A Multilevel Parcel-Based Approach to Change Detection in Very High Resolution Multitemporal Images

Francesca Bovolo, Member, IEEE

Abstract-This letter presents a novel parcel-based contextsensitive technique for unsupervised change detection in very high geometrical resolution images. In order to improve pixel-based change-detection performance, we propose to exploit the spatialcontext information in the framework of a multilevel approach. The proposed technique models the scene (and hence changes) at different resolution levels defining multitemporal and multilevel "parcels" (i.e., small homogeneous regions shared by both original images). Change detection is achieved by applying a multilevel change vector analysis to each pixel of the considered images. This technique properly analyzes the multilevel and multitemporal parcel-based context information of the considered spatial position. The adaptive nature of multitemporal parcels and their multilevel representation allow one a proper modeling of complex objects in the investigated scene as well as borders and details of the changed areas. Experimental results confirm the effectiveness of the proposed approach.

Index Terms—Change detection, multilevel image representation, multitemporal image analysis, very high resolution (VHR) images.

I. INTRODUCTION

► HANGE detection is one of the most important applications of the automatic analysis of multitemporal remote sensing images. In the literature, many techniques have been proposed for change detection in medium-resolution multispectral data [1], [2]. Most of them produce reliable change-detection maps by analyzing pixel-by-pixel two images acquired on the same area at different times. However, the availability of very high resolution (VHR) data (e.g., images acquired by QuickBird, IKONOS, SPOT-5, and WorldView satellites) results in a new set of possible applications, which require the generation of change-detection maps characterized by both a high geometrical precision and the capability of properly modeling the complex objects/areas present in the scene. Classical change-detection techniques become ineffective on VHR images, as they often assume spatial independence among pixels, which is not reasonable in VHR data, where the spatial autocorrelation cannot be approximated as a pulse function.

The author is with the Department of Engineering and Computer Science, University of Trento, 38050 Trento, Italy (e-mail: francesca.bovolo@disi. unitn.it).

Digital Object Identifier 10.1109/LGRS.2008.2007429

For this reason, it is necessary to develop advanced contextsensitive change-detection methods that are capable to properly exploit the specific properties of VHR images. In this context, although a great effort has been devoted to the development of automatic classification techniques, less attention has been devoted to change-detection techniques. In the literature, some unsupervised approaches have been presented based on the comparison between features computed on homogeneous regions that are obtained according to segmentation procedures [4], [5]. Even if these context-sensitive methods are implicitly suitable to analyze VHR images, they model images (and thus changes) at a single scale (resolution) level, while VHR images include information at several scales. This information should be properly exploited to obtain reliable change-detection maps. Some attempts to model the change/no-change information at different scales can be found in [6] and [7]. Nonetheless, the method in [6] does not explicitly consider the spatio-contextual information, while the one in [7] models it by means of a fixed-size neighboring system which is not always appropriate, particularly in border regions. A first tentative to jointly exploit the spatio-contextual and the multiresolution information for change detection is given in [8].

In order to address the aforementioned limitations, in this letter, we present a novel unsupervised approach to change detection in VHR multitemporal images that exploits a multilevel context-based image modeling of multitemporal acquisitions aimed at representing each changed area at the proper resolution level. The scene (and hence changes) is modeled at different resolution levels by defining multilevel and multitemporal "parcels." Multitemporal parcels represent the local adaptive neighborhood of pixels and have the property to be homogeneous on both considered multitemporal images. Change detection is achieved by applying to multitemporal images a specific comparison algorithm based on a multilevel extension of the change vector analysis (CVA). Comparison is performed involving feature vectors which properly represent the multilevel and multitemporal context information of the investigated spatial position. The adaptive nature of parcels and their multiresolution representation allow one to adequately model complex objects in the investigated scene as well as borders and details of the changed areas. Obtained changedetection maps show high accuracy in both homogeneous and border areas as well as high fidelity in modeling the shapes of changes on the ground, outperforming standard pixel-based and single-level object-oriented methods.

1545-598X/\$25.00 © 2008 IEEE

Manuscript received April 17, 2008; revised August 12, 2008 and September 30, 2008. First published November 17, 2008; current version published January 14, 2009.



Fig. 1. General scheme of the proposed approach.

II. PROPOSED MULTILEVEL PARCEL-BASED **CHANGE-DETECTION TECHNIQUE**

Let us consider two coregistered VHR multitemporal and multispectral images, namely, X_1 and X_2 (of size $I \times J$) composed by B spectral bands, acquired over the same area at different times t_1 and t_2 , respectively. Let $\Omega = \{\omega_c, \omega_n\}$ be the set of classes associated with changed and unchanged pixels, respectively. The proposed approach is made up of the following three main steps (see Fig. 1): 1) generation of the hierarchical multitemporal parcel-based model of the spatial context of each pixel; 2) multilevel parcel-based feature extraction; and 3) generation of the final change-detection map according to a context-sensitive multilevel CVA technique.

A. Hierarchical Multitemporal Parcel-Based Modeling of the Spatial Context

The rationale of the first phase is to adaptively generate a model of the spatiotemporal context of each pixel according to a multilevel strategy. Each resolution level is defined according to predefined spectral, spatial, and temporal constraints. This kind of representation allows one to capture and exploit the entire information present in the considered data set by working with adaptive multitemporal neighborhood systems (i.e., parcels) at different levels. At each scale, pixels are characterized by adaptive multitemporal parcels of different size and shape that satisfy homogeneity constraints in the spatial and temporal domains at the same time [5].

In order to characterize the spatiotemporal context of each pixel taking into account a context-based hierarchical multiscale representation, we propose a two-step procedure based on the following: 1) independent hierarchical segmentation of multitemporal images from the pixel level to higher levels of representation of their spatial context and 2) multitemporal fusion of the two segmentations obtained from step 1 [5].

The multilevel context-sensitive image characterization is obtained according to the following formal definition. Let $H^{l}(.)$ be the homogeneity measure in the spatial domain at a generic resolution level l(l = 0, ..., L - 1), and let $P^{l}(\mathbf{X})$ be a segmentation map of an image X obtained by satisfying H^{l} . The segmentation map at a generic level l is a partition $P^{l}(\mathbf{X})$ in a set of N^{l} disjoint regions o_{s}^{l} , with $o_{s}^{l} \subseteq \mathbf{X}$ and $s=1,2,\ldots,N^l$ (i.e., $P^l(\mathbf{X})=\{o_1^l,\ldots,o_{N^l}^l\}$), and such that, $\forall l = 0, \dots, L-1$, the following conditions hold.

- 1) Pixels in o_s^l are spatially contiguous.
- 2) $\bigcup_{s=1}^{N^l} o_s^l = \mathbf{X}.$ 3) $o_s^l \cap o_r^l = \emptyset$ for each pair of adjacent regions o_s^l , o_r^l , $s, r = 1, 2, \dots, N^l$, and $r \neq s.$
- 4) $H^{l}(o_{s}^{l}) =$ true.
- 5) $H^l(o_s^l \cup o_r^l) =$ false for each pair of adjacent regions.
- 6) $\bigcup_{q=1}^{N_k^{l-1}} o_q^{l-1} = o_k^l$, where N_k^{l-1} is the number of objects at level l-1 that compose the *k*th region at level *l*.

Conditions 1)-3) guarantee that, at each resolution level, all pixels in **X** are distributed into N^l spatially adjacent but nonoverlapping regions. Condition 4) determines that homogeneity is satisfied, and condition 5) expresses the maximality of each region (i.e., at each level, l pairs of adjacent regions cannot be merged without relaxing the homogeneity criterion). The constraint in 6) guarantees a precise hierarchy between the context of a pixel defined at different levels. Satisfying 6) means that the region associated to a pixel at level l-1 cannot be included in more than one segment at level l. In other words, a region related to a specific pixel at a given level is completely included in the region of the same pixel at a higher level (i.e., a level with lower resolution).

The aforementioned conditions can be satisfied according to different segmentation strategies (e.g., region growing, watershed, etc.). In this letter, we adopt a bottom-up region-merging segmentation algorithm that operates according to the following iterative procedure [10], [11].

- 1) Define an initial partition (level zero) of the considered image (a common choice is to define one region for each pixel in the scene).
- 2) For each pair of adjacent segments o_s^l and o_r^l (s, r =1,2,..., N^l), evaluate $H^{l+1}(o_s^l \cup o_r^l)$. 3) If $H^{l+1}(o_s^l \cup o_r^l) < H^{l+1}_{\text{TH}}$, merge o_s^l and o_r^l ; otherwise,
- do not.
- 4) Iterate steps 2) and 3) until no more mergence is possible, i.e., until no adjacent pairs of regions exist that satisfy the homogeneity criterion. The process stops when the smallest growth exceeds a threshold defined by the user (the so-called scale parameter that governs H_{TH}^{l+1}).

The use of the described region-growing procedure results in a segmentation map that satisfies at each resolution level the maximality property as the final regions are the largest possible given $H_{\rm TH}^{l+1}$. The higher the value of $H_{\rm TH}^{l+1}$, the higher the dimension of resulting segments (in other words, as H_{TH}^{l+1} increases, the sensibility of the homogeneity measure in the fusion of adjacent regions decreases). Varying the scale parameter, and thus H_{TH}^{l+1} , allows one to obtain the desired hierarchical segmentation.¹ It is worth noting that the basic criteria of the aforementioned segmentation strategy are also implemented in commercial software packages [11].

The choice of the homogeneity measure depends on the specific segmentation algorithm adopted. In this letter, a homogeneity measure has been used that combines spectral and

¹It is worth noting that how to satisfy the hierarchical constraint can become a complex task, depending on the considered segmentation strategy.

spatial properties [11]. Due to space constraints, we refer the reader to [10] and [11] for more details on the definition of the adopted homogeneity measure.

In order to compute the desired hierarchical representation of the spatiotemporal context of the considered scene, the aforementioned hierarchical segmentation technique is applied to images X_1 and X_2 independently using the same homogeneity measure and varying in the same way the scale parameter.² It is worth noting that in order to obtain a reliable representation of the local neighborhood of pixels, oversegmented maps are preferable than undersegmented ones. This also allows one to control the impact of segmentation errors on the change-detection map. Hierarchical segmentation results in two sets of L segmentation maps $\mathbf{P}_1 = \{P^l(\mathbf{X}_1)\}$ and $\mathbf{P}_{2} = \{P^{l}(\mathbf{X}_{2})\} \ (l = 0, \dots, L-1).$ Applying a logical OR to corresponding segmentation levels in \mathbf{P}_1 and \mathbf{P}_2 , a set of hierarchical parcel maps $\mathbf{P}_{1,2} = \{P^l(\mathbf{X}_1, \mathbf{X}_2)\}$ shared by both considered images is obtained. At each resolution level, parcels satisfy both spatial and temporal homogeneities. This approach intrinsically guarantees that each parcel map in $P_{1,2}$ satisfies the hierarchical property as it is satisfied by each single segmentation map in \mathbf{P}_1 and \mathbf{P}_2 .

B. Multilevel Parcel-Based Feature Extraction

The second step of the proposed approach aims at characterizing each pixel in X_1 and X_2 by exploiting the information included in $P_{1,2}$. In particular, the mean vectors (whose components are associated with the different spectral channels of the considered images) computed on the pixels included in parcels at a given resolution level *l* are extracted for each acquisition date. Additional components, like texture measures for characterizing the spatial context of each pixel in the different parcels, can also be extracted. Thus, for each pixel, we define a multidimensional vector $M_h(i, j)$ (h = 1, 2) (which describes the pixels and, through the hierarchical tree, the spatial and temporal contexts (parcels) in which the pixel is included) as

$$\boldsymbol{M}_{h}(i,j) = \left\{ \boldsymbol{x}_{h}(i,j), \boldsymbol{\mu}_{h}^{1}(i,j), \dots, \\ \boldsymbol{\mu}_{h}^{L-1}(i,j), \boldsymbol{t}\boldsymbol{x}_{h}^{1}(i,j), \dots, \boldsymbol{t}\boldsymbol{x}_{h}^{L-1}(i,j) \right\}$$
(1)

where $\boldsymbol{x}_{\boldsymbol{h}}(i,j)$ is the spectral feature vector of the generic pixel at spatial position (i,j), and $\boldsymbol{\mu}_{h}^{l}(i,j)$ and $\boldsymbol{t}\boldsymbol{x}_{h}^{l}(i,j)$ are the mean vector and a generic texture measure, respectively, of the parcel associated with the pixel (i,j) at level l of the hierarchy. Different texture measures can be considered (e.g., standard deviation, entropy, etc.), with the constraint that they should be significant when computed on parcels that may have different size and shape. In order to obtain an effective representation of the change information, it is important to avoid increasing too much the number of components of $\boldsymbol{M}_{\boldsymbol{h}}(i,j)$. This can be achieved by selecting nonredundant and significant (with respect to the investigated change) spectral bands and computing a relatively small number of relevant features. It is worth noting that the selection of significant spectral bands is a common practice in change-detection applications [2].

²Here, the eCognition software package was used for performing hierarchical segmentation [11].

The choice of the number of levels L depends on the geometrical resolution of the considered image and on the maximum scale of the expected changes in the scene. In particular, L should be equal to the segmentation level which completely includes the largest expected changed object in the scene. The reader is referred to [10] for an empirical rule that allows one to identify the number of significant resolution levels. It is worth noting that the choice of L within a reasonable range does not significantly affect the change-detection results.

C. Multilevel CVA

The feature vectors $M_h(i, j)$ that characterize each pixel and the related context information at different levels are given as input to the comparison and change-detection map generation module. The comparison step is carried out extending the wellknown CVA technique [2], [3] to the multilevel hierarchical case. The multilevel CVA is carried out by directly comparing the multilevel context-sensitive feature vectors $M_1(i,j)$ and $M_2(i, j)$ associated with corresponding pixels on the two images with a difference vector operator.³ The rationale is that we can derive a change vector that contains information of changes at different resolution levels, depending on both the differences between the spectral signatures of the analyzed pixel and the different behavior of its spatial neighborhood adaptively modeled according to the hierarchical parcel-based representation. The obtained difference vectors contain information of changes at different resolution levels. From these vectors, we can derive a multilevel difference image M_D associated with the magnitude of change vectors. The value $M_D(i, j)$ of the generic pixel at spatial position (i, j) in the multilevel difference image is given by

$$M_D(i,j) = |M_2(i,j) - M_1(i,j)|.$$
 (2)

Finally, the hard information associated to changes (i.e., a change-detection map) can be produced by thresholding M_D according to either automatic or manual strategies [2], [5].

III. EXPERIMENTAL RESULTS

In order to assess the effectiveness of the proposed approach, we considered a real data set made up of two images acquired by the QuickBird satellite on the city of Trento (Italy) in October 2005 and July 2006. In the preprocessing phase, the two images were treated in the following ways: 1) pan-sharpened by applying the Gram–Schmidt procedure implemented in the ENVI software package [12]; 2) radiometrically normalized (by adjusting the mean value of the two images); 3) coregistered. The final data set was made up of two pan-sharpened multitemporal and multispectral images of 1090 × 740 pixels with a spatial resolution of 0.7 m and a residual misregistration of less than 1 pixel on selected ground control points. Fig. 2 shows the panchromatic channel of the two images X_1 and X_2 .

The goal of the experimental analysis is to detect radiometric changes associated with building and road yards. Thus, only

³According to the kind of features involved in the comparison step, normalization could be necessary to make the comparison operation reliable.





(b)

Fig. 2. Panchromatic channel of images of the city of Trento (Italy) acquired by the QuickBird VHR multispectral sensor in (a) October 2005 and (b) July 2006 (black lines depict changes that occurred between the two acquisition dates).

bands 3 and 4 were considered, as they showed to be the most effective in detecting these kinds of changes for this data set.

In order to permit a quantitative evaluation of the effectiveness of the proposed method, a reference test set was defined according to the available prior knowledge on the considered area. The reference map includes 616 050 unchanged pixels and 44 900 changed pixels [see the black polygons in Fig. 2(b)].⁴ Some 145650 pixels associated to shadow areas in at least one of the two acquisitions that can result in a high number of false alarms were not considered in the change-detection process (such pixels are not highlighted in Fig. 2 to avoid confusion). In our experiments, we considered different resolution levels obtained using the multiscale hierarchical segmentation procedure described in Section II-A with different values of the scale parameter [11]. For each level other than pixel one, the parcel-based mean was extracted for building the multilevel feature vectors M_1 and M_2 (some trials with texture features pointed out that, on this data set, texture measures do not provide significant information). The proposed technique was compared with a standard pixel-based CVA technique [2], [3] and a context-sensitive single-level CVA procedure applied at each parcel level independently [5]. In order to assess the poten-

⁴It is worth noting that although some minor natural changes are visible in Fig. 2, they were not included in the class of changed pixels. This is because the selected combination of spectral bands is less sensitive to them, and they are not of interest for the considered application.

 TABLE I

 Change-Detection Results (in Number of Pixels) Obtained by

 Applying Pixel-Based and Single-Level Parcel-Based

 Techniques and Mean Parcel Area (in Number of Pixels)

Level	False Alarms	Missed Alarms	Overall Error	Mean Parcel Area
Pixel	32811	22389	55200	-
1	23373	19663	43036	99,55
2	23957	6497	30454	242,14
3	19524	10805	30329	398,32
4	16414	7701	24115	604,64

TABLE II
CHANGE-DETECTION RESULTS (IN NUMBER OF PIXELS) OBTAINED BY
USING THE PROPOSED MULTILEVEL PARCEL-BASED TECHNIQUE

Considered Levels	False Alarms	Missed Alarms	Overall Error
Pixel,1	23564	16757	40321
Pixel,1,2	21231	7944	29175
Pixel,1,2,3	16906	6549	23455
Pixel, 1, 2, 3, 4	10112	5598	15710

tialities of the proposed method without any bias introduced by the threshold selection procedure, thresholding was performed according to a manual trial-and-error process that minimizes the overall errors on the reference areas.

Table II reports the numerical results obtained for four different combinations of scales built using five resolution levels (from pixel one to resolution level 4). The first combination includes only the following two resolution levels: the pixel level and level 1. The other combinations are obtained by progressively increasing the number of scales from two to five (see Table II), adding gradually lower resolution levels. By comparing Tables I and II, one can conclude that, in all cases, the proposed technique resulted in higher changedetection accuracy than the pixel-based approach. Moreover, once the lowest resolution level L considered in the hierarchical representation is fixed, the presented multilevel procedure always performs better than a parcel-based single-level CVA. The proposed technique obtains the smallest overall error (i.e., 15710 pixels) when five resolution levels (from the pixel level to level 4) are considered. Adding one more low resolution level resulted in a decrease of performances.

A better understanding of the effectiveness of the proposed technique can be achieved by a visual comparison of changedetection maps. Let us focus on a subimage of 300×300 pixels, including the building yard in the upper left corner of the data set. The change-detection map achieved with the proposed technique resulted in the following: 1) a sharply better representation of homogeneous areas with respect to the map obtained at the pixel level (i.e., standard CVA) (e.g., compare areas in continuous gray circles of Fig. 3(a) and (c) with the corresponding area in the reference map [Fig. 3(d)]) and 2) a more accurate modeling of details with respect to the best map yielded by a single-scale parcel-based approach (i.e., the one obtained at resolution level 4) [e.g., compare areas in dashed gray circles in Fig. 3(b) and (c) with the corresponding area in Fig. 3(d)].

It is worth noting that many other experiments were carried out with different numbers and combinations of resolution levels (e.g., neglecting high resolution levels from the



Fig. 3. Details of the change-detection maps obtained with (a) the standard pixel-based strategy, (b) the single-level parcel-based strategy applied at level 4, and (c) the proposed technique involving levels from pixel one to level 4. (d) Reference map (changed pixels appear in black, while unchanged pixels appear in white).

multiscale hierarchical representation). The results of these experiments (not reported for space constraints) pointed out change-detection accuracies smaller than those obtained with the combinations considered in Table II.

IV. DISCUSSION AND CONCLUSION

In this letter, a novel approach to change detection in VHR multitemporal images has been presented. The approach is based on a representation of the spatiotemporal context of pixels and is ruled by precise hierarchical relationships between each pixel in the multitemporal images and the regions that adaptively define the related context at different levels. Each pixel is characterized by a feature vector, which is obtained by computing the mean on the parcels at different scales for both images. The final change-detection map is computed by extending the well-known CVA procedure to the multilevel parcel-based feature vectors.

Quantitative and qualitative results on the real VHR data set pointed out the effectiveness of the proposed technique. In particular, the proposed technique resulted in a sharply higher accuracy than the standard pixel-based CVA technique, which resulted in an unreliable change-detection map for the considered data set. In addition, once the number L of considered segmentation levels was fixed, the proposed multilevel parcel-based method always obtained higher change-detection accuracy than the single-level parcel-based procedure. The qualitative analysis of the change-detection maps pointed out that the proposed technique also resulted in a high fidelity in both homogeneous and border regions. In greater detail, the proposed multilevel approach resulted in change-detection maps that show both a more accurate modeling of geometrical details of changes than the single-level parcel-based method applied to low resolution level, and a better representation of homogeneous area than the single-level parcel-based methods applied at high resolution level.

The most effective number and combination of resolution levels depend on the considered data set and the desired tradeoff between the accuracy in detail preservation and homogeneous area representation. It is worth noting that neglecting high resolution levels will result in a loss of sensibility to borders and geometrical details, whereas neglecting low resolution levels will result in a poor accuracy in homogeneous areas.

The multilevel nature of the proposed technique allows one to also reduce false alarms that may be introduced from artifacts deriving from the pan-sharpening procedure. In this respect, after a given level, it is also possible to compute the contextual features by directly considering the low-resolution multispectral images, which are not affected from geometric and radiometric distortions.

As future developments of this letter, we plan to do the following: 1) Extend the experimental analysis to VHR multitemporal images acquired by other sensors; 2) introduce in the proposed approach a procedure for the adaptive selection of the number of scales and the kind of features to be considered for an optimal representation of the object in which the analyzed pixel is included; and 3) extend the use of the proposed method to the analysis of long series of multitemporal images.

ACKNOWLEDGMENT

The author would like to thank Prof. L. Bruzzone of the University of Trento, Trento, Italy, for his valuable pieces of advice and useful discussion and comments about this work.

REFERENCES

- R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.
- [2] 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.
- [3] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007.
- [4] G. G. Hazel, "Object-level change detection in spectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 553–561, Mar. 2001.
- [5] L. Bruzzone and D. F. Prieto, "An adaptive parcel-based technique for unsupervised change detection," *Int. J. Remote Sens.*, vol. 21, no. 4, pp. 817–822, Mar. 2000.
- [6] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2963–2972, Dec. 2005.
- [7] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432–1445, May 2007.
- [8] F. Bovolo and L. Bruzzone, "A multilevel parcel-based approach to change detection in very high resolution multitemporal images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2005, vol. 3, pp. 2145–2148.
- [9] K. S. Fu and J. K. Mui, "A survey on image segmentation," *Pattern Recognit.*, vol. 13, no. 1, pp. 3–16, 1981.
- [10] 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.
- [11] eCognition Professional User Guide 4, Definiens Imaging, München, Germany, 2003. [Online]. Available: http://www.definiens-imaging.com
- [12] ENVI User Manual, RSI, Boulder, CO, 2003. [Online]. Available: http:// www.RSInc.com/envi