

# A MULTIVARIATE STATISTICAL MODEL FOR MULTIPLE IMAGES ACQUIRED BY HOMOGENEOUS OR HETEROGENEOUS SENSORS

Jorge Prendes<sup>1,2</sup>, Marie Chabert<sup>1,3</sup>, Frédéric Pascal<sup>2</sup>, Alain Giros<sup>4</sup>, Jean-Yves Tournet<sup>1,3</sup>

<sup>1</sup> TéSA Laboratory, 14-16 Port Saint-Etienne, 31000 Toulouse, France

<sup>2</sup> Supélec - SONDRRA, Plateau du Moulon, 3 rue Joliot-Curie, F-91192 Gif-sur-Yvette Cedex, France

<sup>3</sup> University of Toulouse, INP/ENSEEIH - IRIT, 2 rue Charles Camichel, BP 7122, 31071 Toulouse Cedex 7, France

<sup>4</sup> CNES, 18 Av. Edouard Belin, 31401 Toulouse, France

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

## ABSTRACT

This paper introduces a new statistical model for homogeneous images acquired by the same kind of sensor (e.g., two optical images) and heterogeneous images acquired by different sensors (e.g., optical and synthetic aperture radar (SAR) images). The proposed model assumes that each image pixel is distributed according to a mixture of multi-dimensional distributions depending on the noise properties and on the transformation between the actual scene and the image intensities. The parameters of this new model can be estimated by the classical expectation-maximization algorithm. The estimated parameters are finally used to learn the relationships between the different images. This information can be used in many image processing applications, particularly those requiring a similarity measure (e.g., change detection or registration). Simulation results on synthetic and real images show the potential of the proposed model. A brief application to change detection between optical and SAR images is finally investigated.

**Index Terms**— Image analysis, change detection, remote sensing, multi-temporal images, mixture models, optical images, synthetic aperture radar.

## 1. INTRODUCTION AND PROBLEM FORMULATION

In order to reduce the effect of noise and account for the spatial correlation present in most remote sensing images, a common approach used in image analysis is to consider groups of pixels contained in a sliding window. The distribution of these groups of pixels is clearly interesting for many image processing applications. These applications include change detection [1–3], image segmentation [4, 5], image registration [1, 6], and database updating [7]. Distributions that have been recently considered in the literature include bivariate gamma distributions for two synthetic aperture radar (SAR) images [1] and bivariate Pearson distributions for heterogeneous optical and SAR images [2]. A more specific study was conducted in [7] for the fusion of remote sensing images with cartographic databases. However, a main limitation with these approaches is that they are restricted to a specific pair of sensors. A different approach that overcomes this problem was proposed in [3] where the joint distribution of two images was built using the theory of copulas. However, this method requires to learn the appropriate copula using training samples and it is hardly generalizable to situations where more than two images are available.

This paper proposes to define a new statistical model for multiple remote sensing images. We assume that we have observed

a given scene through a set of  $D$  images denoted as  $\{I_1, \dots, I_D\}$  acquired by  $D$  sensors  $\{S_1, \dots, S_D\}$ . Each sensor has imaged the scene differently since a given sensor captures different physical properties of the objects involved in the scene. Moreover, the kind of noise affecting these objects generally differs from one sensor to another. Consider as an example the case of two optical and SAR images ( $D = 2$ ). The SAR images are very sensitive to the object edges whereas the colorimetry of a scene is clearly an important property contained in optical images. The noise affecting a given area of a homogenous SAR image is classically supposed to be a multiplicative speckle noise with gamma or Weibull distribution [8]. Conversely, the noise affecting optical images has been considered as an additive Gaussian noise in many applications [9].

This paper introduces a new flexible model allowing the physical and statistical properties of images to be captured. The proposed model is flexible in the sense that it can be used for homogeneous or heterogeneous images<sup>1</sup> and for many kinds of sensors. Moreover, the model can be used to describe the pixel intensities contained in a sliding window. In many applications (e.g., change detection, registration), statistical models are used to describe the distribution of the pixels in a sliding window assuming that it remains the same for all pixels [1]. In this case the window is called homogeneous. The model proposed in this paper takes into account possible variations of the statistical model inside the sliding window (due, for instance, to the presence of different objects). In this case the window is called heterogeneous. As an example, its application to change detection will be discussed at the end of this paper.

The paper is organized as follows: Section 2 introduces the model for homogenous and heterogeneous images. Section 3 discusses the parameters estimation of this model using the expectation-maximization (EM) algorithm. In Section 4 a transformation describing the relationships between different modalities is introduced. Section 5 presents results obtained by applying this model to synthetic and real data. Conclusions are reported in Section 6.

## 2. A NEW STATISTICAL MODEL FOR IMAGE ANALYSIS

This section introduces a flexible statistical model for the pixel intensities associated with several images acquired by different sensors. To achieve this, the marginal statistical properties of the pixel intensities contained in a homogeneous area are reviewed in Section 2.1.

<sup>1</sup>Homogeneous images have been acquired by two different sensors of the same kind (e.g., two optical images or two SAR images) contrary to heterogeneous images (e.g., one optical and one SAR image).

Section 2.2 defines the joint distribution of a group of pixels belonging to a homogeneous area contained into the sliding analysis window. An extension to pixels belonging to a non-homogeneous area is introduced in Section 2.3.

### 2.1. Statistical properties of homogeneous areas

A homogeneous area of an image is a region of the image where the pixels have the same physical properties (denoted as  $P$ ). Since the measurements of any sensor  $S$  are corrupted by noise, we propose the following model

$$I_S|P = f_S[T_S(P), \nu_S] \quad (1)$$

where  $P$  is used for the set of physical properties associated with the image intensity  $I_S$ ,  $T_S(P)$  is a deterministic function of  $P$  explaining how an ideal noiseless sensor  $S$  would capture these physical properties  $P$  to form an intensity,  $\nu_S$  is a random variable representing the sensor noise, and  $f_S(\cdot, \cdot)$  describes how the sensor noise interacts with the ideal sensor measurement (which only depends on the kind of sensor  $S$ ). Model (1) indicates that  $I_S$  is a random variable whose distribution depends on the noise distribution but also on  $T_S(P)$ . To clarify this point, the examples of SAR and optical images are considered in what follows.

For SAR images, it is widely accepted that the pixel intensity  $I_{\text{SAR}}$  in a homogeneous area is distributed according to a gamma distribution [8]. For this example, model (1) reduces to

$$I_{\text{SAR}}|P = T_{\text{SAR}}(P) \nu_{\text{SAR}}$$

where  $T_{\text{SAR}}(\cdot)$  is the functional transforming the physical properties of the scene  $P$  to the noiseless radar intensity and  $\nu_{\text{SAR}}$  is a multiplicative speckle noise with gamma distribution, i.e.,  $\nu_{\text{SAR}} \sim \Gamma(L, L^{-1})$ , where  $L$  is the so-called number of looks of the SAR sensor. Using standard results on gamma distributions, we obtain

$$I_{\text{SAR}}|P \sim \Gamma\left[L, \frac{T_{\text{SAR}}(P)}{L}\right].$$

For optical images, we can consider that the pixel intensity  $I_{\text{Opt}}$  is affected by an additive Gaussian noise [9] leading to

$$I_{\text{Opt}}|P = T_{\text{Opt}}(P) + \nu_{\text{Opt}}$$

where  $T_{\text{Opt}}$  is the optical equivalent of  $T_{\text{SAR}}$ , i.e., the functional indicating the true color of the object with physical properties  $P$ , and the random variable  $\nu_{\text{Opt}}$  is the additive Gaussian noise with constant variance  $\sigma^2$ , i.e.,  $\nu_{\text{Opt}} \sim \mathcal{N}(0, \sigma^2)$ . This results in

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

The notations  $\Gamma_P(I_{\text{SAR}})$  and  $\mathcal{N}_P(I_{\text{Opt}})$  will be used to denote the probability density functions (pdfs) of  $I_{\text{SAR}}|P$  and  $I_{\text{Opt}}|P$ .

### 2.2. Distribution for multiple sensors in a homogeneous area

Assume that we have observed  $D$  images acquired by  $D$  different and independent sensors. It makes sense to assume that the  $D$  random variables  $\nu_1, \dots, \nu_D$  (defining the random vector  $\nu$ ) associated with the sensor noises are independent leading to

$$p(\nu) = p(\nu_1, \dots, \nu_D) = \prod_{d=1}^D p(\nu_d).$$

Since the image intensity  $I_d|P$  only depends on  $\nu_d$  for any  $d = 1, \dots, D$ , the joint distribution of the image intensities is

$$p(I_1, \dots, I_D|P) = \prod_{d=1}^D p(I_d|P). \quad (2)$$

For example, in the (interesting) particular case where we have observed one radar and one optical image, we obtain

$$p(I_{\text{SAR}}, I_{\text{Opt}}|P) = \Gamma_P(I_{\text{SAR}}) \mathcal{N}_P(I_{\text{Opt}}).$$

### 2.3. Joint distribution for multiple sensors in a sliding window

Denote as  $p(I_1, \dots, I_D|W)$  the joint pdf of the pixel intensities within a sliding window  $W$ . To obtain this distribution, we propose to assume that the region of interest (located inside the sliding window) is composed of a finite number  $K$  of homogeneous areas with different physical properties  $P_1, \dots, P_K$ . In this case, it makes sense to assume that the physical properties of the region of interest can be described by a discrete random variable with distribution

$$p(P|W) = \sum_{k=1}^K w_k \delta(P - P_k) \quad (3)$$

where  $w_k$  is the weight of  $P_k$  which represents how much of  $W$  is covered by  $P_k$ . Using (2) and the total probability theorem, the joint distribution of the pixel intensity can be expressed as

$$\begin{aligned} p(I_1, \dots, I_D|W) &= \sum_{k=1}^K w_k p(I_1, \dots, I_D|P_k) \\ &= \sum_{k=1}^K w_k \prod_{d=1}^D p(I_d|P_k). \end{aligned} \quad (4)$$

In the particular case of two SAR and optical images (which will be considered in our simulations), we obtain

$$p(I_{\text{SAR}}, I_{\text{Opt}}|W) = \sum_{k=1}^K w_k \Gamma_{P_k}(I_{\text{SAR}}) \mathcal{N}_{P_k}(I_{\text{Opt}}). \quad (5)$$

The expressions (4) and (5) show that the joint distribution of the pixel intensities within a given window is a mixture of distributions. Moreover, using (2) each component of this mixture is the product of densities associated with independent random variables.

## 3. PARAMETER ESTIMATION

Different approaches have been used in the literature to estimate the parameters of a mixture model. Even if the method of moments has received some interest for this estimation problem [10], the EM algorithm has become a reference for mixture models [11, 12]. The EM algorithm has the property to converge to a local maximum of the likelihood function. When applied to the joint distribution (4), the algorithm iteratively optimizes the following  $Q$  function defined as an expectation (E-step)

$$Q(\theta|\theta^{(i)}) = \mathbb{E}_{\mathbf{K}|\mathbf{I}, \theta^{(i)}} [\log p(\mathbf{I}, \mathbf{K}|\theta)] \quad (6)$$

where  $\mathbf{I} = [I_1, \dots, I_N]$  is the vector of observed data ( $N$  is the number of pixels in the sliding window  $W$ ),  $\mathbf{I}_n = [I_{1,n}, \dots, I_{D,n}]$  is the vector of intensities for the  $n$ th pixel (with  $n = 1, \dots, N$ ) acquired by the  $D$  sensors,  $\theta = [\theta_1, \dots, \theta_K]$  is the set of parameters defining the mixture, i.e.,  $\theta_k = [w_k, \theta_{k,1}, \dots, \theta_{k,D}]$  (with  $k = 1, \dots, K$ ) contains the parameters related to  $k$ th homogeneous region and  $w_k$  is the weight of the  $k$ th region in the window  $W$  (note that  $\theta_{k,d}$  is the set of parameters for the  $d$ th sensor that defines the distribution associated with the physical properties  $P_k$ ),  $\theta^{(i)}$  is the set of parameters  $\theta$  resulting from the  $i$ th iteration of the algorithm and  $\mathbf{K} = [k_1, \dots, k_N]$  is the unobserved map of labels indicating

that pixel  $I_n$  results from the observation of the  $k$ th component  $P_k$ . At each iteration the following optimization (M-step) is performed

$$\theta^{(i+1)} = \arg \max_{\theta} Q(\theta | \theta^{(i)}).$$

It can be easily proven that optimizing  $Q(\theta | \theta^{(i)})$  with respect to (wrt)  $\theta$  is equivalent to optimizing  $\log p(\mathbf{I} | \theta)$  wrt  $\theta$  [11]. Throughout this paper, we consider the standard assumption according to which the samples  $[I_1, k_1], \dots, [I_N, k_N]$  are independent (pixel independence in the observation window) leading to

$$\log p(\mathbf{I}, \mathbf{K} | \theta) = \sum_{n=1}^N \log p(I_n, k_n | \theta). \quad (7)$$

After replacing (7) in (6), computations detailed in [13] lead to

$$Q(\theta | \theta^{(i)}) = \sum_{n=1}^N \sum_{k=1}^K \pi_{n,k}^{(i)} \log w_k + \sum_{n=1}^N \sum_{k=1}^K \pi_{n,k}^{(i)} \log p(I_n | \theta_k).$$

where  $\pi_{n,k}^{(i)} = \frac{p(I_n, k_n = k | \theta^{(i)})}{p(I_n | \theta^{(i)})}$  is constant for a given value of  $(i, n, k)$  and  $p(I_n | \theta_k) = p(I_n | k_n = k, \theta)$  is the probability that the observed pixel intensity  $I_n$  has been produced by an object with physical property  $P_k$ . Thus, based on (2), we obtain

$$Q(\theta | \theta^{(i)}) = \sum_{n=1}^N \sum_{k=1}^K \pi_{n,k}^{(i)} \left[ \log w_k + \sum_{d=1}^D \log p(I_{n,d} | \theta_{k,d}) \right]. \quad (8)$$

This last result shows that  $Q$  can be written as the summation of terms depending on the different components  $\theta_{k,d}$  of  $\theta$ . Thus, each of these terms can be maximized independently with respect to  $\theta_{k,d}$  in order to maximize  $Q(\theta | \theta^{(i)})$ . The maximization of  $Q$  with respect to  $w_k$  [11] and  $\theta_{k,d}$  classically yields

$$w_k^{(i+1)} = \frac{1}{N} \sum_{n=1}^N \pi_{n,k}^{(i)} \quad (9)$$

$$\theta_{k,d}^{(i+1)} = \arg \max_{\theta_{k,d}} \sum_{n=1}^N \pi_{n,k}^{(i)} \log p(I_{n,d} | \theta_{k,d}) \quad (10)$$

where (10) is the maximum likelihood estimator (MLE) of the parameters associated with the distribution of the  $d$ th sensor  $S_d$ . It is interesting to note here that (9) and (10) can be solved for various distributions including the Gaussian, generalized Gaussian [14], exponential, gamma [15] and Weibull [16] distributions.

Of course, the number  $K$  of components in the mixture (which corresponds to the number of homogeneous regions or objects contained in the window  $W$ ) is generally unknown and has to be estimated. In this paper, we have used the algorithm introduced in [17] which requires fixing a maximum number of components, and gradually removes the components that do not describe enough samples.

#### 4. LEARNING THE RELATIONSHIPS BETWEEN IMAGES FROM DIFFERENT MODALITIES

Two or more images associated with the same scene (acquired by homogeneous or heterogeneous sensors) have some similarities. These similarities are useful for image analysis, e.g., to detect whether a given object is present in several images or to detect changes between several images. This section introduces a model allowing the relationships between the different images to be analyzed.

An image is usually composed of several homogeneous areas or objects. It is also common to find many objects of the same kind in

an image. Each kind of object has different physical properties  $P$  when it is observed in different images. In the case of  $D$  images, these physical properties can be gathered in the following vector

$$\mathbf{v}(P) = [T_1(P), \dots, T_D(P)]$$

where  $T_d(P)$  is the transformation introduced in (1) for the sensor  $S_d$ . Depending on the value of the vector  $\mathbf{v}(P)$ , the object has a different aspect in the different observed images. The value of  $\mathbf{v}(P)$  can be seen as the fingerprint of a given object: each object produces a cluster of  $\mathbf{v}$  in a  $D$  dimensional subspace. When considering all the possible values of  $P$ , the clusters of  $\mathbf{v}$  describe a functional in this  $D$  dimensional subspace parameterized by  $P$ .

For instance, when considering two optical gray-scale images, we have  $\mathbf{v}(P) = [T_{\text{Opt}_1}(P), T_{\text{Opt}_2}(P)]$ . Since both optical sensors are similar, both components of  $\mathbf{v}$  can be related as follows [18]

$$T_{\text{Opt}_2}(P) = \alpha T_{\text{Opt}_1}(P) + \beta \quad (11)$$

where  $\alpha$  and  $\beta$  compensate for brightness and contrast differences. In this case, the transformation (11) is affine and defines a straight line in a bi-dimensional subspace. When the sensors are different, the transformation between the different images can be very complex and, of course, the relationships between the components of  $\mathbf{v}(P)$  are *a priori* unknown. For every window  $W$ ,  $K$  different values of  $\mathbf{v}$  (corresponding to the different regions/objects contained in  $W$ ) can be estimated using the statistical model defined in Section 3. Clustering the estimated values of  $\mathbf{v}$  (denoted as  $\hat{\mathbf{v}}$ ) can be used to learn the components of  $\mathbf{v}(P)$ , i.e., the relationships between the different images of interest.

### 5. SIMULATION RESULTS

This section shows how the proposed statistical model can be applied to different datasets, including synthetic and real images. We first present examples describing the use of the proposed model to learn the joint distribution of the pixel intensities contained in a moving window. The behavior of  $\mathbf{v}(P)$  is then studied from the estimated parameters of the proposed statistical model.

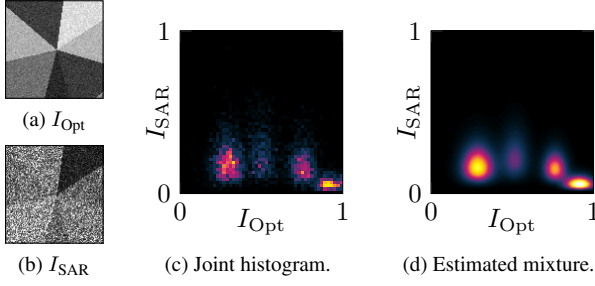
#### 5.1. Synthetic data

A set of two synthetic images was created by generating a synthetic scene  $P$ . The synthetic scene was created from triangular patches representing the different objects contained in the image. The synthetic scene was corrupted by additive Gaussian noise to yield the optical image with  $\text{SNR} = 30\text{dB}$ . To generate the SAR image, first a known transformation was applied to the scene and the result was corrupted by a multiplicative gamma noise with shape parameter equal to  $L = 5$ . The results presented in this paper have been obtained with the following transformation

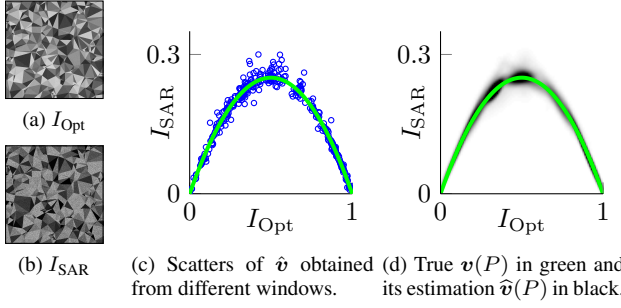
$$\mathbf{v}(P) = [T_{\text{Opt}}(P), T_{\text{SAR}}(P)] = [P, P(1 - P)].$$

where  $\mathbf{v}$  has been chosen to model a non linear, yet relatively simple transformation. Figs. 1(a) and 1(b) show examples of sliding windows extracted from the synthetic optical and SAR images obtained with this method. The histogram of the intensities  $[I_{\text{Opt}}, I_{\text{SAR}}]$  displayed in Fig. 1(c) were obtained by counting the number of samples of the intensities contained in  $50 \times 50$  bins belonging to  $[0, 1] \times [0, 1]$ . This figure shows four main clusters which are clearly captured by the proposed statistical model whose contours are shown in Fig. 1(d).

Figs. 2(a) and 2(b) show the full synthetic optical and SAR images associated with the images of Figs. 1(a) and 1(b). Moving windows of  $20 \times 20$  pixels were then considered to estimate the parameters of the proposed mixture model. These parameters were then



**Fig. 1.** Examples of sliding windows for synthetic optical and SAR images and corresponding joint histogram and estimated mixture.



**Fig. 2.** Full synthetic optical and SAR images. Scatter plots of  $\hat{v}$  and the estimated function  $v(P)$ .

used to estimate the corresponding vector  $v$ . A scatter plot of the resulting  $\hat{v}$  can be observed in Fig. 2(c). The transformation  $v(P)$  was finally learnt by applying a density estimator (based on Parzen windows) on the estimated vectors  $\hat{v}$ . The resulting pdf is superimposed on the desired transformation in Fig. 2(d) showing a very good agreement with the ground truth (green curve).

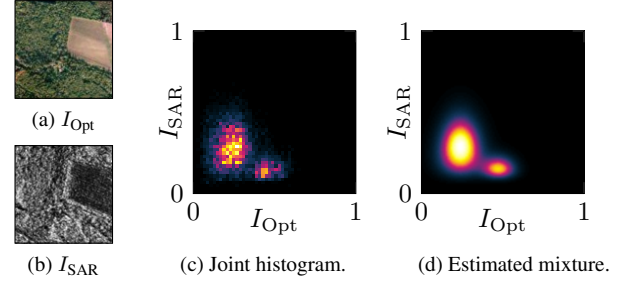
## 5.2. Real data

Figs. 3(a) and 3(b) show a pair of two real optical and SAR images from a rural area containing a forest region and a harvested field. The optical image was converted into a gray-scale image to facilitate the joint pdf representation (which is bivariate), although the proposed model can easily consider each color channel as a different sensor. Figs. 3(c) and 3(d) show the histogram and the estimated mixture resulting from the proposed model, both obtained as for Fig. 1. Two main clusters can be identified representing the forest, and the harvested area, as expected.

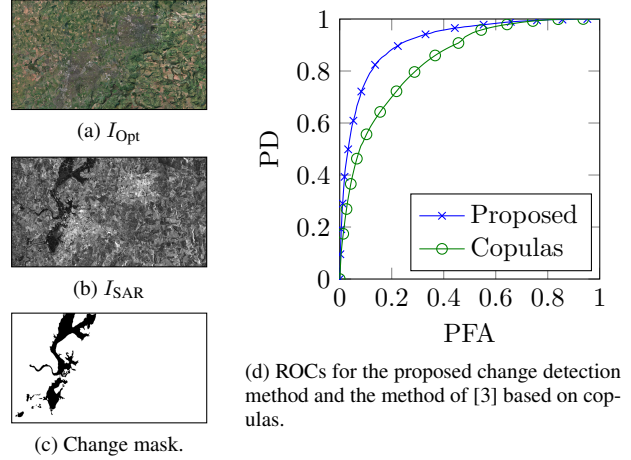
Fig. 4 shows the application of the proposed model to detect changes in real optical and SAR images (note that the two images are heterogeneous). An unchanged training area was used to learn the functional relationship  $v(P)$  using Parzen windows. The estimated density denoted as  $\hat{p}[\hat{v}(P)]$  was used as a change measure, leading to the following change detection strategy

$$\hat{p}[\hat{v}(P)] \underset{H_1}{\overset{H_0}{\geq}} \tau$$

where  $\tau$  is a threshold related to the false alarm probability PFA and the detection probability PD of the detector. The receiver operating characteristic (ROC) curves depicted in Fig. 4(d) (obtained by varying  $\tau$  and estimating the empirical PD and PFA) show that the proposed method provides better performance than the method of [3] based on the theory of copulas.



**Fig. 3.** Window of real optical and SAR images and their corresponding joint histogram and estimated mixture.



**Fig. 4.** Optical image before flooding (a), SAR image after flooding (b), change map (c) and ROCs obtained with the method of [3] and the proposed strategy (d).

## 6. CONCLUSIONS

This paper introduced a new statistical model to describe the distribution of any number of homogeneous and heterogeneous images independently of the kind of sensors used to obtain these images. The proposed model was expressed as a mixture of multi-dimensional distributions whose parameters can be estimated by the expectation-maximization algorithm. This mixture of distributions can be used to determine standard similarity measures such as the mutual information and is thus interesting for many potential applications. As an example, the model was successfully applied to the detection of changes between optical and synthetic aperture radar images. However, it could be interesting for many other applications such as image registration, image indexing or image classification. Moreover, further work should be conducted to validate the proposed model on a larger dataset containing homogeneous and heterogeneous images.

## 7. ACKNOWLEDGMENTS

Part of this work was conducted within the thematic trimester on image processing of the Labex CIMI and was partially supported by ANR-11-LABX-0040-CIMI within the program ANR-11-IDEX-0002-02. The authors would like to thank the CNES of Toulouse for its financial support. They are also very grateful to Grégoire Mercier for providing the change detection results using the method of [3] and for sharing interesting ideas about this work.

## 8. REFERENCES

- [1] F. Chatelain, J.-Y. Tournet, J. Inglada, and A. Ferrari, "Bivariate gamma distributions for image registration and change detection," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1796–1806, July 2007.
- [2] M. Chabert and J.-Y. Tournet, "Bivariate Pearson distributions for remote sensing images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Vancouver, Canada, July 2011, pp. 4038–4041.
- [3] 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.
- [4] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blob-world: image segmentation using expectation-maximization and its application to image querying," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 8, pp. 1026–1038, Aug. 2002.
- [5] M. S. Allili, D. Ziou, N. Bouguila, and S. Boutemedjet, "Image and video segmentation by combining unsupervised generalized gaussian mixture modeling and feature selection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 10, pp. 1373–1377, Oct. 2010.
- [6] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, 1997.
- [7] M. Chabert, J.-Y. Tournet, V. Poulain, and J. Inglada, "Logistic regression for detecting changes between databases and remote sensing images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Honolulu, USA, 2010, pp. 3198–3201.
- [8] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, Artech House, Boston, 1998.
- [9] K. R. Castleman, *Digital image processing*, Prentice Hall, 1996.
- [10] S. Kotz, N. Balakrishnan, and N. L. Johnson, *Continuous Multivariate Distributions, Models and Applications*, Wiley, 2004.
- [11] G. McLachlan and D. Peel, *Finite Mixture Models*, Wiley, 2004.
- [12] O. Schwander, A.J. Schutz, F. Nielsen, and Y. Berthoumieu, "k-MLE for mixtures of generalized gaussians," in *Proc. IEEE Int. Conf. Pattern Recognit.*, Tsukuba, Japan, 2012, pp. 2825–2828.
- [13] J. Prendes, M. Chabert, F. Pascal, A. Giros, and J.-Y. Tournet, "A new multivariate statistical model for change detection in images acquired by homogeneous and heterogeneous sensors," *submitted to IEEE Trans. Image Process.*, 2014.
- [14] F. Pascal, L. Bombrun, J.-Y. Tournet, and Y. Berthoumieu, "Parameter estimation for multivariate generalized Gaussian distributions," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5960–5971, 2013.
- [15] G. J. Hahn and S. S. Shapiro, *Statistical Models in Engineering*, Wiley, 1994.
- [16] W. Q. Meeker and L. A. Escobar, *Statistical Methods for Reliability Data*, Wiley, 1998.
- [17] 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.
- [18] L. Wojnar, *Image Analysis: Applications in Materials Engineering*, Taylor & Francis, 1998.