ISPRS Journal of Photogrammetry and Remote Sensing 141 (2018) 252–264

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

A scale-invariant change detection method for land use/cover change research

Jin Xing [a,*], Renee Sieber [a], Terrence Caelli [b]

[a] Department of Geography, McGill University, Montreal, Canada
[b] Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, Australia

ARTICLE INFO

Article history:
Received 11 August 2017
Received in revised form 31 March 2018
Accepted 19 April 2018

Keywords:
Land use/cover change detection
Scale variance
Scale-invariant feature transformation
Maximally Stable Extremal Region
Hadoop
Cloud computing

ABSTRACT

Land Use/Cover Change (LUCC) detection relies increasingly on comparing remote sensing images with different spatial and spectral scales. Based on scale-invariant image analysis algorithms in computer vision, we propose a scale-invariant LUCC detection method to identify changes from scale heterogeneous images. This method is composed of an entropy-based spatial decomposition, two scale-invariant feature extraction methods, Maximally Stable Extremal Region (MSER) and Scale-Invariant Feature Transformation (SIFT) algorithms, a spatial regression voting method to integrate MSER and SIFT results, a Markov Random Field-based smoothing method, and a support vector machine classification method to assign LUCC labels. We test the scale invariance of our new method with a LUCC case study in Montreal, Canada, 2005–2012. We found that the scale-invariant LUCC detection method provides similar accuracy compared with the resampling-based approach but this method avoids the LUCC distortion incurred by resampling.

1. Introduction

Big data provides us with numerous new sources of data for Land Use/Cover Change (LUCC) but it causes problems related big data's large volume, complex variety, increasing velocity, and challenging verification (Miller and Goodchild, 2015). We now have finer spatio-temporal resolutions of LUCC data, with greater variety in terms of spectra and sensing platforms (Hansen et al., 2013). Larger, faster, and diverse data offers significant potential for LUCC but it quickly exceeds the data handling capacity and capability of existing LUCC algorithms. Among the four "Vs" of big data, volume is predominant focus in LUCC research (e.g., Hampton et al., 2013), although velocity also has attracted interest (Gil-Yepes et al., 2016; Wu et al., 2017a, 2017b). Our paper emphasizes the variety and specifically the various scales that are now available (i.e., different spatial, spectral, and temporal granularities and extents) (Goodchild, 2011). Because LUCC uses two or more datasets to identify changes, big data introduces potential problems in scale variance (Woodcock and Strahler, 1987).

If scales vary, one usually interpolates or re-samples one or more datasets to homogenize spatial granularities (i.e., resolutions) and extents to co-register the images for LUCC detection (e.g., Zhang et al., 2016). These spatial scaling operations can cause various problems like the generation of erroneous artefacts (Kwok and Sun, 1993), loss of information (Sheikh and Bovik, 2006), and distortion of geographic entities (Prashanth and Shashidhara, 2009). As a result of these spatial scaling operations, LUCC accuracy can be significantly degraded by scale variance (Olofsson et al., 2014) particularly if we wish to take advantage of the high resolution characteristic of big data.

To avoid the drawbacks of using spatial interpolation or resampling techniques, research scientists have investigated novel solutions to handle the challenge of scale variance. For example, Chen et al. (2012a) clustered pixels into image objects prior to comparison and then compared the geo-registered objects from datasets at two different scales. Singh (1989) bypassed the comparison of image pixels and explored a post-classification method to extract LUCC by comparing the class label maps. Xiao et al. (2016) combined pixel-based method and object-based approach together to investigate urban LUCC with high-resolution imagery datasets. All these approaches assume that the scale variance in any LUCC would be minor and that geo-registration would be sufficient to compare image objects. Big data does not make these

* Corresponding author at: Department of Geography, Room 428, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec H3A 0B9, Canada.
E-mail addresses: jin.xing@mail.mcgill.ca (J. Xing), renee.sieber@mcgill.ca (R. Sieber), tcaelli@unimelb.edu.au (T. Caelli).

https://doi.org/10.1016/j.isprsjprs.2018.04.013
0924-2716/

assumptions by creating new multi-scale challenges for the study of LUCC.

Computer vision algorithms have been explored to tackle challenges introduced by different kinds of variance (Radke et al., 2005). These algorithms are interesting because they focus on differentiating objects within datasets and do not rely on geo-registration because the objects may be moving image to image. Scale-invariant computer vision algorithms exploit scale by artificially deriving multiple images, each at a different resolution, from a single image. They then extract the stable "scale-invariant" features from these derived images. An example of the utility of scale-invariant computer vision for LUCC can be found in Dellinger et al. (2014). They proposed using the Scale-invariant Feature Transformation (SIFT) (Lowe, 2004) to handle images from diverse sensing platforms. Pham et al. (2016) employed SIFT to study LUCC before and after a volcanic eruption. They used SIFT to reduce variations in each image from illumination, color, or view angle differences. Authors of these two papers compared images at the same resolution. Ye et al. (2014) utilized another computer vision algorithm, Speed Up Robust Features (SURF), with images of different resolutions for LUCC. However, they resampled the images to create scale homogeneity and then extracted the changed regions. We want to exploit computer visions algorithms proven to identify scale-invariant features for LUCC: to detect scale-invariant changes across multiple remote sensing (RS) images of different resolutions.

Our scale-invariant LUCC detection method integrates spatial decomposition, image feature comparisons that are derived from computer vision, change map smoothing, and LUCC labelling. We will show that: (1) LUCC can be extracted by comparing scale-invariant image features directly without spatial interpolation or re-sampling methods; (2) discrimination of scale-invariant image features can be enhanced by the integration of extent, shape, and spectral information for LUCC; and (3) high performance computing can provide significant support in the scale-invariant LUCC detection workflow.

The rest of this paper is organized as follows. Section 2 enumerates the benefits and challenges of scale-invariant algorithms derived from computer vision. Our scale-invariant LUCC detection method is introduced in Section 3, which is based on the integration of SIFT and the Maximally Stable Extremal Region (MSER). Section 4 is a case study in the Greater Montreal Area from 2006 to 2012, which evaluates our scale-invariant LUCC detection algorithm. This paper concludes in Section 5.

2. Handling scale variance with computer vision algorithms

A large body of computer vision algorithms employ scale variance. One widely applied approach is feature detection, expressed in algorithms like SIFT, MSER, and the Gradient Location and Orientation Histogram. Image features are extracted that are stable across various granularities, which are derived as needed from a single original image (Witkin, 1984; Huo et al., 2008). Image fusion is another scale variance handling method in computer vision, which merges relevant information from at least two images at different spectral and spatial granularities to achieve higher granularities (Li et al., 1995). For example, image fusion with multispectral IKONOS (4 m, red, green, blue, and infra-red) and panchromatic (1 m, greyscale) IKONOS images will generate a new image with 1 m resolution and 4 bands of information.

Perona and Malik (1990) and others offered good examples of how computer vision studies differ from LUCC. Although they (Perona and Malik, 1990) explored changes in image object boundaries at different spatial granularities, their study was conducted with everyday object extents (e.g., 1 mm at 1 m^2). LUCC works with larger extents and a broader range of granularities. Their study also was conducted with a single image but LUCC involves comparing images taken at different times. Their study considered changes in image object characteristics; however, LUCC functions at the image level and detects changes throughout the image extents. Ohn-Bar and Trivedi (2014) proposed a temporal interpolation algorithm to model the movement of human hand gestures. They studied a time span of deciseconds (100 ms units or 0.1 of a second). The time span in LUCC datasets may be several years or decades. Non-linear temporal models (e.g., branch, cyclical, and isochronical models) may further complicate temporal scale variance (Jönsson and Eklundh, 2004). Therefore, the scale variance in LUCC requires additional investigation before we can apply the computer vision algorithms.

2.1. Similarity of land use/cover entities

LUCC researchers have expressed considerable interest in SIFT. SIFT is an algorithm designed to detect, describe, and match key points across images. SIFT points are those points (pixels) that persist in the image regardless of various transformations. SIFT points are considered to be invariant to spatial granularity, rotation, affine distortion, translation, and illumination differences. SIFT points are extracted from regions as the minima/maxima of Difference-of-Gaussians (Bundy and Wallen, 1984). Image matching, clustering, and pattern recognition are then performed by matching SIFT points.

Previous work has highlighted the deficiency of SIFT in distinguishing similar land use/cover entities, which mainly occur in the dense urban areas (Tuermer et al., 2013; Sirmacek and Unsalan, 2009). Entities such as those composed of cement (e.g., buildings and roads) can be so similar (Yang et al., 2003) that SIFT cannot adequately discriminate among them. In Fig. 1 two images are carefully geo-registered but SIFT matching largely fails because of a lack of uniqueness in SIFT characteristics (e.g., for corners of roads and buildings).

2.2. Use of shape information

Regions are defined by geometric and topological connections among positions and features. Unsurprisingly, regions are sensitive to spatial granularity changes (Luo and Min, 2010). SIFT points (pixels) can be used to compare images directly and mark clusters of unmatched points as changed regions (Dellinger et al., 2014). Although change information can be represented by individual pixels, our approach considers change as a multi-scale collection of regions, which is more robust and more useful for LUCC. These multi-scale regions represent LUCC areas as clusters composed of different numbers of pixels across scale-heterogeneous remotely sensed images. Regions not only provide more information about LUCC (e.g., change boundaries and areas) but also should prove more resistant to noisy information.

As shown in Fig. 2, numerous changed SIFT key points (red points) are caused by the noise or artefacts, such as shadows, vehicles, trees, and building decorations. Few of these changed points represent actual LUCC. To overcome noise or artefacts, we can combine SIFT with a computer vision algorithm like MSER (Matas et al., 2004). MSER extracts scale-invariant regions (clusters) and matches regions across various images. LUCC can take advantage of MSER's use of shapes because unmatched MSERs represent changed regions.

2.3. Integration of spectral information

SIFT is designed for grey scale images and does not consider spectral information. Since LUCC imagery datasets are being