



**HAL**  
open science

## False discovery rate approach to image change detection

Vladimir Krylov, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia

► **To cite this version:**

Vladimir Krylov, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia. False discovery rate approach to image change detection. IEEE International Conf. on Image Processing (ICIP), Sep 2013, Melbourne, Australia. hal-00841236

**HAL Id: hal-00841236**

**<https://inria.hal.science/hal-00841236>**

Submitted on 21 Sep 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# FALSE DISCOVERY RATE APPROACH TO IMAGE CHANGE DETECTION

Vladimir A. Krylov<sup>1,2</sup>, Gabriele Moser<sup>3</sup>, Sebastiano B. Serpico<sup>3</sup> and Josiane Zerubia<sup>1</sup>

<sup>1</sup> Ayin team, INRIA, 2004 route des Lucioles, BP93, F-06902 Sophia Antipolis, Cedex, France

<sup>2</sup> Dept. Statistical Science, University College London, 1-19 Torrington Place, WC1E 6BT, London, UK

<sup>3</sup> Dept. of Electrical, Electronic, Telecommunications Engineering and Naval Architecture (DITEN),  
University of Genoa, Via Opera Pia 11a, I-16145, Genoa, Italy

## ABSTRACT

In this paper we address the problem of unsupervised change detection on image pairs. We develop a novel patch-based hypothesis testing approach that employs the false discovery rate technique for statistical hypothesis testing. The designed approach can be adopted to specific detection applications via selection of appropriate statistical features. Experiments with still camera imagery demonstrate high performance and flexibility of the proposed method.

**Index Terms**— Change detection, still camera, statistical hypothesis testing, false discovery rate.

## 1. INTRODUCTION

Change detection is one of the fundamental image processing problems and a cornerstone of numerous monitoring and tracking applications. Processing a sequence of images representing the same scene at different time instants, the basic problem of change detection is to determine image areas that have undergone some *significant* changes. The considered types of imagery are very different: from still camera videos with tens of frames per second to satellite acquisitions obtained with a time gap of several months. Each specific change detection problem faces its own set of *disturbance factors*, like, e.g., acquisition noise, coregistration artifacts, illumination changes, etc.

In this paper we focus on change detection on still camera image pairs. Several main classes of change detection methods [1, 2] for still camera images have been developed. The first wide class is given by background subtraction approaches [2–4] that analyse the temporal consistency of image pixels in large image sequences. The second class of methods rely on the spatial support by analysing the local information at two time instants. The disturbance factors are typically modelled as the order-preserving transformations [5–7]. These methods require estimates of variances of the acquisition/camera noise, that are typically obtained from available multiple frames. Note that in this paper we focus on a scenario when only two images are considered for the change detection analysis. Thus, many background-based methods,

e.g., [2, 3], can not be employed. Some methods, like [4–6], can be employed given the noise variance estimates.

In order to avoid a priori knowledge about the camera properties and/or training data, we assume that the amount of significant changes on the image pairs is not more than 25% of the image pixels. This assumption allows to operate on two images in an unsupervised manner and holds for many still camera surveillance sequences, see [2]. We develop a novel statistical change detection approach, that takes pixel-level decisions based on statistical hypothesis testing performed on paired local patch samples. This hypothesis testing relies on the use of test statistics, that are referred hereafter as features. A large choice of such features allows to address various disturbance factors and detection problems. The developed approach belongs to the class of large-scale inference problems, where thousands of individual tests are performed simultaneously, i.e. one per every image pixel. To avoid a difficult problem of adjusting the family-wise error rate (FWER) that arises in such problems [8, 9], we address the multiple inference as a false discovery rate (FDR) problem. This recent multiple hypotheses testing technique [9, 10] results in more powerful testing procedures when addressing large amounts of true positives (changes) [10]. The FDR-formulation also enables to gain robustness with respect to sample dependences that arise in patch-based analysis. This flexible statistical methodology is just emerging in computer vision, see, e.g., [11], and, to the best of our knowledge, has never been previously employed to change detection.

We validate the performance of the developed approach on still camera image pairs. Note that the proposed approach generalizes the previously proposed likelihood ratio-test method [12] and, as such, can be expected to perform well on synthetic aperture radar imagery. The approach can be applied to various change detection problems, like object tracking, monitoring on remotely sensed and medical imagery.

The paper is organized as follows. In Section 2 we formulate the FDR-approach. In Section 3 we propose a novel feature-based change detector. In Section 4 we consider several possible statistical features. We report the experimental results in Section 5, and conclude in Section 6.

## 2. FALSE DISCOVERY RATE APPROACH

Considering pixel-based change detection, we face a large-scale simultaneous hypothesis testing problem [9, 10], in which thousands of simultaneous tests are performed on data patch-based samples. Values of a selected statistical feature  $\mathcal{F}$  are calculated at each pixel location. We employ the statistic features, for which distributions  $D_{\mathcal{F}}$  are known in a no-change scenario. Based on local patch samples, however, the values of the collected features are dependent, at the very least, at neighboring locations. Furthermore, the correction to the individual test significance level in a multi-test is a sophisticated matter [10]. To address the problem, we capitalize on the knowledge of a significant amount of feature observations (one per each pixel of the considered image pair) to learn the actual  $D_{\mathcal{F}}$  distribution. This becomes possible since we assume that the majority (over 75%) of the pixels are unchanged. Below we adopt the false discovery rate (FDR) approach [9, 10] to change detection.

We first introduce the notion of  $z$ -scores which allows a unified formulation for any  $D_{\mathcal{F}}$  distribution. Specifically, if  $D_{\mathcal{F}}$  denotes the cumulative distribution function (CDF) of  $\mathcal{F}$ , then we define a  $z$ -score of the  $i$ -th image pixel as  $z_i = \Phi^{-1}(D_{\mathcal{F}}(f_i))$ , where  $\Phi$  is the standard normal CDF, and  $f_i$  is the observed features  $\mathcal{F}$  at pixel  $i$ .

Suppose that values of the feature statistic fall into two classes corresponding to *unchanged* (we will call it the *null hypothesis*,  $H_0^{\mathcal{F}}$ ) and *changed* (*alternative*  $H_1^{\mathcal{F}}$ ), with prior probabilities  $p_0$  and  $1 - p_0$ , and feature densities  $f_0(z)$  and  $f_1(z)$  respectively. The complete feature density is therefore a mixture  $f(z) = p_0 f_0(z) + (1 - p_0) f_1(z)$ . Thus, according to the Bayes theorem we obtain the a posteriori probability

$$\text{Prob}\{\text{unchanged} \mid z\} = p_0 f_0(z) / f(z). \quad (1)$$

The local false discovery rate is defined as:

$$\text{FDR}(z) \equiv f_0(z) / f(z). \quad (2)$$

By ignoring the factor  $p_0$  in (1),  $\text{FDR}(z)$  can be considered as an upper bound for the probability of observing the unchanged class for an observation with a given value of  $z$ -score. In accordance with the considered change detection problem, we can assume that  $p_0$  is sufficiently close to 1 (say,  $p_0 = 0.8$  or  $p_0 = 0.9$ ), so that  $\text{FDR}(z)$  is a conservatively biased estimator [9] of (1). The FDR-based decision consists in performing the thresholding and reporting  $z$ -scores with  $\text{FDR}(z) \leq \gamma$  as changed.

At this point, the difference between a FDR-approach and more classical FWER-procedures becomes apparent. Both are designed to control the rejection of a true  $H_0^{\mathcal{F}}$  hypothesis probability, and whereas FWER-methods control this quantity directly, FDR-methods control the proportion of false positives among all rejected null hypotheses (*discoveries*). The advantage of the FDR-control is that it leads to fewer type II errors than controlling the FWER at the same level [10].

A sensitive part in the FDR-approach is the choice of the null distribution  $f_0(z)$  for the unchanged  $z$ -scores. If chosen perfectly, i.e., so that all the unchanged scores have  $f_0(z) > 0$  and  $f_1(z) = 0$  and vice versa for the changed  $z$ -scores, then  $p_0$  is an unbiased estimate of the true percentage of the unchanged observations [9]. However, such a choice is impossible and, therefore, an empirical procedure needs to be employed in order to estimate the null distribution  $f_0(z)$ . As has been generally observed and demonstrated in several simple cases [9], under the null hypothesis, the  $z$ -scores demonstrate a pronounced normal behavior even in cases, when substantial dependences are present in the data.

To estimate the null distribution we assume that the changed pixels are those that report  $z$ -scores which lie far from the observed distribution center. In other words, the zero assumption is that the non-null density  $f_1(z)$  is zero near  $z = 0$ , and, thus, the mixture model is identifiable from the central histogram counts. This assumption holds if the chosen statistical feature is appropriate for the considered change detection problem, i.e., if it corresponds well to the significant changes. We employ the central matching approach [9] by assuming that near the central score  $z = 0$  the representation  $f(z) = p_0 f_0(z)$  holds, and that  $f_0(z)$  is normally distributed  $\mathcal{N}(\delta_0, \sigma_0^2)$ . By taking the logarithm we obtain

$$\log f(z) = -\frac{1}{2\sigma_0^2} z^2 + \frac{\delta_0}{\sigma_0^2} z + \left[ \log p_0 - \frac{\delta_0^2}{2\sigma_0^2} + \frac{1}{2} \log(2\pi\sigma_0^2) \right].$$

The central matching strategy consists in estimating  $f_0(z)$  by assuming that  $\log f(z)$  is quadratic near  $z = 0$ , and, therefore, can be presented as:

$$\log f(z) = \beta_2 z^2 + \beta_1 z + \beta_0. \quad (3)$$

The estimates  $(\hat{\delta}_0, \hat{\sigma}_0^2) = \left( -\frac{\beta_1}{2\beta_2}, -\frac{1}{2\beta_2} \right)$  allow to arrive at the empirical null  $f_0(z) = \mathcal{N}(\hat{\delta}_0, \hat{\sigma}_0^2)$ . A critical issue is the choice of the  $z$ -scores near  $z = 0$ , that are used to estimate  $(\beta_1, \beta_2)$ , and correspond only to unchanged observations. In this paper, we employ a conservative strategy by taking  $\Delta = 50\%$  of the data for the fitting of (3). To define the interval of the feature histogram counts that are employed for  $f_0(z)$  estimation, we adopt the following strategy. We start with an empty set of histogram counts and first add the  $z = 0$  score. We then iteratively add the biggest count of the smoothed histogram adjacent to the currently considered histogram interval until it includes 50% of observations. Then, the selected part of the histogram is fitted by the right hand side of (3) by least squares [9]. Finally, we get the estimates  $(\hat{\delta}_0, \hat{\sigma}_0^2)$  as defined above. An example of such an empirical fit with  $f(z)$ ,  $f_0(z)$  and  $\mathcal{N}(0, 1)$  in presented in Fig. 1(a).

## 3. CHANGE DETECTION

In order to accommodate a sample-based approach to construct a pixel-based detection map we employ local patches.

---

**Algorithm 1:** FDR-based change detector
 

---

- 1 calculate feature  $f(i, j)$  for all pixels  $(i, j)$ ;
  - 2 transform  $f(i, j)$  into scores  $z(i, j)$ ;
  - 3 estimate the null density  $f_0(z)$ ;
  - 4 approximate the density  $f(z)$  with splines;
  - 5 calculate  $\text{FDR}(z)$  for all score values;
  - 6 report as changed the pixels with  $\text{FDR}(z) < \gamma$ .
- 

Specifically, for a pixel at location  $(i, j)$ , we consider the sample obtained from the values in a  $S \times S$  local window centered at  $(i, j)$ . It is immediate that such samples cannot be considered independent in themselves, because of the contextual dependences in the image (*interdependence*). Moreover, we consider the change detection problem, in which we assume that the amount of significant changes cannot be large. As such, two patch samples centered at the same location of two images  $X$  and  $Y$  are dependent (*intradependence*).

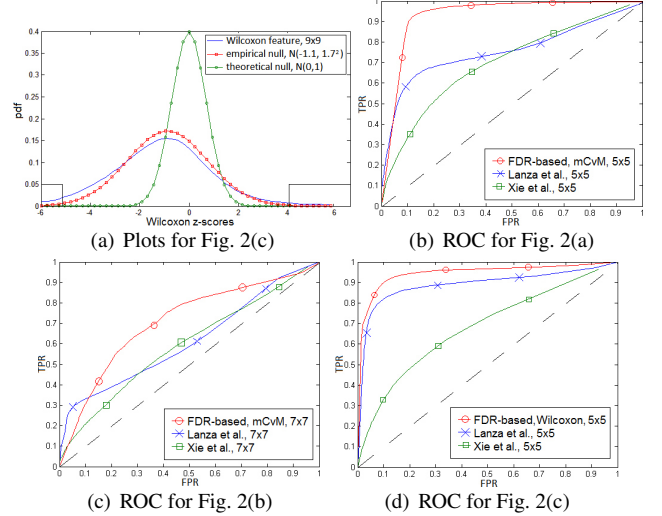
We propose a FDR-based change detector that allows to address: (1) inference based on correlated individual tests (due to overlap of patch-based samples), and (2) interdependence. As has been analysed in [9], the FDR-formulation is robust to a certain level of individual tests' correlation. We will assume that the level of interdependence is moderate, so that the distribution of  $z$ -scores is close to normal. Note that this is a common practical assumption, and that normal behavior of  $z$ -scores is commonly observed [9]. Also note that the majority of existing change detection approaches [4–7] operate under various independency assumptions.

The FDR-based change detector is presented in Algorithm 1. The estimation involved in step 4 can be performed by any interpolation approach, e.g., by natural splines [9]. The choice of  $\gamma$  is similar to that of the significance level in classical inference. Finally, the local patch size  $S$  should be chosen manually with respect to the scale of significant changes, which constitutes the same strategy as in [4–7].

#### 4. STATISTICAL FEATURES

The appropriate features are two-sample statistics with a known distribution under the no-change hypothesis. Encouraged by the positive results obtained in [5–7] we employ several order (rank-based) statistics in the role of features. These non-parametric statistics allow a sufficient level of generality without imposing distribution assumptions.

**Wilcoxon feature.** Given two independent random variables  $X$  and  $Y$ , the paired Wilcoxon statistic [8] tests the null hypothesis  $\{H_0^W : P(X < Y) = 1/2\}$ , which characterizes the comparative central symmetry of two samples. The observed values  $Z_i = |X_i - Y_i|$ ,  $i = 1, \dots, N$ ,  $N = S^2$ , where  $S$  is the local patch size, are ordered as  $Z_{(1)} < \dots < Z_{(N)}$ . The ranks of the positive differences  $(X_i - Y_i)$  in  $Z_{(N)}$  are



**Fig. 1.** (a) The normalized histogram and its estimates for the Wilcoxon feature on Fig. 2(c) image pair. The  $z$ -scores in boxes correspond to the detection areas with  $\text{FDR} \leq 0.1$ . (b)-(d) ROC curves for all detectors from Fig. 2 (a)-(c).

summed to get  $W_+$ . The standardized Wilcoxon statistic

$$W = \left( W^+ - \frac{N(N+1)}{4} \right) \left( \frac{N(N+1)(2N+1)}{24} \right)^{-\frac{1}{2}} \quad (4)$$

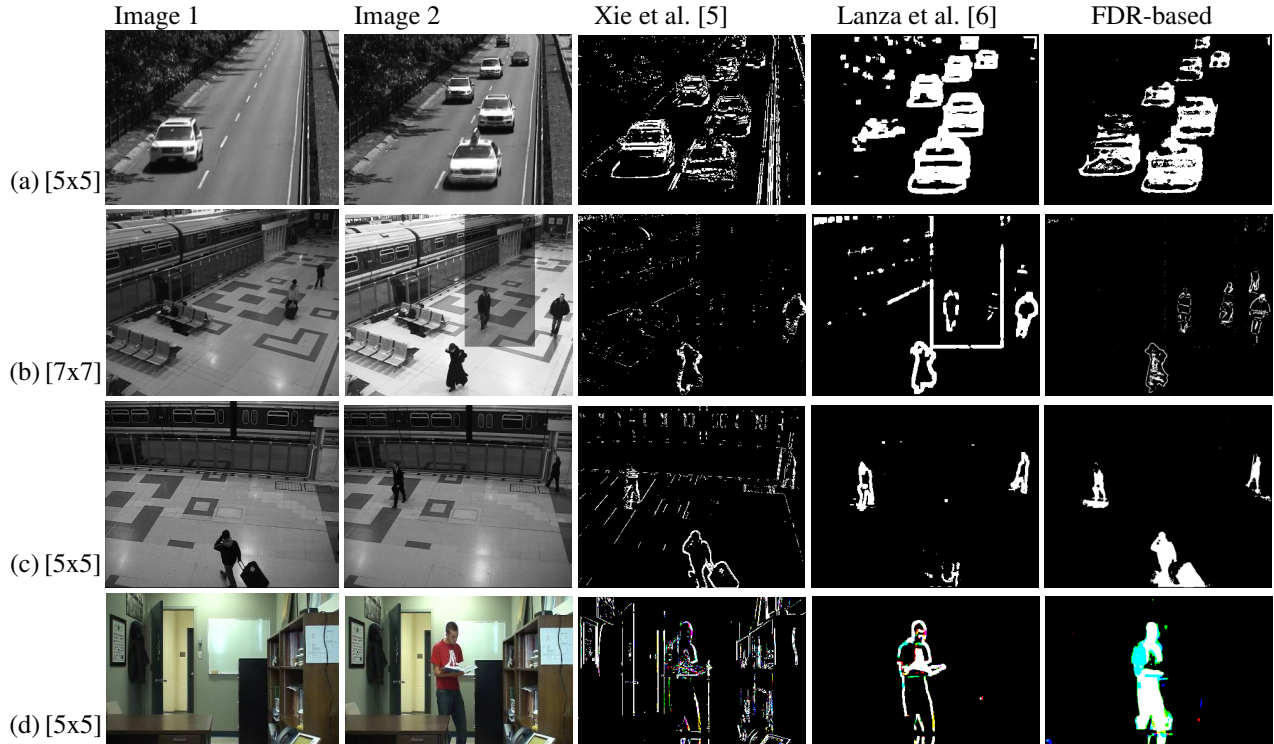
converges to the standard normal distribution  $\mathcal{N}(0, 1)$  [8] when  $N \rightarrow \infty$  under  $H_0^W$  with the independency assumptions. Thus, for large samples ( $N > 20$ , see [8]) the critical values for this test can be drawn from the normal approximation. The rejection region of the test is two-sided [8].

**Cramér-von Mises feature.** The two-sample Cramér-von Mises (CvM) statistic [8, 13] is designed to verify the null hypothesis  $\{H_0^{\text{CvM}} : F_X \equiv F_Y\}$ , for CDFs  $F_X$  and  $F_Y$ . The observations  $X_i$  and  $Y_j$ ,  $i, j = 1, \dots, N$ ,  $N = S^2$ , are concatenated into sample  $Z$  of size  $2N$ , which is then ordered to obtain  $Z_{(1)} < \dots < Z_{(2N)}$ . Each of the initial samples is assumed to be preliminary sorted in the ascending order. Then,

$$T = \frac{\sum_{i=1}^N \left[ (r_i - i)^2 + (s_i - i)^2 \right]}{2N^2} - \frac{4N^2 - 1}{12N} \quad (5)$$

defines the standardized CvM statistic, with ranks  $r_i, s_i$ :  $Z_{(r(i))} = X_i$  and  $Z_{(s(i))} = Y_i$ . The rejection region of the CvM test is right-sided [13].  $T$  follows the Cramér-von Mises distribution with both parameters set to  $N$  [13]. The  $z$ -scores can be obtained using a numerical approximation of  $D_{\mathcal{F}}$  [14].

Along with CvM, we propose the use of a modified CvM (mCvM) feature, that is obtained by subtracting sample medians from both samples before constructing  $T$  (5). This allows focusing on fluctuations of higher order moments, ignoring the first moment changes (like constant illumination shifts). The distribution and critical region are the same as for CvM.



**Fig. 2.** (a),(d)  $360 \times 240$  and (b),(c)  $720 \times 580$  pixels still camera image pairs from dataset [2], and the change detection results by methods [5, 6] and the FDR-based approach with (a),(b) mCvM, (c) Wilcoxon and (d) CvM features.

## 5. EXPERIMENTAL RESULTS

In this section we validate the proposed FDR-based approach in still camera change detection application [2, 5, 6] on the images from the *changedetection.net* dataset [2]. The threshold for the FDR is set to  $\gamma = 0.1$ , see discussion in [9]. Four typical detection results are presented in Fig. 2 along with benchmark methods’ results (with variances estimated on 100 frames of the unobscured background and thresholds chosen manually). The mCvM feature (a), (b), with  $\gamma = 0.05$ , gives a contouring detection, similar to the benchmark methods. The image pair (b) is obtained manually by modifying the brightness of the images in order to demonstrate the robustness of the mCvM feature to the constant intensity shifts. Wilcoxon (c) and CvM (d) features detect the inner-parts of the changed objects as well, yet demonstrate sensitivity to shadows in (c). The colors in (e) represent detection performed separately in R/G/B channels. The ROC curves for image pairs (a)-(c) are presented in Fig. 1(b)-(d). The comparisons demonstrate that the considered features outperform the benchmark techniques. Note that the considered features perform differently and have to be selected taking into account their limitations, e.g., sensitivity to shadows of the Wilcoxon feature.

We have observed only slight deviations of  $z$ -scores from the normal distribution due to interdependence, which ex-

perimentally validates the assumptions of Section 3. Note that the developed method (as well as similar spatial approaches [5–7]) underperforms in case of complicated background motion (waves/foliage), where the temporal approaches [2, 3] are expected to perform well. The computational complexity of the method is  $O(N)$ , same as that of [6].

## 6. CONCLUSIONS

We have proposed a novel statistical approach to patch-based change detection on image pairs. This approach gives a unified statistical thresholding procedure to perform change detection based on statistical features that have a known distribution under the no-change hypothesis. The approach involves only a few parameters and is highly parallelizable. We have presented several rank-based statistical features and the corresponding FDR-detectors positively compare with benchmark techniques on still camera images. Note that further features can be employed to design application-specific change detectors in various applications.

## 7. ACKNOWLEDGMENT

The work of the first author was funded by an INRIA post-doctoral fellowship.

## 8. REFERENCES

- [1] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, “Image change detection algorithms: a systematic survey,” *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, 2005.
- [2] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, “Changetection.net: A new change detection benchmark dataset,” in *Proc. IEEE Workshop on Change Detection (CDW’12), CVPR*, 2012.
- [3] C. Stauffer and W. E. L. Grimson, “Adaptive background mixture models for real-time tracking,” in *Proc. IEEE CVPR*, 1999, pp. 2246–2252.
- [4] N. Ohta, “A statistical approach to background subtraction for surveillance systems,” in *Proc. ICCV*, 2001, vol. 2, pp. 481–486.
- [5] B. Xie, V. Ramesh, and T. E. Boult, “Sudden illumination change detection using order consistency,” *Image Vision Comput.*, pp. 117–125, 2004.
- [6] A. Lanza and L. Di Stefano, “Statistical change detection by the pool adjacent violators algorithm,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 33, no. 9, pp. 1894–1910, 2011.
- [7] M. Singh, V. Parameswaran, and V. Ramesh, “Order consistent change detection via fast statistical significance testing,” in *Proc. IEEE CVPR*, 2008.
- [8] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*, Springer, London, 2005.
- [9] B. Efron, *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*, Cambridge University Press, Cambridge, 2010.
- [10] Y. Benjamini and Y. Hochberg, “Controlling the false discovery rate: A practical and powerful approach to multiple testing,” *J. Royal Stat. Soc. B*, vol. 57, pp. 289–300, 1995.
- [11] H. J. Seo and P. Milanfar, “Action recognition from one example,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 33, no. 5, pp. 867–882, 2011.
- [12] V. Krylov, G. Moser, S. Serpico, and J. Zerubia, “Change detection with synthetic aperture radar images by Wilcoxon statistic likelihood ratio test,” in *Proc. IEEE ICIP*, 2012, pp. 2093–2096.
- [13] T. W. Anderson, “On the distribution of the two-sample Cramer-von Mises criterion,” *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1148–1159, 1962.
- [14] Y. Xiao, A. Gordon, and A. Yakovlev, “A C++ program for the Cramér-von Mises two-sample test,” *J. Stat. Softw.*, vol. 17, no. 8, pp. 1–15, 2006.