CHANGE DETECTION FOR OPTICAL AND RADAR IMAGES USING A BAYESIAN NONPARAMETRIC MODEL COUPLED WITH A MARKOV RANDOM FIELD

Jorge Prendes[1,2], *Marie Chabert* [1,3], *Frédéric Pascal*[2], *Alain Giros*[4], *Jean-Yves Tourneret*[1,3]

[1] TéSA Laboratory, 7 boulevard de la Gare, 31500 Toulouse, France

[2] Supélec - SONDRA, Plateau du Moulon, 3 rue Joliot-Curie, F-91192 Gif-sur-Yvette Cedex, France

[3] University of Toulouse, INP/ENSEEIHT - IRIT, 2 rue Charles Camichel, BP 7122, 31071 Toulouse Cedex 7, France

[4] CNES, 18 Av. Edouard Belin, 31401 Toulouse, France

{jorge.prendes, marie.chabert, jean-yves.tourneret}@enseeiht.fr, frederic.pascal@supelec.fr, alain.giros@cnes.fr

ABSTRACT

This paper introduces a Bayesian non parametric (BNP) model associated with a Markov random field (MRF) for detecting changes between remote sensing images acquired by homogeneous or heterogeneous sensors. The proposed model is built for an analysis window which takes advantage of the spatial information via an MRF. The model does not require any *a priori* knowledge about the number of objects contained in the window thanks to the BNP framework. The change detection strategy can be divided into two steps. First, the segmentation of the two images is performed using a region based approach. Second, the joint statistical properties of the objects in the two images allows an appropriate manifold to be defined. This manifold describes the relationships between the different sensor responses to the observed scene and can be learnt from a training unchanged area. It allows us to build a similarity measure between the images that can be used in many applications such as change detection or image registration. Simulation results conducted on synthetic and real optical and synthetic aperture radar (SAR) images show the efficiency of the proposed method for change detection.

Index Terms— Change detection, Bayesian non parametric, Markov random field, Markov chain Monte Carlo, remote sensing.

1. INTRODUCTION

Many practical applications require the joint analysis of images acquired by multiple sensors. Depending on the situation (particular features to be observed or external constraints), different kinds of sensors can be used for imaging a scene of interest. After acquiring the images, computing an appropriate similarity measure between these images is an important step for change detection, image registration or database updating. The similarity measures available in the literature are mainly devoted to images acquired by the same kind of sensors (denoted as homogeneous images in this paper). For instance, similarity measures designed for optical images include the so-called difference image derived from image pixels [1–4], wavelet coefficients [5, 6], and region based approaches using Markov Random Fields (MRF) [7]. On the other hand, similarity measures proposed for SAR images can be based on the the log-ratio image [8–11] or on multivariate probability distributions [12–14]. Similarity measures that do not target a particular sensor type include the correlation coefficient [15] for homogeneous sensors and the mutual information [15–18] or the copulas theory [19] for heterogeneous sensors. Note that the measure based on copulas initially investigated in [19] cannot be generalized easily to situations where more than two images have to be compared.

A flexible method has been recently proposed to derive a similarity measure between images acquired by homogeneous and heterogeneous sensors [20, 21]. This method considers a manifold that describes the joint behavior of the sensors. Based on the physical properties of the sensors and on the observed scene, a multivariate statistical model was proposed for the pixel intensities within the analysis window. The proposed model was based on a mixture distribution representing each object of the analyzing window by one of the mixture components. The mixture was composed of a finite but unknown number of components associated with the objects included in the analyzing window. The mixture parameters were then estimated by means of a modified expectation-maximization (EM) algorithm [22] allowing the number of components within a predefined range to be estimated. Note that the choice of this range was related to the observation window size and to the kind of observed scene (urban area, coast, etc.).

This paper introduces a new Bayesian model based on specific priors taking advantage of the correlations between adjacent pixels in the estimation window by means of an MRF and mitigating the absence of information about the number of components in this window by using a Bayesian nonparametric distribution. More precisely, a Bayesian non parametric (BNP) model allows us to estimate the number of mixture components through a prior based on a Chinese restaurant process (CRP). The BNP model is combined with an MRF to take into account the image spatial correlation [23]. This is an improvement with respect to [20, 21], which relies on a heuristic estimation of the number of components, and does not account for spatial correlation between pixels. Accounting for the spatial correlation leads to a significant increase in the parameter estimation accuracy and thus in the change detection performance. Moreover, due to its robustness to the number of objects contained in the analysis window, the BNP framework makes it possible to increase the size of the analysis window. As a consequence, a local analysis can now be performed on regions rather than on small sliding windows, resulting in higher resolution change maps.

The paper is organized as follows: Section 2 formulates the problem and recalls previous works. Section 3 defines the new statistical model and the corresponding parameter estimation algorithm while Section 4 defines the change detection strategy. Section 5 evaluates the performance of the proposed change detector for synthetic

The authors would like to thank the French space agency (CNES) in Toulouse for its financial and technical support. Part of this work has been supported by the BNPSI ANR Project n° ANR-13- BS-03-0006-01 and by ANR-11-LABX-0040-CIMI in particular during the program ANR-11-IDEX-0002-02 within the thematic trimester on image processing.

and real data. Conclusions are reported in Section 6.

2. PROBLEM FORMULATION AND PREVIOUS WORK

A flexible change detection strategy for images acquired by heterogeneous sensors was introduced in [20, 21]. This strategy relies on a multivariate statistical model for the pixel intensities observed through multiple sensors. The pixel intensity observed through the sensor S is denoted as

$$I_S = f_S[T_S(P), \eta_S] \tag{1}$$

where $T_S(P)$ is the manifestation, through sensor S, of the physical properties P of the observed scene, whereas $f(\cdot,\cdot)$ models the corruption of $T_S(P)$ with the measurement noise denoted as η_S. For a given object with physical properties P, this model leads to the probability density function (pdf) of $I_S|T_S(P)$ (which is equivalent to $I_S|P$). For instance, in the case of optical images

$$I_{\text{Opt}}|P \sim \mathcal{N}\big[T_{\text{Opt}}(P), \sigma^2\big]. \tag{2}$$

This model can be extended to a set of D sensors, $S_1, \ldots, S_D$, assuming that the measurement noises affecting the different sensors $\eta_{S_1}, \ldots, \eta_{S_D}$ are independent. The joint distribution of the set of observed pixel intensities $\boldsymbol{I} = \{I_{S_d}\}_{d=1}^{D}$ is thus the product of their marginal distributions

$$p(\boldsymbol{I}|P) = p(I_{S_1}, \ldots, I_{S_D}|P) = \prod_{d=1}^{D} p(I_{S_d}|P). \tag{3}$$

The model is applied on an analysis window W to obtain $p(\boldsymbol{I}|W)$. Taking into account that the number of objects in the window (and thus the amount of different values of P) is finite, $P|W$ can be modeled as a discrete random variable. Consequently, $p(\boldsymbol{I}|W)$ is a mixture distribution [20, 21] whose components correspond to the values of P denoted as $\{P_k\}_{k=1}^{K}$ where K is the number of objects in the window W

$$p(\boldsymbol{I}|W) = \sum_{k=1}^{K} w_k\, p(\boldsymbol{I}|P_k). \tag{4}$$

The vector $\boldsymbol{v}(P) = [T_{S_1}(P), \ldots, T_{S_D}(P)]$ can be estimated for each mixture component (i.e., for each object in the analysis window) from the observed pixel intensities. As explained in [20, 21], this vector belongs to a manifold that relates the properties of the different sensors. The distance between an estimated vector $\boldsymbol{v}(P)$ denoted as $\widehat{\boldsymbol{v}}(P)$ and the manifold can be used as a similarity measure to detect changes. This distance is *a priori* unknown. However it can be estimated as the density of $\widehat{\boldsymbol{v}}(P)$ obtained from a learning set of unchanged samples.

In [20, 21], $\boldsymbol{v}(P)$ was estimated from the observed data using a modified version of the EM algorithm [22] that estimated the number of components K within a predefined range $[K_{\min}, K_{\max}]$. This range was fixed *a priori* as functions of the size of the analysis window and the observed scene. Moreover, a limitation of the algorithm proposed in [20,21] is that it was not exploiting the spatial correlation between adjacent pixels of the estimation window. For instance, a random permutation of the pixels within the analysis window produced the same estimates, in presence or absence of spatial correlation.

3. A NEW BAYESIAN NONPARAMETRIC MODEL BASED ON A MARKOV RANDOM FIELD

3.1. Bayesian non parametric model

As already mentioned, the number of components of the mixture distribution associated with an analysis window is unknown but finite since there is a limited number of objects (although selected from an infinite set of possible objects) within the analysis window. This property can be considered through a Dirichlet process mixture (DPM) [24] defined for each set of intensities $\{\boldsymbol{I}_n\}_{n=1}^{N}$ in the window W as follows

$$\boldsymbol{I}_n|P_n \sim M(\boldsymbol{\theta}_n) \tag{5}$$

$$\boldsymbol{\theta}_n \sim \Theta \tag{6}$$

$$\Theta \sim \text{DP}(\Theta_0(\boldsymbol{\psi}), \alpha) \tag{7}$$

where $\boldsymbol{\theta}_n$ is the parameter vector of the distribution M, whose density is (3), P_n is the value of P that produced I_n, $\Theta_0(\boldsymbol{\psi})$ is the prior distribution of the parameters $\boldsymbol{\theta}_n$ with hyperparameters vector $\boldsymbol{\psi}$ and α is the DP concentration parameter. Each realization Θ of $\text{DP}(\Theta_0(\boldsymbol{\psi}), \alpha)$ is a discrete distribution with pdf

$$p(\boldsymbol{\theta}) = \sum\nolimits_{k=1}^{K} w_k \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_k) \tag{8}$$

where $\delta(\cdot)$ is the Dirac delta function, w_k are the weights in (4) and K is a random variable with support $\mathbb{N}^+$. It has been shown [25] that $\boldsymbol{\theta}_n$ exhibits a clustering property and that (5), (6) and (7) are equivalent to the so-called Chinese restaurant process (CRP) mixture

$$\boldsymbol{I}_n|z_n = k \sim M(\boldsymbol{\theta}_k) \tag{9}$$

$$z_n \sim \text{CRP}(\alpha) \tag{10}$$

$$\boldsymbol{\theta}_k \sim \Theta_0(\boldsymbol{\psi}) \tag{11}$$

where z_n is a discrete random variable representing the label of a partition of $\boldsymbol{\theta}_n$ into an unbounded number K of classes. Because of (10), the distribution of the label z_n is defined as

$$p(z_n = k) = \begin{cases} \frac{\alpha}{N+\alpha}, & \text{if } z_n = 0 \\ \frac{N_k}{N+\alpha}, & \text{otherwise} \end{cases} \tag{12}$$

where N_k is the number of samples $\boldsymbol{I}_n$ assigned to the class k, and $z_n = 0$ means that a new class is created. This approach has been widely studied [24,26], and presented as a BNP approach to consider an unknown number of components in a mixture model.

3.2. Markov random field

In [23], a DP mixture is coupled with an MRF prior on $\boldsymbol{z} = \{z_n\}_{n=1}^{N}$ to impose spatial smoothness for image segmentation. The joint distribution of the group of random variables $\boldsymbol{z}$, $p(\boldsymbol{z})$ with a neighborhood graph G is called an MRF if the distribution of a particular variable z_n conditional to all the other variables, denoted $\boldsymbol{z}_{-n}$, is the same as the the distribution conditioned to a reduced set of variables $z_{\delta(n)}$, where $\delta(n)$ denotes the neighborhood of n, i.e.,

$$p(z_n|\boldsymbol{z}_{-n}) = p\big(z_n\big|\boldsymbol{z}_{\delta(n)}\big). \tag{13}$$

However, this relation is usually difficult to verify. The Hammersley-Clifford theorem [27] states that $p(\boldsymbol{z})$ satisfies (13) if and only if it can be factorized over the cliques $\mathcal{C}$ of G. If we consider a cost function $H(\boldsymbol{z})$ such that $p(\boldsymbol{z}) \propto \exp[H(\boldsymbol{z})]$, this condition is

$$H(\boldsymbol{z}) = \sum\nolimits_{C \subset \mathcal{C}} H_C(\boldsymbol{z}_C) \tag{14}$$

where $\boldsymbol{z}_C = \{z_n\}_{n \in C}$ and $H_C(\cdot)$ is a local cost function. If we denote as $H(z_n|\boldsymbol{z}_{-n})$ all the terms in $H(\boldsymbol{z})$ involving z_n, i.e.,

$$H(z_n|\boldsymbol{z}_{-n}) = \sum\nolimits_{C \subset \mathcal{C} | n \in C} H_C(\boldsymbol{z}_C) \tag{15}$$

then the conditional pdf $p(z_n|\boldsymbol{z}_{-n})$ can be expressed as

$$p(z_n|\boldsymbol{z}_{-n}) \propto \exp[H(z_n|\boldsymbol{z}_{-n})]. \tag{16}$$

Since z_n is a discrete random variable representing a partition label, we are interested in evaluating whether two samples $\boldsymbol{I}_n$ and $\boldsymbol{I}_m$ belong to the same class, i.e., if $z_n = z_m$. This can be evaluated

with a cost function satisfying (14) as follows

$$H(z_n|z_{-n}) = \sum\nolimits_{m\in\delta(n)} \omega_{nm} \mathbb{1}_{z_n}(z_m) \tag{17}$$

where $\mathbb{1}_{z_n}(z_m)$ is an indicator function taking the value of 1 when $z_n = z_m$ and 0 otherwise, ω_{mn} is a weight relating the samples $\boldsymbol{I}_n$ and $\boldsymbol{I}_m$, and $\delta(n)$ defines a neighborhood of the n-th pixel constructed as the group of samples $\boldsymbol{I}_m$ with a spatial L^∞ distance smaller than a certain threshold (i.e., a square region around the pixel of interest). In this paper, we propose to consider the following weight

$$\omega_{nm} \propto \exp\left[-\frac{(x_n - x_m)^2 + (y_n - y_m)^2}{d^2}\right] \tag{18}$$

where d is a parameter related the neighborhood size (indicating how fast the influence between two pixels decrease with the distance) and $[x_n, y_n]$, $[x_m, y_m]$ are the 2D coordinates of the pixels associated with the intensities $\boldsymbol{I}_n$ and $\boldsymbol{I}_m$, while the neighborhood $\delta(n)$ is limited to pixels with a L^∞ distance of less than $5d$.

This framework can be integrated with a DPM to produce a DPM-MRF mixture. In [23] it is proved that the product of (12) and (16) satisfies the condition (13). A direct consequence of (12) and (16) is the following conditional distribution

$$p(z_n|\boldsymbol{z}_{-n}) \propto \begin{cases} \alpha, & \text{if } z_n = 0 \\ \lambda\, N_k \exp\left(\sum_{m\in\delta_k(n)} \omega_{nm}\right), & \text{otherwise} \end{cases} \tag{19}$$

where $\delta_k(n)$ is the group of samples in the neighborhood of n belonging to the partition k, and λ is a parameter weighting the influence of the MRF, chosen by cross-validation in our simulations.

3.3. A Collapsed Gibbs Sampler

To obtain the vector $\widehat{\boldsymbol{v}}(P)$ associated with each pixel, we must identify the object corresponding to this pixel, i.e., find $\boldsymbol{z} = \arg\max_{\boldsymbol{z}} p(\boldsymbol{z}|\boldsymbol{I})$. Markov chain Monte Carlo (MCMC) algorithms can be used to sample jointly from a group of random variables by sampling each variable conditionally to the others. When the sampled variables are discrete, the mode or maximum of a distribution can be estimated by computing the mode or maximum with the generated samples. Since $\boldsymbol{z}$ has a multivariate discrete distribution, MCMC algorithms are particularly adapted to maximize $p(\boldsymbol{z}|\boldsymbol{I})$ with respect to $\boldsymbol{z}$ by iteratively sampling $p(z_n|\boldsymbol{z}_{-n}, \boldsymbol{I})$ for $1 \leq n \leq N$. We can express this conditional distribution up to a normalizing constant as

$$p(z_n|\boldsymbol{z}_{-n}, \boldsymbol{I}) \propto p(\boldsymbol{I}|z_n, \boldsymbol{z}_{-n}) p(z_n|\boldsymbol{z}_{-n}) \tag{20}$$

where $p(z_n|\boldsymbol{z}_{-n})$ can be obtained from (19), and $p(\boldsymbol{I}|\boldsymbol{z})$ is obtained after marginalizing out $\boldsymbol{\theta}$,

$$p(\boldsymbol{I}|\boldsymbol{z}, \boldsymbol{\psi}) = \int p(\boldsymbol{I}|\boldsymbol{z}, \boldsymbol{\theta})\, p(\boldsymbol{\theta}|\boldsymbol{\psi})\, d\boldsymbol{\theta} \tag{21}$$

$$= \prod\nolimits_{k=1}^{K} \int p\big(\boldsymbol{I}_{\{k\}}|\boldsymbol{\theta}_k\big)\, p(\boldsymbol{\theta}_k|\boldsymbol{\psi})\, d\boldsymbol{\theta}_k \tag{22}$$

$$= \prod\nolimits_{k=1}^{K} p\big(\boldsymbol{I}_{\{k\}}|\boldsymbol{\psi}\big) \tag{23}$$

where $I_{\{k\}} = \{\boldsymbol{I}_n\}_{z_n=k}$ is the group of samples $\boldsymbol{I}_n$ assigned to class k, and $p(\boldsymbol{\theta}_k)$ is obtained from (11). Note that marginalizing out $\boldsymbol{\theta}$, which leads to a collapsed Gibbs sampler, increases the convergence speed, and can reduce the amount of computation at each iteration [28].

By choosing appropriate priors $p(\boldsymbol{\theta}_k|\boldsymbol{\psi})$, the integral can be analytically solved. However, in order to avoid the computation of (22), we can choose a conjugate prior as follows

$$\boldsymbol{\theta} \sim \Theta_0(\boldsymbol{\psi}) \tag{24}$$

$$\boldsymbol{\theta}\big|\boldsymbol{I}_{\{k\}} \sim \Theta_0\Big[\boldsymbol{\psi}'_{\boldsymbol{I}_{\{k\}},\boldsymbol{\psi}}\Big] \tag{25}$$

where the hyperparameters update rule $\boldsymbol{\psi}'_{\boldsymbol{I}_{\{k\}},\boldsymbol{\psi}}$, yields

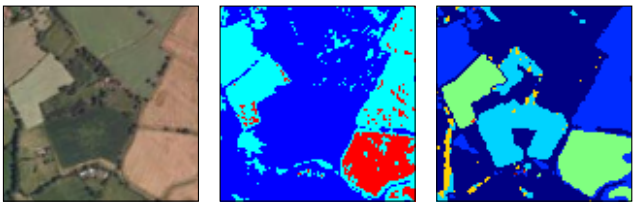
$$p\big(\boldsymbol{I}_{\{k\}}|\boldsymbol{\psi}\big) = \frac{p\big(\boldsymbol{I}_{\{k\}}|\boldsymbol{\theta}_k\big)\, p(\boldsymbol{\theta}|\boldsymbol{\psi})}{p\Big(\boldsymbol{\theta}_k\Big|\boldsymbol{\psi}'_{(\boldsymbol{I}_{\{k\}},\boldsymbol{\psi})}\Big)} \tag{26}$$

where all terms depending on θ have been simplified.

Since we are only interested in $p(z_n|\boldsymbol{z}_{-n}, \boldsymbol{I})$ up to a proportional constant we can replace the term $p(\boldsymbol{I}|\boldsymbol{z}, \boldsymbol{\psi})$ with

$$\frac{p(\boldsymbol{I}|\boldsymbol{z}, \boldsymbol{\psi})}{p(\boldsymbol{I}_{-n}|\boldsymbol{z}_{-n}, \boldsymbol{\psi})} = \frac{p\big(\boldsymbol{I}_{\{k\}}|\boldsymbol{\psi}\big)}{p\big(\boldsymbol{I}_{\{k\}-n}|\boldsymbol{\psi}\big)} \tag{27}$$

where $\boldsymbol{I}_{-n}$ is the set of all samples except $\boldsymbol{I}_n$ (i.e., $\{\boldsymbol{I}_m\}_{m\neq n}$), and $\boldsymbol{I}_{\{k\}-n}$ is the set of all samples in the k-th class excepting $\boldsymbol{I}_n$ (i.e., $\{\boldsymbol{I}_m\}_{m\neq n, z_m\neq k}$). This is particularly useful when $p(\boldsymbol{\theta}|\boldsymbol{\psi})$ belongs to an exponential family, since (27) can be expressed in terms of its sufficient statistics.



(a) Analysis window. (b) Detection by EM [20, 21]. (c) Detection by DPM-MRF (this paper).

Fig. 1. Objects detected in the image (a) using EM (b) and DPM-MRF (c).

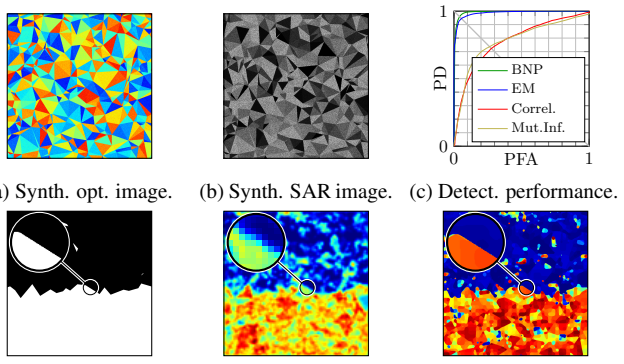
4. REGION BASED CHANGE DETECTION

In [20, 21], a similarity measure was computed for each analysis window. However, the resolution of the change map was directly linked to the size of the estimation window. Moreover, the estimation algorithm was based on a soft pixel classification: each pixel was assigned to different classes with particular weights. Conversely, the method proposed in this paper produces a hard classification linking each pixel to a particular object.

Figure 1 illustrates the difference between both algorithms on a typical example. Fig. 1(a) is an image with 100×100 pixels and 3 spectral bands obtained from Google Earth with a pixel resolution of 7.27m. Figure 1(b) displays the posterior probabilities of the labels estimated using the EM algorithm. As a consequence of EM soft clustering, only 3 classes are identified, while the mixture contains 9 components. Moreover, since the parameter estimation algorithm does not account for spatial correlations, some isolated pixels are miss-classified. In Figure 1(c), 6 classes are identified. The hard classification obtained with the MCMC-DPM approach solves the first issue and the MRF reduces the number of isolated pixels.

This improvement enables the use of the similarity measure at the object level. In [20, 21], each analysis window produced a set of estimated vectors $\{\widehat{\boldsymbol{v}}(P_k)\}_{k=1}^{K}$, with their corresponding similarity measures $\{d_k\}_{k=1}^{K}$. The similarity measure d_n corresponding to a pixel $\boldsymbol{I}_n$ was shared by all the pixels belonging to the same analysis window W, i.e.,

$$d_n = d_W = \sum\nolimits_{k=1}^{K} w_k d_k \tag{28}$$



(a) Synth. opt. image. (b) Synth. SAR image. (c) Detect. performance.

(d) Change mask. (e) Change map (EM). (f) C. map (DPM-MRF).

Fig. 2. Change detection on synthetic heterogeneous images.

With the proposed approach, the similarity measure assigned to each pixel depends on the region it belongs to, so that its similarity measure is simply defined as

$$d_n = d_{z_n} \tag{29}$$

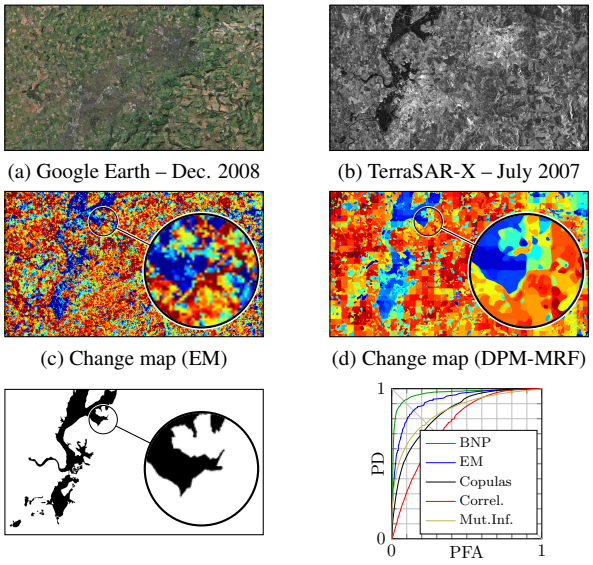
where z_n depends not only on the window, but also on the neighborhood of $\boldsymbol{I}_n$ and on $\boldsymbol{I}_n$ itself. This strategy produces a pixel-level instead of a window-level change detection. This change results on a more detailed estimation of the change map, improving the performance for windows containing adjacent changed and unchanged areas. Note that it would be possible to run the algorithm on the whole image instead of an analysis window. However, this would increase considerably the computational cost of the algorithm.

5. SIMULATION RESULTS

To evaluate the performance of the proposed algorithm we have run different tests on synthetic and real images, and have compared the performance of the DPM-MRF algorithm with the EM algorithm studied in [20, 21] and change detectors based on the correlation coefficient, the mutual information and, in the case of real data, the copulas theory. The performance is evaluated by means of the receiver operating characteristics (ROC) [29] curves.

5.1. Synthetic images

Figures 2(a) and 2(b) shows two synthetic optical and SAR images affected by changes in the upper part of the image as shown in the change mask (d). Both images were generated with the image generation model described in [20, 21]. Figure 2(d) was computed from the ground truth, where black areas indicate the changes. The change map in Fig. 2(e) shows the estimated change maps using our previous algorithm [20, 21], with an analysis window of 20×20 pixels, while Fig. 2(f) shows the change map for the approach described in this paper using a window size of 100×100 pixels. The circular zoom area highlights a typical resolution enhancement obtained by the proposed change detector. The resolution obtained in Fig. 2(e) is limited by the window based approach, where each pixel represents the measured distance for the whole analysis window. This is particularly visible on the edges between changed and unchanged areas, where the analysis window might contain changed and unchanged regions. However, the DPM-MRF algorithm can detect changes using a pixel by pixel approach, increasing the accuracy of the resulting change map as observed in Fig. 2(f). The ROC curves in Fig. 2(c) display the performance improvement obtained with the proposed approach compared with the previous work, dropping the equal error rate (EER, i.e., where the probability of detection PD and the probability of false alarm PFA coincide) from 5.5% to 4.4%, which represents a 20% improvement.



(a) Google Earth – Dec. 2008 (b) TerraSAR-X – July 2007

(c) Change map (EM) (d) Change map (DPM-MRF)

(e) Change mask (f) Detection performance

Fig. 3. Change detection results on heterogeneous images

5.2. Real optical and SAR images

Figures 3(a) and 3(b) show one optical image from Google Earth and one SAR image from a TerraSAR-X satellite, acquired during and after a big flooding in Gloucester (UK) respectively. The change mask in Fig. 3(e) represents our ground truth provided by a photo interpreter, where the black areas indicate the flooded areas. The change maps in Figs. 3(c) and 3(d) show the estimated change maps using our previous algorithm [20, 21] with a 10×10 analysis windows, and the approach described in this paper with a 200×200 analysis window (note that the window size was determined to optimize the processing time). Both figures display a circular zoom area emphasizing the resolution enhancement obtained by the proposed approach. The ROC curves in Fig. 3(f) show more quantitative results. As it can be observed, the use of a BNP model combined with an MRF allows significant performance improvement. The EER drop from EM to DPM-MRF is from 14% to 8%.

6. CONCLUSION

This paper introduced a new Bayesian model for change detection based on a Dirichlet process mixture combined with a Markov random field. This new model allowed to detect changes between heterogeneous images based on the strategy presented in our previous work [20, 21]. The main properties of the proposed change detection strategy are to account for an unknown number of objects in the analysis window using a Chinese restaurant prior and to the presence of spatial correlations between adjacent pixels in the window using a Markov random field. This strategy enables the detection of changes with a region based detection approach mitigating the impact of the size of the estimation window. The results obtained with synthetic and real data are very promising. Future work includes validating the proposed method on a wider range of datasets, reducing the dependency of the manifold estimation on selected training region and estimating the DPM-MRF parameters from the sampled data.

7. REFERENCES

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

[2] T. Fung, "An assessment of TM imagery for land cover change detection," *IEEE Trans. Geosci. and Remote Sensing*, vol. 28, no. 12, pp. 681–684, 1990.

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

[4] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and k-means clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 772–776, 2009.

[5] T. Celik, "Multiscale change detection in multitemporal satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 820–824, 2009.

[6] T. Celik and M. Kai-Kuang, "Unsupervised change detection for satellite images using dual-tree complex wavelet transform," *IEEE Trans. Geosci. and Remote Sensing*, vol. 48, no. 3, pp. 1199–1210, 2010.

[7] T. Sziranyi and M. Shadaydeh, "Segmentation of remote sensing images using similarity measure based fusion-MRF model," *IEEE Geosci. and Remote Sensing Letters*, vol. 11, no. 9, pp. 1544–1548, Sept 2014.

[8] R. Touzi, A. Lopés, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *IEEE Trans. Geosci. and Remote Sensing*, vol. 26, no. 6, pp. 764–773, 1988.

[9] Eric J. M. Rignot and J. Van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Trans. Geosci. and Remote Sensing*, vol. 31, no. 4, pp. 896–906, 1993.

[10] R. Fjortoft, A. Lopés, P. Marthon, and E. Cubero-Castan, "An optimal multiedge detector for SAR image segmentation," *IEEE Trans. Geosci. and Remote Sensing*, vol. 36, no. 3, pp. 793–802, 1988.

[11] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images," *IEEE Trans. Geosci. and Remote Sensing*, vol. 43, no. 4, pp. 874–887, 2005.

[12] F. Chatelain, J.-Y. Tourneret, A. Ferrari, and J. Inglada, "Bivariate gamma distributions for image registration and change detection," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1796–1806, 2007.

[13] G. Quin, B. Pinel-Puyssegur, J.-M. Nicolas, and P. Loreaux, "MIMOSA: An automatic change detection method for SAR time series," *IEEE Trans. Geosci. and Remote Sensing*, vol. 52, no. 9, pp. 5349–5363, 2014.

[14] L. Giustarini, R. Hostache, P. Matgen, G. J.-P. Schumann, P. D. Bates, and D. C. Mason, "A change detection approach to flood mapping in urban areas using TerraSAR-X," *IEEE Trans. Geosci. and Remote Sensing*, vol. 51, no. 4, pp. 2417–2430, 2013.

[15] J. Inglada and A. Giros, "On the possibility of automatic multisensor image registration," *IEEE Trans. Geosci. and Remote Sensing*, vol. 42, no. 10, pp. 2104–2120, Oct. 2004.

[16] J. Inglada, "Similarity measures for multisensor remote sensing images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Toronto, Canada, 2002, vol. 1, pp. 104–106 vol.1.

[17] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis," *IEEE Trans. Geosci. and Remote Sensing*, vol. 45, no. 5, pp. 1432–1445, May 2009.

[18] H. M. Chen, P. K. Varshney, and M. K. Arora, "Performance of mutual information similarity measure for registration of multitemporal remote sensing images," *IEEE Trans. Geosci. and Remote Sensing*, vol. 41, no. 11, pp. 2445–2454, Nov. 2003.

[19] G. Mercier, G. Moser, and S. B. Serpico, "Conditional copulas for change detection in heterogeneous remote sensing images," *IEEE Trans. Geosci. and Remote Sensing*, vol. 46, no. 5, pp. 1428–1441, May 2008.

[20] J. Prendes, M. Chabert, F. Pascal, A. Giros, and J.-Y. Tourneret, "A multivariate statistical model for multiple images acquired by homogeneous or heterogeneous sensors," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc. (ICASSP)*, Florence, Italy, May 2014.

[21] 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, March 2015.

[22] M. A. T. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 381–396, March 2002.

[23] P. Orbanz and J. M. Buhmann, "Nonparametric Bayesian image segmentation," *Int. J. Comput. Vision*, vol. 77, no. 1–3, pp. 25–45, May 2008.

[24] S. J. Gershman and D. M. Blei, "A tutorial on bayesian nonparametric models," *J. Math. Psychol.*, vol. 56, pp. 1–12, 2012.

[25] T. S. Ferguson, "A Bayesian analysis of some nonparametric problems," *Ann. Statist.*, vol. 1, no. 2, pp. 209–230, March 1973.

[26] Y. W. Teh and M. I. Jordan, "Hierarchical Bayesian nonparametric models with applications," Cambridge University Press, 2010.

[27] G. R. Grimmett, "A theorem about random fields," *Bulletin of the London Mathematical Society*, vol. 5, no. 1, pp. 81–84, 1973.

[28] D. A. Van Dyk and T. Park, "Partially collapsed Gibbs samplers: Theory and methods," *J. Amer. Stat. Soc.*, vol. 103, no. 482, pp. 790–796, 2008.

[29] W. W. Peterson, T. G. Birdsall, and W. Fox, "The theory of signal detectability," *IRE Trans. Inf. Theory*, vol. 4, no. 4, pp. 171–212, Sept. 1954.