Feature-Level Change Detection Using Deep Representation and Feature Change Analysis for Multispectral Imagery

Hui Zhang, Maoguo Gong, *Senior Member, IEEE*, Puzhao Zhang, Linzhi Su, and Jiao Shi

***Abstract*—Due to the noise interference and redundancy in multispectral images, it is promising to transform the available spectral channels into a suitable feature space for relieving noise and reducing the redundancy. The booming of deep learning provides a flexible tool to learn abstract and invariant features directly from the data in their raw forms. In this letter, we propose an unsupervised change detection technique for multispectral images, in which we combine deep belief networks (DBNs) and feature change analysis to highlight changes. First, a DBN is established to capture the key information for discrimination and suppress the irrelevant variations. Second, we map bitemporal change feature into a 2-D polar domain to characterize the change information. Finally, an unsupervised clustering algorithm is adopted to distinguish the changed and unchanged pixels, and then, the changed types can be identified by classifying the changed pixels into several classes according to the directions of feature changes. The experimental results demonstrate the effectiveness and robustness of the proposed method.**

***Index Terms*—Change detection, change vector analysis (CVA), cosine angle distance (CAD), deep belief networks (DBNs), multispectral images.**

I. Introduction

CHANGE detection is a hot topic with a wide range of applications, including urban planning [1], agricultural surveys, environment monitoring, and hazard assessment [2]. With the development of remote sensing technology, multispectral data over the same scene at different times are now available, with high spatial and spectral resolutions. The focus of this letter is to highlight the changes where the materials within have been replaced and suppress the meaningless changes caused by the environments, such as illumination, calibration, parallax, registration, diurnal, and seasonal variations [3].

In the literature, many methods have been proposed to identify the changes between bitemporal images. Generally, these change detection methods can be summarized into four categories.

1) Image arithmetical operations, such as differencing, ratioing, and change vector analysis (CVA) [4].
2) Image transformation or feature extraction, such as principal component analysis (PCA) [5] and iteratively reweighed multivariate alteration detection (IR-MAD) [6].
3) Image classification.
4) Other advanced methods using wavelet and Markov random field.

Among them, image transformation is very promising in change detection tasks due to the fact that it transforms multispectral images into a specific feature space to highlight the changed pixels and suppress unchanged ones. PCA is a widely used image transform algorithm, but it greatly depends on the statistical property of the considering images and easy to be affected by unbalanced data [5]. As a classic bitemporal image analysis method, IR-MAD applies different weights to the observations, and larger weights are assigned to the observations that show little changes [6]. However, IR-MAD ignores the significant inner relationship between bitemporal bands.

In the field of change detection, the detection of multiple changes (i.e., different kinds of changes) in bitemporal multispectral images is a complex task. In 2012, Bovolo *et al.* proposed an unsupervised framework named the compressed CVA (C^2VA) to detect multiple changes from bitemporal images [7]. C^2VA maps spectral change information into a 2-D polar domain, where pixels with high magnitude values are more likely to be changed ones, and the change type is determined by the direction of spectral change vectors (SCVs). The magnitude of SCVs is calculated by measuring the Euclidean distance (EUD) between the spectral vectors at different times. However, criticism has been presented against EUD but in favor of the cosine angle distance (CAD), since EUD ignores several vector components that are equal or close to zero, which may cause EUD to work not very well at separating changed and unchanged pixels, whereas the CAD seems to be more effective [8]. Additionally, the CAD has the ability to normalize the influence of shading and accentuate the nature of feature changes, which is important for the change detection task, for it is common to have illumination variation and shadows [9].

Manuscript received April 5, 2016; revised July 22, 2016; accepted August 13, 2016. Date of publication September 2, 2016; date of current version October 12, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 61273317 and Grant 61422209, by the National Program for the Support of Top-Notch Young Professionals of China, and by the Specialized Research Fund for the Doctoral Program of Higher Education under Grant 20130203110011.

H. Zhang is with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, China, and also with the Department of Integrated Circuit Design and Integrated System, School of Microelectronics, Xidian University, Xi'an 710071, China.

M. Gong, P. Zhang, and L. Su are with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, China (e-mail: gong@ieee.org).

J. Shi is with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, China, and also with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2016.2601930

However, C^2VA is often applied to the available spectral channels of multispectral images acquired at different times; here, the spectral channels denote the calibrated radiance data. Nevertheless, the spectral channels are easy to be affected by noise interference, and the data are often redundant. Therefore, it is necessary to transform the available spectral channels into an abstract feature space due to their limited representation power and noise resistance ability. Recently, the deep learning-driven unsupervised feature learning has become popular; and deep models, such as the deep belief network (DBN) [10] and the stacked denoising autoencoders, have achieved the state-of-the-art performance in many image recognition tasks. As for the change detection task, deep learning offers a flexible tool to transform a bitemporal image into a desired feature space, facilitating to capturing the key information for discrimination and suppressing meaningless variations caused by environment or registration.

To address the problem previously described, in this letter, we present a framework which combines the DBN and the feature change analysis (FCA). In this framework, a DBN is established to learn robust and abstract representation from the available spectral channels, facilitating to relieving noise and reducing the redundancy. For the change detection task, the features learned by the DBN would capture key discriminative information and suppress irrelevant variations, such as illumination, parallax, and shadow. Unlike the standard C^2VA, the change information is extracted from the DBN feature space instead of the original spectral channels, and the CAD is selected to measure the magnitude of feature changes instead of EUD, which ignores the vector components close to zero.

This letter is organized into four sections. The next section briefly introduces the proposed change detection framework. Section III presents the experimental results on the three real multispectral data sets. Finally, Section IV draws the conclusions of this letter.

II. Feature-Level Change Detection Framework for Multispectral Imagery

Here, we briefly introduce the DBN-based unsupervised feature learning and the CAD-based FCA.

A. DBN-Based Unsupervised Feature Learning

Recently, the breakthrough in deep learning through unsupervised per-training schemes has led to the resurgence and massive interest in deep neural networks (DNNs), significantly outperforming the shallow ones [10]. The conventional machine-learning techniques are limited in their ability to process natural data in their raw forms. In the past decades, careful engineering and considerable domain expertise are often required to design a feature extractor that transforms the raw data into a suitable internal representation for pattern recognition tasks. However, DNN has the ability to learn directly from the raw data and automatically discover hierarchical representations needed for some specific applications, such as detection or classification. For the change detection task, higher layers of representation have the ability to amplify some aspects of the input which are key for discrimination and suppress irrelevant variations, highlighting main changes. The merits of the deep models lay in the fact that these layers of features are not designed by human engineers but are learned from the raw data using a general-purpose learning procedure.

In 2006, Hinton and Salakhutdinov proposed the DBN model [10], which adopts an unsupervised layerwise learning strategy to pretrain the initial weights of the deep networks, and then fine-tunes the entire network using backpropagation. When a local spectral patch is fed into the visible layer of a restricted Boltzmann machine (RBM), the hidden layer of the RBM detects the pattern hidden in data by minimizing the energy function. Then, when the hidden layer is used as the visible layer of the next RBM, the corresponding hidden layer learns the feature of features, i.e., the higher level feature. The pretrained RBMs are unrolled into a DBN, where the connectional weights are fine-tuned using a backpropagation algorithm [11]. As for the change detection task, DBN offers a flexible tool to transform a bitemporal image into a desired feature space, facilitating to capturing the key information for discrimination and suppressing irrelevant variations, such as noise interference and illumination. In addition, studies have demonstrated that the DBN shows significant advantages over the other handcrafting techniques in learning useful features with rotation, translation, and scale invariance [12]. The hierarchical nature of deep networks allows them to increasingly learn abstract and invariant features of the perceptual input. What's more, the feature is straightforwardly learned from the data rather than handcrafted, and deep networks could discover invariant features that are difficult to specify or build by hand. Fig. 1 shows the workflow of the proposed feature-level change detection approach, where the DBN is used to learn higher level representations, suppressing noise and reducing redundancy, and then, the CAD-based FCA is adopted to highlight changes and distinguish change types.

B. CAD-Based FCA

Through DBN, we can obtain the feature maps of multispectral images $\mathbf{X}_1$ and $\mathbf{X}_2$, denoted by $\mathbf{V}_1$ and $\mathbf{V}_2$, respectively, of size $w \times h \times d$, where d is the dimension of features. The spectral channels are mapped into a multidimension feature space, and each patch pair $(\mathbf{x}_{ij}^{t_1}, \mathbf{x}_{ij}^{t_2})$ is mapped from the original spectral space into the feature space, forming the corresponding feature pair $(\mathbf{v}_{ij}^{t_1}, \mathbf{v}_{ij}^{t_2})$. For the change detection task, we need to analyze the magnitude of feature changes to highlight changes and estimate the direction of feature changes to distinguish the types of feature changes. In the standard C^2VA, EUD is used to evaluate the change magnitude of the available spectral channels. However, criticism has been presented against EUD since it ignores the vector components close to zero, which implies that the EUD poorly estimates the change magnitude. Nevertheless, studies have demonstrated that the CAD works better than the EUD, because it can normalize the influence of shading and accentuate the nature of feature change. Therefore, the CAD is used to measure the similarity between learned feature pairs.

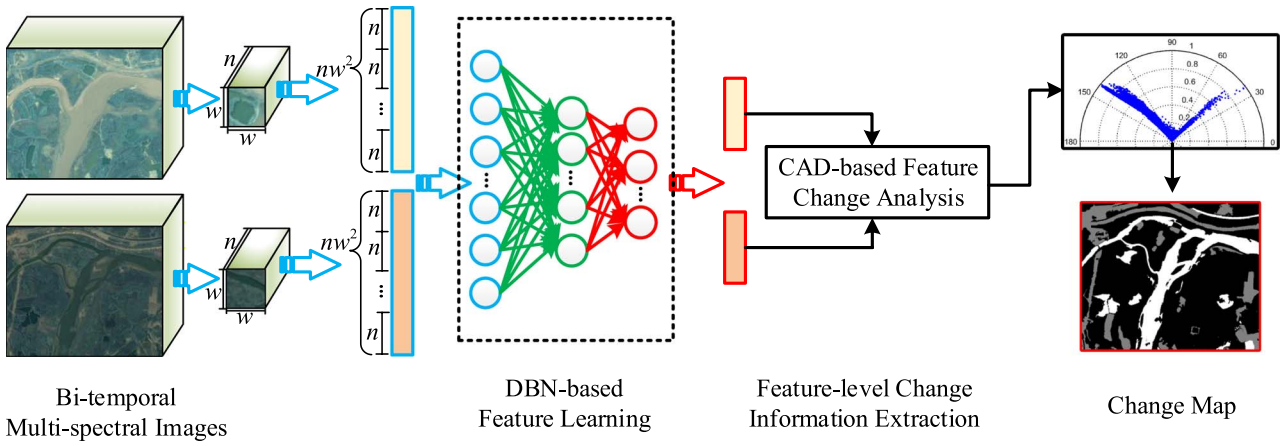


Fig. 1. Illustration on the proposed unsupervised change detection approach for bitemporal multispectral images with N-bands (w is the size of considered neighborhood windows). Each patch is sampled from these two multispectral images pixelwise, and then, these patches are fed into the DBN for its training. After training, the DBN automatically discovers a better representation needed for the change detection task. Then, the change information of these feature pairs are extracted into a 2-D polar representation by the CAD-based FCA. Finally, the change map is generated by the unsupervised fuzzy local information C-means (FLICM) according to the 2-D polar representation.

1) Magnitude of FCVs: The cosine measure is well known as the most frequently applied similarity measure in many applications. It measures the angle similarity between two vectors. The cosine has the range [0, 1] since all components of feature vectors are nonnegative due to the exploitation of sigmoid activation in DBN. According to the natural interpretation that paralleled vectors oriented in the opposite and same directions are considered as the least and most similar, respectively, we defined CAD by subtracting the cosine measure from its maximum as follows:

$$\rho = 1 - \frac{\mathbf{v}_{t_1}(\mathbf{v}_{t_2})'}{\sqrt{\mathbf{v}_{t_1}(\mathbf{v}_{t_1})'}\sqrt{\mathbf{v}_{t_2}(\mathbf{v}_{t_2})'}}, \quad \rho \in [0,1] \tag{1}$$

where $\mathbf{v}_{t_i}$ denotes the DBN feature of the given pixel at time t_i ($i = 1, 2$).

2) Direction of FCVs: As the magnitude of FCVs only measures the changing degree of each feature pair, without any information about different kinds of changes, a proper measure is needed to distinguish the change types. Thus, the improved version of the angular distance measure proposed in [7] is used to measure the direction of FCVs. As is known, higher level representation has the power to enhance some aspects of the input which are important for discrimination and suppress irrelevant variations. Therefore, we analyze the change types in the DBN feature space, instead of the original spectral space. The feature difference map $\mathbf{V}_D$ of multispectral images $\mathbf{X}_1$ and $\mathbf{X}_2$ can be computed by (2). In the N-D feature space, the desired angle measure for distinguishing change types can be rewritten as (3)

$$\mathbf{V}_D = \mathbf{V}_2 - \mathbf{V}_1 \tag{2}$$

$$\theta = \arccos\left[\frac{1}{\sqrt{N}}\left(\frac{\sum_{b=1}^{N} V_{b,D}}{\sqrt{\sum_{b=1}^{N} V_{b,D}^2}}\right)\right] \tag{3}$$

where $V_{b,D}$ represents the bth channel of the feature difference maps $\mathbf{V}_D$.

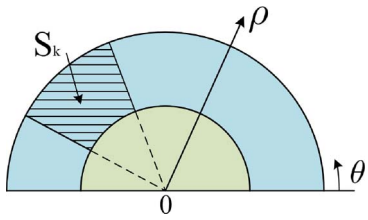


Fig. 2. Polar representation of change information. ρ denotes the magnitude of FCVs, θ shows the direction of FCVs, and S_k represents the kth change type according to θ.

Fig. 3. Bahe River (Bahe bridge section, Xi'an, China) data set, of size 662 × 686 × 4. (a) Multispectral image with 4-bands, acquired by GF-1 satellite on August 19, 2013. (b) Multispectral image with 4-bands, acquired by GF-1 satellite on August 29, 2015. (c) The ground truth map.

As shown in Fig. 2, based on the magnitude ρ and direction θ of feature change vectors (FCVs), change information can be compressed into a 2-D polar domain. First, changed and unchanged pixels are distinguished from each other by applying FLICM [13], a widely used unsupervised clustering technique, on the change magnitude ρ. Then, the changed pixels can be clustered into several different change types, according to the change direction θ.

III. Experimental Study

A. Data Set Description

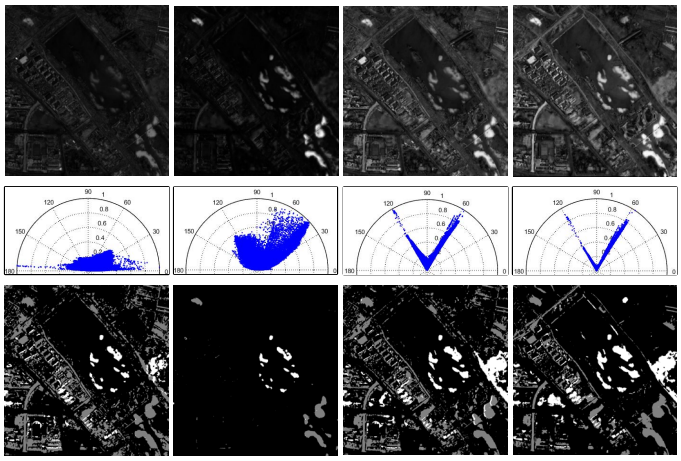
In our experiments, three groups of multispectral data sets, including Bahe River, Weihe River, and Lake Lotus data sets, are used to assess the effectiveness of the proposed change detection method. As shown in Figs. 3–5, these three data sets

(a) (b) (c)

Fig. 4. Weihe River (Madong village section, Xi'an, China) data set, of size $718 \times 592 \times 4$. (a) Multispectral image (4-bands) acquired by GF-1 satellite on August 19, 2013. (b) Multispectral image (4-bands) acquired by GF-1 satellite on August 29, 2015. (c) Ground truth map.

(a) (b) (c)

Fig. 5. Lake Lotus (Xi'an, China) data set, of size $440 \times 398 \times 4$. (a) Multispectral image (4-bands) acquired by GF-1 satellite on August 19, 2013. (b) Multispectral image (4-bands) acquired by GF-1 satellite on August 29, 2015. (c) Ground truth map.



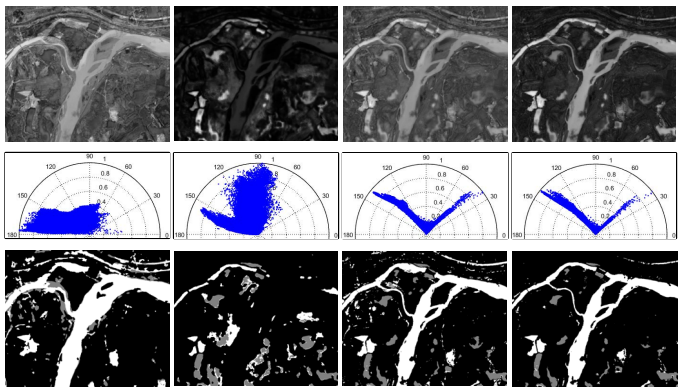
(a) (b) (c) (d)

Fig. 6. Polar representation of change information on the Bahe River data set, under different configurations with $w = 3$, and the spectral channels are denoted by SC. (a) SC+EUD. (b) SC+CAD. (c) DBN(36-100-20)+EUD. (d) DBN(36-100-20)+CAD.

are acquired by the GF-1 satellite on August 19, 2013 and August 29, 2015, respectively, over different scenes in Xi'an, China. Each image has 4-bands (R, G, B, and near-infrared), and they share a 2-m spatial resolution. Each pair of images has been radiometrically corrected and coregistered to make them as more comparable as possible. The overall accuracy (OA) and Kappa coefficient are used to quantitatively evaluate the performance of the proposed method.

B. Results and Analysis

Fig. 6 shows the change detection results by different methods on the Bahe River data set. The top row lists the change maps, the middle row shows the 2-D polar representations of change information, and the bottom row displays the final change detection results. From the visual comparison between Fig. 6(a) and (b), the CAD highlights changes better than the EUD at the bottom right corner, but it fails to label the changed regions at the bottom left corner. Fig. 6(c) and (d) shows the results by applying the EUD and CAD measures on the learned DBN features, respectively. Based on the DBN features, DBN+CAD almost detects main changes, whereas the result by DBN+EUD looks noisy and has many false positive pixels. In the polar representation of the SC+EUD, the magnitudes of the spectral changes are limited in a very small range, mostly in [0, 0.4], which makes it hard to identify changed pixels from unchanged ones. The SC+CAD seems to work a little better than the SC+EUD, but it is still limited. In the polar representation of DBN+EUD and DBN+CAD, the change magnitudes are extended into [0, 1]; and the change directions are concentrated into two directions, i.e., 60° and 120°, which makes it easier to identify the changes and distinguish different change types.



(a) (b) (c) (d)

Fig. 7. Polar representation of change information on the Weihe River data set, under different configurations with $w = 5$, and the spectral channels are denoted by SC. (a) SC+EUD. (b) SC+CAD. (c) DBN(36-100-20)+EUD. (d) DBN(36-100-20)+CAD.

Fig. 7 shows the change detection results on the Weihe River data set. As shown in Fig. 7(a) and (b), EUD and CAD work not so well at highlighting changes based on the available spectral channels. However, both EUD and CAD achieve much better change maps based on DBN features [see Fig. 7(c) and (d)]. In addition, DBN+CAD constructs a robust change map with high contrast, which facilitates identifying changes. Based on the DBN features, the change directions are concentrated into two directions, i.e., 45° and 135°; and the change magnitudes are distributed in the range of [0 1], making it easy to highlight changes and identify different change types. Compared with the ground truth map, most of the water-related changes are detected by DBN+EUD/CAD. However, both of them fail to label some artificial changes, such as road works.

Fig. 8 illustrates the influence of the neighbor window size w on the accuracy of change detection over different data sets. This table is obtained by taking $w = 1, 3, 5,$ and 7. When w varies from 1 to 3, the OA and Kappa values increase, but a larger w $(w > 3)$ leads a decrease in the accuracy. The reason lays in the fact that a larger w would cause higher spectral

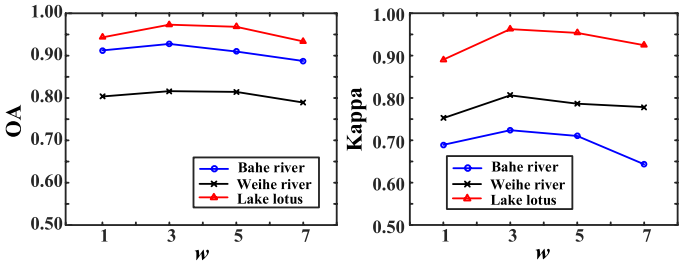


Fig. 8. Influence of w on the accuracy of change detection over three different data sets.

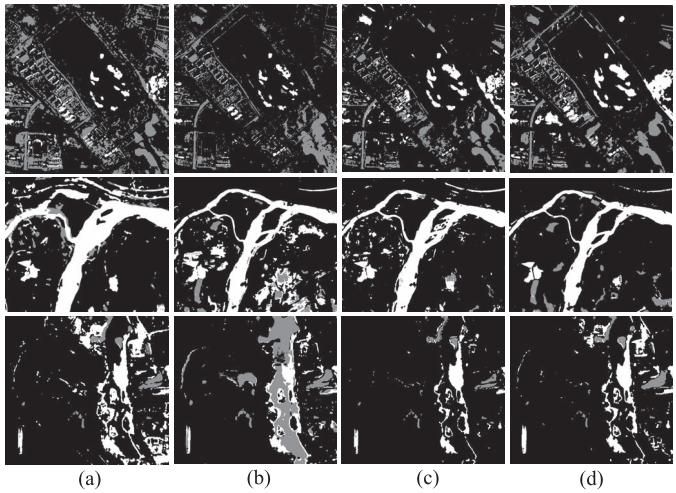


Fig. 9. Change detection results by different methods on three multispectral data sets. From top to bottom, the data sets are, in order, the Bahe River, the Weihe River, and the Lake Lotus data set, respectively. (a) C^2VA. (b) PCA. (c) IR-MAD. (d) Ours.

repetition rate between the two adjacent pixels, resulting in more similar change information between them. Therefore, a proper w value should be selected to enhance the robustness to noise by introducing local neighborhood information.

C. Performance Comparison

The proposed method is compared with three state-of-the-art methods, such as C^2VA [7], PCA [5], and IR-MAD [6]. Fig. 9 shows the final change detection results by different methods over the three real data sets. For the Bahe River, C^2VA wrongly classifies many unchanged pixels into changed ones, resulting in the noisy map. PCA works well at suppressing noise, but it fails to correctly label some unchanged regions in three testing data. IR-MAD produces good detection results both in the Bahe River and Lake Lotus data sets. The proposed method achieves the best visual results among the compared methods, and it detects main changes and suppresses unchanged pixels well.

Table I summarizes the OA and Kappa values with different methods on the three real multispectral data sets. The quantitative comparison shows that the proposed method achieves higher OA and Kappa values, which demonstrate that the feature learned by the DBN has the ability to amplify something key for identifying changes and suppress the irrelevant variations caused by the environment. The experimental results also indicate the effectiveness of CAD on the DBN feature.

TABLE I
COMPARISON WITH THE STATE-OF-THE-ART METHODS ON THREE EXPERIMENTAL DATA SETS

Data Sets	Metrics	C^2VA	PCA	IR-MAD	Ours
Bahe River	OA(%)	81.02	84.88	89.73	92.77
	Kappa	0.3840	0.3721	0.5816	0.7238
Weihe River	OA(%)	71.60	75.69	79.76	81.62
	Kappa	0.5538	0.6482	0.7555	0.8062
Lake Lotus	OA(%)	91.65	84.72	94.46	97.32
	Kappa	0.6807	0.4242	0.7225	0.9625

IV. CONCLUSION

In this letter, an unsupervised change detection method has been proposed for detecting multiple changes from bitemporal multispectral images. The proposed method successfully detects multiple changes from multispectral images by combining the DBN and the FCA. Through the DBN, the available spectral channels are transformed into an abstract feature space to capture the key information for discrimination and suppress the irrelevant variations. Then, we apply the FCA on the feature learned by DBN to identify different types of changes. The experimental studies on the three real multispectral data sets have demonstrated the effectiveness and robustness of the proposed method and indicated that CAD works better than the EUD on the DBN features in a multispectral change detection task. In the near future, we will try to learn the difference representation directly from the stacked bitemporal images through deep models.

REFERENCES

[1] K. Ding, C. Huo, Y. Xu, Z. Zhong, and C. Pan, "Sparse hierarchical clustering for VHR image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 577–581, Mar. 2015.
[2] A. P. Tewkesbury, A. J. Comber, N. J. Tate, A. Lamb, and P. F. Fisher, "A critical synthesis of remotely sensed optical image change detection techniques," *Remote Sens. Environ.*, vol. 160, pp. 1–14, 2015.
[3] M. T. Eismann, J. Meola, and R. C. Hardie, "Hyperspectral change detection in the presence of diurnal and seasonal variations," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 237–249, Jan. 2012.
[4] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007.
[5] J. S. Deng, K. Wang, Y. H. Deng, and G. J. Qi, "PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data," *Int. J. Remote Sens.*, vol. 29, no. 16, pp. 4823–4838, 2008.
[6] N. A. Aasbjerg, "The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 463–78, Feb. 2007.
[7] F. Bovolo, S. Marchesi, and L. Bruzzone, "A framework for automatic and unsupervised detection of multiple changes in multitemporal images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2196–2212, Jun. 2012.
[8] T. Korenius, J. Laurikkala, and M. Juhola, "On principal component analysis, cosine and Euclidean measures in information retrieval," *Inf. Sci.*, vol. 177, no. 22, pp. 4893–4905, 2007.
[9] O. A. C. Junior, R. F. Guimaraes, A. R. Gillespie, N. C. Silva, and R. A. T. Gomes, "A new approach to change vector analysis using distance and similarity measures," *Remote Sens.*, vol. 3, no. 11, pp. 2473–2493, 2011.
[10] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
[11] T. Kuremoto, S. Kimura, K. Kobayashi, and M. Obayashi, "Time series forecasting using a deep belief network with restricted Boltzmann machines," *Neurocomputing*, vol. 137, no. 15, pp. 47–56, 2014.
[12] O. K. Oyedotun, E. O. Olaniyi, and A. Khashman, "Deep learning in character recognition considering pattern invariance constraints," *Int. J. Intell. Syst. Technol. Appl.*, vol. 7, no. 7, pp. 1–10, 2015.
[13] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1328–1337, May 2010.