

A Fractal Projection and Markovian Segmentation-Based Approach for Multimodal Change Detection

Max Mignotte 

Abstract—Change detection in heterogeneous bitemporal satellite images has become an emerging, important, and challenging research topic in remote sensing for rapid damage assessment. In this article, we explore a new parametric mapping strategy based on a modified geometric fractal decomposition and a contractive mapping approach allowing us to project the before image on any after imaging modality type. This projection exploits the fact that any satellite image data can be approximatively encoded in terms of spatial self-similarities at different scales and this property remains quite invariant to a given imaging modality type. Once the projection is performed and that a pixelwise difference map between the two images (presented in the same imaging modality) is then binarized in the unsupervised Bayesian framework. At this stage, we will test several parameter estimation procedures combined with several segmentation strategies based on different Bayesian cost functions. The experiments for change detection, with real images showing different multimodalities and changed events, indicate that this new fractal-based projection method, which is entirely based on a series of structural and spatial information, is an interesting alternative to classical regression-based projection methods (based only on luminance transformation). Besides, the experiments also show that the difference map, resulting in this novel projection strategy, is also particularly amenable for an unsupervised Markovian binarization approach.

Index Terms—Change detection, contractive mapping, fractal projection, heterogeneous sensors, Markov random field (MRF), multimodal, multisource.

I. INTRODUCTION

IN RECENT years, the advent of new sensor technologies and the improvement of their capabilities have stimulated the emergence of a new field of research in satellite remote sensing imagery, namely the multimodal change detection (MCD).

MCD [1] is defined as a procedure for identifying any land cover changes (such as the emergence or disappearance of artificial architectures or another type of landscape) that occurred between two (or possibly several) satellite images acquired at different times, in the same geographical area but under heterogeneous conditions.

Manuscript received November 8, 2019; revised March 6, 2020; accepted March 31, 2020. Date of publication April 30, 2020; date of current version October 27, 2020.

The author is with the Vision Lab, Département d'Informatique et de Recherche Opérationnelle (DIRO), Université de Montréal, Faculté des Arts et des Sciences, Montreal, QC H3C 3J7, Canada (e-mail: mignotte@iro.umontreal.ca)

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

Digital Object Identifier 10.1109/TGRS.2020.2986239

It generally includes methods based on heterogeneous images, i.e., acquired by different satellite sensors, either with the same sensor type, but with two different optical, SAR or other systems (cross-sensor or multisensor images) or with different sensor types such as SAR/optical images (multisource images) or possibly with the same satellite sensor but with either different looks or specification (multilooking images) or outer (weather or light) conditions that may also influence the imaging effect such as the discrepancy in brightness, contrast, and the noise distribution (information unbalanced data [2]).

MCD is thus a challenging task since the change detection procedure must be powerful and flexible enough to deal with image data that follow different statistical behavior in order to solve the same problems usually processed by monomodal CD techniques [3] such as natural, land, or environmental monitoring, damage monitoring (earthquake, flooding, landslides, etc.), or urban planning, to name a few.

In the monomodal case, two main strategies are usually considered. In the pixel-based CD (PBCD) approach, the two coregistered images are used to generate a difference image by either differencing, (log-)rationing or by considering that the second image is in fact a linear function of the first (regression method) or finally by classification-based CD techniques which presume that the two images are first segmented and classified and then compared with each other (with supervised or automatic methods). At a higher level of abstraction the object-based CD (OBCD) category operates with extracted objects on the two images which can be classified (postclassification framework) before comparison [4].

MCD has recently received a growing interest in the remote sensing community since this technique imposes less restrictive conditions about the origin and characteristics of the acquired data than the classical monomodal CD technique. As a result, this technique allows us to exploit the enormous amount of heterogeneous data that we can now obtain from various archives or different types of (ever-increasing) existing number of Earth observation satellites, including the most advanced systems equipped with the latest sensor technologies. Finally, let us add that the different imaging modalities may be complementary and by this fact, this imaging modality fusion technique could potentially be exploited (not only in geoscience imaging [5]) for further improving the change detection and analysis of land surfaces

with complex properties subject to extreme conditions (e.g., temperature, fire, ice, etc.) [6].

Despite its undeniable potential, there is relatively little research on heterogeneous or MCD approaches in geoscience. Nevertheless, it can be clearly seen, from an exhaustive review of the literature on this problem, four different (nonexclusive) types of strategies: firstly, the simplest methods are those provided by algorithms using similarity measures [7]–[10] or local descriptors [9] with assumed invariance properties according to the imaging modality. Secondly, we can find the class given by the parametric models in which a set (or mixture) of multivariate or meta-Gaussian distributions are generally used to model the joint statistics or the dependencies between the two imaging modalities or the different types of multisensor data [11]–[17]. Thirdly, we can consider the class of nonparametric methods that have the property of not explicitly assuming a specific parametric distribution for the data. This includes learning machine algorithms which learn directly from examples (training data) [2], [18]–[27] or unsupervised nonparametric-based procedures, that do not require a training step, such as the least-squares energy-based model introduced in [6] and aiming at satisfying a set of overdetermined constraints, expressed for each pair of pixels existing in the before and after images. Finally, the last class is given by procedures mainly based on a transformation or projection of the two multimodal images into a common domain or feature space, in which the two heterogeneous images have the property to share the same statistics and on which classical monomodal CD methods can then be applied [26], [28]–[34]. In the same spirit, an interesting common space can be provided by a segmentation or a clustering map [35] with which it is easier to detect the different changes.

A subcategory of the latter strategy groups together all the procedures mapping one of the two image data to the domain (i.e., imaging modality) of the other image. This subcategory, which may also be common to the aforementioned parametric or nonparametric method classes, will be now described in more detail since the proposed MCD model described in this article fits into this category. In [36], parametric and nonparametric regression methods based respectively on Gaussian processes, support vector machines, random forests, and homogeneous pixel transformation [25], [37] are evaluated. Such mapping can also be learned by a neural network [20] and a training database. In the same vein, a nonparametric projection map that aligns data from different domains in a nonlinear way through kernelization (and called Kernel Manifold Alignment) is also proposed in [38]. A parametric method is used in [11] and [12] in which a quantile regression is applied, according to the copula theory, to perform an estimation of the local statistics that would have been observed if the acquisition conditions of the first image had been similar to the ones of the second image. Brunner *et al.* [8] have presented another parametric MCD method to quantify the earthquake damages in urban areas using a pre-event optical image and a postevent SAR image. To this end, the parameters of each building, estimated from the optical scene and combined with the acquisition parameters of the actual postevent SAR scene are both used to predict (via

simulation) the expected SAR signature of the building which is then subsequently compared, with a similarity measure, to the actual SAR scene in order to quantify the structural damages caused by the earthquake.

In this article, we explore a new parametric mapping strategy and, more precisely, we investigate the truth and test the validity of the following hypothesis: any satellite image data can be encoded in terms of spatial self-similarities at different scales (or, equivalently, can be geometrically represented by a collection of spatial transforms) and this property remains quite invariant to a given imaging modality type. This property allows us to project the before image on the after imaging modality type and to obtain, with a simple pixelwise difference, a relevant change probability map which is then binarized with an unsupervised Markovian approach.

In this context and in summary, this article makes the following main contributions.

- Unlike the different regression-based projection methods [20], [25], [36], [37], [39] which aim to learn/estimate (and then apply) a luminance transformation to map the first image to the domain of the other image, we herein propose to exploit the inherent and structural fractal property of any satellite (or natural) image and more especially the fact that this appealing propriety remains quite invariant to a given imaging modality type. This gives us a new parametric strategy (or a new set of transformations, entirely based on a series of structural spatial information), allowing us to efficiently map the pre-event image to the domain of the postevent image.
- In addition, we will show that, once the mapping is done by our fractal decoding strategy, the pixel-by-pixel difference map between the two images (presented in the same imaging modality) has good properties in terms of likelihood distributions associated with the two classes (“change” and “no change”); i.e., first exhibiting a mixture of sufficiently Gaussian (or quasi-Gaussian) distributions which are also not too mixed between them and that are suitable for that an unsupervised Markovian segmentation (binarization) approach to be effective, regardless the very different imaging modalities that could be encountered in remote sensing.

Let us note that a possible and more adequate way to estimate (or learn) a luminance transformation (or a collection of luminance transformations) that maps the first image to the domain of the other image (and/or vice versa) remains the different regression-based projection methods already proposed in the literature [20], [25], [36], [37]. An interesting MCD approach would be to combine a projection method based on regression (luminance transformation) with our method based solely on spatial transformations.

The remainder of this article is organized as follows. Section II describes the proposed fractal-based mapping strategy and the unsupervised Markovian approach used to binarize the change map. Section III presents a set of experimental results and comparisons with existing MCD algorithms. Finally, Section IV concludes the article.

II. PROPOSED CHANGE DETECTION MODEL

A. Introduction

In multimodal case, it is well known that the local statistics of the input (before and after change) images acquired by the two different imaging modalities may be radically different from one image to the other one and this characteristic prevents that a simple pixelwise difference [or (log-)ratio] model from being effective in highlighting the different changes. Nevertheless, there is an important feature of any satellite image data (also common to all natural images) that is quite well preserved across the different types of imaging modality; any kind of images show self-similar structures at different scales. Otherwise said, all images are (approximately) formed of properly transformed parts of itself (or are made up of reduced copies of themselves) and thus exhibits this peculiar type of self-similarities across a range of scales. In fact, this piecewise self-similarity property of the image is the redundancy that fractal image compression schemes take advantage and attempt to eliminate [40], [41] (in this case, at two successive scales to be precise). Similarly, the high degree of redundancy expressed by the self-similarity property, at the same scale, of any images (thus simply expressing that “*every small window in a natural image has many similar windows in the same image*”) has been first successfully exploited in texture synthesis [42], [43] and then has demonstrated all its effectiveness as regularity assumption in image denoising algorithms with the so-called NL-means introduced by Buades *et al.* [44]. This intrinsic self-similarity property was also successfully used in zooming [45], deconvolution/restoration [46], inpainting [47], super-resolution [48], and demosaicking [49] procedures (whose final aim is not to denoise, but instead to infer high frequency information by transportation from known pixels to unknown pixels) to name a few.

B. Fractal Projection Model

1) *Fractal Encoding Step*: The original image is partitioned into a set a range blocks $\mathbf{R} = \{R_1, R_2, \dots, R_{N_R}\}$ of $N_W \times N_W$ pixel nonoverlapping squared subwindows. Let now $\mathbf{D} = \{D_1, D_2, \dots, D_{N_D}\}$ be the domain blocks, i.e., the set of all $Sc \cdot N_W \times Sc \cdot N_W$ overlapping squared subwindows existing in the image (with $Sc = 2$ in conventional fractal encoding, see Fig. 1). First, each of the domain blocks is contracted by pixel averaging or down sampling to match the size of the range block. Second, eight symmetrical transformations (rotations and flips) are applied to all contracted domain blocks to bring out an extended domain pool (see Fig. 2) that we now call \mathbf{D}^\dagger . For each range block R_i , we search through all of the extended domain pool \mathbf{D}^\dagger to find the N_i best matched domain blocks $D_i \in \mathbf{D}^\dagger$, i.e., the N_i domain blocks that most look like the subimage R_i in the LS sense.

At this stage, it is important to mention that there are two main differences with the conventional fractal encoding proposed by Barnsley [40] or Fisher [41]. Firstly, we consider only spatial transformations. We do not consider any transformation of luminance or gray level such as the brightness and

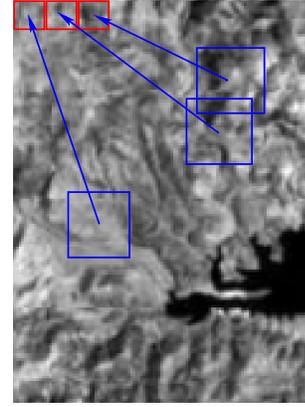


Fig. 1. In a fractal encoding (or any partitioned iterated function) system, self-similarity is sought between (larger and possibly overlapping) parts of the image also called domain blocks (herein represented in blue and mathematically defined by $\{D_i\}$ in the following) and smaller (and nonoverlapping) parts of the image or range blocks (represented in red and defined by $\{R_i\}$). In this example, we illustrate this kind of matching for the three first (red) range blocks existing at the beginning of the image. Comparison between domain and range blocks is done in the least square (LS) sense, by first extending the domain pool (or set of domain blocks) to the eight possible symmetrical transformations (isometry) for each domain block (rotations and flips, see Fig. 2), and second, by contracting each domain block (by down sampling) to match the size of the range block.

contrast transformation¹ that could be used to determine (as in conventional fractal encoding) the most similar subimages R_i to a given scale factor (contrast) and offset (brightness) factor. Secondly, in our encoding scheme, we do not consider the best domain block (that most looks like the subimage R_i in the LS sense) but rather the N_i best domain blocks (and this search is ensured in our application by a kd-tree search method [50]). The usefulness of this feature will be explained in detail in Section II-B2. The position of the N_i best domain blocks along with the right isometric transformation (rotations and flip) corresponding to each range block is then stored in an appropriate data structure called the fractal code ($\mathcal{F}_{\text{frct}}$) that will be used in the following step.

2) *Fractal-Based Projection Step*: The goal of this step is to use the previous fractal mapping/encoding procedure to express the pre-event satellite image content (\mathbf{y}^{t_1}) according to the imaging modality of the postevent image (\mathbf{y}^{t_2}) so that a simple pixelwise difference between the two images, presented in the same imaging modality, can highlight the different changes.

To this end, we can use the iterative decoding process proposed by Barnsley [40] that consists of iterating, for all range blocks, the inverse transforms, listed in the fractal code ($\mathcal{F}_{\text{frct}}$). More precisely, this means that, for every range block R_i , the right domain block D_i is taken from an initial image

¹Indeed, the luminance transformation closely depends on the considered imaging modality type. Let us note that a possible and more adequate way to estimate (or learn) a luminance transformation (or a collections of luminance transformations) that maps the first image to the domain of the other image (and/or vice versa) remains the different regression-based projection methods already proposed in the literature [20], [25], [36], [37]. An interesting MCD approach would be to combine a projection method based on regression (luminance transformation) with our method based solely on spatial transformations.

(that will be specified later), the inverse isometry is applied and is then mapped to the range block. This completes an iteration of the decoding process. The resultant image is thus made of the collage of the entire subsampled transformed domain blocks or otherwise said, the decoded image is made up of reduced copies of itself (see Fig. 1). This image is then used back as the initial image of the second decoding iteration until the sequence of images converges, i.e., the iterates get attracted to a fixed set also called the attractor. This convergence is ensured, after a few number of iterations, by the fractal theory of iterated contractive image transformations also called partial iterated function system (PIFS) [40], [41].

In our application, since we represent image data as collections of spatial transformations without luminance (brightness and contrast) transformations, we need to make the following changes in the decoding process in order to converge and to represent the *before* image with the imaging system generating the *after* image.

a) First change: We take for the first iteration of the decoding process, the *after* image as an initial image and this for several reasons. Indeed, this image is closest to our goal (excepted for the few regions in this image where there has been a change) and consequently the convergence of the decoding process is thus greatly facilitated. In fact, to be precise, we must start the iterative decoding process with an initial image that has been generated with the *after* imaging modality if we want to reach our goal (and this contrary to the conventional fractal decoding procedure which can start with a random image). Indeed, the constraint used in our application of using only spatial transformations and thus spatial collages of the entire domain block (eventually flipped and/or rotated) without luminance (brightness and contrast) transformations allows to recreate the image content but with the rendering or textural characteristics intrinsically dependent on the imaging modality (or imaging system and data acquisition) of the domain blocks (taken from the initial image).

b) Second change: During the fractal decoding (or the reconstruction process) of the *before* image with the imaging modality belonging to the *after* image, some domain blocks could be contaminated by the postevent change (see Fig. 3). This could induce some noise when copying/mapping these blocks to their respective range block. That is why we consider, in our application, not one, but N_i best domain blocks (with the right isometry) corresponding to each range block (see Section II-B1). In order to eliminate the most contaminated blocks, we then select, among these N_i blocks, the ρ percent that most look like the block R_i (in the LS sense) and in order to eliminate even more noise, we average them and mapped the averaging result to the corresponding range block (see Figs. 3 and 4).

C. Probability Change Map Estimation

In its simplest form, the fractal coding is applied over fixed blocks for the range domain, typically of size 8×8 or 16×16 ($N_W \times N_W$) with domain blocks generally twice larger than the range block ($Sc = 2$). In fact, for these two choices of range

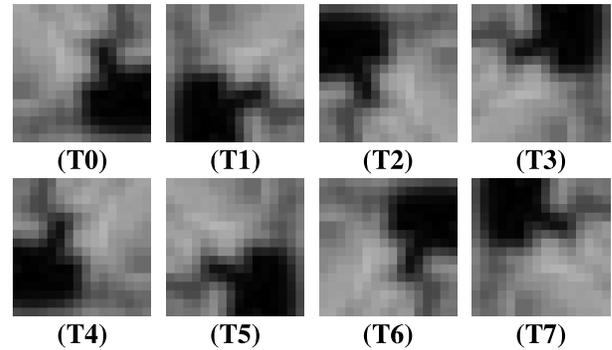


Fig. 2. The eight symmetrical transformations (rotations 0° , 90° , 180° , 270° , and their vertical flip).

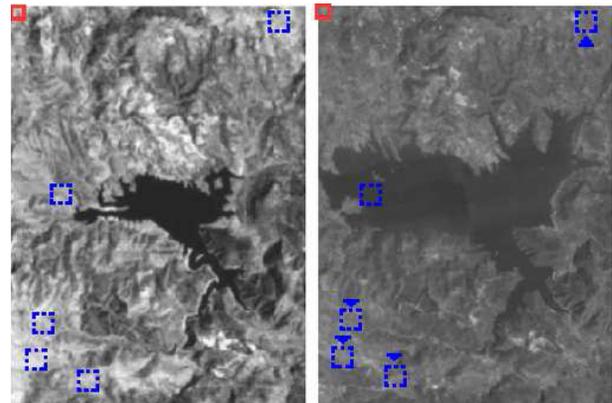


Fig. 3. From left to right; before (pre-event) y^1 and after (postevent) y^2 images. Let us assume that we have found the $N_i = 5$ best matched 16×16 domain blocks D_i (blue dotted frame) existing in the before image which most look like (in this example with the T_0 isometry [see Fig. 2 at top left]) the squared block or 8×8 subimage R_i (at the top left of the before image drawn in red and with solid lines) in the LS sense. Among these $N_i = 5$ domain blocks D_i , once reported on the after image, one domain block (the middle left) turns out to be clearly contaminated by the postevent change (a flood). In order to express the pre-event satellite image content according to the imaging modality of the postevent image, we take for the first iteration of the decoding process, the *after* image as initial image and we then select, among these $N_i = 5$ domain blocks (taken in the postevent image) a percentage ρ (or subset) of domain blocks that most looks like the R_i block of the postevent image (as, for example, those who are pointed by an arrow and who are the least contaminated by the “change” area). We average them (after that the corresponding inverse isometry is applied) and we finally mapped the averaging result to the location corresponding to R_i . This is repeated for each range block and for several (N_k) iterations.

block sizes, one generally obtains a substantially identical compression error (depending on the image) but related to a different resulting contractive mapping which then generates a different spatially distributed reconstruction error map. Moreover, let us note that this reconstruction error map appears as random noise instead of visible ringing or smoothing or block-boundary artifacts which are typical in JPEG, discrete cosine transform (DCT), or wavelet compressed image data [41]. Here, we can exploit this appealing property and reduce the noise of our final difference or probability change map by averaging several difference maps obtained by a pixelwise absolute value of the difference between the after image and the decoded image (i.e., the before image expressed according to the imaging modality of the after image).

TABLE I
DESCRIPTION OF THE TEN HETEROGENEOUS DATA SETS

no	Date	Location	Size (pixels)	Event (& Spatial resolution)	Sensor
1	Sept. 1995 - Jul. 1996	Sardinia, It	412 × 300	Lake overflow (30 m.)	Landsat-5 (NIR band) / Optical
2	July 2006 - July 2007	Gloucester, UK	2325 × 4135	Flooding (0.65 m.)	TerraSAR-X / QuickBird 2
3	Feb. 2009 - July 2013	Toulouse, Fr	4404 × 2604	Construction (2 m.)	TerraSAR-X / Pleiades
4	May 2012 - July 2013	Toulouse, Fr	2000 × 2000	Construction (0.52 m.)	Pleiades / WorldView 2
5	2001 - Jan. 2002	Congo, Africa	400 × 800	Nyiragongo volcano eruption (10 m.)	Radarsat (3 / 5-looks)
6	Jan. 2017 - Feb. 2017	Sutter, CA, USA	502 × 876	Flooding (≈ 15 m.)	Landsat 8 / Sentinel-1A
7	June 2008 - June 2013	Island town, China	415 × 403	Urbanization (8 m.)	Radarsat-2 / Google Earth
8	June 2008 - Sept. 2012	Shuguang village, China	419 × 342	Building construction (8 m.)	Radarsat-2 / Google Earth
9	June 2008 - June 2009	Yellow river, China	409 × 367	River's drying up (8 m.)	Radarsat-2 (1 / 4-looks)
10	1999 - 2000	Gloucester area, UK	554 × 990	Flooding (≈ 25 m.)	Spot / NDVI

Algorithm 1 Fractal Encoding & Projection Algorithm

Input: Pair of bitemporal satellite images: $(\mathbf{y}^{t_1}, \mathbf{y}^{t_2})$
Output: $\mathbf{y}^{t_1 \triangleright t_2}$: \mathbf{y}^{t_1} with the imaging modality of \mathbf{y}^{t_2}

N_W Size of the square range blocks
 S_c Scale of the contractive transformation
 N_i Number of (closest) domain blocks to a given range block
 ρ Percentage of the closest domain blocks used in the projection/decoding process
 N_k Number of iterations of the projection/decoding

$\mathbf{R} = \{R_i\}$ Set of all nonoverlapping range blocks of size $N_W \times N_W$ covering the image
 $\mathbf{D} = \{D_i\}$ Set of all overlapping domain blocks of size $S_c \cdot N_W \times S_c \cdot N_W$ covering the image
 \mathbf{D}^\uparrow Extended domain block considering for each D_i the 8 existing symmetrical transformations (see Fig. 2)
 $\mathbf{D}_\downarrow^\uparrow$ Set of all downsampled domain blocks of \mathbf{D}^\uparrow by factor S_c
 $\mathcal{F}_{\text{frct}}$ Fractal encoding data (structure) file

1. Fractal Encoding Step

▷ Initialization: build $\mathbf{D}_\downarrow^\uparrow$ from \mathbf{y}^{t_1}
for each R_i (lexicographically indexed) of \mathbf{y}^{t_1} **do**
◦ Search through all of $\mathbf{D}_\downarrow^\uparrow$ (built from \mathbf{y}^{t_1}) to find the N_i domain blocks which most look like R_i in the LS sense and save the location & symmetrical transformation (associated with these N_i blocks) in the file $\mathcal{F}_{\text{frct}}$

2. Fractal Projection/Decoding Step

▷ Initialization: $k \leftarrow 0$; $\mathbf{y}^{t_1 \triangleright t_2} \leftarrow \mathbf{y}^{t_2}$; build $\mathbf{D}_\downarrow^\uparrow$ from \mathbf{y}^{t_2}
while $k < N_k$ **do**
for each R_i (lexicographically indexed) of $\mathbf{y}^{t_1 \triangleright t_2}$ **do**
◦ From $\mathcal{F}_{\text{frct}}$ file, find the location and the right isometry of the N_i domain blocks associated with R_i
◦ Select the ρ percent of N_i domain blocks that most look like R_i (in the LS sense) and average them and map the averaging result to the corresponding range block in $\mathbf{y}^{t_1 \triangleright t_2}$
 $k \leftarrow k + 1$

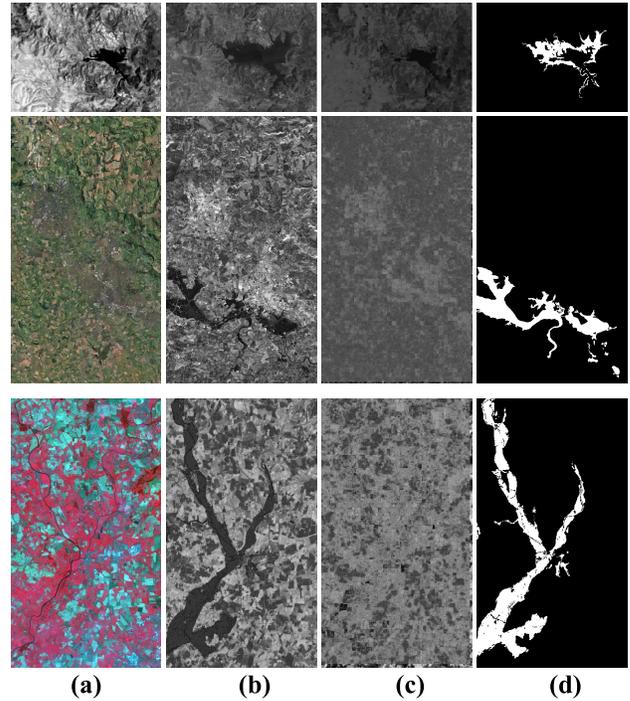


Fig. 4. Heterogeneous data sets #1, #2, #10 (see Table I). (a, b) Pre-event image t_1 (\mathbf{y}^{t_1}), postevent image t_2 (\mathbf{y}^{t_2}), (c) $\mathbf{y}^{t_1 \triangleright t_2}$, pre-event satellite image reconstructed according to the imaging modality of the postevent image ($N_w = 8$, $N_i = 5$, and $\rho = 50\%$) (see Section II-B2), and (d) binary change map (ground truth).

In our application, this averaging is achieved by averaging three difference maps, respectively, generated for $N_W = 8$, $N_W = 12$, and $N_W = 16$ (the gain of this averaging strategy will be quantified in Section III). The result is then filtered by a 3×3 classical Gaussian filter to further remove the remaining noise.

The fractal encoding and projection algorithm including the averaging process above explained is summarized in pseudocode in Algorithm 2.

D. Two-Class Unsupervised Segmentation

Finally, in order to automatically separate the change and no change areas from the previously obtained probability change map, we have formulated the binarization problem

in the unsupervised Bayesian framework with a two-step process. First, a parameter estimation step is conducted to infer the likelihood model parameters in the maximum likelihood (ML) sense. At this stage, we consider a simple mixture of two Gaussian likelihood distributions (for the two possible class labels; “change” and “no change”) and to estimate the parameters of this distribution mixture, we will implement and test the expectation maximization (EM) [51] (see Fig. 8) and the stochastic expectation maximization (SEM) estimation procedure [52]. Then, a second step is devoted to the two-class segmentation or change detection itself based on the value of estimated parameters [53]. In this step, we will test successively the simple binarization scheme using the Markov random field (MRF) segmentation model, called iterated conditional modes (ICM), introduced by Besag [54] using the maximum a posteriori (MAP) estimator, a Monte-Carlo algorithm using the Gibbs sampler to find the maximizer of the posterior marginal (MPM) estimator proposed by Marroquin *et al.* [55] (with a Potts MRF pairwise interaction in the two latter cases), and finally the multiscale MRF segmentation algorithm presented by Bouman and Shapiro using the sequential MAP (SMAP) estimator [56].

The overall proposed change detection algorithm is finally summarized in pseudocode in Algorithm 2.

Algorithm 2 Fractal Projection & Markovian Segmentation-Based Algorithm for MCD

Input: Pair of bitemporal satellite images: $(\mathbf{y}^{t_1}, \mathbf{y}^{t_2})$
Output: A binary CD segmentation map: \mathbf{x}

N_i Number of (closest) domain blocks to a given range block
 ρ Percentage of the closest domain blocks used in the projection/decoding process

$\mathbf{y}^{t_1 \triangleright t_2}$ \mathbf{y}^{t_1} with the imaging modality of \mathbf{y}^{t_2}
 \mathbf{y}^d Difference or probability change map
 $\Phi_{\mathbf{y}^d}$ Estimation of the distribution mixture of \mathbf{y}^d

1. Difference Map Estimation Step

- ▷ For Each $k \in \{0, 1, 2\}$ do
 - $\mathbf{y}_1^{t_1 \triangleright t_2} \leftarrow \text{FRCT_ENCOD\&PROJ_ALGO}(\mathbf{y}^{t_1}, \mathbf{y}^{t_2}, \dots, N_W = 8 + 4k, N_i = 5, \rho = 50\%, S_c = 2)$ (SEE ALGORITHM 1.)
 - $\mathbf{y}_k^d \leftarrow |\mathbf{y}_k^{t_1 \triangleright t_2} - \mathbf{y}^{t_2}|$ & linear contrast stretch \mathbf{y}_k^d from 0–255
 - $\mathbf{y}^d \leftarrow \frac{1}{3}|\mathbf{y}_1^d + \mathbf{y}_2^d + \mathbf{y}_3^d|$ & linear contrast stretch \mathbf{y}^d from 0–255
 - ▷ $\mathbf{y}^d \leftarrow [3 \times 3]$ GAUSSIANFILTER(\mathbf{y}^d)

2. Segmentation Parameter Estimation Step

- ▷ $\Phi_{\mathbf{y}^d} \leftarrow \text{EM OR SEM algorithms applied on } \mathbf{y}^d$

3. Segmentation Step

- ▷ $\mathbf{x} \leftarrow \text{ICM OR MPM OR SMAP algorithms applied on } \mathbf{y}^d$ and based on $\Phi_{\mathbf{y}^d}$
-

III. EXPERIMENTAL RESULTS

A. Heterogeneous Data Set Description

To validate our approach, we present in this section a series of tests conducted on ten real and very different

TABLE II
PERCENTAGE FOR THE TEN HETEROGENEOUS DATA SETS IN TERMS OF TN, TP, FP, AND FN AND F-MEASURE (SEE TABLE I) FOR THE EM/ICM COMBINATION

Multimodal pair	TN	TP	FP	FN	F_m
Thermic/Optical	87.330	5.482	6.433	0.755	0.604
Optical/SAR	91.578	5.495	1.997	0.931	0.790
SAR/Optical	86.182	4.679	5.863	3.276	0.506
Optical/Optical	80.366	3.412	4.417	11.805	0.296
SAR (3/5-looks)	75.678	7.363	4.839	12.120	0.465
Optical/SAR	93.319	1.844	1.944	2.892	0.433
SAR/Optical	92.878	1.084	3.029	3.009	0.264
SAR/Optical	88.197	6.051	5.709	0.043	0.678
SAR/SAR	94.995	2.858	0.846	1.301	0.727
Optical/Optical	86.036	10.202	1.341	2.421	0.844

heterogeneous (multimodal) data sets, illustrating the complexity of the addressed problem. This data set both includes different multimodalities (belonging to the three different classes of heterogeneity existing in this research domain; namely multisensor, multisource, and multilooking), reflecting different change detection conditions, showing a wide diversity of changed events (river flooding, lake overflow, river’s drying up, construction, urbanization volcano eruption) and provided at different resolution levels (varying from 0.52 to 30) and for different image sizes (varying from 300 to 4135 pixels in length or width) (see Table I).

More precisely, this data set includes 3 (#1, #4, #10) multisensor optical image pairs (i.e., same sensor type) but with two different optical sensors (#4) or same satellite sensor but with different specifications or bands (#1, #10), 5 (#2, #3, #6² [39], #7, #8) multisource image pairs (i.e., different sensor types), respectively optical/SAR (#2, #6) and SAR/optical (#3, #7, #8) and 2 (#6, #9) multilooking image pairs (i.e., with the same satellite sensor but with different looks or specification).

This allows us to compare the performance of the proposed method with different state-of-the-art MCD algorithms recently proposed in this field [12], [13], [31], [57], [15]–[17], [6], [10], [20], [24]–[26], [34], [36], [39], [58]–[60] in different MCD conditions. In this benchmark, all the ground-truth images (change detection mask) were provided by an expert photograph interpreter.

B. Results and Evaluation

In all the experimental results, we have considered the simple gray level of the image (and thus converted), when necessary, the optical color (or three-bands) satellite image to grayscale (by simply averaging the three color channels) and reduce the size of the image such that its maximal size (length or width) does not exceed 500 pixels (for reducing the computational load of our procedure).

As already said in Section II-C, we average three difference maps respectively obtained by a pixelwise absolute value of the difference between the after image and the before image reconstructed according to the imaging modality of the

²This heterogeneous data set is kindly available online at http address <https://sites.google.com/view/luppino/data>

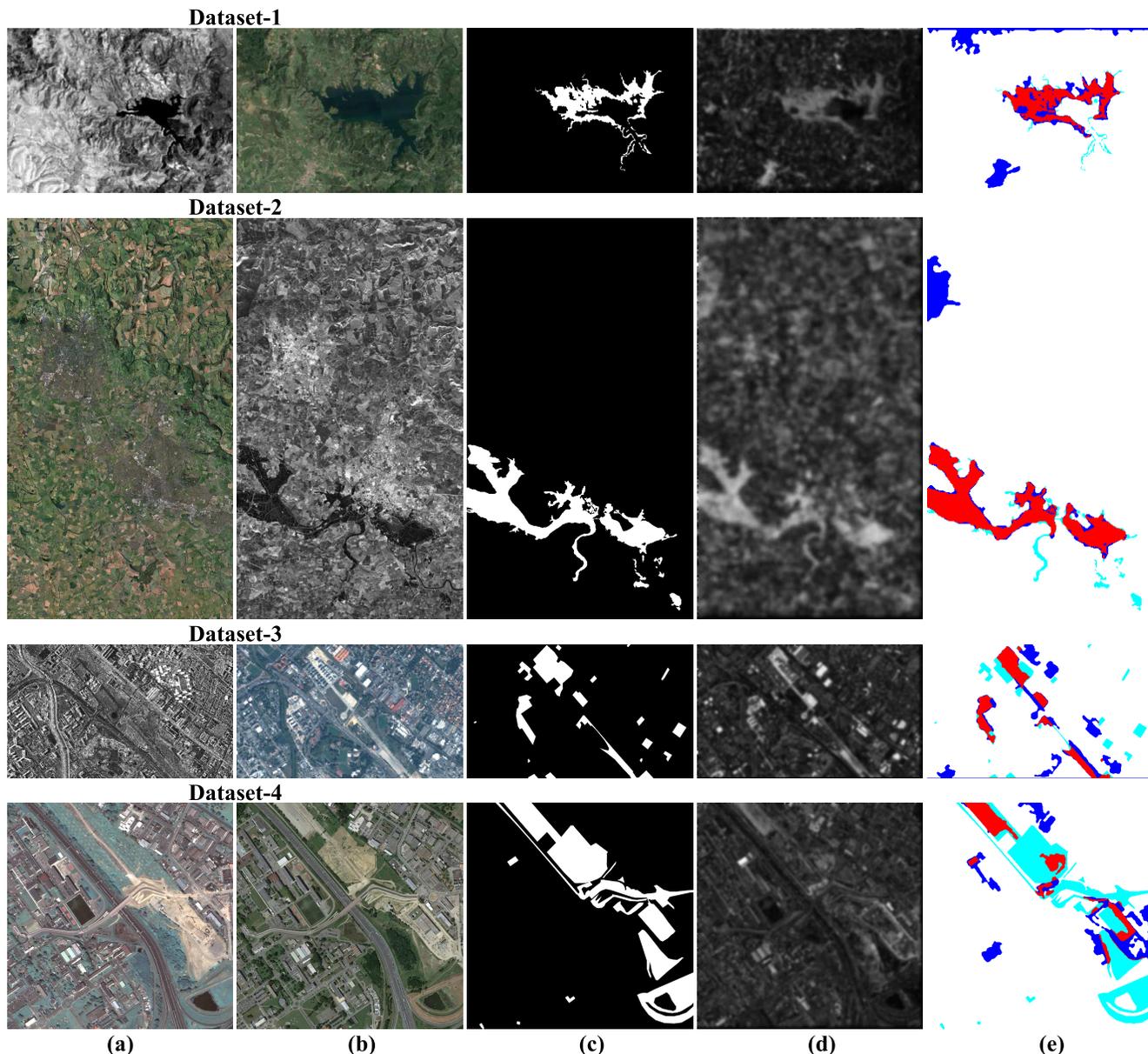


Fig. 5. Heterogeneous data sets (see Table I). (a–c) Image t_1 , t_2 , ground truth; (d) difference map, and (e) confusion map (white: TN, red: TP, blue: FP, cyan: FN) obtained by the proposed approach [Fractal-(EM/ICM) Markovian CD model].

after image, and generated (encoded/decoded) respectively for $N_W = 8$, $N_W = 12$, and $N_W = 16$.

The internal parameters of our model are N_i , the number of best-matched domain blocks D_i that, once subsampled, most looks like the subimage R_i block of the pre-event image (see Section II-B1 and Fig. 3) and ρ , the percentage of domain blocks, reported on the after image, that most looks like the R_i block of the postevent image. For all the experimental results, we use $N_i = 5$ and $\rho = 50\%$. For the parameter estimation step, we initialize the EM and the SEM algorithms with a mixture of two Gaussians with equal proportion and equal variance ($\sigma^2 = 100$) and with two distinct means $m_0 = 100$ and $m_0 = 200$. For the Markovian segmentation, we have considered an *a priori* Potts model relying on a second-order neighborhood system with every clique potential setting to

$\beta = 1$ in the case of the ICM and 6 level of pyramids for the SMAP with a hyperparameter equals to $= 0.995$.

In order to discuss and compare obtained results, a quantitative study is realized by computing first the classification rate accuracy that measures the percentage of the correct changed and unchanged pixels: $PCC = (TP+TN)/(TP+TN+FN+FP)$ and the F-measure: $Fm = (2TP)/(2TP+FP+FN)$, where TP, TN, FN, FP designate classically the true positives, negatives, false negatives and positives.

First, we have indicated in Table III the score (mean accuracy rate, mean F-measure, and the overall computation time in seconds) for the CD estimation of the ten heterogeneous data sets considered in our study and obtained by our fractal reconstruction process followed by an unsupervised Markovian segmentation with a two step-procedure combining

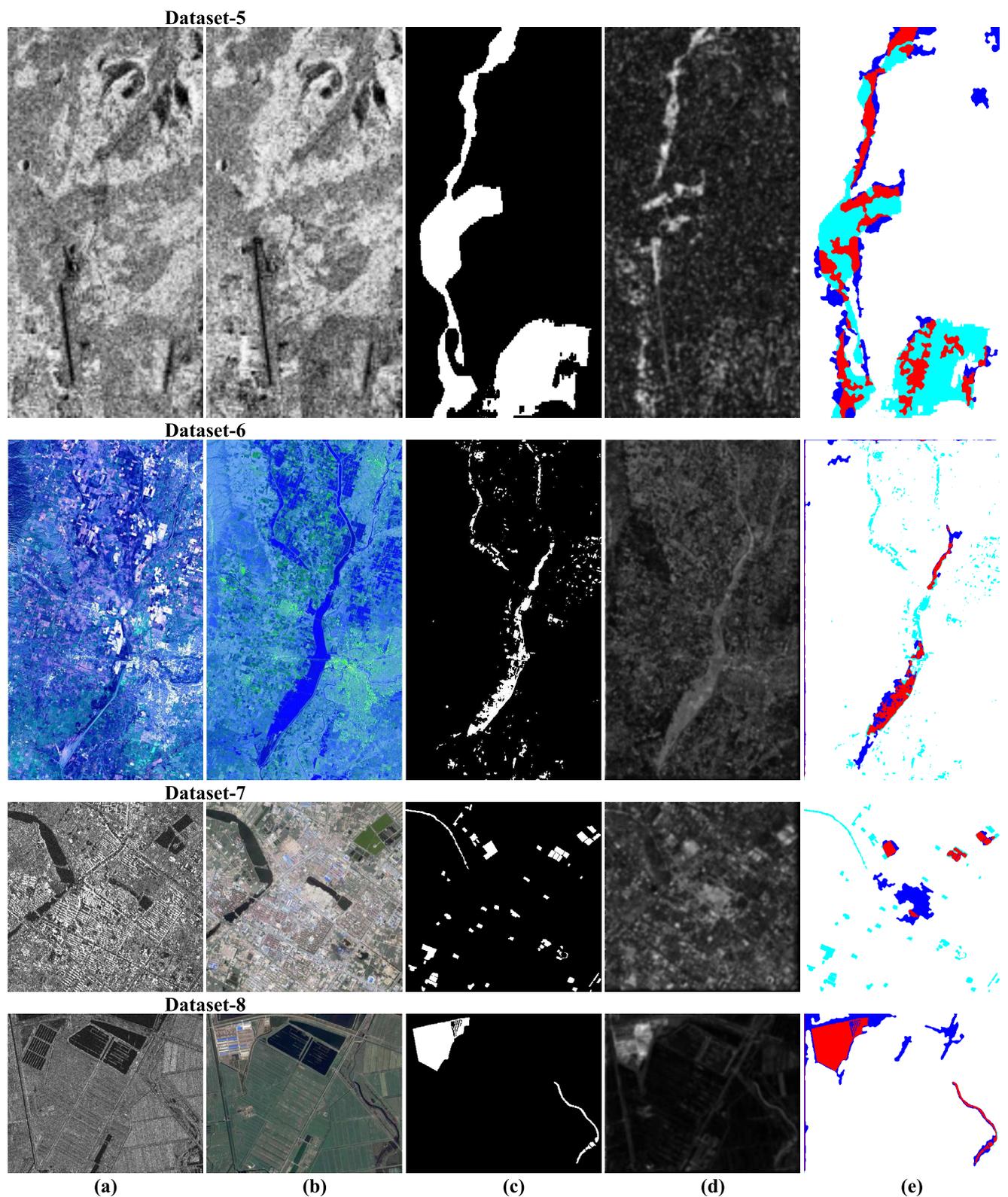


Fig. 6. Heterogeneous data sets (see Table I). (a–c) Image t_1 , t_2 , ground truth; (d) difference map, and (e) confusion map (white: TN, red: TP, blue: FP, cyan: FN) obtained by the proposed approach [Fractal-(EM/ICM) Markovian CD model].

either the parameter estimation step provided by the EM [51] or SEM [52] estimation procedure followed by either

algorithm), or the ICM [54], MPM [55], or finally the SMAP [56] segmentation model. The obtained results are comparable with less than 1% for the accuracy rate and less than 0.8% for

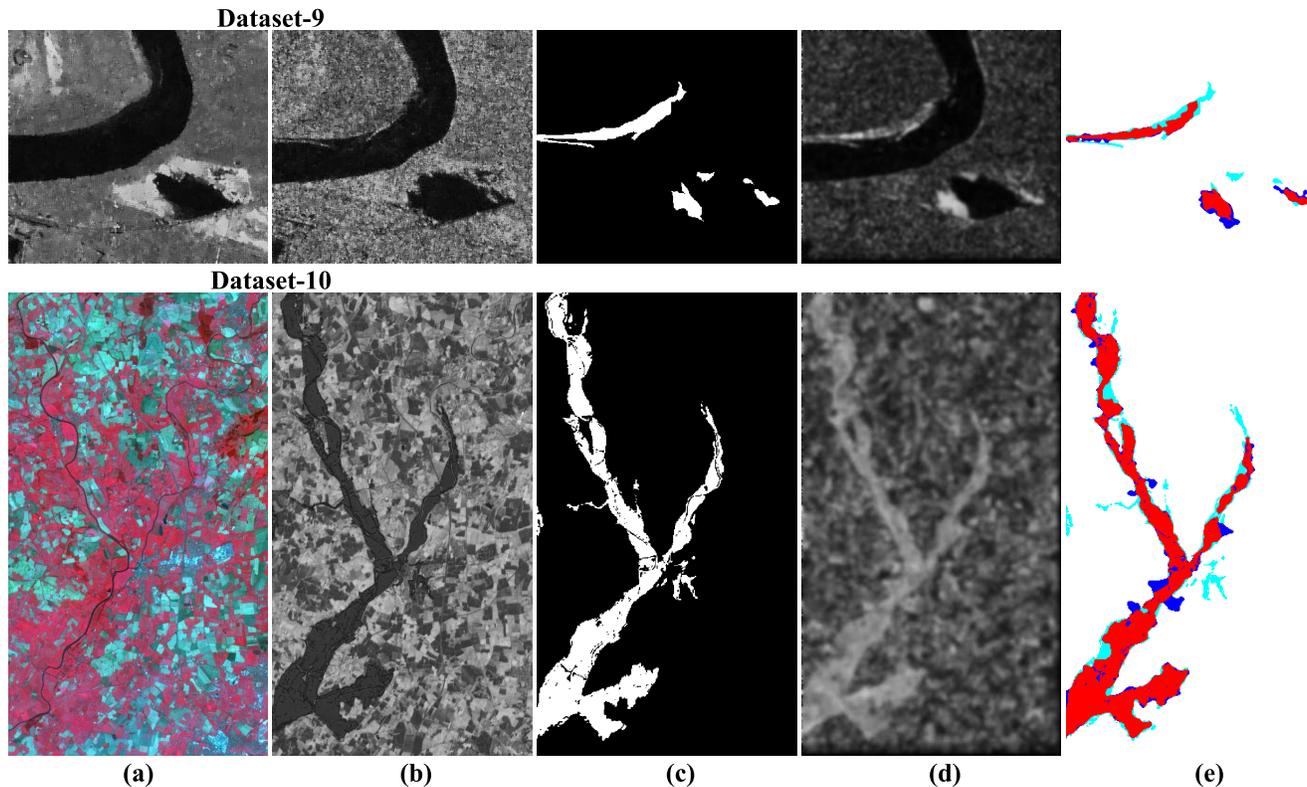


Fig. 7. Heterogeneous data sets (see Table I). (a–c) Image t_1 , t_2 , ground truth; (d) difference map, and (e) confusion map (white: TN, red: TP, blue: FP, cyan: FN) obtained by the proposed approach [Fractal-(EM/ICM) Markovian CD model].

the F-measure for each estimation/segmentation combination (which, in fact, means that the difference map is reliable, stable over different multimodalities and has also good properties in terms of noise [Gaussian-distributed and not too mixed], see Fig. 8). In terms of implementation and computational complexity, the EM/ICM gives us a good compromise between simplicity of implementation and precision. A comparison with different state-of-the-art approaches [6], [10], [12], [15], [34], [39] is summarized in Table IV. We have also summarized in Table II the percentage, obtained by our proposed Fractal- (EM/ICM) Markovian CD model, in terms of TN, TP, FP, FN. From Figs. 5–7 and Table IV, we can see that the rate accuracy of our method performs very well and is able to give consistent good results comparatively to the other state-of-the-art approaches. The average accuracy rate obtained on the ten multimodal data set based on our fractal and Markovian CD approach is 92.50% with an F-measure equals to 0.560 thus with well-balanced confusion matrices.

C. Discussion

We have compared our MCD model (including the averaging of three difference maps respectively generated for $N_W \in \{8, 12, 16\}$, see Section II-C) and the score obtained with the averaging of two difference maps (generated for $N_W = \{8, 16\}$) or the one obtained for a single difference map generated for $N_W = 8$. For two maps, we obtain: Accuracy = 92.298 and a mean F-measure = 0.551 and for one map, Accuracy = 91.99 and a mean F-measure = 0.534. A slight

TABLE III
SCORE (MEAN ACCURACY RATE, MEAN F-MEASURE, AND THE OVERALL COMPUTATION TIME IN SECONDS) FOR THE CD ESTIMATION OF THE TEN HETEROGENEOUS DATA SETS CONSIDERED IN TABLE I WITH DIFFERENT BAYESIAN OR MARKOVIAN ESTIMATION/SEGMENTATION STRATEGIES

Estimation/Segmentation	Accuracy	F-Measure	Time
EM/Bayes Minimum Error	91.70	0.499	504
SEM/Bayes Minimum Error	91.69	0.498	516
EM/ICM	92.50	0.560	507
EM/MPM	92.60	0.567	555
EM/SMAP	92.63	0.567	551
SEM/ICM	92.46	0.554	522
SEM/MPM	92.57	0.565	537
SEM/SMAP	92.63	0.567	558

increase in gain is observed in the case of a combination of several difference maps.

An interesting experimentation is to apply the proposed approach to two multimodal images without any “change” zone and to visualize the difference (or probability change) map with the result of detection obtained. This can be done easily by extracting (for example, from data set-1 and data set-2) the same part of the image showing no changed area (see Fig. 9). This experiment gives us, from the difference map, the fractal encoding/decoding error which is, for the same part of the image, quite similar to (the same part of) the difference map shown in Fig. 5 and also, from the obtained

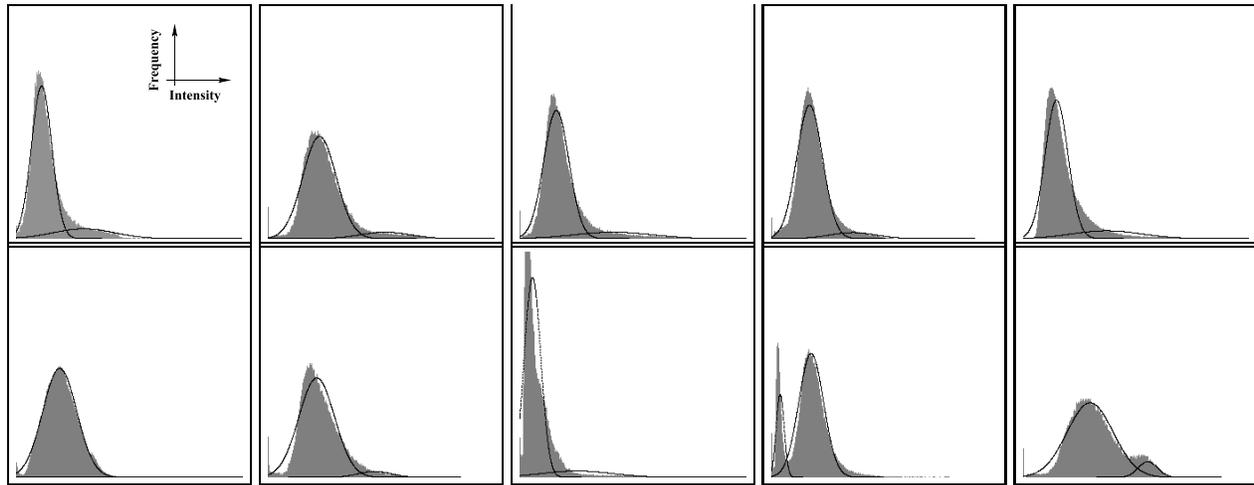


Fig. 8. Histograms (with luminance intensity on the x-axis versus the number (or proportion) of pixels on the y-axis) of the probability change (or difference) map and mixture of two Gaussian likelihood distributions for the two possible class labels; “change” and “no change” estimated by the EM algorithm, on the ten heterogeneous data sets.

TABLE IV

ACCURACY RATE OF CHANGE DETECTION ON THE TEN HETEROGENEOUS DATA SETS OBTAINED BY THE PROPOSED METHOD (EM/ICM) AND THE STATE-OF-THE-ART MULTIMODAL CHANGE DETECTORS (FIRST UPPER PART OF EACH TABLE) AND MONO-MODAL CHANGE DETECTORS (SECOND LOWER PART OF EACH TABLE) [12], [13], [31], [57], [15]–[17], [6], [10], [20], [24]–[26], [34], [39], [58]–[60] IN DECREASING ORDER OF ACCURACY WITH THE BEST RESULT INDICATED IN FIRST AND IN BOLD

NIR Thermic/Optical [#1]	Accuracy
Zhang et al. [20]	0.975
Touati et al. [60]	0.964
Touati et al. [34]	0.942
Proposed method	0.928
PCC [20]	0.882
Touati et al. [58]	0.847

Optical/SAR [#2]	Accuracy
Proposed method	0.971
Touati et al. [60]	0.955
Touati et al. [6]	0.949
Touati et al. [34],[58]	0.943
Touati et al. [10]	0.932
Touati et al. [59]	0.918
PrenDES et al. [16]	0.844
Correlation [16]	0.670
Mutual Inf. [16]	0.580

Optical/Optical [#4]	Accuracy
Touati et al. [34]	0.877
Touati et al. [10]	0.870
Touati et al. [60]	0.862
Touati et al. [6]	0.853
PrenDES et al. [16],[17]	0.844
Proposed method	0.838
Mutual Inf. [16],[17]	0.759
Pixel Dif. [17],[57]	0.708
Correlation [16],[17]	0.679
Pixel Ratio [17],[57]	0.661

SAR 3/5-looks [#5]	Accuracy
Proposed method	0.830
Touati et al. [59]	0.823
Chatelain et al. [13]	0.749
Ratio edge [13]	0.737
Correlation [13]	0.713

SAR/Optical [#3]	Accuracy
PrenDES et al. [17],[61]	0.918
Proposed method	0.909
Touati et al. [60]	0.909
Touati et al. [59]	0.897
Touati et al. [34],[58]	0.881
Touati et al. [34]	0.878
Touati et al. [6]	0.867
PrenDES et al. [15]	0.854
Copulas [12],[15]	0.760
Pixel Ratio [15],[57]	0.813
Pixel Dif. [15],[57]	0.782
Mutual Inf. [12],[15]	0.768
Correlation [12],[15]	0.688

Optical/SAR [#6]	Accuracy
Proposed method	0.952
Luppino et al. [39]	0.933

SAR/Optical [#7]	Accuracy
Zhan et al. [26]	0.979
Proposed method	0.940

SAR/Optical [#8]	Accuracy
Liu et al. [24]	0.976
Touati et al. [34]	0.967
Proposed method	0.942
Touati et al. [58]	0.884
PCC [24]	0.821

SAR/SAR [#9]	Accuracy
Zhan et al. [26]	0.991
Proposed method	0.979

Optical/Optical [#10]	Accuracy
Liu et al. [25]	0.957-0.964
Proposed method	0.962
Touati et al. [59]	0.858
Liu et al. [31]	0.818

detection, almost the same false positives and true negatives (respectively in blue and white in the confusion map) obtained in Fig. 5(e). This demonstrates us that the proposed algorithm

is quite stable, numerically speaking, knowing, that in these two cases, the set of self-similarities between each range and domain blocks (and stored in the fractal file) as well as the

selection of the best “noncontaminated” domain blocks are necessarily different.

Most (nine-tenths) of the time of our procedure is taken by its first fractal encoding/decoding step but it worths mentioning that this part could be easily parallelizable on graphics processing units (GPU). Indeed, block matching on GPU is an extensively studied problem since this one is useful for many vision, graphic or robotic applications (stereovision, tracking, compression, robotic localization, 3D reconstruction, graphic rendering, etc.).

Let us underline that if the image has more than three bands, our strategy requires to use any dimensionality reduction model such as [62] and [63] to reduce beforehand the image into one or three bands. Nevertheless, it should be mentioned that our fractal projection strategy remains easily generalizable to accept input multispectral (or hyperspectral or multiband) raw data without requiring a dimensional reduction step. In this case, the matching process for self-similar parts of the image (between range and domain data blocks) at the first fractal encoding step can be done simply from (possibly spectral) data volumes (possibly always in the LS sense) but at a higher computational cost. The second (fractal projection/decoding) step is straightforward and simple since it essentially consists of averaging several volumes of data rather than blocks of image data.

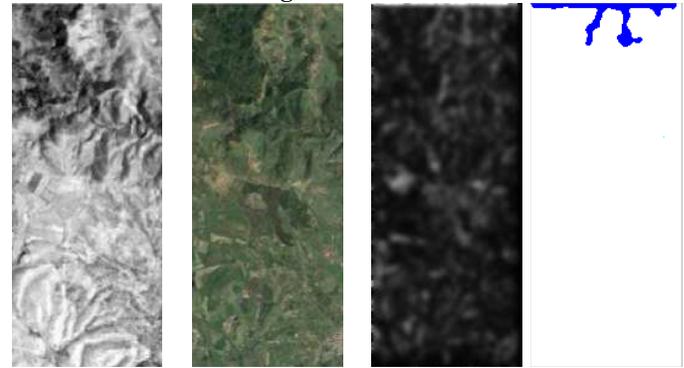
Let us finally add that our model is perfectible, maybe by using more difference maps to be averaged and more likely if we improve the fractal coding/decoding model; especially the partitioning step of the range domain. As proposed in Section II-B1, we can replace the simple (nonadaptive) and regular partitioning by a more complex (and adaptive to the image content) Quad-tree, HV-partitioning, or triangular partitioning scheme [41]. Finally, and as already said in Section II-B1, another interesting improvement to be considered would be to combine a projection method based on regression (luminance transformation) with our fractal-based projection method based solely on spatial transformations.

The C++ code of the overall MCD proposed model (and outlined in pseudocode in Algorithms 1 and 2 running on Linux, data, and all that is necessary for reproduction of the results shown in this article is freely accessible at <http://www.iro.umontreal.ca/~mignotte/ResearchMaterial>

IV. CONCLUSION

In this article, we have addressed the problem of change detection in heterogeneous remote sensing. To this end, we have explored a new parametric mapping strategy based on a spatial fractal decomposition and a contractive mapping approach, allowing us to project the before image on any after imaging modality type. This projection exploits the fact that any satellite image data can be approximately encoded in terms of properly spatially transformed parts of itself. We have seen that this piecewise self-similarity property of any image, which is the redundancy that fractal image compression schemes takes advantage, also remains quite invariant to a given imaging modality type and can be used, after some modification and without supervised learning step, to map the before image on

Dataset-1 without change area



Dataset-2 without change area

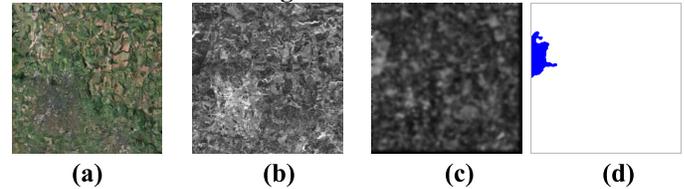


Fig. 9. Heterogeneous data sets (a, b) image t_1 , t_2 (cropped image extracted from the “no-change” area in data sets 1 and 2, see Fig. 5), (c) difference map, and (d) confusion map (white: TN, red: TP, blue: FP, cyan: FN) obtained by the proposed approach [Fractal-(EM/ICM) Markovian CD model].

the imaging modality of the after image. Once the projection is performed and that a pixelwise difference map between the two images (presented in the same imaging modality) is then binarized in the unsupervised Bayesian framework. At this stage, we have noticed that the EM/ICM gives us a good compromise between implementation simplicity and precision for a wide diversity of heterogeneities, changed events and for images acquired at different resolution levels.

ACKNOWLEDGMENT

The author is very grateful and thanks all the researchers who kindly shared their *multimodal* data set and made their databases available for the study and the comparisons made in this article.

The author would also like to thank the anonymous reviewers for their many valuable comments and suggestions that helped to improve both the technical content and the presentation quality of this article.

REFERENCES

- [1] N. Longbotham *et al.*, “Multi-modal change detection, application to the detection of flooded areas: Outcome of the 2009–2010 data fusion contest,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 331–342, Feb. 2012.
- [2] L. Su, M. Gong, P. Zhang, M. Zhang, J. Liu, and H. Yang, “Deep learning and mapping based ternary change detection for information unbalanced images,” *Pattern Recognit.*, vol. 66, pp. 213–228, Jun. 2017.
- [3] R. Hedjam, M. Kalacska, M. Mignotte, H. Z. Nafchi, and M. Chériet, “Iterative classifiers combination model for change detection in remote sensing imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 6997–7008, Dec. 2016.
- [4] M. A. Lebedev, Y. V. Vizilter, O. V. Vygodov, V. A. Knyaz, and A. Y. Rubis, “Change detection in remote sensing images using conditional adversarial networks,” *ISPRS-Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 42, no. 2, pp. 565–571, 2018. [Online]. Available: <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2/565/2018/>

- [5] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: An overview of methods, challenges, and prospects," *Proc. IEEE*, vol. 103, no. 9, pp. 1449–1477, Aug. 2015.
- [6] R. Touati and M. Mignotte, "An energy-based model encoding nonlocal pairwise pixel interactions for multisensor change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 1046–1058, Feb. 2018.
- [7] V. Alberga, "Similarity measures of remotely sensed multi-sensor images for change detection applications," *Remote Sens.*, vol. 1, no. 3, pp. 122–143, 2009.
- [8] D. Brunner, G. Lemoine, and L. Bruzzone, "Earthquake damage assessment of buildings using VHR optical and SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2403–2420, May 2010.
- [9] G. Liu, J. Delon, Y. Gousseau, and F. Tupin, "Unsupervised change detection between multi-sensor high resolution satellite images," in *Proc. 24th Eur. Signal Process. Conf. (EUSIPCO)*, Budapest, Hungary, Aug. 2016, pp. 2435–2439.
- [10] R. Touati, M. Mignotte, and M. Dahmane, "A new change detector in heterogeneous remote sensing imagery," in *Proc. 7th Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Montreal, QC, Canada, Nov. 2017, pp. 1–6.
- [11] G. Mercier, G. Moser, and S. Serpico, "Conditional copula for change detection on heterogeneous SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2007, pp. 2394–2397.
- [12] G. Mercier, G. Moser, and S. B. Serpico, "Conditional copulas for change detection in heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1428–1441, May 2008.
- [13] F. Chatelain, J.-Y. Tourneret, and J. Inglada, "Change detection in multisensor SAR images using bivariate gamma distributions," *IEEE Trans. Image Process.*, vol. 17, no. 3, pp. 249–258, Mar. 2008.
- [14] B. Storvik, G. Storvik, and R. Fjørtoft, "On the combination of multisensor data using meta-Gaussian distributions," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2372–2379, Jul. 2009.
- [15] J. Prendes, M. Chabert, F. Pascal, A. Giros, and J.-Y. Tourneret, "A new multivariate statistical model for change detection in images acquired by homogeneous and heterogeneous sensors," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 799–812, Mar. 2015.
- [16] J. Prendes, M. Chabert, F. Pascal, A. Giros, and J.-Y. Tourneret, "Performance assessment of a recent change detection method for homogeneous and heterogeneous images," *Revue Française Photogrammétrie et Télédétection*, vol. 209, pp. 23–29, Jan. 2015.
- [17] J. Prendes, "New statistical modeling of multi-sensor images with application to change detection," Ph.D. dissertation, Univ. Toulouse, Toulouse, France, 2015.
- [18] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. L. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1822–1835, Jun. 2008.
- [19] P. Du, S. Liu, J. Xia, and Y. Zhao, "Information fusion techniques for change detection from multi-temporal remote sensing images," *Inf. Fusion*, vol. 14, no. 1, pp. 19–27, Jan. 2013.
- [20] P. Zhang, M. Gong, L. Su, J. Liu, and Z. Li, "Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 116, pp. 24–41, Jun. 2016.
- [21] M. Gong, P. Zhang, L. Su, and J. Liu, "Coupled dictionary learning for change detection from multisource data," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7077–7091, Dec. 2016.
- [22] W. Zhao, Z. Wang, M. Gong, and J. Liu, "Discriminative feature learning for unsupervised change detection in heterogeneous images based on a coupled neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7066–7080, Dec. 2017.
- [23] N. Merkle, P. Fischer, S. Auer, and R. Müller, "On the possibility of conditional adversarial networks for multi-sensor image matching," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Fort Worth, TX, USA, Jul. 2017, pp. 1–4.
- [24] J. Liu, M. Gong, K. Qin, and P. Zhang, "A deep convolutional coupling network for change detection based on heterogeneous optical and radar images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 545–559, Mar. 2018.
- [25] Z. Liu, G. Li, G. Mercier, Y. He, and Q. Pan, "Change detection in heterogeneous remote sensing images via homogeneous pixel transformation," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1822–1834, Apr. 2018.
- [26] T. Zhan, M. Gong, J. Liu, and P. Zhang, "Iterative feature mapping network for detecting multiple changes in multi-source remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 146, pp. 38–51, Dec. 2018.
- [27] R. Touati, M. Mignotte, and M. Dahmane, "Anomaly feature learning for unsupervised change detection in heterogeneous images: A deep sparse residual model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 588–600, 2020.
- [28] M. Xu *et al.*, "Change detection of the tangjiashan barrier lake based on multi-source remote sensing data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 4, Jul. 2009, pp. IV-303–IV-306.
- [29] M. Volpi, F. de Morsier, G. Camps-Valls, M. Kanevski, and D. Tuia, "Multi-sensor change detection based on nonlinear canonical correlations," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2013, pp. 1944–1947.
- [30] C. Wu, B. Du, and L. Zhang, "Slow feature analysis for change detection in multispectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2858–2874, May 2014.
- [31] Z.-G. Liu, G. Mercier, J. Dezert, and Q. Pan, "Change detection in heterogeneous remote sensing images based on multidimensional evidential reasoning," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 168–172, Jan. 2014.
- [32] M. Volpi, G. Camps-Valls, and D. Tuia, "Spectral alignment of multi-temporal cross-sensor images with automated kernel canonical correlation analysis," *ISPRS J. Photogramm. Remote Sens.*, vol. 107, pp. 50–63, Sep. 2015.
- [33] X. Chen, J. Li, Y. Zhang, and L. Tao, "Change detection with multi-source defective remote sensing images based on evidential fusion," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. III-7, pp. 125–132, Jun. 2016.
- [34] R. Touati, M. Mignotte, and M. Dahmane, "Change detection in heterogeneous remote sensing images based on an imaging modality-invariant MDS representation," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Athens, Greece, Oct. 2018, pp. 3998–4002.
- [35] L. T. Luppino *et al.*, "A clustering approach to heterogeneous change detection," in *Proc. 20th Scand. Conf. Image Anal. (SCIA)*, Tromsø, Norway, Jun. 2017, pp. 181–192.
- [36] L. T. Luppino, F. M. Bianchi, G. Moser, and S. N. Anfinsen, "Remote sensing image regression for heterogeneous change detection," *Sep. 2018*, pp. 1–6, doi: [10.1109/MLSP.2018.8517033](https://doi.org/10.1109/MLSP.2018.8517033).
- [37] Z.-G. Liu, L. Zhang, G. Li, and Y. He, "Change detection in heterogeneous remote sensing images based on the fusion of pixel transformation," in *Proc. 20th Int. Conf. Inf. Fusion (Fusion)*, Xi'an, China, Jul. 2017, pp. 1–6.
- [38] D. Tuia, D. Marcos, and G. Camps-Valls, "Multi-temporal and multi-source remote sensing image classification by nonlinear relative normalization," *ISPRS J. Photogramm. Remote Sens.*, vol. 120, pp. 1–12, Oct. 2016.
- [39] L. T. Luppino, F. M. Bianchi, G. Moser, and S. N. Anfinsen, "Unsupervised image regression for heterogeneous change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9960–9975, Dec. 2019, doi: [10.1109/TGRS.2019.2930348](https://doi.org/10.1109/TGRS.2019.2930348).
- [40] M. Barnsley, *Fractals Everywhere*. San Diego, CA, USA: Academic, 1988.
- [41] Y. Fisher, Ed., *Fractal Image Compression: Theory and Application*. London, U.K.: Springer-Verlag, 1995.
- [42] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Kerkyra, Greece, Sep. 1999, pp. 1033–1038.
- [43] M. Mignotte, "Nonparametric multiscale energy-based model and its application in some imagery problems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 184–197, Feb. 2004.
- [44] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, Jan. 2005.
- [45] M. Ebrahimi and E. R. Vrscay, "Solving the inverse problem of image zooming using 'self-examples,'" in *Image Analysis and Recognition*. Berlin, Germany: Springer, 2007, pp. 117–130.
- [46] M. Mignotte, "A non-local regularization strategy for image deconvolution," *Pattern Recognit. Lett.*, vol. 29, no. 16, pp. 2206–2212, Dec. 2008.
- [47] A. Wong and J. Orchard, "A nonlocal-means approach to exemplar-based inpainting," in *Proc. 15th IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 2600–2603.
- [48] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 36–51, Jan. 2009.

- [49] A. Buades, B. Coll, J.-M. Morel, and C. Sbert, "Self-similarity driven color demosaicking," *IEEE Trans. Image Process.*, vol. 18, no. 6, pp. 1192–1202, Jun. 2009.
- [50] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," *J. ACM*, vol. 45, no. 6, pp. 891–923, Nov. 1998.
- [51] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc., B, Methodol.*, vol. 39, no. 1, pp. 1–22, 1977.
- [52] P. Masson and W. Pieczynski, "SEM algorithm and unsupervised statistical segmentation of satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 3, pp. 618–633, May 1993.
- [53] M. Mignotte, "A segmentation-based regularization term for image deconvolution," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1973–1984, Jul. 2006.
- [54] J. Besag, "On the statistical analysis of dirty pictures," *J. Roy. Stat. Soc., B, Methodol.*, vol. 48, no. 3, pp. 259–279, Jul. 1986.
- [55] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *J. Amer. Stat. Assoc.*, vol. 82, no. 397, pp. 76–89, 1987.
- [56] C. A. Bouman and M. Shapiro, "A multiscale random field model for Bayesian image segmentation," *IEEE Trans. Image Process.*, vol. 3, no. 2, pp. 162–177, Mar. 1994.
- [57] ORB Development Team. (2014). *The ORFEO Toolbox Software Guide*. Accessed: Apr. 20, 2010. [Online]. Available: <http://orfeo-toolbox.org/>
- [58] R. Touati, M. Mignotte, and M. Dahmane, "A reliable mixed-norm-based multiresolution change detector in heterogeneous remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3588–3601, Sep. 2019.
- [59] R. Touati, M. Mignotte, and M. Dahmane, "Multimodal change detection using a convolution model-based mapping," in *Proc. 9th Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Istanbul, Turkey, Nov. 2019, pp. 1–6.
- [60] R. Touati, M. Mignotte, and M. Dahmane, "Multimodal change detection in remote sensing images using an unsupervised pixel pairwise-based Markov random field model," *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 757–767, Jan. 2020.
- [61] J. Prendes, M. Chabert, F. Pascal, A. Giros, and J.-Y. Tourneret, "Change detection for optical and radar images using a Bayesian nonparametric model coupled with a Markov random field," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Brisbane, QLD, Australia, Apr. 2015, pp. 1513–1517.
- [62] M. Mignotte, "A multiresolution Markovian fusion model for the color visualization of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4236–4247, Dec. 2010.
- [63] M. Mignotte, "A bicriteria-optimization-approach-based dimensionality-reduction model for the color display of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 501–513, Jan. 2012.



Max Mignotte received the DEA (postgraduate degree) in digital signal, image and speech processing from the INPG University, Grenoble, France, in 1993 and the Ph.D. degree in electronics and computer engineering from the University of Bretagne Occidentale (UBO), Brest, France, and the Digital Signal Laboratory (GTS) of the French Naval Academy, France, in 1998. He was an INRIA Post-Doctoral Fellow with the University of Montreal (DIRO), Montreal, QC, Canada, from 1998 to 1999.

He is currently with DIRO at the Computer Vision & Geometric Modeling Lab as a Professor with the University of Montreal. He is also a member of LIO (Laboratoire de recherche en imagerie et orthopedie, Center de recherche du CHUM, Hopital Notre-Dame) and Researcher at CHUM. His research interests include statistical methods, Bayesian inference, and energy-based models (especially encoding nonlocal pairwise pixel interactions) for solving diverse large-scale high-dimensional ill-posed inverse problems in imaging.