Title: An Approach to Multiple Change Detection in VHR Optical Images Based on Iterative Clustering and Adaptive Thresholding

Author(s): Yady Tatiana Solano-Correa, Francesca Bovolo and Lorenzo Bruzzone

# An Approach to Multiple Change Detection in VHR Optical Images Based on Iterative Clustering and Adaptive Thresholding

Yady Tatiana Solano-Correa, *Member, IEEE*, Francesca Bovolo, *Senior Member, IEEE,* and Lorenzo Bruzzone, *Fellow, IEEE*

*Abstract*—One of the most common approaches to unsupervised change detection in multispectral images is Change Vector Analysis (CVA). CVA computes the multispectral difference image and exploits its statistical distribution in (hyper-)spherical coordinates by means of two steps: i) magnitude and ii) direction thresholding. The two steps require assumptions on: i) the model of class distributions: and ii) the number of changes. However, both assumptions are seldom satisfied or difficult to formulate, especially when considering VHR images. Thus, we propose an approach to multiple change detection in VHR optical images based on iterative clustering and adaptive thresholding in (hyper-) spherical coordinate. The proposed approach: i) is distribution free, ii) is unsupervised, iii) automatically identifies the number of changes; and iv) is robust to noise. Results obtained on two multitemporal single-sensor and multisensor datasets, including images from WorldView-2 and QuickBird, corroborate the effectiveness of the proposed approach.

*Index Terms*— Very High Resolution images, Change detection, Multitemporal, Clustering, Adaptive Thresholding.

## I. INTRODUCTION

WHEN dealing with VHR optical images, one of the most common approaches to multiple Change Detection (CD) is the computation of the multispectral difference image by means of Change Vector Analysis (CVA) [1]–[3]. CVA uses two or more spectral channels to model both the magnitude and the direction of spectral change vectors. To separate changed from unchanged samples, sub-optimal solutions are often employed that identify decision thresholds along the magnitude and direction variables independently [1]–[3]. A threshold over the magnitude variable provides separation between changed and unchanged samples. Whereas multiple thresholds along the direction variable(s) provide separation among kinds of change. However, such thresholds are complex to define. The accuracy of these sub-optimal methods depends on the a priori knowledge of both class statistical models and expected number of changes in order to

define the thresholds [1]–[3]. Moreover, when the number of dimensions increases, sub-optimal solutions become less effective as they do not fully exploit the correlation among variables. Two main limitations arise from the above-mentioned assumptions. The first one is given by the definition of the threshold value along the magnitude, which is defined by considering all kinds of change as a single large metaclass [2], [4]–[6]. This implies that this threshold is defined according to the average properties of the kinds of change and does not gather the peculiarities of each of them. The second one is given by the definition of thresholds along the direction variable(s), which are usually defined by assuming a prior knowledge on the number of classes and on the statistical distribution model [1]–[3].

In order to overcome the two main limitations, methods based on adaptive thresholding or clustering have been designed [3], [6]–[11]. According to the literature, and under the assumption of independence between magnitude and direction variable(s), two main steps are usually applied [2], [5]. The first step aims at distinguishing changed from unchanged pixels by defining a single threshold value along the magnitude variable. The second one separates kinds of change from each other by defining threshold values along the direction(s). Both steps can be addressed by using thresholding techniques available in the literature [3], but specific solutions are required depending on the type of data (e.g., optical, SAR) and application (e.g., bi-temporal or multitemporal analysis, short or long term analysis). Sometimes [6], a third step is considered that adapts the threshold along the magnitude to the characteristics of different kinds of change identified in step 2. Nevertheless, the problem remains about how to properly detect the number of changes and how to separate them without prior information on their number and statistical distribution.

To detect and separate the kinds of change in an automatic way, methods based on clustering can be considered [4], [7], [8], [12]. There are mainly two kinds of clustering algorithms [8]: i) partitioning; and ii) hierarchical methods. The former relies on the centroid-based model and partition the data into $k$ clusters (e.g., k-means). Such methods often exploit a predefined statistical model for the classes and assume the knowledge of the number of clusters. Thus, they show similar drawbacks to the sub-optimal thresholding solutions. Whereas the latter (e.g., Ordering Points to Identify the Clustering Structure (OPTICS) [13]) build a cluster hierarchy based on a

Y. T. Solano-Correa and F. Bovolo are with the Fondazione Bruno Kessler, Center for Information and Communication Technology, 38123 Trento, Italy (e-mail: solano@fbk.eu, bovolo@fbk.eu)

L. Bruzzone is with the Department of Information Engineering and Computer Science Trento, University of Trento, I-38123 Trento, Italy (e-mail: bruzzone@ing.unitn.it).

tree structure. Tree construction follows: i) bottom-up; or ii) top-down paradigms. Hierarchical clustering identifies clusters with different size and densities and is sensitive to outliers.

One of the most well-known hierarchical clustering is the OPTICS [13], and was proposed in the literature in order to overcome two specific drawbacks: i) to allow to search for clusters with different densities among the same set of points; and ii) to be able to deal with sparse data. OPTICS performs well when there are large differences in cluster densities and therefore is able to deal with sparse data. Accordingly, the use of OPTICS for solving the problem of multiple CD in VHR images in the (hyper-) spherical domain is promising. However, the overlapping of change clusters reduces OPTICS performance and increases computational time. This leads to the need of defining novel efficient methods in which threshold values (in the magnitude domain) and change sectors (in the direction domain) are adaptively defined.

In this letter, we propose an approach for automatic multiple CD that jointly exploits the histogram distribution of the magnitude and direction variables in (hyper-) spherical coordinates. The approach is distribution free (in the direction domain) and thus particularly suitable for VHR images. It first performs a rough discrimination between unchanged and changed samples according to the properties of the magnitude variable. Then, it automatically identifies the sectors along direction variables associated to different kinds of change by iteratively performing density-based clustering and region growing [8], [12], while adding samples from around the joint magnitude-directions histogram peaks. Finally, it tunes the magnitude threshold by adapting it according to the characteristics of kinds of change. CD is achieved in an unsupervised and application-independent way.

The rest of this letter is organized as follows. Section II describes the proposed method giving details on each of its three steps. Section III presents the datasets and experimental results. Finally, Section IV draws the conclusions and provides future developments.

## II. PROPOSED METHOD FOR MULTIPLE CD BY ITERATIVE CLUSTERING AND ADAPTIVE THRESHOLDING

Let us consider two VHR optical images, $X_1$ and $X_2$, acquired over the same geographical area by sensors $S_1$ and $S_2$ at time $t_1$ and $t_2$, respectively. $S_1$ and $S_2$ can be the same or different sensors (e.g., GeoEye, QuickBird, WorldView). Let $X_{1,a}$ and $X_{2,b}$ be the images associated to spectral bands $a$ ($a = 1,2, ..., A$) and $b$ ($b = 1,2, ..., B$) for $S_1$ and $S_2$. $a$ and $b$ may be associated to the same wavelength or not depending on the sensor properties. Let $I_1 \times J_1$ and $I_2 \times J_2$ be the size of $X_1$ and $X_2$, respectively. Let $X_D$ be the multispectral difference image obtained by subtracting on a pixel base the spectral feature vectors. Let $\omega_n$ be the class of unchanged pixels and $\Omega_c = \{\omega_{c_1}, \omega_{c_2}, ..., \omega_{c_K}\}$ the metaclass of $K$ possible kinds of change. Thus, $\Omega = \omega_n \cup \Omega_c$ is the set of classes to be identified by the proposed approach.

The approach assumes, $X_1$ and $X_2$ to be radiometrically and geometrically corrected to mitigate differences that are not related to changes on the ground [1]. This step is relevant for any change detection approach and particularly critical when dealing with multisensor images [1], [3], [12], [14]. To

properly model and automatically detect multiple changes in VHR an adequate feature extraction is required [1]. CVA feature space is used to represent/highlight the presence/absence of changes. For sake of simplicity and visualization purposes, CVA is often applied to a 2D space [2], [5], [10]. Nevertheless, employing a higher number of dimensions increases the change representation capabilities, but also the complexity due to density and sparsity of clusters. Here we define a 3D feature space, thus $X_D$ ca be represented in spherical coordinates by eq. (1)-(3) [1].

$$\rho = \sqrt{X_{D,1}^2 + X_{D,2}^2 + X_{D,3}^2} \tag{1}$$

$$\theta = arctan\left(\frac{X_{D,2}}{X_{D,1}}\right) \tag{2}$$

$$\varphi = arcos\left(\frac{X_{D,3}}{\rho}\right) \tag{3}$$

where $\rho$ represents the magnitude variable, and $\theta$ and $\varphi$ the direction variables (the number of direction variables increases as the space dimensionality does). Once the (hyper-) spherical space is defined, the proposed approach considers 3 steps (see Fig. 1). Step 1 separates unchanged from changed samples by means of the magnitude variable (binary CD). Step 2 iteratively identifies multiple changes by means of a hierarchical algorithm applied to the direction variables (iterative multiple CD). And step 3 adaptively adjusts the magnitude threshold values by considering each change class separately (adaptive magnitude thresholding).
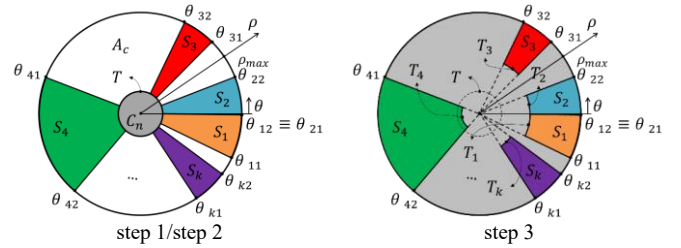
### A. Step 1: Binary CD



Fig. 1 Illustration of the 3-step procedure for the detection of multiple changes (zenithal view of the azimuth and magnitude variables of CVA) [6].

Step 1 separates unchanged from changed samples. In the literature, it has been demonstrated that in a (hyper-) spherical domain unchanged pixels tend to cluster around the origin of the coordinate system, i.e., they show a small magnitude $\rho \leq T$, where $T$ is a decision threshold separating changed from unchanged samples [2], [5]. $T$ can be calculated by means of the Bayesian decision theory by estimating class statistical parameters using the Expectation-Maximization (EM) algorithm [15]. This can be done by following a Gaussian ([15]), Rayleigh-Rice ([5]) or compound multiclass statistical model ([16]). In step 1, $\rho$ statistical distribution is exploited to separate the feature space of multispectral change vectors in: 1) the circle of unchanged samples ($C_n$) and 2) the annulus of changed samples ($A_c$) (Fig. 1, step 1).

### B. Step 2: Iterative Multiple CD

In step 2, we focus on $A_c$, where: i) changed pixels are located far from the origin, ii) kinds of change may show stronger or weaker magnitude ($\rho$) and specific preferred direction ($\theta, \varphi$); and iii) pixels affected by the same kind of

change tend to cluster [2]. Thus, for each kind of change a cluster exists with a peak surrounded by a spread and irregular cloud of points. Change clusters may overlap. To distinguish among change clusters, an iterative clustering is applied.

The proposed method jointly exploits the capabilities of OPTICS and the well-known region growing algorithm [8], [12] in an iterative configuration, such as that drawbacks are addressed. In the specific case of OPTICS, main problems are related to the overlap of different clusters in a cloud of points. For the region growing, it requires the definition of seed points (from the user) to determine whether a neighbor pixel belongs to a given region. Once the seeds are defined, it starts to add samples from surrounding pixels in the cloud of points, which results in problems with overlapping clusters. Here our set of cloud points is defined by the $\rho, \theta, \varphi$ histogram resulting from the CVA (hyper-) spherical representation of the feature space (see Fig. 2 (a)). In our iterative clustering, instead of applying the OPTICS/region growing directly to the whole cloud of points (as in the literature), we exploit the different frequency levels ($n$ - see Fig. 2 (b)) of the magnitude-direction variables. This configuration allows us to reduce the computational time, by starting from a reduced number of samples (green dots/line in Fig. 2), and adding them as we move to lower frequency levels (pink, blue and purple dots/lines in Fig. 2).
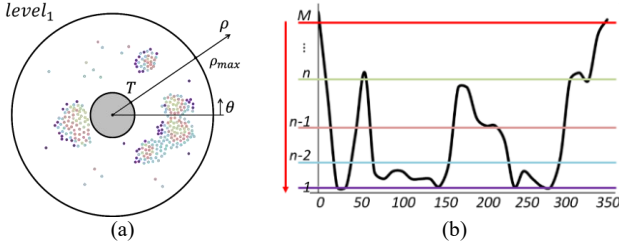


Fig. 2. Illustration of (a) the cloud points histogram in a CVA coordinate system (zenithal view of the azimuth and magnitude variables) and (b) its corresponding azimuth histogram.

Let $M$ be the maximum number of samples in a bin of the histogram in the (hyper-) spherical domain. First, iterative clustering finds a set of bins showing a number of samples higher than $M$-$n$ ($n = 1, ..., M$). In this way at least one cloud of samples can be defined which is dense and isolated enough to be considered as a cluster. Whether the cluster is dense enough or not is decided by the initial conditions of the density-based clustering algorithm. Let us assume that these conditions are reached at the level $n$ of the histogram (where $0 < n < M$), and that $k$=3 ($k \geq 1$) clusters are differentiated. Each of the clusters is expected to represent a different kind of change. The most populated bins are used to apply the OPTICS density-based algorithm (Fig. 3 (a)). OPTICS requires some initial parameters: i) the minimum number of points (*MinPts*) to form a cluster, and ii) the reachability (*Eps*), i.e., the maximum distance (radius) to which a point can be considered of the same cluster. Once the first $k$ clusters are found, the algorithm evolves to the next frequency level (*level n-1*), i.e., to histogram bins with less number of samples (see Fig. 2 (b)). Thus, new samples are added (black dots in Fig. 3 (b)). Each of the existing $k$ clusters is grown by means of the region growing [8], [12] (Fig. 3 (c)). To apply the region growing algorithm, the number of cycles and the neighborhood connectivity need to be defined. Then the OPTICS algorithm is applied (with the same parameters) to

search for new possible clusters (Fig. 3 (d)). The same process is iterated until the algorithm reaches the *level* 1 and all the histogram bins are processed. At the end, the optimal clusters are found, each of them corresponding to one specific kind of change, together with outliers. $S_k$ sectors are defined according to the region associated to the $k$th cluster (Fig. 1, step 2). Fig. 3 (a)-(d) illustrates the working process of the proposed step 2 by considering a 2D zenithal view of the spherical space. Given the nature of the proposed iterative multiple CD step, we name it as Growing OPTICS (GO).
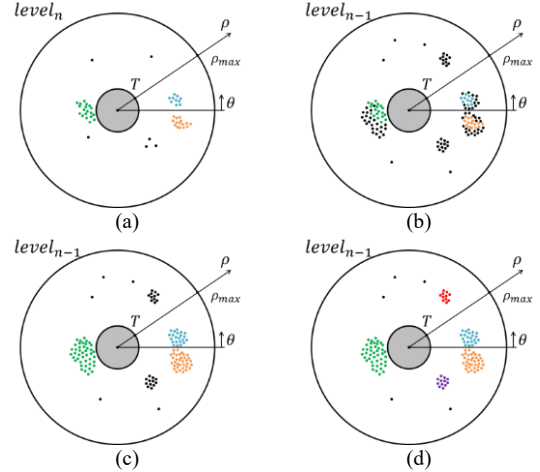


Fig. 3. Proposed step-2: zenithal view of the $\theta$ and $\varphi$ variables.

### C. Step 3: Iterative Multiple CD

In step 1 a single global threshold value $T$ along $\rho$ is computed that exploits the properties of the meta-class $\Omega_c$. However, each $\omega_{c_k} \in \Omega_c$ may have a different overlapping grade with the distribution of $\omega_n$. Thus, the optimal threshold $T_k$ along the magnitude variable that separates a given kind of change $\omega_{c_k}$ from $\omega_n$ might be different from the global threshold $T$ defined in step 1. Thresholds $T_k$ ($k = 1, ..., K$) can be different from each other. To properly exploit the peculiarities of each $\omega_{c_k}$, the last step tunes the threshold value along $\rho$ by considering only the pixels in each sector $S_k$, $k = 1, ..., K$. Therefore, $K$ binary problems are defined that differ from the problem described in step 1 because of the different balance between change and no-change class prior probabilities and conditional distributions. Each optimal threshold value $T_k$ can be computed by applying the same methods as for step 1. The residual gray light area in Fig. 1, step 3 is associated to unchanged pixels, whereas the sectors $S_1, S_2, S_3, S_4, ..., S_k$ are associated to changed areas.

### III. EXPERIMENTAL RESULTS

#### A. Dataset description and design of experiments

To validate the proposed approach, two datasets made up of both single-sensor and multisensor VHR optical images were selected. The study areas are located in the Trentino region, north Italy. For dataset 1, images were acquired by the QuickBird (QB) and the WorldView-2 (WV-2) satellites in 2006 and 2010, respectively. Whereas for dataset 2, images were acquired by the WV-2 satellite in 2010 and 2011. The images were pre-processed by absolute radiometric and geometric corrections, as per [1]. This step is critical and may

affect CD performance if not accurate. However experimental analysis showed that the proposed method is robust to pre-processing accuracy. Pre-processed images show 0.6m spatial resolution and 640×640 pixels size for dataset 1 (Fig. 4, left column) and 0.5m spatial resolution and 1230×1230 pixels size for dataset 2 (Fig. 4, right column).

The selected areas of interest contain mainly agricultural fields in the case of dataset 1, plus a new road for dataset 2 (Fig. 4). In order to properly highlight these kind of changes, Tasseled-Cap (TC) features were extracted [1]. CVA is applied to the brightness, wetness and greenness features and the information is represented in spherical coordinates [1]. A reference map of each study area was built by photointerpretation, and 5 and 4 kinds of change were found for datasets 1 and 2, respectively. For comparison we use OPTICS clustering since it does not require a priori knowledge of the number of clusters and is distribution free, whereas most of the clustering methods do (therefore they provide poor performance when used for multi-class change detection with VHR multitemporal images). This choice also allows to explicitly prove how GO improves OPTICS performance.
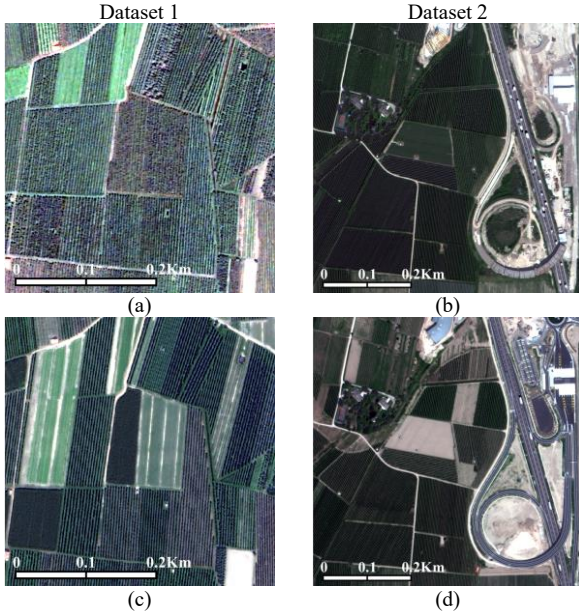

Dataset 1 — Dataset 2

Fig. 4. Left column, true color composition of dataset 1: (a) QB 2006 and (c) WV-2 2010 images; and Right column dataset 2: (b) WV-2 2010 and (d) WV-2 2011 images.

## B. Experimental Results

Once pre-processing and feature selection have been carried out, we proceed with the multiple CD. In step 1 the threshold $T$ is calculated by means of [5]. Results are reported in Table I for both datasets. Step 2 is applied only after masking unchanged samples. For both datasets, the minimum number of clusters $k$ for the initial step of GO was set to 2. *MinPts* was set to 20 and the number of iterations of the region-growing algorithm was set to 5 with a neighborhood of connectivity equal to 26. The selection of these parameters is critical for the performance of OPTICS and proposed approach, but they are easy to fix by means of quick error and trial tests. In this set up GO was able to detect all the changes in 11.9s for dataset 1 and 26.8s for datasets 2, using MATLAB® on a standard

workstation. Hardware is Intel(R) Core(TM) i7-3630QM CPU @2.40 GHz, 16.00GB RAM. Whereas OPTICS required 140s for dataset 1 and 225s for dataset 2. Step 3 identifies $T_k$ values (Table I). As expected $T_k$ threshold values are different among them and from $T$.

TABLE I. MAGNITUDE THRESHOLD VALUES FOR DATASETS 1 AND 2.

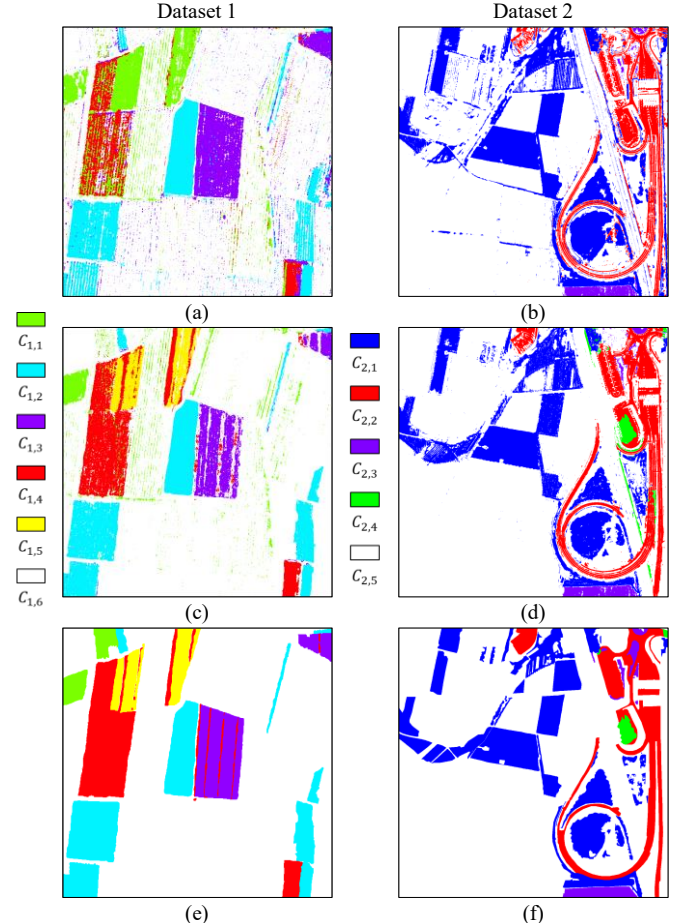| Dataset | $T$ | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_5$ |
|---|---|---|---|---|---|---|
| 1 | 0.030 | 0.027 | 0.035 | 0.047 | 0.032 | 0.026 |
| 2 | 0.070 | 0.070 | 0.068 | 0.045 | 0.073 | - |



Fig. 5. CD maps obtained by applying the 3-step configuration with: (a) and (b) OPTICS; (c) and (d) proposed approach; and (e) and (f) Reference Maps. Left column dataset 1 and right column dataset 2.

The CD maps obtained are shown in Fig. 5. It is possible to see that the proposed approach was able to properly detect all the kinds of change for both datasets. A visual comparison of the two CD maps with the reference map (Fig. 5 (e) and (f)), points out the improvement achieved while working with the GO algorithm. In dataset 1, GO detects $C_{1,1}$ and $C_{1,4}$ properly, whereas OPTICS tends to confuse those classes with other changes, resulting in several false alarms (FA). In both experiments, one part of $C_{1,1}$ is misdetected, maybe because of TC features. GO is further able to detect and separate class $C_{1,5}$ from $C_{1,1}$ (which correspond to overlapped clusters in the spherical domain), whereas OPTICS confuses it with class $C_{1,1}$. Concerning $C_{1,2}$ and $C_{1,3}$ the proposed approach results in a slightly better detection (Table II). In dataset 2, improvements are mainly achieved in the separation of $C_{2,2}$ and $C_{2,4}$, as well as in a better detection of unchanged samples

with a reduction of FA (Table III). The proposed approach outperforms OPTICS of about 7% and 4% for datasets 1 and 2, respectively, without using any prior knowledge about the number of changes.

TABLE II. CONFUSION MATRIX: (A) OPTICS AND (B) GO (DATASET 1).

(A)

| Predicted | Actual | | | | | | |
|---|---|---|---|---|---|---|---|
| | $C_{1,1}$ | $C_{1,2}$ | $C_{1,3}$ | $C_{1,4}$ | $C_{1,5}$ | $C_{1,6}$ | Rel. (%) |
| $C_{1,1}$ | **3243** | 8 | 3 | 10249 | 11976 | 8143 | 9.65 |
| $C_{1,2}$ | 24 | **41087** | 291 | 0 | 12 | 4023 | 90.43 |
| $C_{1,3}$ | 102 | 1011 | **21300** | 4782 | 2 | 5445 | 65.25 |
| $C_{1,4}$ | 39 | 3 | 196 | **18069** | 41 | 1569 | 90.72 |
| $C_{1,5}$ | 0 | 0 | 0 | 0 | **0** | 0 | 0.00 |
| $C_{1,6}$ | 3331 | 6112 | 1234 | 4119 | 1438 | **261748** | 94.16 |
| Acc. (%) | 48.12 | 85.21 | 92.51 | 48.55 | 0.00 | 93.17 | |
| **Overall Accuracy** | | | | 84.34% | | | |

(B)

| Predicted | Actual | | | | | | |
|---|---|---|---|---|---|---|---|
| | $C_{1,1}$ | $C_{1,2}$ | $C_{1,3}$ | $C_{1,4}$ | $C_{1,5}$ | $C_{1,6}$ | Rel. (%) |
| $C_{1,1}$ | **3893** | 23 | 3 | 995 | 60 | 8754 | 28.36 |
| $C_{1,2}$ | 24 | **44671** | 83 | 0 | 10 | 12 | 99.71 |
| $C_{1,3}$ | 5 | 358 | **16992** | 480 | 0 | 11 | 95.21 |
| $C_{1,4}$ | 41 | 6 | 658 | **29885** | 1034 | 1480 | 90.28 |
| $C_{1,5}$ | 0 | 0 | 0 | 781 | **11242** | 660 | 88.64 |
| $C_{1,6}$ | 2776 | 3163 | 5288 | 5078 | 1123 | **270011** | 93.94 |
| Acc. (%) | 57.77 | 92.64 | 73.80 | 80.30 | 83.47 | 96.11 | |
| **Overall Accuracy** | | | | 91.97% | | | |

TABLE III. CONFUSION MATRIX: (A) OPTICS AND (B) GO (DATASET 2).

(A)

| Predicted | Actual | | | | | |
|---|---|---|---|---|---|---|
| | $C_{2,1}$ | $C_{2,2}$ | $C_{2,3}$ | $C_{2,4}$ | $C_{2,5}$ | Rel. (%) |
| $C_{2,1}$ | **213918** | 2156 | 54 | 630 | 80413 | 71.98 |
| $C_{2,2}$ | 1665 | **90079** | 1007 | 6302 | 27361 | 71.26 |
| $C_{2,3}$ | 41 | 232 | **16506** | 0 | 783 | 93.99 |
| $C_{2,4}$ | 0 | 0 | 0 | **0** | 0 | 0.00 |
| $C_{2,5}$ | 32662 | 38665 | 6755 | 1059 | **992593** | 92.62 |
| Acc. (%) | 86.16 | 68.68 | 67.86 | 0.00 | 90.14 | |
| **Overall Accuracy** | | | 86.79% | | | |

(B)

| Predicted | Actual | | | | | |
|---|---|---|---|---|---|---|
| | $C_{2,1}$ | $C_{2,2}$ | $C_{2,3}$ | $C_{2,4}$ | $C_{2,5}$ | Rel. (%) |
| $C_{2,1}$ | **215985** | 138 | 1 | 73 | 41828 | 83.71 |
| $C_{2,2}$ | 441 | **108541** | 2602 | 251 | 27257 | 78.04 |
| $C_{2,3}$ | 40 | 517 | **17224** | 1 | 1595 | 88.89 |
| $C_{2,4}$ | 951 | 2030 | 0 | **6912** | 5041 | 46.28 |
| $C_{2,5}$ | 30869 | 19926 | 4495 | 754 | **1025428** | 94.82 |
| Acc. (%) | 86.99 | 82.76 | 70.82 | 86.50 | 93.12 | |
| **Overall Accuracy** | | | 90.82% | | | |

## IV. CONCLUSION

In this letter, an approach to multiple CD in VHR optical images based on iterative clustering and adaptive thresholding has been proposed. The adaptive GO works in (hyper-) spherical coordinates, is distribution free (in the direction domain), unsupervised and automatically identifies the number of changes. The effectiveness of the adaptive GO, with respect to OPTICS thresholding method, has been validated in two pairs of VHR single-sensor and multi-sensor images. Such datasets showed complex and noisy set-ups. GO demonstrated to be robust in handling noisy samples and overlapping clusters and required a reduced computational time when compared to SoA. The adaptive thresholding improved the overall accuracy thanks to the correct separation of overlapping clusters. Open issues regarding the selection of initial clustering algorithm parameters remain (for both OPTICS and region growing). Though the parameters are easy to setup (correlated to data spatial resolution and density of cloud points), they are not fixed in an automatic way. Future works will consider the applicability and adaptation of the method for CD in images with lower spatial resolution as well as possible extensions to other applications (e.g., classification).

## REFERENCES

[1] Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "An Approach for Unsupervised Change Detection in Multitemporal VHR Images Acquired by Different Multispectral Sensors," *Remote Sens.*, vol. 10, no. 4, p. 533, Mar. 2018.

[2] 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.

[3] 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.

[4] Z. Yetgin, "Unsupervised Change Detection of Satellite Images Using Local Gradual Descent," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1919–1929, May 2012.

[5] M. Zanetti, F. Bovolo, and L. Bruzzone, "Rayleigh-Rice Mixture Parameter Estimation via EM Algorithm for Change Detection in Multispectral Images," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5004–5016, Dec. 2015.

[6] F. Bovolo and L. Bruzzone, "An adaptive thresholding approach to multiple-change detection in multispectral images," in *2011 IEEE Int. Geoscience and Remote Sensing Symposium*, 2011, pp. 233–236.

[7] Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "An approach to multiple Change Detection in multisensor VHR optical images based on iterative clustering," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 5149–5152.

[8] L. Rokach and O. Maimon, "Clustering Methods," in *Data Mining and Knowledge Discovery Handbook*, Springer, Boston, MA, 2005, pp. 321–352.

[9] D. Y. Huang, T. W. Lin, and W. C. Hu, "Automatic multilevel thresholding based on two-stage Otsu's method with cluster determination by valley estimation," in *Int. Journal of Innovative Computing, Information and Control*, 2011, vol. 7, pp. 5631–5644.

[10] K. Ding, C. Huo, Y. Xu, Z. Zhong, and C. Pan, "Sparse Hierarchical Clustering for VHR Image Change Detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 577–581, Mar. 2015.

[11] T. Leichtle, C. Geiß, M. Wurm, T. Lakes, and H. Taubenböck, "Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment," *Int. J. Appl. Earth Obs. Geoinformation*, vol. 54, no. Suppl. C, pp. 15–27, Feb. 2017.

[12] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Pearson Education, 2009.

[13] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "OPTICS: Ordering Points to Identify the Clustering Structure," in *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*, New York, NY, USA, 1999, pp. 49–60.

[14] G. Hong and Y. Zhang, "A comparative study on radiometric normalization using high resolution satellite images," *Int. J. Remote Sens.*, vol. 29, no. 2, pp. 425–438, Jan. 2008.

[15] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 47–60, Nov. 1996.

[16] M. Zanetti and L. Bruzzone, "A Theoretical Framework for Change Detection Based on a Compound Multiclass Statistical Model of the Difference Image," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 1129–1143, Feb. 2018.