Feature Fusion based Unsupervised Change Detection in Optical Satellite Images

Neha Gupta, Pooja Singh, and Samit Ari
Department of Electronics and Communication Engineering,
National Institute of Technology, Rourkela, India 769008.
Email: neha27brs@gmail.com, psnitr@gmail.com, samit@nitrkl.ac.in

***Abstract*—This paper proposes a feature fusion technique for unsupervised change detection. Features extracted from two different techniques are fused to get the final feature vectors. The first technique utilizes the Gabor wavelet at multiple orientations and scales, where maximum magnitude over all orientation in each scale is taken to create features of two multitemporal satellite images. The second technique applies canonical correlation analysis (CCA) on the combination of original multispectral bands and extracted local neighborhood information from all the bands. Next, the difference feature vectors obtained from individual techniques are fused to generate the final feature vectors. Furthermore, to get the binary change map, fuzzy c-means clustering is applied on final extracted features. In this feature fusion, the local neighborhood information from Gabor wavelet kernel is combined with joint change information from group of pixels extracted by CCA to produce more discriminant features. Experiments conducted on optical satellite images, which are collected by two sensors of Landsat satellite, and it shows the better performance of the proposed technique compared to earlier stated techniques.**

***Index Terms*—Binary change map, Canonical correlation analysis (CCA), Change detection, fuzzy c-means clustering, Gabor wavelet kernel, multitemporal satellite image.**

I. Introduction

With the advancement of sensor technologies, continuous monitoring of earth surface became more feasible by capturing the remote sensing images through sensors. Earth surface monitoring involves analysis of the changes that have occurred over a period of time. The detection of these changes can be achieved by identifying the differences in the images taken at two different times [1]. These remote sensing images are acquired by passive optical sensors that detect natural energy reflected by the scene or object being observed. Now a days, optical multi- or hyperspectral images are easily accessible that allows us to detect changes more effectively [2].

Recently, unsupervised change detection techniques are widely used in the field of remote sensing as it does not require any prior information about the data being analyzed [1]. Some of the earlier unsupervised techniques are applied directly either on multitemporal images or on difference image. In [3], binary descriptors are generated by analyzing the multitemporal images separately. These descriptors are compared by using Hamming distance for each pixel, and Lloyd−Max's algorithm used to generate the binary change map. It is represented as binary descriptor and Lloyd Max's (BDLM) technique. By using inter image information, in [4]–[6], binary descriptor based technique is again applied to get the binary feature vectors. A combination of thresholds to generate the binary feature vectors is taken in [4]. In the same way, combination of thresholds based technique is again applied on satellite image in [7]. In difference image based techniques, the difference image is examined in both domains i.e. original domain and feature space domain [8]. This feature space is obtained by transforming the original data to reproducing kernel Hilbert space (RKHS), and applying kernel k-means (KKM) clustering in this feature space provides more discriminant information. In [9], Gabor wavelet kernels are applied on difference image to obtain the feature vectors, which will be clustered by using two level clustering. Now, by jointly exploiting the relationship between two groups of data in [10], [11], canonical correlation analysis (CCA) based multivariate alteration detection (MAD) technique is used to get change detection map of multitemporal multispectral satellite images. In MAD technique, the canonical variates obtained by CCA are subtracted from each other to get the change information between two images.

At present, fusion based techniques have gained much attention in the field of change detection. In [12], the complementary information from log-ratio and mean-ratio images are fused to generate the difference image. This will take the advantage of restraining the unchanged areas or background and enhancement of the changed areas that are obtained by log-ratio image and mean-ratio image, respectively. In addition, the source images are fused based on the two different fusion rules i.e. minimum standard deviation and weight averaging, which are applied on wavelet coefficients for high- and low-frequency bands, respectively. Moreover, a multiple wavelet fusion kernel (MWF kernel) method is proposed by again fusing the complementary information of two types of difference images i.e. the ratio image and the subtraction image [13]. The smooth and homogeneous changed areas obtained by subtraction image and suppression of complex background obtained by ratio image are fused using MWF kernel to provide less missed detections and less false detections, respectively.

Information fusion plays an important role in data or image processing [14]. Fusion in pattern recognition is done by three different ways: 1) pixel level fusion, 2) feature level fusion, and 3) decision level fusion. Among these three methods, feature level fusion provides more discriminative information [15]. Therefore, this work mainly focuses on feature level fusion technique. Some of the above mentioned techniques fuses information at difference image level, whereas, the proposed work fuses the feature vector obtained from the two different tech-

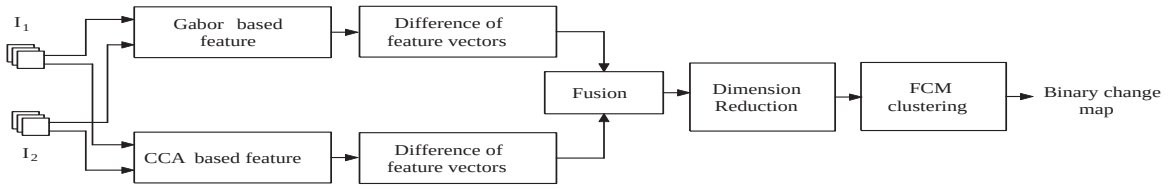


Fig. 1. Overall procedure of the proposed change detection technique.

niques. First, features are extracted by applying Gabor wavelet kernel [9], [16] and CCA techniques [11], [17] on multitemporal multispectral images. Then, the difference feature vectors are generated individually for both methods. Now, these two difference feature vectors are fused serially to get the final feature vector. Fuzzy c-means (FCM) clustering [18] is applied on the final feature vectors to separate the pixels into unchanged and changed classes.

II. Proposed Methodology

Let multitemporal multispectral optical satellite images I_1 and I_2 are taken at two different times t_1 and t_2. Both the images are of same geographical area and of same size $m \times n$. The primary focus of the work is to generate the binary map of any geographical region to indicate change and no change details within acquired acquisition dates. The proposed technique procedure is shown in Fig. 1 through block diagram. The proposed approach consists of the following steps: feature extraction, feature fusion, and clustering. In feature extraction, two methodologies are used viz., Gabor based feature extraction and CCA based feature extraction. For feature fusion, serial fusion strategy is used. Then, extracted features are clustered by using FCM algorithm to generate the binary map. The detailed explanation of the proposed method is as follows.

A. Feature extraction

Here, features are extracted by using two different methodologies i.e. Gabor wavelets and CCA technique. Gabor wavelets are applied individually on two multitemporal multispectral satellite images, whereas, CCA takes both the images at a time and finds the correlation based features. These methodologies are explained below.

1) Gabor based feature extraction: Gabor wavelet captures the salient visual properties such as orientation selectivity, spatial localization, spatial frequency characteristics. Also, due to the biological relevance and computational properties, and extension of Gabor wavelets from 1-D to the 2-D case, it has been successfully applied to image analysis. 2-D Gabor wavelet kernel is defined as follows:

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2}\exp\left(-\frac{\|k_{u,v}\|^2\|z\|^2}{2\sigma^2}\right)\left[\exp(ik_{u,v}z) - \exp\left(-\frac{\sigma^2}{2}\right)\right] \quad (1)$$

where u, v represent the orientation and scale of the Gabor kernels, respectively, $z = (m, n)$ denotes the pixel position, $\|\bullet\|$ is the norm operator, $k_{u,v}$ denotes the wave vector and is defined as follows:

$$k_{u,v} = k_v \exp(j\phi_u) \quad (2)$$

where $k_v = k_m / f^v$ and $\phi_u = \pi u/8$. k_m is the maximum frequency, and f is the spacing factor between kernels in the frequency domain.

Since multispectral data is used here that contains number of spectral bands, therefore, to get the Gabor feature representation, each of the single channel of multispectral data is convolved with Gabor wavelet kernel.

The convolution output for each band of each image is defined as

$$Y_{u,v}(z) = I_{t_B} * \psi_{u,v}(z) \quad (3)$$

where $Y_{u,v}(z)$ is the convolution result corresponding to the Gabor kernel at orientation u and scale v. I_{t_B} is the single band image where t denotes the acquisition time and B corresponds to the number of spectral bands. $*$ denotes the convolution operator. The total number of orientations $U = 4$ and scales $V = 3$ are used in Gabor kernel.

The convolution result is a complex-valued quantity in which $\mathrm{Re}(Y_{u,v}(z))$ and $\mathrm{Im}(Y_{u,v}(z))$ are the real and imaginary parts, respectively. Then, $Y_{u,v}(z)$ can be written as

$$Y_{u,v}(z) = M_{u,v}(z)\ \exp(\mathrm{j}\theta_{u,v}(z)) \quad (4)$$

where

$$M_{u,v}(z) = \sqrt{\mathrm{Re}(Y_{u,v}(z))^2 + \mathrm{Im}(Y_{u,v}(z))^2} \quad (5)$$

$$\theta_{u,v}(z) = \arctan\left(\frac{\mathrm{Im}(Y_{u,v}(z))}{\mathrm{Re}(Y_{u,v}(z))}\right) \quad (6)$$

Here, magnitude $M_{u,v}(z)$ is considered for the analysis as it integrates the smoothing and edge detection properties from its real and imaginary part, respectively.

In this paper, multitemporal and multispectral data have been considered, which has six spectral bands. Now, if magnitude response is generated for all bands in all orientations and scales then the generated features will contain high dimensional data. To overcome the dimensionality problem, maximum magnitude over all orientation is taken instead of taking all magnitude responses, i.e.,

$$f_v = \max_{\mathrm{u}\in[0,\mathrm{U}-1]} M_{u,v}(z) \quad (7)$$

The obtained compact feature vector F for each pixel and corresponding scales is defined as

$$F_t^B = [f_0, f_1, \cdots, f_v, \cdots, f_{V-1}] \quad (8)$$

Similarly, the feature vector F is obtained for all bands of both multitemporal images and the new feature vector of each multitemporal image is represented as follows:

$$\chi_1 = [F_1^1, F_1^2 ....., F_1^6] \tag{9}$$

$$\chi_2 = [F_2^1, F_2^2 ....., F_2^6]. \tag{10}$$

2) Canonical correlation based feature extraction: CCA is the way of investigating the relationship between two groups of multidimensional variables. It provides the basis vectors, which are linearly combined with the original variables, such that the correlation between the transformed variables is mutually maximized. The linear combination of the basis vectors with original variables is called canonical variates that is found corresponding to the largest correlation, and these canonical variates are orthogonal to each other. Moreover, CCA can be used to find the correlation between the two multidimensional group of variables having different dimensions.

Here, the multidimensional variables are created by taking the spatial neighborhood information around each pixel for each band of multitemporal images. To get the spatial neighborhood information, each image of multispectral bands is partitioned into overlapping patches or blocks. After that, these blocks are converted into vectors for each pixel. These vectors of all bands of multitemporal images are concatenated separately.

Suppose, vectors for single band of each image is represented as

$$\Gamma = [\alpha_1, \alpha_2, ..., \alpha_p] \tag{11}$$

$$\Lambda = [\beta_1, \beta_2, ..., \beta_p] \tag{12}$$

where α and β are the pixels of the block corresponding to both the multispectral images I_{1_B} and I_{2_B}, respectively. p is the number of pixels in a block.

The multidimensional variables on which CCA will be applied are represented as

$$P = [\Gamma_1, \Gamma_2, ..., \Gamma_B] \tag{13}$$

$$Q = [\Lambda_1, \Lambda_2, ..., \Lambda_B]. \tag{14}$$

CCA algorithm determines the basis vectors

$$\gamma = [a_1, a_2, ..., a_m] \qquad \lambda = [b_1, b_2, ..., b_m] \tag{15}$$

where $m = p \times B$. The new vectors of the projected data are obtained using these basis vectors are: $R = \gamma^T P,\ S = \lambda^T Q$. The elements of the vectors R and S are referred to as the canonical variates. The linear combinations R and S of two sets of multivariate data P and Q are obtained with maximum correlation

$$\rho = \mathrm{Corr}\{\gamma^T P, \lambda^T Q\} = \frac{\mathrm{Cov}\{\gamma^T P, \lambda^T Q\}}{\sqrt{\mathrm{Var}\{\gamma^T P\}\mathrm{Var}\{\lambda^T Q\}}} \tag{16}$$

Without loss of generality, let us assume that P and Q are zero mean, i.e., $\mathrm{E}\{P\} = 0$ and $\mathrm{E}\{Q\} = 0$. Since $\mathrm{Cov}\{\gamma^T P, \lambda^T Q\} = \gamma^T \sum_{12} \lambda$, where $\sum_{12}$ is the covariance between P and Q, and the variances $\mathrm{Var}\{\gamma^T P\} = \gamma^T \sum_{11} \gamma$ and $\mathrm{Var}\{\lambda^T Q\} = \lambda^T \sum_{22} \lambda$, now the obtained correlation is

$$\rho = \frac{\gamma^T \sum_{12} \lambda}{\sqrt{\gamma^T \sum_{11} \gamma \cdot \lambda^T \sum_{22} \lambda}} \tag{17}$$

The maximum of ρ with respect to coefficients γ and λ is referred as the maximum canonical correlation.

B. Fusion of feature vectors

Now a days, fusion in pattern recognition is done by three different ways: 1) pixel level fusion, 2) feature level fusion, and 3) decision level fusion. This work mainly focuses on feature level fusion as it has grabbed much attention in some recent researches. The main advantage of feature level fusion is that it provides effective discriminant information after combining feature sets derived from multiple techniques and it also improves the accuracy. Here, multiple feature vectors are fused in end to end manner so as to obtain the final feature vector.

Herein, the final feature vectors contains the Gabor features and CCA based features. As mentioned above, Gabor features are extracted for individual multispectral images, and linear combinations of the multivariable data that contain the spatial neighborhood information is obtained by CCA technique. Generally, change detection analysis of satellite image is performed by directly taking the difference image, while taking the difference in transformed feature space provides more effective results. Therefore, difference is taken after transforming the data into new feature space.

Now, difference between the features of both techniques are taken and obtained difference feature vectors are combined into the final feature vectors. The difference feature vectors are as follows:

$$\phi = \chi_1 - \chi_2 \tag{18}$$

$$\eta = R - S \tag{19}$$

where ϕ and η are difference feature vectors obtained from Gabor and CCA based feature extraction, respectively. Final feature vector is obtained by fusing both the difference feature vectors and denoted as

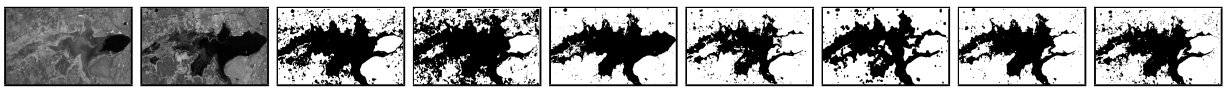
$$Z = [[\phi_1, \phi_2, \cdots \phi_{mn}]^T, [\eta_1, \eta_2, \cdots \eta_{mn}]^T]. \tag{20}$$

For the simplicity of notations, suppose $z = [\phi, \eta]$ then Z can be written as

$$Z = [z_1, z_2, ..., z_{mn}]^T. \tag{21}$$

C. PCA based dimensionality reduction

The Gabor wavelet kernel are derived for multiple scales and orientations and convolving them with data provides redundant results. In addition, using multispectral images may provide repetitive information along with the discriminative information. To overcome this issue, PCA technique is used to reduce the dimensionality of the feature vector, this in turn will increase the efficiency of the clustering process.



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Fig. 2. Illustration of the binary change map of optical satellite images of dataset I. (a) and (b) are multitemporal images of Landsat 5. (c) ground truth. (d), (e), (f), (g) and (h) are binary change map of BDLM, KKM, GaborTLC, NTIRC, OTLNI and proposed methods, respectively.

(a) (b) (c) (d) (e) (f) (g) (h) (i)

Fig. 3. Illustration of the binary change map of optical satellite images of dataset II. (a) and (b) are multitemporal images of Landsat 7. (c) ground truth. (d), (e), (f), (g) and (h) are binary change map of BDLM, KKM, GaborTLC, NTIRC, OTLNI and proposed methods, respectively.

D. Clustering and binary change map generation

Once the final feature vectors are generated, binary change map can be obtained by classifying the feature vector of each pixel into changed and unchanged classes by using the FCM algorithm.

FCM is a soft clustering algorithm in which every data point may belong to more than one cluster rather than belonging to just one cluster as in k-means clustering. Therefore, points in the center of the cluster, may be in the cluster to greater degree than points in the edge of cluster.

The final feature vector Z corresponding to all pixels are given as input to FCM clustering. To find fuzzy partitioning of all pixels, the FCM algorithm is applied on Z by minimizing the objective function

$$\mathcal{J}_d(\mathcal{U},\mathcal{V}) = \sum_{i=1}^{c}\sum_{j=1}^{mn} \mu_{ij}^{d} \left\| z_j - \nu_i \right\|^2 \tag{22}$$

where d is the degree of fuzziness and it is taken as $d \in [1, +\infty)$. μ_{ij} is the membership grade that belongs to ith cluster for jth pixel and represented with the partition matrix $\mathcal{U} = [\mu_{ij}]_{c\times mn}$. $\mathcal{V} = [\nu_1, \nu_2]$ denotes the cluster's centroids.

Fuzzy partitioning is performed by updating the membership grade μ_{ij} and the cluster centroids ν through an iterative optimization of the objective function $\mathcal{J}_d$. Here, $d = 2$, number of clusters, $c = 2$ are taken, and the iteration procedure is initialized at $l = 0$.

The centroids and membership grades are calculated as follows:

$$\nu_i^{(l+1)} = \frac{\sum_{j=1}^{mn} \left(\mu_{ij}^{(l)}\right)^d z_j}{\sum_{j=1}^{mn} \left(\mu_{ij}^{(l)}\right)^d} \tag{23}$$

$$\mu_{ij}^{(l+1)} = \frac{\left\| z_j - \nu_i^{(l+1)} \right\|^{-2/(d-1)}}{\sum_{r=1}^{c} \mu_{ij}^{d} \left\| z_j - \nu_r^{(l+1)} \right\|^{-2/(d-1)}} \tag{24}$$

$$w_i = \arg \max_{i=1,2} \{\mu_{ij}\} \tag{25}$$

where w represents the class. According to the maximum value of membership grade, the pixels are assigned to the corresponding classes and hence, the binary change map is generated.

III. Experimental Results

The experiments are carried out on two real data sets to check the performance of the proposed method.

A. Dataset Description

1)$DatasetI$: This pair of multitemporal images is captured by Landsat 5 Thematic Mapper (TM) sensor. The observing site is upper lake in Bhopal city, Madhya Pradesh, India. They are captured on May 29, 2009, and November 09, 2011 [4], [7], [19]. The size of the images is 206×424 pixels. The dried lake is the main land change that occurs due to insufficient rains in 2009.

2)$DatasetII$: This pair of multitemporal images is captured by Landsat 7 Enhanced Thematic Mapper Plus (ETM+) sensor. The observing site is Natural Lake, Rajasthan, India. They are captured on February 09, 2001 and September 21, 2001 [4], [7], [19]. The size of the images is 220×550 pixels. The main land change is the dried lake due to summer.

B. Qualitative and quantitative results

Both qualitative and quantitative results are shown to analyze binary change map. In case of qualitative results, visual results are taken. The performance of the proposed technique is compared with KKM [8], BDLM [3], GaborTLC [9], descriptor based technique (or NTIRC) [4], and combination of threshold based method called Otsu's thresholding of local neighborhood information (OTLNI) [7] methods. The visual results of dataset I and II are shown in Fig. 2 and 3, respectively, where white pixels represent the unchanged areas and black pixels represent the change areas. By visualizing the qualitative results, it can be observed that BDLM techniques generate noisy results with more false alarms, and KKM technique produces noise free visual results with more false alarms. Comparatively, result of GaborTLC technique is less noisy with less false alarms rate, and NTIRC method detects more changed pixels with high false alarms. OTLNI technique yields better results results compared to other technique. The proposed technique detects changes more accurately with less noise and less false alarms comparatively to other methods.

In case of quantitative results, some predefined parameters are taken to compare the performance with ground truth [3], [9]. The predefined measures are: Overall accuracy (OA), False alarms (P_{FA}), Total error (P_{TE}) and Kappa (κ). All these parameters are calculated in

Table I
QUANTITATIVE PERFORMANCE (IN PERCENTAGE)

Dataset	Method	OA	P_{FA}	P_{TE}	κ
Dataset I (Bhopal City)	BDLM [3]	82.92	25.63	17.08	66.31
	KKM [8]	82.48	15.26	17.52	64.37
	GaborTLC [9]	85.46	1.32	14.54	69.42
	NTIRC [4]	85.81	8.73	14.19	70.83
	OTLNI [7]	87.75	1.23	12.25	74.40
	Proposed	89.35	2.23	10.65	77.89
Dataset II (Natural Lake)	BDLM [3]	80.80	25.00	19.20	60.74
	KKM [8]	80.85	8.36	19.15	53.21
	GaborTLC [9]	87.99	0.24	12.01	69.63
	NTIRC [4]	91.08	7.72	8.92	79.85
	OTLNI [7]	91.24	1.07	8.76	78.69
	Proposed	91.69	1.86	8.31	80.04

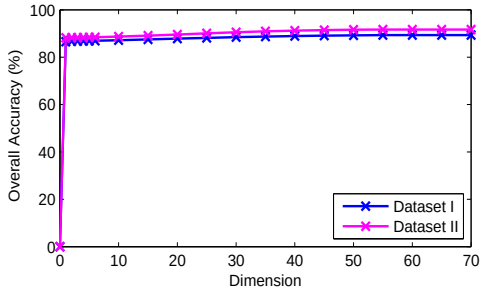


Fig. 4. Overall Accuracy of the proposed method versus dimensions for dataset I and dataset II.

percentage. The quantitative measures for the methods are shown in Table 1. In this experiment, the maximum frequency (k_m) and spacing factor (f) of Gabor wavelet kernel are taken as 6.28 and $\sqrt{2}$, respectively. The square blocks considered in CCA are of size 3×3. In comparison to earlier reported techniques as stated in Table 1, the proposed technique yields better performance in terms of overall accuracy and kappa value. Moreover, the proposed methods provides less total error with reduced false alarms.

The merits of the proposed method lies in utilizing the feature fusion of Gabor feature vectors and CCA feature vectors. Generally, Gabor wavelet extracts the local spatial neighborhood information that reduces the noise effects. Moreover, applying Gabor wavelet kernel on individual band images instead of applying it into difference image, provides abundant spatial neighborhood information that produces more distinguished features. In addition, the difference vectors created by CCA technique provides maximum details on joint change in all spectral bands, and also considered local neighborhood in the analysis of CCA effectively utilizes the joint information present in the group of pixels. Hence, more discriminant features are produced.

Fig. 4 shows the plot of overall accuracy versus variation of dimensions. It can be observed that even though reducing the dimensions of features, there is not much significant changes in the overall accuracy of the proposed method.

IV. CONCLUSION

In this paper, feature fusion based an unsupervised change detection technique is proposed, where features are extracted by Gabor wavelet kernel and CCA techniques. Here, the difference is taken in new feature space instead of taking in original domain, which provides more discriminative results. Finally, FCM clustering is applied to the obtained feature vectors. The effectiveness of the technique is demonstrated by experiments that are carried out on real multispectral satellite images. The proposed method yields improved performance compared to KKM, BDLM, NTIRC, OTLNI, and GaborTLC techniques.

ACKNOWLEDGMENT

This Publication is an outcome of the R&D work undertaken in the project under the Visvesvaraya PhD Scheme of Ministry of Electronics & Information Technology, Government of India, being implemented by Digital India Corporation (formerly Media Lab Asia). [grant number PhD-MLA/4(13)/2015-16].

REFERENCES

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

[2] M. J. Canty and A. A. Nielsen, "Linear and kernel methods for multivariate change detection," *Computers & geosciences*, vol. 38, no. 1, pp. 107–114, 2012.

[3] A. Radoi and M. Datcu, "Automatic change analysis in satellite images using binary descriptors and lloyd–max quantization," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 6, pp. 1223–1227, 2015.

[4] N. Gupta, G. V. Pillai, and S. Ari, "Change detection in optical satellite images based on local binary similarity pattern technique," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 3, pp. 389–393, 2018.

[5] N. Gupta, G. V. Pillai, and S. Ari, "Unsupervised change detection in optical satellite images using binary descriptor," in *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. IEEE, mar 2017.

[6] G. V. Pillai, N. Gupta, and S. Ari, "Descriptors based unsupervised change detection in satellite images," in *2017 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, apr 2017.

[7] N. Gupta, G. V. Pillai, and S. Ari, "Change detection in landsat images based on local neighbourhood information," *IET Image Processing*, vol. 12, no. 11, pp. 2051–2058, nov 2018.

[8] M. Volpi, D. Tuia, G. Camps-Valls, and M. Kanevski, "Unsupervised change detection with kernels," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 6, pp. 1026–1030, 2012.

[9] H.-C. Li, T. Celik, N. Longbotham, and W. J. Emery, "Gabor feature based unsupervised change detection of multitemporal sar images based on two-level clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2458–2462, 2015.

[10] A. A. Nielsen, K. Conradsen, and J. J. Simpson, "Multivariate alteration detection (mad) and maf postprocessing in multispectral, bitemporal image data: New approaches to change detection studies," *Remote Sens. Environ.*, vol. 64, no. 1, pp. 1–19, 1998.

[11] A. A. Nielsen, "The regularized iteratively reweighted mad method for change detection in multi-and hyperspectral data," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 463–478, 2007.

[12] J. Ma, M. Gong, and Z. Zhou, "Wavelet fusion on ratio images for change detection in sar images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 6, pp. 1122–1126, 2012.

[13] L. Jia, M. Li, P. Zhang, Y. Wu, L. An, and W. Song, "Remote-sensing image change detection with fusion of multiple wavelet kernels," *IEEE J. Sel. Topics. Appl. Earth Observ. Remote Sens.*, vol. 9, no. 8, pp. 3405–3418, 2016.

[14] Q.-S. Sun, S.-G. Zeng, Y. Liu, P.-A. Heng, and D.-S. Xia, "A new method of feature fusion and its application in image recognition," *Pattern Recognit.*, vol. 38, no. 12, pp. 2437–2448, 2005.

[15] Y.-H. Yuan, Q.-S. Sun, Q. Zhou, and D.-S. Xia, "A novel multiset integrated canonical correlation analysis framework and its application in feature fusion," *Pattern Recognit.*, vol. 44, no. 5, pp. 1031–1040, 2011.

[16] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 467–476, 2002.

[17] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.

[18] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media, 2013.

[19] [Online]. Available: http://earthexplorer.usgs.gov/