “Bivariate Gamma Distributions for Image Registration and Change Detection,” *IEEE Transactions on Image Processing*, Vol. 16, pp. 1796-1806, 2007.
[65] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli and C. Zoppetti, “Nonparametric Change Detection in Multitemporal SAR Images Based on Mean-Shift Clustering,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 51, pp. 2022-2031, 2013.
[66] E. Erten, A. Reigber, L. Ferro-Famil and O. Hellwich, “A New Coherent Similarity Measure for Temporal Multichannel Scene Characterization,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 50, pp. 2839-2851, 2012.
[67] M. Meila, “Comparing Clustering by the Variation of Information,” in *Proceedings of the 16th Annual Conference of Computational Learning Theory*, Springer, 2003.
[68] L. Gueguen, M. Datcu, “Mixed Information Measure: Application to Change Detection in Earth Observation,” *Fifth International Workshop on the Analysis of Multi-Temporal Remote Sensing Images*, (MultiTemp 2009).
[69] M. Gong, Z. Zhou, J. Ma, “Change Detection in Synthetic Aperture Radar Images based on Image Fusion and Fuzzy Clustering,” *IEEE Transactions on Image Processing*, Vol. 21, pp. 2141 – 2151, 2011.
[70] W. Dierking, H. Skriver, “Change detection for thematic mapping by means of airborne multitemporal polarimetric SAR imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 40, pp. 618-636, 2002.
[71] S. Cloude and E. Pottier, “An entropy based classification scheme for land applications of polarimetric SAR,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 35, pp. 68–78, 1997.
[72] J.V. Zyl, H. Zebker, and C. Elachi, “Imaging radar polarization signatures: theory and observation,” *Radio Science*, Vol. 22, pp. 529–543, 1987.
[73] R. Touzi, S. Goze, T.L. Toan, A. Lopes, and E. Mougin, “Polarimetric discriminators for SAR images,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, pp. 973–980, 1992.
[74] M. Molinier, J. Laaksonen, and T. Häme, “Detecting man-made structures and changes in satellite imagery with a content-based information retrieval system built on self-organizing maps,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, pp. 861–874, 2007.
[75] P. Kersten, J. Lee, and T. Ainsworth, “A comparison of change detection statistics in POLSAR images,” in *Proceedings IEEE 2005 International Geoscience and Remote Sensing Symposium* (IGARSS’05), pp. 4836–4839, 2005.
[76] K. Conradsen, A.A. Nielsen, J. Schou, and H. Skriver, “A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, pp. 4–19, 2003.
[77] T. Fung and E. LeDrew, “The determination of optimal threshold levels for change detection using various accuracy indices,” *Photogrammetric Engineering and Remote Sensing*, Vol. 54, pp. 1449-1454, 1988.
[78] A.P. Dempster, N.M. Laird, and D.B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of Royal Statistics Society*, Vol. 39, pp. 1-38, 1977.
[79] L. Bruzzone and D. Fernández Prieto, “A minimum-cost thresholding technique for unsupervised change detection,” *International Journal of Remote Sensing*, Vol. 21, pp. 3539-3544, 2000.

[80] L. Bruzzone and R. Cossu, “Analysis of multitemporal remote-sensing images for change detection: bayesian thresholding approaches,” in Geospatial Pattern Recognition, Eds: E. Binaghi, P.A. Brivio and S.B. Serpico, Research Signpost/Transworld Research, Kerala, India, 2002, chap. 15.
[81] M. Hao, W. Shi, H. Zhang and C. Li, “Unsupervised Change Detection With Expectation-Maximization-Based Level Set,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 11, pp. 210-214, 2014
[82] Y. Bazi, L. Bruzzone, F. Melgani, “Automatic identification of the number and values of decision thresholds in the log-ratio image for change detection in SAR images,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 3, pp. 349-353, 2006.
[83] G. Moser and S.B. Serpico, “Generalized Minimum-Error Thresholding for Unsupervised Change Detection From SAR Amplitude Imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 44, pp. 2972-2982, 2006.
[84] G. Moser, S.B. Serpico, “Unsupervised Change Detection From Multichannel SAR Data by Markovian Data Fusion,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 44, pp. 2114-2128, 2009.
[85] L. Bruzzone, 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, 1997, pp. 3883-3888.
[86] N. Otsu, “A threshold selection method from gray-level histogram,” *IEEE Transactions on Systems, Man and Cybernetics*, 9, 62-66, 1979.
[87] J. Kittler and J. Illingworth, “Minimum error thresholding,” *Pattern Recognition*, 19, 41-47, 1986.
[88] L.K. Huang and M.J. Wang, “Image thresholding by minimizing the measures of fuzziness,” *Pattern Recognition*, 28, 41-51, 1995.
[89] F. Melgani, G. Moser, S.B. Serpico, “Unsupervised change detection methods for remote sensing images,” *Optical Engineering*, 41, 12, 3288-3297, 2002.
[90] F. Bovolo, L. Bruzzone, and M. Marconcini, “A Novel Approach to Unsupervised Change Detection Based on a Semisupervised SVM and a Similarity Measure,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, pp. 2070 – 2082, 2008.
[91] J. Muñoz-Marí, F. Bovolo, L. Gómez-Chova, L. Bruzzone and G. Camps-Valls, “Semisupervised One-Class Support Vector Machines for Classification of Remote Sensing Data,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, pp. 3188-3197, 2010.
[92] T. Celik, “Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-Means Clustering,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 6, pp. 772-776, 2009.
[93] M. Volpi, D. Tuia, Camps-Valls and M. Kanevski, “Unsupervised Change Detection With Kernels,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 9, pp. 1026-1030, 2012.
[94] A. Ghosh, B. Narayan Subudhi and L. Bruzzone, “Integration of Gibbs Markov Random Field and Hopfield-Type Neural Networks for Unsupervised Change Detection in Remotely Sensed Multitemporal Images,” *IEEE Transactions on Image Processing*, Vol. 22, pp. 3087-3096, 2013.
[95] P.L. Rosin and E. Ioannidis, “Evaluation of global image thresholding for change detection,” *Pattern Recognition Letters*, Vol. 24, pp. 2345-2356, 2003.
[96] S.U. Lee, S.Y. Chung, and R.H. Park, “A comparative performance study of several global thresholding techniques for segmentation,” *Computer Vision, Graph. Image Processing*, Vol. 52, pp.171-190, 1990.
[97] L. Bruzzone and D. Fernández Prieto, “An adaptive parcel-based technique for unsupervised change detection,” *International Journal on Remote-Sensing*, 21, 4, 817-822, 2000.
[98] S. Ghosh, L. Bruzzone, S. Patra, F. Bovolo, and A. Ghosh, A Context-Sensitive Technique for Unsupervised Change Detection Based on Hopfield-Type Neural Networks,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, pp. 778-789, 2007.
[99] F. Dell'Acqua, P. Gamba, G. Lisini: “Co-registration of multi-angle fine spatial resolution SAR images,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 1, n. 4, pp. 237-241, 2004.
[100] A. Della Ventura, A. Rampini, and R. Schettini: “Image Registration by recognition of corresponding structures,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, pp. 305-314, 1990.
[101] H. Li, B.S. Manjunath, and S.K. Mitra: “A Contour based approach to multisensor image registration,” *IEEE Transactions on Image Processing*, Vol. 4, pp. 320-334, 1995.
[102] F. Dell’Acqua, P. Gamba, G. Lisini: “Change Detection of Multi-Temporal SAR Data in Urban Areas Combining Feature-Based and Pixel-Based Techniques”, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 44, n.10, pp. 2820-2827, 2006.
[103] G.R. Byrne, P.F. Crapper, K.K. Mayo, “Monitoring Land Cover changes by Principal Components Analysis of Multitemporal Landsat Data,” *Remote Sensing of the Environment*, Vol. 10, pp. 175-184, 1980.
[104] J.A. Richards,”Thematic mapping from multitemporal image data using principal components transformation,” *Remote Sensing of the Environment*, Vol. 16, pp. 35-46, 1984.
[105] J.A Richards and A.K. Milne, “Mapping fire burns and vegetation regeneration using principal components analysis,” in *Proceedings IEEE* 1983 *International Geoscience and Remote Sensing Symposium* (IGARSS ’83).
[106] J.B. Collins and C.E. Woodcock, “Change detection using the Gram-Schmidt transformation applied to mapping forest mortality,” *Remote Sensing of the Environment*, Vol. 50, pp. 267-279, 1994.
[107] A.A. Nielsen and K. Conradsen, “Multivariate alteration detection (MAD) in multispectral, bitemporal image data: a new approach to change detection studies,” IMM, Technical University of Denmark, Lyngby, Technical Report 1997-11, 1997.
[108] A.A. Nielsen, K. Conradsen, and J.J. Simpson, “Multivariate Alteration Detection (MAD) and MAF Postprocessing in Multispectral, Bitemporal Image Data: New Approaches to Change Detection Studies,” *Remote Sensing of Environment*, Vol. 64, pp. 1-19, 1998.

[109] L. Bruzzone, D.F. Prieto, “Unsupervised change detection in multisource and multisensor remote sensing images,” in *Proceedings IEEE 2000 International Geoscience and Remote Sensing Symposium* (IGARSS ‘00), Vol. 6, pp. 2441-2443, 2000.
[110] B. Jeon and D. A. Landgrebe, "Classification with spatio-temporal interpixel class dependency contexts", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, pp. 663-672, 1992.
[111] A.H.S. Solberg, T. Taxt and A. K. Jain, "A Markov Random Field model for classification of multisource satellite imagery", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, pp. 100-113, 1996.
[112] L Bruzzone and D. Fernández Prieto, “Unsupervised retraining of a maximum-likelihood classifier for the analysis of multi-temporal remote-sensing images,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, pp. 456-460, 2001.
[113] L. Bruzzone and S.B. Serpico, “An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 35, pp. 858-867, 1997.
[114] L. Bruzzone, D. Fernández Prieto and S.B. Serpico, “A neural-statistical approach to multitemporal and multisource remote-sensing image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, pp. 1350-1359, 1999.
[115] M. Cebrian, M. Alfonseca, A. Ortega, “The Normalized Compression Distance Is Resistant to Noise,” *IEEE Transactions on Information Theory*, Vol. 53, pp. 1895 – 1900, 2007.
[116] B. Kosko, Neural Networks and Fuzzy Systems, Prentice Hall, 1992.
[117] V.N. Vapnik, The Nature of Statistical Learning Theory, Berlin: Springer-Verlag, 1995.
[118] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge University Press, 2000.
[119] P.J. Howarth and G.M. Wickware, “Procedure for change detection using Landsat,” *International Journal of Remote Sensing*, Vol. 2, pp. 277-291, 1981.
[120] F.G. Hall, D.B. Botkin, D.E. Strebel, K.D. Woods and S.J. Goetz, “Large scale patterns of forest succession as determined by remote sensing,” *Ecology*, Vol. 72, pp. 628-640, 1991.
[121] R.A. Schowengerdt, “Techniques for image processing and classification in remote sensing,” New York: Academic Press, 1983
[122] L. Bruzzone, R. Cossu and G. Vernazza, “Detection of land-cover transitions by combining multidate classifiers”, *Pattern Recognition Letters*, Vol. 25, pp. 1491-1500, 2004.
[123] D. Fernández Prieto and O. Arino, “A partially supervised change detection technique”, in *Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium* (IGARSS’01), Vol. 2, pp. 858-860, 2001.
[124] R. Cossu, S. Chaudhuri and L. Bruzzone, “A spatial-contextual partially supervised classifier based on Markov random fields”, *IEEE Geoscience and Remote Sensing Letters*, Vol. 2, pp. 352-356, 2005.
[125] T. Häme, I. Heiler, J.S. Miguel-Ayanz, “An unsupervised change detection and recognition system for forestry,” *International Journal of Remote Sensing*, Vol. 19, 6, p. 1079-1099, 1998.
[126] T. Macrì Pellizzeri, P. Lombardo P. Gamba, F. Dell’Acqua, “Multisource Urban Classification: Joint Processing of Optical and SAR Data for Land Cover Mapping,” in *Proceedings of the IEEE 2003 International Geoscience and Remote Sensing Symposium* (IGARSS’03), Vol. 2, pp. 1044–1046, 2003.
[127] P.C. Smits and A. Annoni, "Updating Land-Cover Maps by Using Texture Information from Very High-Resolution Space-Borne Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, pp. 1244-1254, 1999.
[128] F. Ling, W. Li, Y. Du and X. Li, “Land Cover Change Mapping at the Subpixel Scale With Different Spatial-Resolution Remotely Sensed Imagery,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 8, pp. 182-186, 2011.
[129] F. Pacifici, F. Del Frate, C. Solimini and W.J. Emery, “An Innovative Neural-Net Method to Detect Temporal Changes in High-Resolution Optical Satellite Imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, pp. 2940-2952, 2007.
[130] P. Serra, X. Ponsab, and D. Sauría, “Post-classification change detection with data from different sensors: Some accuracy considerations,” *International Journal of Remote Sensing*, Vol. 24, pp. 3311-3340, 2003.
[131] P. Kempeneers, F. Sedano, P. Strobl, D. O. McInerney and J. San-Miguel-Ayanz, “Increasing Robustness of Postclassification Change Detection Using Time Series of Land Cover Maps,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 50, pp. 3327-3338, 2012.
[132] M. Volpi, D. Tuia, F. Bovolo, M. Kanevski, L. Bruzzone, “Supervised Change Detection in VHR Images Using Contextual Information and Support Vector Machines,” *International Journal of Applied Earth Observation and Geoinformation*, Vol. 20, pp. 77-85, 2013.
[133] B. Demir, F. Bovolo, L. Bruzzone, “Classification of Time Series of Multispectral Images with Limited Training Data,” *IEEE Transactions on Image Processing*, Vol. 22, 2013, pp. 3219-3233.
[134] B. Demir, F. Bovolo and L. Bruzzone, “Detection of Land-Cover Transitions in Multitemporal Remote Sensing Images With Active-Learning-Based Compound Classification,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 50, pp. 1930-1941, 2012
[135] B. Demir, F. Bovolo and L. Bruzzone, “Updating Land-Cover Maps by Classification of Image Time Series: A Novel Change-Detection-Driven Transfer Learning Approach,” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 51, pp. 300-312, 2013.
[136] Proceedings First International Workshop on the Analysis of Multitemporal Remote-Sensing Images, Eds: L. Bruzzone and P.C. Smits, World Scientific Publishing Co.: Singapore, 2002, ISBN: 981-02-4955-1, Book Code: 4997 hc.
[137] Proceedings Second International Workshop on the Analysis of Multitemporal Remote-Sensing Images, Eds: P.C. Smits and L. Bruzzone, World Scientific Publishing Co.: Singapore, 2004, ISBN: 981-238-915-6, Book Code: 5582 hc.
[138] Proceedings Third International Workshop on the Analysis of Multitemporal Remote-Sensing Images, Eds: R.L. King and N. H. Younan, IEEE: USA, 2005, ISBN: 0-7803-9118-7.

[139] Proceedings Fourth International Workshop on the Analysis of Multitemporal Remote-Sensing Images, Eds: G.J.M. De Lannoy, R. Hoeben, N.E.C. Verhoest, W.W. Verstraeten, S. Bruneel and P. Coppin, IEEE: USA, 2007, ISBN: 1-4244-0846-6.
[140] Proceedings Fifth International Workshop on the Analysis of Multitemporal Remote-Sensing Images, Eds: D.L. Civco, Curran Associate, 2011, ISBN: 978-161-839-194-0.
[141] Proceedings Sixth International Workshop on the Analysis of Multitemporal Remote-Sensing Images, Eds: L. Bruzzone and F. Bovolo, IEEE, 2011, ISBN: 978-1-4577-1203-6.
[142] Proceedings Seventh International Workshop on the Analysis of Multitemporal Remote-Sensing Images, Eds: M. Hall-Beyer and G. McDermid, IEEE, 2013, ISBN: 978-1-4599-2413-4.
[143] Special issue on Analysis of Multitemporal Remote Sensing Images, *IEEE Transactions on Geoscience and Remote Sensing*, Guest Eds: L. Bruzzone, P.C. Smits, J.C. Tilton, Vol. 41, 2003.
[144] Special issue on Multi-temporal Imagery Analysis: Methods and Applications for Earth Observation, *IEEE Journal of Selected Topics in Applied Remote Sensing*, Guest Eds: R.S. Lunetta, D.L. Civco, R.L. King, L. Bruzzone, J.G. Lyon, Vol. 4, 2011.
[145] Special issue on Analysis of Multitemporal Remote Sensing Data, *IEEE Transactions on Geoscience and Remote Sensing*, Guest Eds: F. Bovolo, L. Bruzzone, R.L. King, Vol. 51, 2013.