

Automatic Change Analysis in Satellite Images Using Binary Descriptors and Lloyd-Max Quantization

Anamaria Radoi, *Student Member, IEEE*, and Mihai Datcu, *Fellow, IEEE*

Abstract—In this letter, we present a novel technique for unsupervised change analysis, that leads to a method of ranking the changes that occur between two satellite images acquired at different moments of time. The proposed change analysis is based on binary descriptors and uses the Hamming distance as a similarity metric. In order to render a completely unsupervised solution, the obtained distances are further classified using vector quantization methods (i.e., Lloyd’s Algorithm for optimal quantization). The ultimate goal in the change analysis chain is to build Change Intensity Maps (CIM) that provide an overview of the severeness of changes in the area under analysis. In addition, the proposed analysis technique can be easily adapted for change detection by selecting only two levels for quantization. This discriminative method (i.e., between changed/unchanged zones) is compared with other previously developed techniques, that use Principal Component Analysis, or Bayes theory as starting points for their analysis. The experiments are carried on Landsat images at 30 meters spatial resolution, covering an area of approximately $59 \times 51 \text{ km}^2$ over the surroundings of Bucharest, Romania, and containing multispectral information.

Index Terms—Change analysis, binary descriptors, Hamming distance, Lloyd-Max quantization, multitemporal satellite images.

I. INTRODUCTION

THE past few years have witnessed an increased interest towards unsupervised change detection techniques. Unlike supervised methods, the latter techniques directly compare two multitemporal images, without having any other type of information regarding the contained classes or *a priori* distribution of the change/unchanged states. Under these assumptions, a challenging step is to find the optimal threshold to discriminate between change and no change.

The majority of the change detection methods are developed based on the analysis of the difference image. Other techniques use Change Vector Analysis to represent the change in the

n -dimensional spectral space and to show that each class of change has a distinct spectral signature [1].

In [2], two techniques based on the Bayes theory are proposed for the analysis of the difference image. The first approach assumes that the pixels in the difference image are independent of one another, and computes the optimal threshold by minimizing the error probability, which, in Bayesian terms, translates into maximizing the posterior conditional probability. The second technique uses the Besag’s iterated conditional modes for solving a MRF¹-based model, which considers that a pixel belonging to a certain class (change/no change) is likely to be surrounded by pixels belonging to the same class. Although the spatial contextual hypothesis is likely to be true in most cases, the problem of convergence becomes problematic when changes appear in too many scattered places. However, the exploitation of contextual information with MRF models is not well suited for on-line applications due to their high computational complexity. Both techniques are based on the estimation of statistical parameters using the Expectation - Maximization (EM) algorithm.

Other approaches consist of linear transformations [3]. For example, Principal Component Analysis (PCA) is applied in [4] over non-overlapping blocks of the difference image in order to extract the main directions of change. In the same paper, the challenging problem of searching the threshold between change and no change is solved by introducing a K-Means step ($K = 2$).

In this paper, we propose a multi-level change detection method to assess the degree of change suffered by an area between two moments of time. The system provides a fast and unsupervised diagnosis of change, that does not require either the availability of any type of *a priori* information or any learning phase. To this end, we start our analysis from two images acquired at different moments of time over the same area. For each pixel inside each image, we build a binary descriptor representing the gradients in the neighborhood around that pixel. Next, the similarity for each pixel is measured by applying the Hamming distance over the two temporal binary descriptors computed at the same location. Finally, Lloyd Max’s algorithm is used to form the change map representing the intensity of change for each location. Furthermore, the method can be extended to multispectral images, by a simple concatenation of the extracted binary descriptors for all channels.

The rest of the paper is organized as follows. Sections II and III present the proposed method for unsupervised change detection. The results obtained for synthetic and real data are

Manuscript received August 5, 2014; revised November 29, 2014; accepted December 24, 2014. This work was supported in part by the Romanian Ministry of European Funds under the Financial Agreement POS-DRU/159/1.5/S/134398.

A. Radoi is with the Department of Applied Electronics and Information Engineering, Faculty of Electronics, Telecommunications and Information Technology, University Politehnica of Bucharest, 060032 Bucharest, Romania, and also with the Research Center for Spatial Information, CEOSpaceTech, 011061 Bucharest, Romania (e-mail: rdi.anamaria@gmail.com).

M. Datcu is with the Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany, and also with the Research Center for Spatial Information, Department of Applied Electronics and Information Engineering, Faculty of Electronics, Telecommunications and Information Technology, University Politehnica of Bucharest (UPB), 060032 Bucharest, Romania (e-mail: mihai.datcu@dlr.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LGRS.2015.2389144

¹Markov Random Field

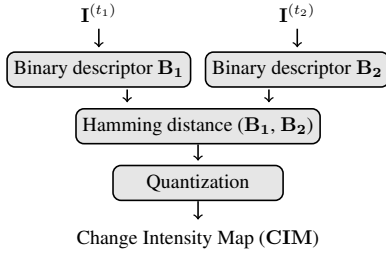


Fig. 1. Block diagram of the proposed change analysis method.

reported in Section IV, whilst the conclusions are summarized in the last section.

II. PROPOSED CHANGE ANALYSIS METHOD

The main objective of the present paper is to build a simple and effective unsupervised method for change analysis. In order to solve this problem, we propose binary descriptors, that are computed in each point of the image. The construction of these descriptors imply simple bit-strings, which can be easily compared by computing the Hamming distance. This operation can be performed in a fast manner using a XOR and a counter, instructions that are done extremely fast on modern CPUs. The last step towards an automatic assignment of change level is the separation of the values yield by the Hamming distance into categories of variation. This is achieved by applying the Lloyd's algorithm for quantization. The procedure is summarized in Fig. 1.

Let us denote by $\mathbf{I}^{(t_1)}$ and $\mathbf{I}^{(t_2)}$ two gray-level images of size $N_x \times N_y$, acquired at two different times, t_1 and t_2 .

A. The Fast Way Towards Constructing Binary Descriptors

The idea behind binary descriptors is that of finding a simple and fast way to compare feature points while requiring a small amount of memory. In order to construct local binary descriptors at pixel-level, we start from the binary test used also to build BRIEF (i.e., Binary Robust Independent Elementary Features), in [5]. However, BRIEF is designed especially for object retrieval in a test image, and finding points of interest (e.g., FAST² corner detection [6]) in the test image figures strongly in the success of this technique.

In our setup, the binary descriptors are calculated separately in each image for each pixel, as detailed below.

1) *Binary descriptor*: In the following, let us consider $S \times S$ square neighborhoods around each test pixel. For simplicity, let us consider S to be odd. For one pixel, a binary test τ_t applied on a test pixel \mathbf{P} and its neighbor \mathbf{O} , at time t , is defined as:

$$\tau_t(\mathbf{O}, \mathbf{P}) = \begin{cases} 1 & \text{if } I^{(t)}(\mathbf{O}) < I^{(t)}(\mathbf{P}) \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where $I^{(t)}(\mathbf{O})$ and $I^{(t)}(\mathbf{P})$ are the pixel intensities at points $\mathbf{O} = (O_x, O_y)^T$ and $\mathbf{P} = (P_x, P_y)^T$, respectively, at time t .

In order to assess the differences in all points for precise change detection, we perform binary tests considering $S \times S$

patches around each pixel, in both images. More precisely, considering a pixel $\mathbf{O} = (O_x, O_y)^T$, with $\lfloor \frac{S}{2} \rfloor < O_x < N_x - \lfloor \frac{S}{2} \rfloor$, $\lfloor \frac{S}{2} \rfloor < O_y < N_y - \lfloor \frac{S}{2} \rfloor$, but placed in different images (e.g., two co-registered satellite images), we construct two n -bits descriptors, one for each image.

There are many options for placing the test locations, i.e., the pairs $(\mathbf{O}, \mathbf{P}_i)$, in a patch around a pixel. Among all choices, the following is convenient because it has a physical interpretation (i.e., it represents the signs of the gradients in all directions around each pixel):

$$\begin{aligned} \mathbf{O} &= (O_x, O_y)^T \\ \mathbf{P}_i &= (P_{x_i}, P_{y_i})^T, \end{aligned}$$

where $\mathbf{O} = (O_x, O_y)^T$ is the pixel around which the patch is considered, $O_x - \lfloor \frac{S}{2} \rfloor \leq P_{x_i} \leq O_x + \lfloor \frac{S}{2} \rfloor$, and $O_y - \lfloor \frac{S}{2} \rfloor \leq P_{y_i} \leq O_y + \lfloor \frac{S}{2} \rfloor$.

The number of binary tests, and, thus, the length of the descriptor for each point is equal to $n = S^2$. As a remark, we could have taken $n = S^2 - 1$, by excluding the test between point \mathbf{O} and itself, but taking this test into consideration will not influence the future results (i.e., the Hamming distances), as $\tau_t(\mathbf{O}, \mathbf{O}) = 0$.

2) *Preprocessing step*: The binary tests take the information at pixel level and are, therefore, very noisy. Pre-filtering the images yields an increase in the stability of the change detection algorithm. Moreover, in areas where there is no change, the elements of both descriptors at the same point are likely to repeat more often if the images are smoothed before performing the binary tests. This phase is the most time-consuming part of the proposed method depending on the type of filters chosen for smoothing. Using a simple box filter, which is less time-consuming, serves our purposes well and can be regarded as a good compromise if compared to Gaussian-like low-pass filters. In the experimental part of the paper, we show the results for both filters.

B. The Hamming Distance as a Similarity Metric

In information theory, the Hamming distance between two binary vectors of equal length is the number of positions at which the corresponding symbols are different. For two binary descriptors \mathbf{B}_1 and \mathbf{B}_2 , of length n , the Hamming distance is equal to:

$$H(\mathbf{B}_1, \mathbf{B}_2) = \sum_{i=1}^n B_1(i) \oplus B_2(i), \quad (2)$$

where \oplus is the modulo 2 addition between two binary values, whilst $B_1(i)$ and $B_2(i)$ are the i^{th} elements of the descriptors.

In our case, the Hamming distance comes as a natural choice because it measures the number of times when the binary tests differ for each patch representing the same geographical area, at different moments of time. In fact, the distance measure assesses the changes in gradient, in time, along all directions around the analyzed point (or, pixel). This is the rationale upon which the proposed change analysis algorithm is based.

Additionally, for a fixed length n of the descriptors, the Hamming distance is a metric for the space $\{0, 1\}^n$. Besides its simplicity, the metric property of the Hamming distance gives it added values if compared to other similarity measures.

²Features from Accelerated Segment Test

III. CIMs

In order to assess the level of change in the bitemporal images, Change Intensity Maps (CIM) are built using the Lloyd-Max's algorithm for quantization [7], [8], which is briefly reviewed below.

Lloyd-Max's algorithm answers to the following problem of quantization: For a signal x with a probability density function p_X , find a quantizer with M representative levels that minimize the mean squared error:

$$MSE = E \left[(X - \hat{X})^2 \right], \quad (3)$$

where \hat{X} is the quantized version of the random variable X .

The solution to this problem is obtained by successively applying the following formulas until convergence:

$$\begin{cases} t_q = \frac{1}{2} [\hat{x}_{q-1} + \hat{x}_q] & , \quad q = 1, 2, \dots, M-1 \\ \hat{x}_q = \frac{\int_{t_q}^{t_{q+1}} x p_X(x) dx}{\int_{t_q}^{t_{q+1}} p_X(x) dx} & , \quad q = 0, 1, 2, \dots, M-1 \end{cases} \quad (4a)$$

where $\{t_q\}_{q=1, \dots, M-1}$ are the $M-1$ decision thresholds, and $\{\hat{x}_q\}_{q=0, \dots, M-1}$ are the M representative levels between two successive decision thresholds. For an unknown distribution p_X , \hat{x}_q in equation (4b) can be approximated with $\hat{x}_q = 1/\|T_q\| \sum_{x \in T_q} x$, where $T_q = \{x | t_q \leq x < t_{q+1}\}$ and $\|\cdot\|$ represents the number of elements in a set.

As already mentioned, the similarity between the two images is measured using the Hamming distances. In order to build the change maps, these distances are quantized with M levels ($M \geq 2$). In particular, if $M = 2$, the distinction is made between insignificant versus medium or important changes, whilst, if $M = 3$, we divide the changes into three groups that take into account the distribution of changes inside the images. The number of change levels is established by the user, depending on the application.

The choice for the Lloyd-Max's quantization is argued by the fast and simple manner to find optimal thresholds (in the sense of minimizing the expected quantization error) based on the distribution of the distances on the area of analysis.

IV. EXPERIMENTS

For the experimental part, we used the Landsat dataset available online at [9]. Two experiments are conducted, namely: binary change detection on synthetic images and multi-level change detection on real datasets.

A. Detection of Synthetic Changes

In what follows, we consider two gray-level images of 200×200 pixels, namely, an original image and a synthetically generated one. The former is obtained by inserting land-cover changes in the original image, in random places, and, also, by slightly modifying the illumination level. In order to assess the effectiveness of the method with respect to noise, we additionally add white Gaussian noise, with different signal-to-noise ratios (SNRs). The performance of the proposed

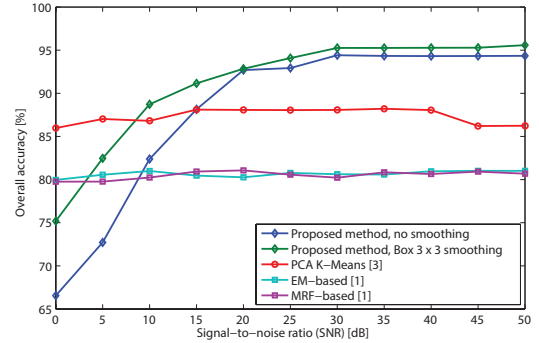


Fig. 2. Change detection performance comparison between the proposed algorithm and the algorithms proposed in [2] and [4], for different levels of noise.

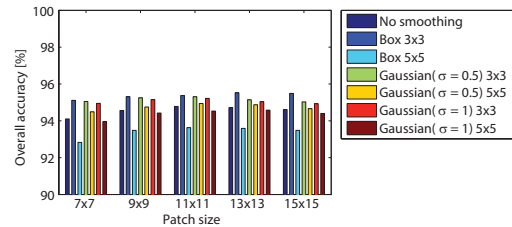


Fig. 3. Change detection performance for binary descriptors and Lloyd Max's two-levels quantization under no noise conditions, different patch sizes and various smoothing filters.

algorithm with $M = 2$ levels of quantization is compared against the methods proposed in [2] and [4]. The results are reported in Fig. 2. In almost all noise scenarios, the proposed method outperforms the others, whilst a better performance is obtained when the images are smoothed. Smoothing can be skipped when dealing with a low level of noise, but it is mandatory when working with noisy data, and this happens in most of the real cases. If the level of noise is very high, we observe a slightly better performance of the other methods that report a high number of true negatives (i.e., correct *no change* detection) under the assumption that the number of changed pixels in the image is smaller than the number of unchanged pixels. Fig. 3 shows the change detection performance when different low-pass filters are considered for pre-filtering and various patch sizes are used for building the binary descriptors. The best performance is achieved for a Box 3×3 smoothing filter and a patch size of 9×9 pixels. These settings are used for change detection in Fig. 4.

B. CIMs

The goal of this experiment is to build Change Intensity Maps that show the degree of modification on a scale that uses M ($M > 2$) levels of quantization. Two cases are shown in Fig. 5, namely the construction of the Palace of Parliament (173×119 pixels) and Morii Lake (191×191 pixels) in Bucharest, Romania, between 1984 and 1992. All the experiments are carried on multispectral Landsat 7 images, for which only the first three bands are shown.

The aim of the proposed multi-level approach for change detection is to discover all changes and to determine the opti-

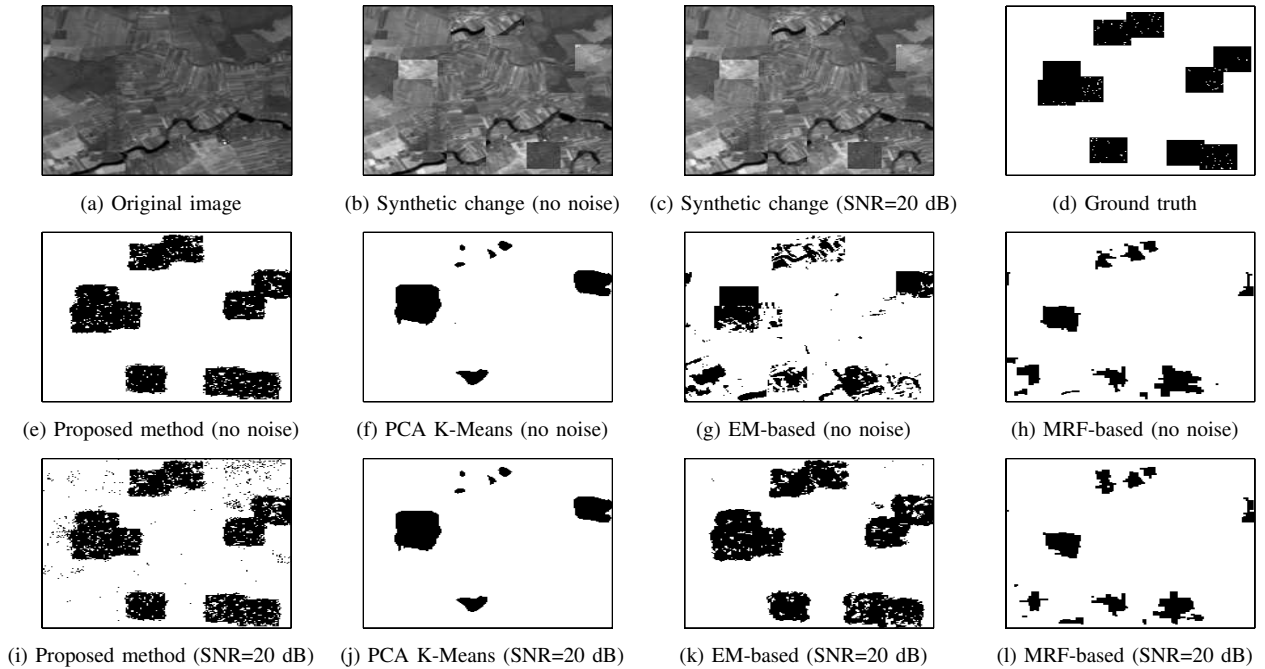


Fig. 4. Binary change detection of synthetic changes on Landsat 7 images (NIR band). The second and third rows show the outputs of the proposed method with 2 levels of quantization and also the results obtained by applying the methods described in [4] and [2], in two cases: no noise and 20 dB additive white Gaussian noise.

mal threshold values to be applied on the similarity measures. Assigning levels of change to each pixel, the method is able to provide more information for a user regarding the areas that were effected by major, medium, or minor changes. As shown in Fig. 5, as the number of quantization levels increases, the areas of major changes are better distinguished (see the red colored zones), whilst the other changed areas confirm the fact that transformations occurred, but they are not as important (see the yellow colored zones).

In a particular binary change detection scenario (i.e., $M = 2$ quantization levels), the major and major-to-medium changes interfere, resulting in a higher number of changes detected by the proposed method. In order to evaluate the binary change detection methods, the following measures are computed: true positives (TP), true positive rate ($P_{TP} = TP/N_1 \times 100$), true negatives (TN) and true negative rate ($P_{TN} = TN/N_0 \times 100$), correct classification rate or overall accuracy ($P_{CC} = (TP + TN)/(N_0 + N_1) \times 100$), false alarms (FP), false alarm rate ($P_{FP} = FP/N_1 \times 100$), missed alarms (FN), missed alarm rate ($P_{FN} = FN/N_0 \times 100$), and error rate ($P_E = (FP + FN)/(N_0 + N_1) \times 100$). The results for the binary change detection are shown in Fig. 5 and in Table I.

C. Average Running Time

Furthermore, the method proposed in this paper is efficient in terms of execution time if compared to EM, MRF, and PCA K-Means based methods. The average running times are shown in Table II for gray-level images of approximately 200×200 pixels. The method can be designed to make use of parallel architecture for computing the binary descriptors of the images in parallel and this will further decrease the computation time.

All the experiments were carried on a 2.3GHz 5-core Intel HP EliteBook computer running Windows 7 operating system.

V. CONCLUSIONS

This paper proposes a novel method for unsupervised multi-level change detection in satellite images. The technique is based upon the optimal quantization of the Hamming distances between binary descriptors built at pixel level, rendering a suitable method for building Change Intensity Maps that assess the level of change in an area. In terms of binary change detection, the experiments show that the proposed algorithm performs well, even under noisy conditions and variations of illumination.

There are several advantages of using this technique for change analysis. Firstly, from a methodological viewpoint, the method is local (i.e., it takes into account the changes that occur in the area around each pixel). In the same time, it has the capability of reporting these local changes to the changes that happen in the whole area of analysis. Secondly, from the implementation viewpoint, the method implies simple binary descriptors, that are fast to compute and to compare with modern CPUs. These facilities are beneficial in terms of on-line change analysis and unsupervised observation of the Earth's surface.

REFERENCES

- [1] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 1, pp. 218–236, Jan 2007.
- [2] L. Bruzzone and D. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 38, no. 3, pp. 1171–1182, May 2000.

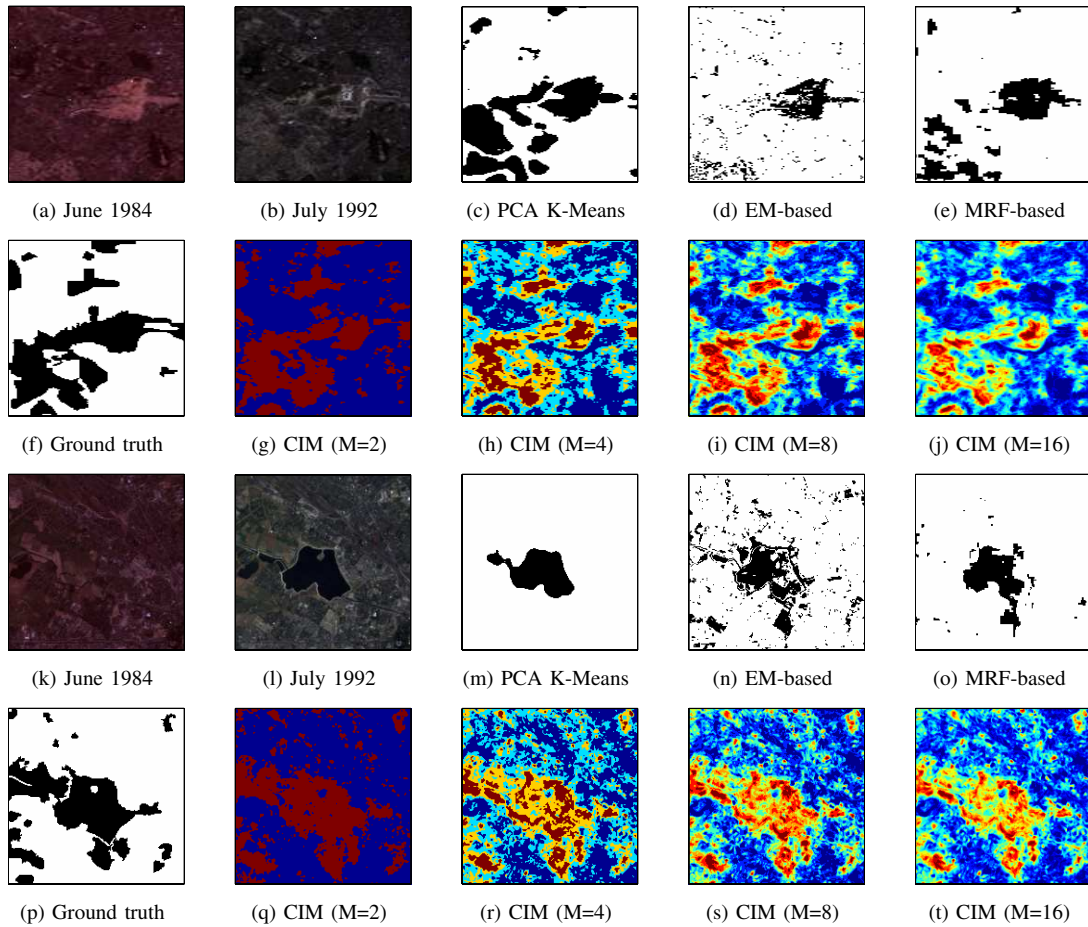


Fig. 5. Multi-level change detection on real data. The Landsat 7 images correspond to the periods before (a), (k) and after (b), (l) the construction of the Palace of Parliament and Morii Lake in Bucharest, Romania. The visible changes are manually labeled in (f), (p) (i.e., ground truth). (c), (d), (e) and (m), (n), (o) show the outcome of the change detection algorithm presented in [2] and [4], whilst (g), (h), (i), (j) and (q), (r), (s), (t) show the Change Intensity Maps as presented in this paper. The red and yellow pixels correspond to more important changes, whereas the shades of blue correspond to less important changes or no changes. For the binary change detection case, red represents major-medium changes, whilst no change is marked with blue.

TABLE I
COMPARISONS OF BINARY CHANGE DETECTION METHODS.

Dataset	Method	True Positives		True Negative		Correct	False alarms		Missed alarms		Errors
		Pixels	P_{TP}	Pixels	P_{TN}	P_{CC}	Pixels	P_{FP}	Pixels	P_{FN}	P_E
Palace of Parliament	Proposed	4603	79.73	12903	87.10	85.03	1911	33.10	1170	20.27	14.97
	PCA K-Means	3751	64.97	13699	92.47	84.76	1115	19.31	2022	35.03	15.24
	EM-based	1171	20.28	12772	86.22	67.73	351	6.08	4602	79.72	32.27
	MRF-based	1974	34.19	12574	84.88	70.67	549	9.51	3799	65.81	29.33
Morii Lake	Proposed	7563	87.38	24101	86.61	86.80	3725	13.39	1092	12.62	13.20
	PCA K-Means	2543	29.38	27725	99.64	82.97	101	0.36	6112	70.62	17.03
	EM-based	4406	50.91	26726	96.05	85.34	1100	3.95	4249	49.09	14.66
	MRF-based	2756	31.84	27617	99.25	83.26	209	0.75	5899	68.16	16.74

TABLE II
AVERAGE RUNNING TIME COMPARISON [IN SECONDS].

Method	Number of quantization levels			
	M = 2	M = 4	M = 8	M = 16
Proposed	0.21	0.30	0.31	0.33
PCA K-Means	1.58	–	–	–
EM - based	0.55	–	–	–
MRF - based	4.18	–	–	–

[3] J. Théau, "Change detection," in *Springer Handbook of Geographic Information*, 2012, pp. 75–94.
 [4] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and k-means clustering," *Geoscience and Remote Sensing Letters, IEEE*, vol. 6, no. 4, pp. 772–776, Oct. 2009.

[5] M. Calonder, V. Lepetit, M. Ozuysal, T. Trzcinski, C. Strecha, and P. Fua, "Brief: Computing a local binary descriptor very fast," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 7, pp. 1281–1298, July 2012.
 [6] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *European Conference on Computer Vision*, vol. 1, 2006, pp. 430–443.
 [7] S. Lloyd, "Least squares quantization in pcm," *Information Theory, IEEE Transactions on*, vol. 28, no. 2, pp. 129–137, March 1982.
 [8] J. Max, "Quantizing for minimum distortion," *Information Theory, IRE Transactions on*, vol. 6, no. 1, pp. 7–12, March 1960.
 [9] [Online]. Available: <http://earthexplorer.usgs.gov/>