

# Information theoretical similarity measure for change detection

Shiyong Cui, Mihai Datcu

Remote Sensing Technology Institute (IMF)  
German Aerospace Center  
Oberpfaffenhofen, Germany  
[shiyong.cui\\_mihai.datcu@dlr.de](mailto:shiyong.cui_mihai.datcu@dlr.de)

Lionel Gueguen

ISFEREA - Globesec - IPSC  
Joint Research Centre - European Commission  
Ispra, Italy  
[lionel.gueguen@jrc.ec.europa.eu](mailto:lionel.gueguen@jrc.ec.europa.eu)

**Abstract**—In this paper, mixed information similarity measure and a multidimensional density estimation method based on multivariate Edgeworth series expansion are proposed and assessed for the task of multi-temporal change detection. To unify mutual information and variational information, mixed information is proposed to quantify the degree of dependence between two random variables, which are intuitively appropriate for multi-temporal change detection. In the literature, Edgeworth series expansion is widely used in statistics and various engineering fields for one-dimensional density estimation. To compute the mixed information measure, multidimensional density estimation based on multivariate Edgeworth series expansion is proposed and evaluated. Two experiments on real SAR images and optical images are carried out to evaluate the performance of change detection. Experimental results confirm the promising capability of mixed information and the multivariate density estimation based on Edgeworth series expansion.

## I. INTRODUCTION

Multi-temporal analysis of Synthetic Aperture Radar (SAR) images and optical satellite images has been recognized to play a fundamental role in the field of Earth observation applications. As an important part of multi-temporal analysis, change detection in images of a given scene acquired at two different times has received increasing interest in Earth science, especially in environmental monitoring, land-use/land-cover dynamics, damage assessment, and forest or vegetation monitoring. In the literature, a large amount of different techniques have been proposed for supervised change detection and unsupervised change detection [1], [2]. For supervised change detection, samples based on the prior knowledge of a research scenario are selected and used to train the method, which is usually a classification method, and then use the learning method after training to classify each pixel as changed or unchanged pixel. Usually, two steps are involved in unsupervised change detection, the first of which is to compare two images resulting in a change map, and the second is to threshold or label the change map to derive a binary change map consisting of two classes associated with changed and unchanged pixels [1]. Therefore, in the sense of sample selection, unsupervised approaches are more preferable to supervised approach in practice.

Among the available methods in the literature, a kind of most promising methods for multi-temporal change detection is

based on information measures due to their efficiency and simplicity. In this kind of methods, an information measure is usually used to assess the similarity between two images to generate a change map by quantifying the dependence between the two random variables associated with the two images. In information theory, mutual information is widely used for multi-temporal image analysis, especially for image registration and content based image trivial. In [3], Kullback-Leibler divergence was used for change detection by comparing the local distributions of the two images. A new method for Kullback-Leibler divergence estimation was proposed based on cumulant expansion, which is actually the so-called one-dimensional Edgeworth series expansion. In [4], several information theoretical similarity measures including distance to independence, mutual information, cluster reward algorithm, Woods criterion and correlation ratio, were proposed and compared for change detection. Among them, the mutual information based similarity measure proved to be rather efficient. Taking advantage of the mutual information, a pixel-based approach comparing the localized mutual information shared by two pixels was proposed in [5]. Intuitively, when the two pixels share little mutual information, it is reasonable to assume change happens in their locations. Based on this idea, another new information measure derived from the mutual information was originally introduced in [6]: the mixed information which unifies mutual information and variational information by introducing a new parameter to control the balance between the common information and the different information. In this paper, we propose an unsupervised change detection approach based on mixed information similarity measures in the frame of information theory which extends the previous work in [7]. In that work, the mixed information was computed using the marginal histogram and the joint histogram associated with the two images. In this paper, we propose to estimate the multidimensional density by Edgeworth series expansion and replace the histogram used in [7] by this method and further evaluate the performance of this novel information measure in multi-temporal analysis.

In Section II, we briefly recall information measures, while density estimation for the computation of information measures based on multivariate Edgeworth series expansion is presented in Section III. Experiments on SAR and optical image are carried out in Section IV. Finally, a conclusion is given in Section V.

## II. INFORMATION MEASURES

In this section, information similarity measures, mutual information, variational information, and mixed information are briefly reviewed.

As a powerful similarity measure, mutual information is widely used in image analysis, especially for multi-temporal image registration. The mutual information between two random variables originally proposed by Shannon [8] is defined as

$$I(X_1, X_2) = \iint f(x_1, x_2) \log \frac{f(x_1, x_2)}{f(x_1)f(x_2)} dx_1 dx_2 \quad (1)$$

where  $f(x_1), f(x_2), f(x_1, x_2)$  are the marginal probability density functions and the joint probability density function of  $X_1$  and  $X_2$ . The variational information proposed in [9] is defined as

$$\begin{aligned} VI(X_1, X_2) &= -\iint f(x_1, x_2) \log \frac{f(x_1, x_2)^2}{f(x_1)f(x_2)} dx_1 dx_2 \\ &= H(X_1, X_2) - I(X_1, X_2) \end{aligned} \quad (2)$$

where  $H(X_1, X_2)$  is the joint entropy of  $X_1$  and  $X_2$ . While mutual information quantifies the common information between  $X_1$  and  $X_2$ , variational information quantifies the different information transmitted through  $X_1$  and  $X_2$ . To unify the mutual information and variational information, a new measure termed mixed information [6]  $I_\alpha$  which trades off between common and different information through a trade-off parameter  $\alpha$  taking a value within the interval  $[0, 1]$  was defined as

$$\begin{aligned} I_\alpha(X_1, X_2) &= \iint f(x_1, x_2) \log \frac{f(x_1, x_2)^{1+\alpha}}{f(x_1)f(x_2)} \\ &= I(X_1, X_2) - \alpha H(X_1, X_2). \end{aligned} \quad (3)$$

As  $\alpha$  may vary between 0 and 1,  $I_\alpha$  can be viewed as the mixture of common and variational information measures. In particular, when  $\alpha = \frac{I(X_1, X_2)}{H(X_1, X_2)}$ , the mixed information is null and  $I_\alpha = 0$ .

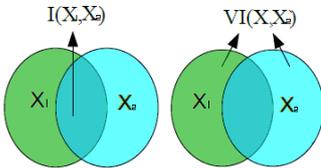


Figure 1. Mutual information and variational information

If the two random variables representing two images share a lot of information, it is reasonable to assume that there is no change; otherwise there is change, which is demonstrated in

Fig.1. Therefore, the change map can be intuitively generated by the following equation:

$$CD(i) = (1 + \alpha) \log(X_1(i), X_2(i)) - \log(X_1(i)) - \log(X_2(i)) \quad (4)$$

where  $i$  is the position in the image grid,  $CD(i)$  is the change detection map and  $X_1(i), X_2(i)$  are the pixel values of image  $X_1$  and  $X_2$  at location  $i$ .

## III. MULTIDIMENSIONAL DENSITY ESTIMATION BY MULTIVARIATE EDGEWORTH SERIES EXPANSION

Edgeworth series expansion is an asymptotic expansion technique widely used in statistics, especially for density estimation which is an important issue in engineering fields. Edgeworth series expansion is achieved by inverse transform of the moment generating function which is the integral transform of a random vector. Usually, the components of the random vector are correlated which raises the computational complexity of density estimation. Therefore, we first introduce decorrelation applied prior to density estimation. After that the detailed procedure of multivariate Edgeworth series expansion for density estimation is presented.

### A. Decorrelation transform

Decorrelation is a linear operation  $T$  which removes the correlation between components of a random vector  $X$  and gives rise to a new random vector  $Y$  with a covariance matrix being an identity matrix. Decorrelation is defined as

$$T = \Lambda^{-1/2} V^T \quad (5)$$

where  $V$  is a unitary matrix and  $\Lambda$  is a diagonal matrix such that

$$\Sigma = V \Lambda V^{-1} \quad (6)$$

and  $\Sigma$  is the covariance matrix of observation samples.

### B. Edgeworth series expansions

Edgeworth series expansion is an asymptotic expansion and is defined as [10], [11]

$$\begin{aligned} f_X(x; k) &= \phi(x; \lambda) \left[ 1 + \frac{1}{3!} k^{i,j,k} h_{ijk}(x) + \frac{1}{4!} k^{i,j,k,l} h_{ijkl}(x) \right. \\ &\quad \left. + \frac{[10]}{6} k^{i,j,k} k^{l,m,m} h_{ijklmn}(x) + \dots \right] \end{aligned} \quad (7)$$

where  $\phi(x; \lambda)$  is a multivariate Gaussian distribution and  $h$  is a multivariate Hermite polynomial,  $k$  is the multivariate cumulant. If the components of  $X = (X_1, X_2, \dots, X_n)$  are uncorrelated and have unit variance, then  $k^{i,i} = 1, k^{i,j} = 0$ . The Hermite polynomials for the multivariate distribution of  $X$  are then formed by taking all possible products of the Hermite polynomials.

In the case of two dimensions ( $n = 2$ ) with uncorrelated components and unit variance

$$f_x(x; k) = \phi(x; k) \left[ 1 + \frac{1}{3!} v_1(x; k) + \frac{1}{4!} v_2(x; k) + \frac{10}{6!} v_3(x; k) \right] + O(n^{-\frac{3}{2}}) \quad (8)$$

where  $v_1(x; k)$ ,  $v_2(x; k)$  and  $v_3(x; k)$  are defined as:

$$\begin{aligned} v_1(x; k) &= k^{1,1,1} h_{111}(x) + 3k^{1,1,2} h_{112}(x) + 3k^{1,2,2} h_{122}(x) \\ &\quad + k^{2,2,2} h_{222}(x) \\ v_2(x; k) &= k^{1,1,1,1} h_{1111}(x) + 4k^{1,1,1,2} h_{1112}(x) + 6k^{1,1,2,2} h_{1122}(x) \\ &\quad + 4k^{1,2,2,2} h_{1222}(x) + k^{2,2,2,2} h_{2222}(x) \\ v_3(x; k) &= k^{1,1,1,1,1} h_{11111}(x) + 6k^{1,1,1,1,2} h_{11112}(x) \\ &\quad + 15k^{1,1,1,2,2} h_{11122}(x) + 20k^{1,1,1,2,2,2} h_{111222}(x) \\ &\quad + 15k^{1,1,2,2,2} h_{12222}(x) + 6k^{1,2,2,2,2} h_{22222}(x) \\ &\quad + k^{2,2,2,2,2} h_{22222}(x). \end{aligned} \quad (9)$$

The cumulants are calculated from moments which are estimated from observation samples. The relationship between multivariate cumulants and moments is

$$\begin{aligned} k^{i,j} &= k^{ij} - k^i k^j \\ k^{i,j,k} &= k^{ijk} - k^i k^{jk} [3] + 2k^i k^j k^k \\ k^{i,j,k,l} &= k^{ijkl} - k^i k^{jkl} [4] - k^{ij} k^{kl} [3] + 2k^i k^j k^{kl} [6] - 6k^i k^j k^k k^l \\ &\dots \end{aligned} \quad (10)$$

The sum is over all partitions of the indices and the coefficient is  $(-1)^{v-1} (v-1)!$ , where  $v$  is the number of blocks of the partition. Based on the marginal and the joint densities estimated by Edgeworth series expansion, the mixed information can be computed for each pixel according to Eq. (4), which is an indicator of change.

#### IV. EXPERIMENTS AND DISCUSSION

In this section, two experiments are carried out on SAR and optical images to assess and evaluate the performance of the proposed method for change detection. In the first experiment, SAR images are used while optical images are used in the second experiment.

##### A. Experiment with SAR images

Test images used in [3] are shown in Fig. 2, which were acquired before and after an eruption of the Nyragongo volcano. As there is a lot of noise in the two images, filtering has to be performed prior to the application of the proposed change detection method. The first step in this experiment is to estimate the marginal and the joint densities of the two test images. The estimated marginal densities of the two images are shown in Fig. 3. For the convenience of comparison, the histogram and the estimated Gaussian density are simultaneously displayed. From the estimated density shown in Fig. 3, we can see that the density estimated by Edgeworth

series expansion fit much better than the Gaussian density by considering the long tail of the distribution, which directly improves the performance of change detection.

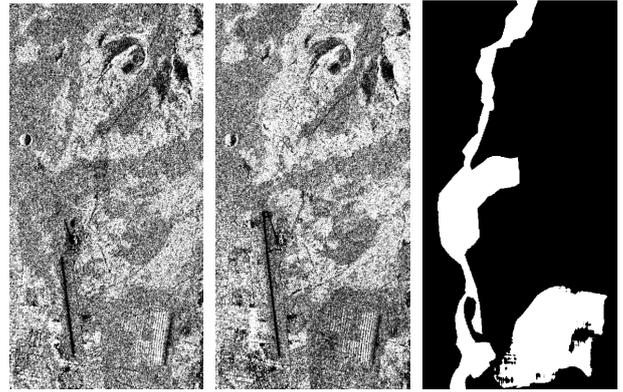


Figure 2. Test images acquired before and after an eruption and the change mask.

As the estimated marginal and joint densities are continuous and smooth, there is less noise in the final change map, as we can see from Fig. 5. The colors represent different degrees of changes. To compare the performance of the change detection based on histogram to those derived by the density estimated by Edgeworth series expansion, the ROC curve based on ground truth is plotted on the left in Fig. 4. As can be seen in Fig. 5, the change map derived by the density estimated using Edgeworth series expansion is much better than that by histogram.

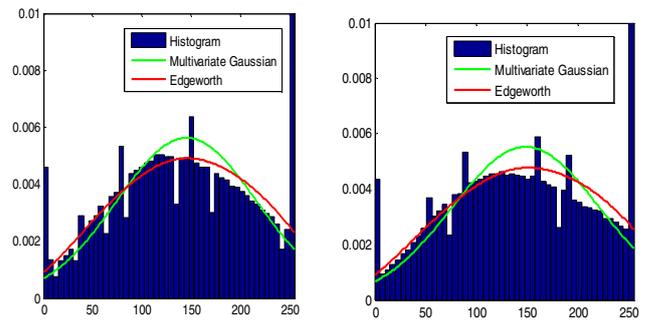


Figure 3. Marginal density estimation of the two images.

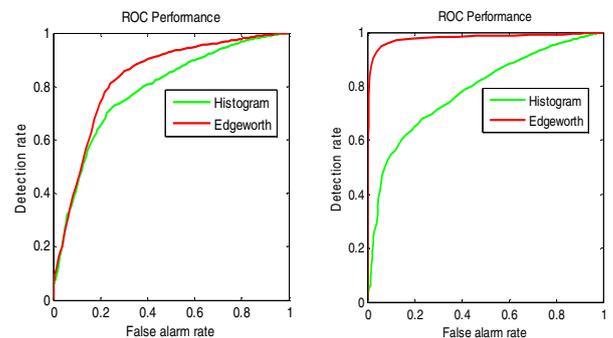


Figure 4. ROC curve for the performance of change detection via histogram and the density derived by Edgeworth series expansion: (left) the ROC curve for SAR change detection; (right) the ROC curve for optical change detection.

### B. Experiment with optical images

This approach was also tested on the SPOT images shown in Fig. 6 acquired before and after a flooding event. Following the proposed procedure, the change map is derived and shown in Fig. 7. Different colors represent different degrees of change.

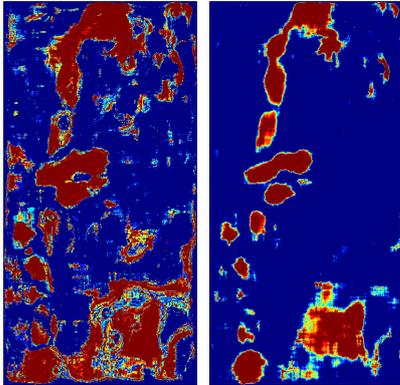


Figure 5. Change detection results: (left) result derived by histogram; (right) change map derived by density estimated with Edgeworth series, which is much better than the change map on the left

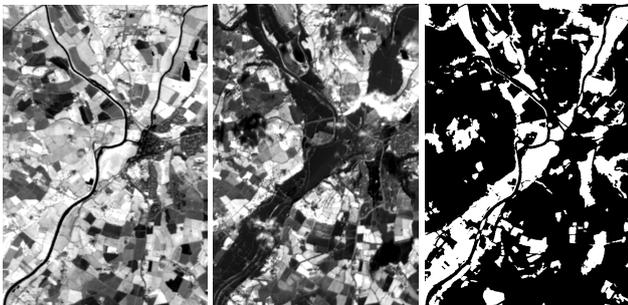


Figure 6. (left) SPOT image before flooding; (middle) after flooding; (right) change mask.

Higher pixel values in the change map represent higher degrees of changes. As a result of the discontinuity of the histogram, the change map produced by histogram contains a lot of noise. In this sense, the change map derived by the Edgeworth series method is much better. As we can see in Fig. 6, the change map is smooth and very accurate compared to the change map derived by histogram. The ROC curve of this experiment is plotted on the right in Fig. 4 based on ground truth data derived by manual interpretation. As can be seen in Fig. 7, the performance of the change detection by Edgeworth series, denoted by the red curve, is much better than the performance of the histogram technique denoted by the green curve.

### V. CONCLUSION

In this paper, a new information measure and a new method for multidimensional density estimation are proposed. This approach was tested on SAR and optical images. The promising results show the capability of mixed information measure in multi-temporal change detection and Edgeworth series expansion for multidimensional change detection.

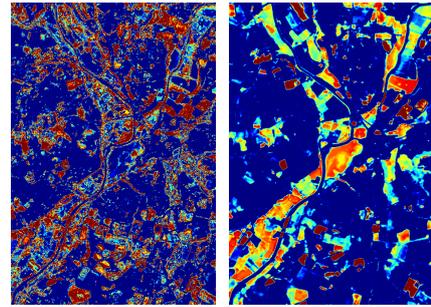


Figure 7. Optical experiment: (left) change map derived by histogram; (right) change map derived with density estimated by Edgeworth series

It turned out that quantification of the dependence of the two images is a better choice for change detection. It is worth to note that the mixed information measure can capture nonlinear dependencies between two random variables if nonlinear joint probabilities are employed.

### ACKNOWLEDGMENTS

The author would like to thank DAAD to provide financial support to carry out this research in the German Aerospace Center and his colleague Gottfried Schwarz for valuable suggestions in writing this paper.

### REFERENCES

- [1] L. Bruzzone, D. Fernandez Prieto, "An adaptive semi-parametric and context-based approach to unsupervised change detection in multitemporal remote sensing images", *IEEE Transactions on Image Processing*, Vol. 1, No. 4, 2002, pp.452-466.
- [2] L. Bruzzone and D. Fernandez Prieto, "Automatic analysis of the difference image for unsupervised change detection", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, No.3, 2000 pp. 1171-1182.
- [3] J. Inglada and G. Mercier, "A New Similarity Measure for Change Detection in Multitemporal SAR Images and Its Extension to Multiscale Change Analysis", *IEEE Trans. Geosci. Remote Sensing*, Vol. 45, No. 5, May 2007, pp.1432-1445.
- [4] V. Alberga, "Similarity Measures of Remotely Sensed Multi-sensor Images for Change Detection Applications", *Remote Sensing*, Vol. 1, No. 3, 2009, pp.122-143.
- [5] A. Winter, H. Maitre, N. Cambou, and E. Legrand, "Entropy and Multiscale Analysis: A New Feature Extraction Algorithm for Aerial Images", In: *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 4, April 1997, pp. 2766-2768.
- [6] L. Gueguen and M. Datcu, "Mixed Information Measure: Application to Change Detection in Earth Observation", to appear in *MultiTemp* 2009.
- [7] L. Gueguen, S. Cui, G. Schwarz and M. Datcu, "Multi-temporal analysis of multisensor data: information theoretical approaches", In: *Proc. IGARSS conference*, 2010, to appear.
- [8] C. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, Vol. 27, 379-423, July & October 1948, pp. 379-423.
- [9] M. Meila, "Comparing Clustering by the Variation of Information," In: *Proceedings of the 16th Annual Conference of Computational Learning Theory*, Springer, 2003.
- [10] O.E. Barndorff-Nielsen and D.R. Cox, *Asymptotic Techniques for Use in Statistics*. Chapman and Hall, London, 1989.
- [11] P. McCullagh, *Tensor Methods in Statistics*. Chapman and Hall, London, 1987.