ISPRS Journal of Photogrammetry and Remote Sensing 116 (2016) 24–41

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

# Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images

Puzhao Zhang, Maoguo Gong *, Linzhi Su, Jia Liu, Zhizhou Li

*Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi Province 710071, China*

ARTICLE INFO

*Article history:*
Received 14 July 2015
Received in revised form 9 December 2015
Accepted 12 February 2016
Available online 12 March 2016

*Keywords:*
Change detection
Spatial-resolution
Denoising autoencoder
Stacked denoising autoencoder
Deep neural networks
Feature transformation

ABSTRACT

Multi-spatial-resolution change detection is a newly proposed issue and it is of great significance in remote sensing, environmental and land use monitoring, etc. Though multi-spatial-resolution image-pair are two kinds of representations of the same reality, they are often incommensurable superficially due to their different modalities and properties. In this paper, we present a novel multi-spatial-resolution change detection framework, which incorporates deep-architecture-based unsupervised feature learning and mapping-based feature change analysis. Firstly, we transform multi-resolution image-pair into the same pixel-resolution through co-registration, followed by details recovery, which is designed to remedy the spatial details lost in the registration. Secondly, the denoising autoencoder is stacked to learn local and high-level representation/feature from the local neighborhood of the given pixel, in an unsupervised fashion. Thirdly, motivated by the fact that multi-resolution image-pair share the same reality in the unchanged regions, we try to explore the inner relationships between them by building a mapping neural network. And it can be used to learn a mapping function based on the most-unlikely-changed feature-pairs, which are selected from all the feature-pairs via a coarse initial change map generated in advance. The learned mapping function can bridge the different representations and highlight changes. Finally, we can build a robust and contractive change map through feature similarity analysis, and the change detection result is obtained through the segmentation of the final change map. Experiments are carried out on four real datasets, and the results confirmed the effectiveness and superiority of the proposed method.

© 2016 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

## 1. Introduction

Change detection is the process for identifying the changed regions between the given image-pair observing the same scene at different times (Kit and Lüdeke, 2013; Wang et al., 2015). Widely used in remote sensing (Tian et al., 2013; Zhu and Woodcock, 2014), land cover monitoring (Chen et al., 2013; Hulley et al., 2014), natural disaster damage assessment (Stramondo et al., 2006), medical diagnosis and treatment (Bosc et al., 2003), etc., change detection has attracted many research interests. With increasing multi-temporal and multi-spatial-resolution data available from remote sensing platforms, the efficient exploitation of this unprecedented wealth of data is a critical issue at present (Camps-Valls et al., 2008).

Change detection aims at identifying the set of pixels that are "significantly different" between the given image-pair, and these pixels compose the change map (Radke et al., 2005). Generally, change detection often consists of three main procedures: change extraction (CE), segmentation of change map and accuracy assessment (Hecheltjen et al., 2014). According the unit of analysis, CE can be divided into pixel-based and object-based (Hebel et al., 2013; Hussain et al., 2013), but the key lays in the design of the comparison method. To identify change, the input image-pair are compared and a decision is made based on the changing degree. When a comparison framework is established, then the analysis units can be compared to highlight change (Tewkesbury et al., 2015). Chen et al. (2012) suggests that pixels have limited features, tone or radiance and they cannot offer an adequate framework to model spatial structure and contextual information. On the other hand, it is easy to produce spurious, noisy change pixels when the pixel is taken as an analysis unit (Hussain et al., 2013). Later, a pixel kernel filter or moving window is used to introduce

* Corresponding author.
*E-mail address:* gong@ieee.org (M. Gong).

http://dx.doi.org/10.1016/j.isprsjprs.2016.02.013
0924-2716/© 2016 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

contextual information, and it does help to filter noise and identify 'true' change by considering the local neighborhood of the given pixel (Klaric et al., 2013). However, the use of kernel filter may lead to blurred boundaries and the removal of smaller features. Tewkesbury et al. (2015) classify the comparison methods into six broad categories, i.e., layer arithmetic (Desclée et al., 2006; Jin et al., 2013), post-classification change (Demir et al., 2013; Teo and Shih, 2013), direct classification (Gao et al., 2012; Schneider, 2012), transformation (Bovolo et al., 2012; Carvalho Júnior et al., 2011), change vector analysis (CVA) (Xian and Homer, 2010) and hybrid change detection (Doxani et al., 2012).

Most of the literatures on change detection focuses on the comparison between the given image-pair with the same spatial resolution, acquired by exactly the same sensor or produced by sensors of the same type (Bruzzone and Prieto, 2000; Inglada and Mercier, 2007) on different times. For example, Yousif and Ban (2013) adapted a nonlocal means denoising algorithm to improve change detection in urban area using synthetic aperture radar (SAR) images with the same spatial-resolution, where the denoising weights depends on the similarity of the local neighborhoods, and then principal component analysis (PCA) was used to reduce the dimensionality of the neighborhood feature. Multipolarization SAR change detection problem was formulated as a binary hypothesis test, and the principle of invariance has been applied to design decision rules exhibiting a special symmetry (Carotenuto et al., 2015). However, the further extension of this framework relaxing Gaussian requirement for the data is needed. On the other hand, the use of coarse spatial resolution remote sensing images also has attracted some researchers. Focusing on the use of coarse spatial resolution remote sensing data, Hgarat-Mascle et al. (2005) proposed a land cover change detection approach at coarse spatial scales based on iterative estimation and previous state information. To get rid of low spatial-resolution limits of coarse-resolution remote sensing images on land cover maps updating, Li et al. (2015) proposed a land cover updating method, which involves the use of coarse-resolution images to update fine-resolution land cover maps by integrating change detection and sub-pixel mapping methods. In the near future, as the updating of operational airborne and spaceborne sensors, the demanding for developing technical tools to exploit multi-spatial-resolution data in a combined way by taking advantages of their characteristics would increase.

As for multi-source change detection, Camps-Valls et al. (2008) proposed a kernel-based framework for multi-source remote sensing data, which develops nonlinear kernel classifiers for the well-known difference and ratio change detection methods by formulating the input space into a high-dimensional feature space. But it is not easy to select proper kernels and the regularization parameter. Another multi-source images change detection method based on similarity measures was proposed in Alberga (2009), where a series of measures is employed for multi-source change detection. Though several change detection approaches have been proposed for multi-source remote sensing images, as far as we know, so far there is no a general methodological framework for multi-spatial-resolution change detection, which combines different sources of information that involve different sensors, time instants or contextual extracted features efficiently. Different change detection algorithms have their own merits and no single approach is optimal and applicable to all cases. Previous literature has shown that image differencing, PCA and post-classification comparison (PCC) are the most common methods used for change detection (Lu et al., 2004).

In this paper, we present a novel framework for multi-spatial-resolution change detection, which incorporates deep-architecture-based feature learning and mapping-based feature change analysis (FCA). Multi-spatial-resolution image-pair share a common basis since they are the different representations of the same reality (Alberga, 2009). The different representations are often incommensurable due to the different modalities and properties of multi-resolution data. Therefore, the same reality motivates us to build a mapping neural network (MNN) to find the inner relationship between the different representations, making them be directly comparable. And the salient development of deep architecture gives us a chance to learn more robust and abstract representations for the neighborhood gray information of the given pixel. In Tang et al. (2015), deep architecture has been successfully exploited for high-level feature representation and classification in ship detection. And Xu et al. (2013) applied the autoencoder to the change detection for the very-high-resolution images.

In the field of neural network, there has been a long held belief that the composition of several levels of nonlinearity would be key to efficiently model complex relationship between variables and achieve better generalization performance on classical recognition tasks (Vincent et al., 2010). This standpoint is strongly supported by the layered architecture of regions of the human brain. However, due to the problematic non-convex optimization of multi-layer neural networks beyond one or two hidden layers, it is hard to achieve the expected solutions. The publications of several novel works (Bengio, 2009; Bengio et al., 2007; Hinton et al., 2006; Hinton and Salakhutdinov, 2006; Lee et al., 2008) make it possible to successfully train deep architecture, and the outstanding and seminal work on deep belief networks (DBN) was finished in Hinton and Salakhutdinov (2006).

Deep architecture uses a local unsupervised criterion to pre-train each layer in turn, and then stacks them, aiming to learn a useful and robust representation. The output of the previous layer is fed as the input of the next one, thus building a higher-level representation based on the lower-level representation. A fine-tuning process across the entire deep network, which treats all layers as a single model, is often helpful to greatly improve the performance of the deep network. Autoencoder, as the bricks of deep architecture, consists of an encoder and a decoder, which learns an abstract representation of the input by minimizing the reconstruction error (Vincent et al., 2010). Without any restrictions, it is difficult to learn an interesting and meaningful representation. By adding some restrictions, autoencoder has some variants, such as sparse autoencoder (SAE) (Coates et al., 2011), denoising autoencoder (DAE) (Vincent et al., 2008) and contractive autoencoder (CAE) (Rifai et al., 2011), which have the ability to learn different representations respectively. SAE is produced by restricting the number of hidden units or directly suppressing activation number of hidden units (Zou et al., 2011). DAE, which is generated by imposing the function of denoising, has the ability to reconstruct the original data from the corrupted version of the input (Vincent et al., 2010). Due to the good performance of DAE on representation learning and denoising, and that remote sensing images are usually corrupted by noise, DAE is stacked to learn robust and useful representation for each pixel.

The contributions of this work can be concluded in three aspects as follows:

(1) This paper present a novel framework for solving multi-spatial-resolution change detection problem, which has the ability to utilize different spatial-resolution images for change detection.
(2) Deep architecture such as SDAE is integrated in this novel framework for learning high-level features from raw data, and SDAE is demonstrated effective on change detection for remote sensing images.
(3) The relationship between different spatial representations of the same reality is established by learning a mapping, which bridges different observation spaces and highlights changes, constructing a robust and contractive final change map.

Download English Version:

https://daneshyari.com/en/article/6949322

Download Persian Version:

https://daneshyari.com/article/6949322

Daneshyari.com