An Unsupervised Technique Based on Morphological Filters for Change Detection in Very High Resolution Images

Mauro Dalla Mura, *Student Member, IEEE*, Jon Atli Benediktsson, *Fellow, IEEE*, Francesca Bovolo, *Member, IEEE*, and Lorenzo Bruzzone, *Senior Member, IEEE*

***Abstract*—An unsupervised technique for change detection (CD) in very high geometrical resolution images is proposed, which is based on the use of morphological filters. This technique integrates the nonlinear and adaptive properties of the morphological filters with a change vector analysis (CVA) procedure. Different morphological operators are analyzed and compared with respect to the CD problem. Alternating sequential filters by reconstruction proved to be the most effective, permitting the preservation of the geometrical information of the structures in the scene while filtering the homogeneous areas. Experimental results confirm the effectiveness of the proposed technique. It increases the accuracy of the CD process as compared with the standard CVA approach.**

***Index Terms*—Change detection (CD), morphological filters, remote sensing, very high geometrical resolution images.**

I. Introduction

HUMAN development and natural forces continuously alter landscapes. The analysis of these variations is necessary in many tasks such as monitoring land use, risk assessment, and the analysis of worldwide population growth and development. For this reason, change detection (CD) has an increasing importance in the field of remote sensing. The images acquired by periodical passes of remote sensing satellites over the same areas permit a regular analysis of the changes that occurred on the ground. The large amount of available satellite data has led the remote sensing community to focus its attention on unsupervised CD techniques, where ground-truth information is not necessary.

In this scenario, with the launch of a new generation of optical satellites, such as IKONOS, Quickbird, and Eros A1, very high resolution (VHR) images have been commercially available, and their diffusion will further increase with the future World View satellites. The VHR images are characterized by a submetric resolution; thus, the acquired scenes show many details (e.g., small trees, particulars of buildings, etc.) that were not observable by the previous-generation sensors. In addition, the high resolution in representing the surveyed scene makes the contextual information a predominant feature in the VHR images. In fact, unlike in low and medium spatial resolution images, the relations between adjacent pixels become a fundamental information source for the understanding of the scene. The high geometrical resolution and the contextual information are features that are particularly important in the urban scenes, opening new perspectives for CD applications.

A widely used unsupervised CD technique for medium-resolution images is Change Vector Analysis (CVA). CVA can be divided into three phases [1]: 1) preprocessing, where the multitemporal images are made comparable through coregistration, geometric correction, and radiometric calibration; 2) image comparison, where the spectral differences between the two images are represented by computing the spectral change vectors (SCVs); and 3) analysis of the results of the comparison, which aims at extracting the changed regions by generating a map where each pixel is associated to the class of changed ω_{c} or unchanged ω_{u} patterns.

In light of the properties of multitemporal VHR images (i.e., presence of a relevant amount of geometrical details, shadows, residual misregistration, and multiscale objects), unsupervised CD in these data is a complex task [2]. Most of the CD methodologies presented within the last 30 years [1], [3] for low or medium geometrical resolution imagery cannot handle the geometric and textural information present in VHR images. In particular, standard pixel-oriented techniques based on the thresholding of the magnitude of the SCVs result in CD maps showing a great number of false alarms (FAs) and artifacts [2].

According to the specific characteristics of VHR images, object-oriented approaches are more suitable to exploit the context relations. These approaches permit us to drive the analysis of the multitemporal images by the spatial information extracted from the objects in the scene. In [4], an object-oriented method was presented, which decomposes the images at different resolution levels and exploits the multiscale structure for the classification of VHR data. This method was extended to multilevel CD in [5] by defining a multiscale CVA technique. In greater detail, the decomposition of the image into different levels permits us to adaptively take into account the scales of the structures in the scene and to progressively reduce the complexity of the magnitude image. However, only relatively few CD techniques are available for VHR images, and further research is necessary on this topic.

A very promising approach to the analysis of VHR remote sensing images is based on morphological filters.

Manuscript received January 14, 2007; revised December 29, 2007. This work was supported in part by the University of Iceland Research Fund.

M. Dalla Mura is with the Department of Electrical and Computer Engineering, University of Iceland, 101 Reykjavik, Iceland, and also with the Department of Information Engineering and Computer Science, University of Trento, 38050 Trento, Italy.

J. A. Benediktsson is with the Department of Electrical and Computer Engineering, University of Iceland, 101 Reykjavik, Iceland.

F. Bovolo and L. Bruzzone are with the Department of Information Engineering and Computer Science, University of Trento, 38050 Trento, Italy.

Digital Object Identifier 10.1109/LGRS.2008.917726

Morphological filters are nonlinear operators defined in the mathematical morphology (MM) framework and widely applied to image processing problems. MM is based on the operators of erosion ε and dilation δ. The other morphological operators can be constructed by combining these fundamental operators [6]. In image processing, morphological filters are defined by a structuring element (SE) (which designs the shape and size of the filter) and a neighborhood transformation (which defines how the values of the pixels included in the SE are processed). The output of the morphological transformation shows how the image interacts with the size and shape of the SE. In this framework, morphological filters are formally defined as image transformations that are idempotent and increasing [6]. Since opening and closing satisfy these properties, they are morphological filters. The effect of an opening or a closing is basically to simplify the input image by erasing bright and dark objects (in the meaning of brighter and darker than surrounding regions), respectively, in the scene while preserving other structures in the image. Morphological filters are intrinsically object-oriented transformations because they focus the processing of the image on areas with shape and size defined by the SE. These properties and the nonlinearity of the morphological operators result to be very important for the analysis of VHR remote sensing images, particularly for reducing the noise components by preserving the geometrical features of the objects into the scene. For these reasons, the application of morphological filters to VHR images was investigated in the context of segmentation [7] and classification [8] problems with convincing results. However, despite its potential effectiveness, MM was not used in CD problems on the VHR images.

In this letter, we present a CD technique based on the integrated use of morphological filters and the CVA technique. In particular, we define a processing scheme that jointly exploits the SCV information and the capabilities of MM in properly separating the changes in multitemporal images from the sources of noise. In addition, we present an analysis aimed at choosing the morphological filter that is more suitable for the analysis of SCVs derived from multitemporal VHR images.

II. Proposed Change Detection Technique Based on Morphological Filters and CVA

A. Architecture of the Proposed CD Technique

Let us consider two coregistered multispectral VHR images with B spectral channels, i.e., $\mathbf{X}_1 = \{X_1^b, 1 \leq b \leq B\}$ and $\mathbf{X}_2 = \{X_2^b, 1 \leq b \leq B\}$, acquired over the same area at different times t_1 and t_2, respectively. Here, each spectral image X_k^b, with $k = 1, 2$, has a size of $I \times J$ pixels: $X_k^b = \{X_k^b(i,j), 1 \leq i \leq I, 1 \leq j \leq J\}$. Let $\Omega = \{\omega_c, \omega_u\}$ be the set of classes associated with changed and unchanged pixels. It is assumed that an adequate preprocessing phase has been applied to the multitemporal images in order to make them as more comparable as possible. In particular, geometric corrections, coregistration, and radiometric corrections should be applied to the data.

The proposed CD technique is made up of three steps: 1) computation of the SCVs; 2) morphological filtering; and 3) generation of the map of changes (Fig. 1). The comparison between the multitemporal multispectral images $\mathbf{X}_1$ and $\mathbf{X}_2$

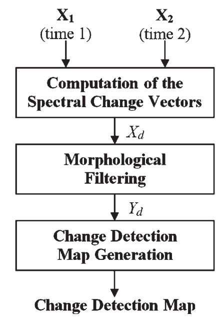


Fig. 1. General scheme of the proposed technique.

is obtained through the CVA technique. The magnitude of the SCVs (which is associated to an image called difference image X_d) contains information about the changes that occurred between the two acquisitions. As any grayscale image, the difference image can be thought as a topographic map, where the grayscale intensity of each pixel is associated to a measure of elevation. According to this interpretation, a higher intensity value (i.e., a brighter pixel in the image) corresponds to a higher elevation in the related topographic map. Hence, it is possible to characterize the structures present in the image in terms of convexity and concavity, where these properties are associated to areas of high and low intensities, respectively, w.r.t. the surroundings. The presence of convex regions (areas made up of pixels with high intensity) is due to a difference in the spectral signature of the correspondent regions on the ground (which can be associated to a change occurred between two acquisitions [1]). Different phenomena during the acquisitions (e.g., illumination of the scene, sensor view angle, etc.) and the presence of noise (mainly due to the residual misregistration between the multitemporal images) lead to the presence of convex portions in the magnitude image associated with unchanged areas. In general, the residual misregistration appears in the difference image as thin and elongated convex regions, which corresponds in the scene to the borders of well-defined structures (e.g., buildings, streets, rivers, etc.) or textured areas (e.g., the patterns on the rural fields) [2]. Differences in the scene illumination and sensor view angle result in shadows with different positions and shapes in the multitemporal images. These differences appear in the magnitude image as convex regions that are made up of pixels with intensity comparable with those of patterns that belong to changed areas. If no prior information about the shape and the size of the changed objects is available, it is not possible to distinguish a real change from these regions by analyzing only the magnitude of the SCVs. These effects are the main source of errors in the generation of the CD map.

In the proposed technique, after CVA, the complexity of X_d is reduced with the filtering phase, carried out by a morphological filter, which is applied to the magnitude of the SCVs. If the decision phase is based on the analysis of the image statistics computed on the magnitude of the SCVs, the filtering process should not alter the distribution of the data. Thus, the morphological filter should fulfill the property of self-duality, guaranteeing that convex and concave structures are equally processed. The transformation resulting from the application of a morphological filter depends on the shape and the size of the SE. Filtering with relatively small SEs results in the following: 1) removing structures that are brighter and darker than the surroundings; 2) reducing noise present into the image; and 3) flattening light textures on the object surfaces. In other words, the size of SE defines the grade of the simplification reached on the resulting image, determining how many details or small changes will be deleted from the image. Obviously, the larger the SE size, the greater the simplification and the coarser the resulting image. The shape of the SE is another important parameter. For CD problems without any *a priori* information about the shape of the changed areas, it is reasonable to choose an isotropic SE that processes the image without preferring any specific direction. In this scenario, the choice of an SE with a disk shape can be considered the most general because the pixels in the perimeter of the neighborhood have the same distance to the center. On the contrary, if information about the investigated changes is available, the shape and the size of the SE can be selected according to the geometrical characteristics of the objects in the scene. For example, in an urban area, a rectangular-shaped SE would better match the building shapes; in this scenario, the size of the SE could be chosen by taking into account the average dimension of the analyzed structures.

Finally, the CD map can be generated according to one of the unsupervised CD techniques proposed in the literature for the analysis of X_{d} [3]. Each pixel, in the final CD map, belongs to the class of changed ω_{c} or unchanged ω_{u} patterns.

B. *Morphological Filters for CD in VHR Images*

The morphological filters applied to the magnitude of the SCVs permit us to erase the convex and concave regions of a defined shape and size. This leads to a simplification of the difference image and a reduction of the noise components. Nevertheless, by considering the SCVs computed from two multispectral VHR images, the selection of a proper morphological filter is not a trivial task. The complexity of X_{d} needs to be reduced, preserving its geometrical information. If we consider standard morphological filters (e.g., morphological opening γ, which is the dilation of an eroded image, and morphological closing ϕ, which is the erosion of a dilated image), the simplification of the magnitude image would be reached with a partial loss of the geometrical information.

In order to properly exploit the VHR representation of the details in the scene, we propose the use of self-dual reconstruction filters (SDRFs) and alternating sequential filters (ASFs), which are both based on morphological operators by reconstruction. The morphological filters by reconstruction, which are defined by new advances in MM, are based on a non-Euclidean metric [9]. This family of filters is effective in applications where the geometrical information has to be preserved. These operators simplify the difference image by erasing the structures that interact with the SE, but they preserve the shape of those that are not canceled. Their use permits us to also avoid typical drawbacks of the morphological classical operators such as the shape noise (i.e., presence in the output image of patterns with the same shape as the SE used) and the shift of the object borders [6], which can compromise the analysis of VHR images.

SDRFs are operators based on the self-dual reconstruction. SDRF could be designed as a median filter, followed by a self-dual reconstruction using the original image as a mask (this operation retrieves the geometry of structures that are degraded but not erased by the median filter). The application of the median filter and the self-dual reconstruction guarantees the equal processing of concave and convex structures, satisfying the property of self-duality. According to Soille [6], the self-dual reconstruction of the median filter of the difference image X_{d} can mathematically be expressed as

$$Y_{\mathrm{d}}^{\mathrm{SDRF}} = R_{X_{\mathrm{d}}}\left[\zeta^{(i)}(X_{\mathrm{d}})\right] \tag{1}$$

where R is the self-dual reconstruction operator, and $\zeta^{(i)}$ is the median-filter operator with SE of size i (defining the SDRF size). By increasing the size of the median filter, a greater simplification of the image (enlarging the flat zones) is achieved.

ASFs, which are a sequential composition of opening and closing by reconstruction, are defined as follows:

$$Y_{\mathrm{d}}^{\mathrm{ASF}_m} = M_i = m_i, \ldots, m_1(X_{\mathrm{d}}), \quad \text{with} \quad m_i = \gamma_R^{(i)}\phi_R^{(i)} \tag{2}$$

$$Y_{\mathrm{d}}^{\mathrm{ASF}_n} = N_i = n_i, \ldots, n_1(X_{\mathrm{d}}), \qquad \text{with} \quad n_i = \phi_R^{(i)}\gamma_R^{(i)} \tag{3}$$

where m_i is the sequence of a closing by reconstruction followed by the dual opening with an SE of size i, whereas n_i is the combination of an opening followed by a closing. The opening and closing by reconstruction with size i of a general image F are based on the reconstruction by dilation R^{δ} and the reconstruction by erosion R^{ε}, respectively [6]

$$\gamma_R^{(i)}(F) = R_F^{\delta}\left[\varepsilon^{(i)}(F)\right] \quad \phi_R^{(i)}(F) = R_F^{\varepsilon}\left[\delta^{(i)}(F)\right]. \tag{4}$$

The reconstruction performed by using the original image as a mask leads to the iterative degradation of the image but restores partially its geometrical information. The processing driven by using the original image as a mask consists of progressively "flattening" the object surfaces while preserving their borders. The robustness of ASF against the noise is well known in the literature [10], where it has been used mostly in applications on SAR imagery (e.g., despeckling). Nonetheless, ASF has already been used on VHR imagery for classification. Chanussot *et al.* [11] have shown how this operator is well suited for a progressive simplification of the VHR image. Therefore, the idea of applying this operator to the CD task is very promising. The sequence of an opening and closing (or its dual) can be iterated several times, increasing the size of the SE at each iteration. By iterating the filter, larger regions are processed, involving a reduction of the image complexity. It is possible to filter the difference image through the sequences of open–close or close–open. These operators are dual with respect to the set complementation but not self-dual, i.e., they lead to different results when one is applied instead of the other. By choosing, for instance, the open–close sequence n, narrow bright structures would be suppressed by the starting opening. By duality, if closing is selected first, as for m, small darker structures would be removed first. Considering the VHR images, in the first iteration of the ASF, the SE has the smallest dimension (if we consider a disk, it would have the diameter of three pixels equivalent to about 2 m at nadir

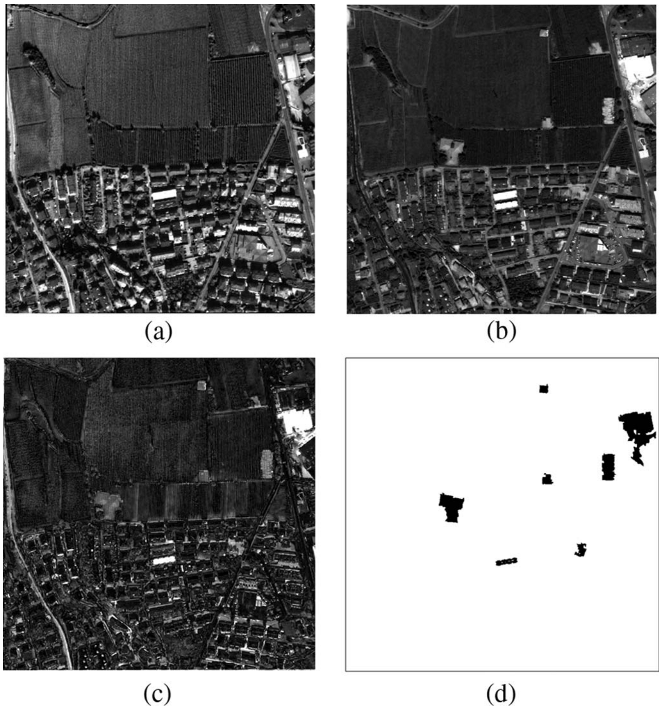


Fig. 2. Grayscale representation of the pan-sharpened multispectral and multitemporal images acquired in (a) October 2005 and (b) July 2006. (c) Magnitude of the SCVs computed from (a) and (b) (the image is visually enhanced). (d) Reference map.

on a pancromatic Quickbird image having a resolution of 0.61 m). The first iteration of the ASF (with the smallest filter size) filters the noise present into the image. After that, the choice of the open and close sequence is not really significant because the geometrical resolution is much higher than the minimum target size [11].

III. Experimental Results

The proposed method was evaluated on a data set of multitemporal and multispectral VHR images acquired by the Quickbird sensor on the Trentino area, Italy, in October 2005 and July 2006. In the preprocessing phase, the two images were pan-sharpened, generating a new set of multispectral images with the same geometrical resolution of the panchromatic band. The pan-sharpening was carried out by applying the Gram–Schmidt procedure implemented in the ENVI software package [12] to the panchromatic channel and the four bands of the multispectral image. Moreover, the multitemporal images were radiometrically corrected and coregistered. The registration process was carried out by using a polynomial function of the second order according to 14 ground control points (GCPs) and applying the nearest neighbor interpolation. The final data set was made up of two pan-sharpened multitemporal and multispectral images of 992×992 pixels with a spatial resolution of 0.61 m, which have a residual misregistration of about 1 pixel on GCPs (Fig. 2). Between the two acquisition dates, two kinds of changes occurred: 1) new houses were built in rural area, and 2) some roofs in industrial and urban areas were rebuilt. Different illumination of the scene (due to different acquisition seasons) and different acquisition angle during the imaging are the reasons for the presence of a great number of

TABLE I
Change Detection Errors (in Number of Pixels and Percentages) Obtained by Using the Proposed Technique

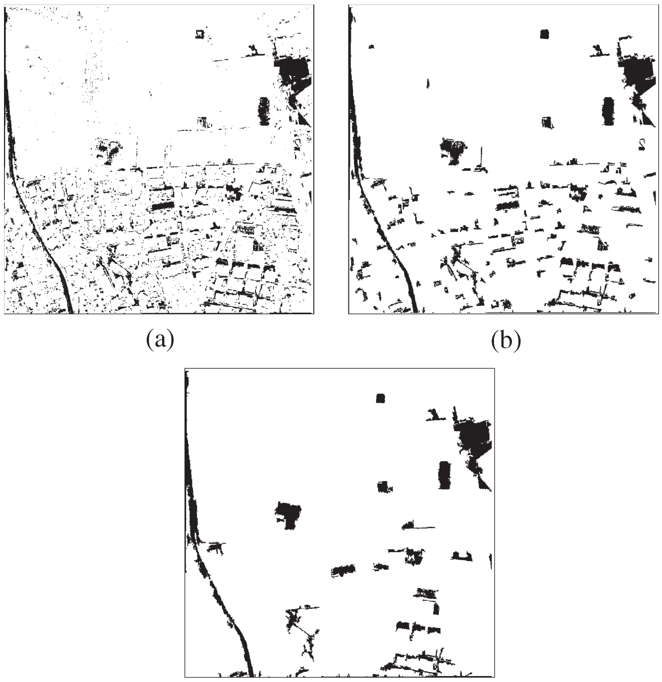
Method	Filter Size	Correct Detections pixels	Correct Detections %	False Alarms pixels	False Alarms %	Missed Alarms pixels	Missed Alarms %	Total Errors pixels	Total Errors %
Standard CVA		15396	73.60	64502	6.70	5522	26.40	70024	7.12
CVA with SDRF	9	17595	84.11	67682	7.03	3323	15.89	71005	7.22
	13	18692	89.36	69181	7.18	2226	10.64	71407	7.26
	17	18689	89.34	61510	6.39	2229	10.66	63739	6.48
	21	18680	89.30	53963	5.60	2238	10.70	56201	5.71
	25	18680	89.30	47867	4.97	2238	10.70	50105	5.09
CVA with ASF	9	19465	93.05	62605	6.50	1453	6.95	64058	6.51
	13	19682	94.09	47688	4.95	1236	5.91	48924	4.97
	17	19851	94.90	42898	4.45	1067	5.10	43965	4.47
	21	18779	89.77	16382	1.70	2139	10.23	18521	1.88
	25	17880	85.48	15679	1.63	3038	14.52	18717	1.90

not correspondent shadows in the two images. It is worth noting that the different illumination in the two multitemporal images modifies the spectral response of some unchanged areas.

In applying the proposed CD technique, the comparison between the two multitemporal images was carried out by computing the magnitude of the SCVs obtained by CVA after the data normalization. The difference image was processed with an SDRF and an ASF by reconstruction using, for both the operators, an SE with disk shape. In ASF, the close–open sequence m was chosen. The final map of changes was obtained by thresholding the filtered image with the automatic technique based on the Kittler–Illingworth method [13] under the Gaussian assumptions. However, other thresholding techniques can be used [14].

In order to allow a quantitative evaluation of the effectiveness of the presented method, the CD map generated by the proposed technique was compared with a reference map (which includes 20 918 changed pixels and 963 146 unchanged pixels) defined according to the available prior knowledge on the considered area. The results presented in Table I permit us to assess the effectiveness of the proposed technique with respect to the standard pixel-based CD procedure based on the CVA. In particular, when considering a disk SE of diameter 17 pixels, the number of missed alarms (MAs) sharply decreases to 15.74% and 21.30% for the SDRF and the ASF, respectively; in both cases, the FAs decrease to less than 3%.[1] Furthermore, Table I shows that by increasing the size of the filters, the number of FAs (mostly due to both the different acquisition conditions at the two dates and the residual registration noise) is progressively reduced. The ASF by reconstruction leads to a significant reduction of the number of FAs, but the MA rate is strongly dependent on the filter size. In greater detail, changed regions with the same size of the SE are removed (this implicitly defines a minimum bound on the size of detectable changed objects). If we compare the results obtained by the ASF with those yielded by the SDRF in similar conditions on the SE size, we can observe that with the SDRF, the complexity of the difference image is not reduced as much as applying the ASF, but the MA rate is less sensitive to the size of the filter. This was expected from the SDRF definition, which

[1]MAs and FAs are also referred in the literature as false negatives and false positives, respectively.



(a) (b) (c)

Fig. 3. CD maps obtained (a) by applying the standard pixel-based method and the proposed technique (b) with an SDRF and (c) with an ASF by reconstruction with the open–close sequence. Both filters were applied with a disk SE diameter of 17 pixels.

permits the preservation of the structures that are not removed by the median filter even if they are smaller than the SE.[2]

The qualitative analysis of the CD maps shown in Fig. 3 confirms the effectiveness of the proposed technique. The comparison of the maps generated by the standard CVA [Fig. 3(a)] and the proposed method [Fig. 3(b) and (c)] with the reference map [Fig. 2(d)] leads us to conclude that the use of morphological filters attenuates significantly the noise associated with FAs in the CD process. In particular, the effects of the residual registration noise are strongly reduced by the proposed approach. Moreover, the use of both considered morphological filter types leads to a better exploitation of the spatial correlation of the adjacent pixels in the images, thus increasing the detection of changed structures and reducing the residual noise.

IV. Conclusion

In this letter, a technique for CD in VHR images based on morphological filters has been proposed. The method is based on the integration of morphological filters by reconstruction (ASFs and SDRFs, which are specifically selected and tuned for the processing of VHR images) with the CVA technique. This technique exploits the strong nonlinearity characteristic of the morphological operators for filtering the VHR images while preserving their geometrical information and exploiting the contextual relations.

The aforementioned method was evaluated on a data set that is made up of two real multitemporal and multispectral VHR images. From the analysis of the obtained results, the proposed method confirms to be effective in detecting the changed areas in a more accurate and precise way with respect to the standard pixel-based CVA technique. Moreover, the use of morphological filters by reconstruction, particularly the ASFs, permits us to decrease the error rate by exploiting the high geometrical resolution of the data with a limited computational effort. In fact, the details of the changed structures are extracted by preserving their geometrical properties.

An important topic of our future research is the development of a strategy for choosing the best filter size in order to fully automatize the CD process.

References

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

[2] F. Bovolo, L. Bruzzone, and S. Marchesi, "A multiscale technique for reducing registration noise in change detection on multitemporal VHR images," in *Proc. IEEE 4th Int. Workshop MultiTemp*, 2007, pp. 1–6.

[3] P. Coppin, I. Jonckheere, K. Nackaerts, and B. Muys, "Digital change detection methods in ecosystem monitoring: A review," *Int. J. Remote Sens.*, vol. 25, no. 9, pp. 1565–1596, May 2004.

[4] L. Bruzzone and L. Carlin, "A multilevel context-based system for classification of very high spatial resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2587–2600, Sep. 2006.

[5] F. Bovolo and L. Bruzzone, "A multilevel parcel-based approach to change detection in very high resolution multitemporal images," in *Proc. IGARSS*, 2005, vol. 3, pp. 2145–2148.

[6] P. Soille, *Morphological Image Analysis, Principles and Applications*, 2nd ed. Berlin, Germany: Springer-Verlag, 2003.

[7] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.

[8] J. A. Benediktsson, M. Pesaresi, and K. Arnason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.

[9] J. Crespo, J. Serra, and R. Schafer, "Theoretical aspects of morphological filters by reconstruction," *Signal Process.*, vol. 47, no. 2, pp. 201–225, Nov. 1995.

[10] P. Soille and M. Pesaresi, "Advances in mathematical morphology applied to geoscience and remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 2042–2055, Sep. 2002.

[11] J. Chanussot, J. A. Benediktsson, and M. Pesaresi, "On the use of morphological alternated sequential filters for the classification of remote sensing images from urban areas," in *Proc. IGARSS*, 2003, vol. 1, pp. 473–475.

[12] *Envi User Manual*, RSI, Boulder, CO, 2003. [Online]. Available: http://www.RSInc.com/envi

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

[14] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1171–1182, May 2000.

[2] As a general guideline, the maximum value of the filter size should be not higher than the expected minimum size of changed areas.