Remote Sensing of Environment 141 (2014) 155–167

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

# Improving the impervious surface estimation with combined use of optical and SAR remote sensing images

Yuanzhi Zhang [a,b], Hongsheng Zhang [a,*], Hui Lin [a]

[a] Institute of Space and Earth Information Science & Shenzhen Research Institute, The Chinese University of Hong Kong (CUHK), Shatin, New Territories, Hong Kong
[b] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China

ARTICLE INFO

*Article history:*
Received 18 June 2013
Received in revised form 22 October 2013
Accepted 26 October 2013
Available online xxxx

*Keywords:*
Impervious surface
SAR
Random Forest
Fusion

ABSTRACT

Accurate mapping of urban impervious surfaces is important but challenging due to the diversity of urban land covers. This study presents an effort to synergistically combine optical and SAR data to improve the mapping of impervious surfaces. Three pairs of optical and SAR images, Landsat ETM+ and ENVISAT ASAR, SPOT-5 and ENVISAR ASAR, and SPOT-5 and TerraSAR-X, were selected in three study areas to validate the effectiveness of the methods in this study. The potential of Random Forest (RF) was evaluated with parameter optimization for combining the optical and SAR images. Experiment results demonstrate some interesting findings. Firstly, the built-in out-of-bag (OOB) error is insufficient for accuracy assessment, and an assessment with additional reference data is required for combining optical and SAR images using RF. Secondly, the optimal number of variables ($m$) for splitting the decision tree nodes in RF should be some different from the principles reported previously, and an empirical relationship was given for determining the parameter $m$. Thirdly, the optimal number of decision trees ($T$) in RF is not sensitive to the resolutions and sensor types of optical and SAR images, and the optimal $T$ in this study is 20. Fourthly, the combined use of optical and SAR images by using RF is effective to improve the land cover classification and impervious surface estimation, by reducing the confusions between bright impervious surface and bare soil and dark impervious surface and bare soil, as well as shaded area and water surface. Even though the easily-confused land classes tend to be different in different resolutions of images, the effectiveness of combining optical and SAR images is consistent. This improvement is more significant when combing lower resolution optical and SAR images. The conclusions of this study could serve as an important reference for further applications of optical and SAR images, and as a potential reference for the applications of RF to the fusion of other multi-source remote sensing data.



## 1. Introduction

Urban impervious surfaces, such as transport related land (e.g., roads, streets, and parking lots) and building roof tops (commercial, residential, and industrial areas), have been widely recognized as important indicator for urban environments (Arnold & Gibbons, 1996; Hurd & Civco, 2004; Lu & Weng, 2006; Weng, 2001; Weng, Lu, & Liang, 2006; Yuan & Bauer, 2007). Remote sensing has become the major technique to estimate impervious surfaces due to its low cost and convenience for the impervious surface mapping from local to global scales. Numerous methods have been proposed to estimate impervious surfaces from remotely sensed images, including sub-pixel approaches (e.g., the spectral mixture analysis method (Wu & Murray, 2003; Wu, 2004), the classification and regression tree model (Yang, Xian, Klaver, & Deal, 2003), the artificial neural network (Weng & Hu, 2008), and the support vector machine (Sun, Guo, Li, Lu, & Du, 2011)) and per-pixel approaches such as conventional classification methods (Weng, 2012). Recently, a biophysical composition index (BCI) was proposed to extract urban impervious surfaces following the VIS conceptual model (Deng & Wu, 2012). However, most of these approaches were proposed with optical remote sensing images, and accurate estimation of impervious surfaces remains challenging due to the diversity of urban land covers, leading to difficulties of separating different land covers with similar spectral signatures (Weng, 2012). For instance, dry soils or sands are reported to be confused with bright impervious surfaces due to their high reflectance, while water and shades tend to be confused with dark impervious surfaces.

The use of multi-satellite images is considered as one promising approach to improve the accuracy of impervious surfaces (Weng, 2012). SAR is able to provide useful information about urban areas as it is sensitive to the geometric characteristics of urban land surfaces (Calabresi, 1996; Henderson & Xia, 1997; Leinenkugel, Esch, & Kuenzer, 2011; Soergel, 2010; Tison, Nicolas, Tupin, & Maitre, 2004; Zhang, Zhang, & Lin, 2012), and thus SAR has been identified as an important source to help extract impervious surfaces with optical data (Jiang, Liao, Lin, & Yang, 2009; Weng, 2012; Yang, Jiang, Lin, & Liao, 2009). Fusion between optical and SAR data can be performed on three different levels: pixel-level, feature-level, and decision-level. Pixel-level fusion is reported inappropriate for SAR images because of speckle noises (Zhang, Yang, Zhao, Li, & Zhang, 2010). For feature-level

* Corresponding author. Tel.: +852 39434280; fax: +852 26037470.
*E-mail address:* zhangstream@gmail.com (H. Zhang).

0034-4257/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.10.028

fusion, several approaches have been proposed including layer-stacking and ensemble-learning methods (e.g., bagging, boosting, AdaBoost & Random Forest (Hall & Llinas, 1997; Rokach, 2010)). The ensemble-learning methods can be combined with different classifiers (e.g., ANN and SVM (Rokach, 2010)). For decision-level fusion, various weighting methods (e.g., majority voting, entropy weighting, and performance weighting) and the Dempster–Shafer theory have been applied. However, conventional classifiers with a layer-stacking technique are not appropriate in this case as optical reflectance and SAR backscattering data do not correlate (Zhang et al., 2010). Among these methods, the decision-tree (DT) method will be given more attention, while the Random Forest (RF) algorithm has been reported to perform excellently in the fusion of optical and SAR data (Waske & van der Linden, 2008). However, the potential and effectiveness of RF on the fusion between optical and SAR images needs to be explored, especially in terms of the estimation of urban impervious surfaces.

This study aims to evaluate the effectiveness of RF to synergistically combine the optical and SAR data in terms of impervious surface estimation. A combination of pixel level and feature level fusion method is adopted. Additionally, the Kappa coefficient based on confusion matrix and the OOB error built-in the RF are compared to assess the effectiveness of fusing optical and SAR images.

## 2. Study area and data sets

### 2.1. Study area

Three cities, Guangzhou, Shenzhen and Hong Kong, located in the Pearl River Delta (PRD) are selected as the study areas to evaluate the effectiveness of the proposed approach. The region is located on the Pearl River Estuary (PRE), known as the third largest metropolitan area in China, experiencing tremendously fast development during the past 30 years. The region has rapidly become urbanized, with a population of over 19 million in an area of over 21 thousand km$^2$ (Fan, Wang, & Wang, 2008). However, due to significant interactions between human population and environment, environmental issues quickly emerged causing a series of problems, including air and water pollution (Zhang et al., 2008). As key indicators of urban growth and related environmental problems, impervious surfaces and their distribution are being paid greater attention by the local government. Thus, the accurate estimation of urban impervious surfaces is of high significance for the environment studies of PRD.

### 2.2. Satellite data and coregistration

Three different combinations of optical and SAR satellite data sets are selected for the three cities (Fig. 1). For Guangzhou, a scene of Landsat ETM+ image and a scene of ENVISAT Advanced Synthetic Aperture Radar (ASAR) Wide Swath Mode (WSM) image are employed. The ETM+ images have one panchromatic band at 15 m resolution and 6 bands at 30 m resolution. In this study, only the 30-m data were used. The ETM+ image was acquired on 31 December 2010. As there are stripes on the eastern and western edges of each scene because of the footprints (location and spatial extent) of each band due to the Scan Line Corrector (SLC) failure, a process should be applied to get rid of these stripes. For this reason, the study area is located in the middle of each scene where there are no stripes, and thus no stripe removal operation should be applied. We assume that the atmospheric condition is clear and homogeneous and the small part of clouds would not significantly impact the whole scene of image, and thus no atmospheric correction was performed (Wu & Murray, 2003). The ENVISAT ASAR WSM data were obtained on 23 September, 2010, on the descending direction with V/V polarization and a pixel size of 75 × 75 m.

For the Shenzhen City, a scene of SPOT-5 image and a scene of ENVISAT ASAR ASA_IMP_1P data are used. The SPOT-5 image is a precision 2A level data, and was obtained on 21 November 2008, with a spatial resolution of 10 m. The ASAR data were obtained on 19 November 2008, on the ascending direction, Track-25 of ENVISAT, with V/V polarization. The spatial resolution of the ASAR IMP data is 12.5 m. Due to the speckle noises, Enhanced Lee filter is selected to filter the speckle noises in the ASAR data. Enhanced Lee filter is an improved version of Lee filter which was designed to better and preserve texture information, edges, linear features and point targets in SAR images (Lee, 1983). Enhanced Lee filter is an adaptive filter which was proved to be more suitable for preserving radiometric and textural information than other speckle filters (Lopes, Touzi, & Nezry, 1990; Xie, Pierce, & Ulaby, 2002).

For Hong Kong, a SPOT-5 and a SAR image from TerraSAR-X are employed. The SPOT-5 image is also a precision 2A level data, and was obtained on 21 November 2008, with a spatial resolution of 10 m. TerraSAR-X (TSX) is a German earth observation satellite launched on 15 June 2007 and is still on operation. TSX operates in the X band (9.6 GHz) and has three main imaging modes, SpotLight, StripMap and ScanSAR. The TSX image used in this study was obtained in the StripMap mode, on 16 November 2008, with a spatial resolution of 3 × 3 m, and the scene size is 30 km (width) × 50 km (length). The TSX image was geocoded with the NEST (Next ESA SAR Toolbox) 4C-1.1 software developed by ESA, under the coordinate system of WGS84 and UTM (Zone 50N). Moreover, geometric correction was also conducted by the Range-Doppler Terrain Correction in NEST with Digital Elevation Model (DEM) data. Additionally, due to the uncertainty of speckle noises in SAR images, Enhanced Lee filter is selected to filter the speckle noises.

After preprocessing all the satellite images, both optical images and SAR images were coregistered to the same geo-reference system of the Universal Transverse Mercator (UTM) projection (Zone 50N) and Datum World Geodetic System 84 (WGS84). Over 20 control points were manually selected for each pair of optical and SAR images, and the linear transformation approach is used to conduct the coregistration. The spatial resolutions of the final registered images are determined by the corresponding optical image which is clearer for human perception and easier for the manual selection of control points (Table 1). The Root Mean Square Error (RMSE) of the coregistration for each pair of optical and SAR images is less than half pixel.

## 3. Methods

### 3.1. Feature extraction

According to previous literatures, segmentation methods are superior over pixel by pixel methods as segmentation methods take the texture characteristics into account (Dell'Acqua & Gamba, 2003; Stasolla & Gamba, 2008). Texture is important for the interpretation of SAR data because the speckles in SAR data result in difficulties for the pixel by pixel approach. Therefore, in order to extract complementary information for urban impervious surfaces from optical and SAR images, texture feature extraction is necessary and important. In this paper, the popular gray level co-occurrence matrix (GLCM) approach (Haralick, Shanmuga, & Dinstein, 1973) is employed to analyze the texture features. For the application of GLCM, the size of image block and the texture measures with GLCM have been a major issue (Marceau, Howarth, Dubois, & Gratton, 1990). In terms of the classification of remote sensing images in urban areas, it is reported that a window size of 7 × 7 pixels is suitable with a test on the resolutions from 2.5 m × 2.5 m to 10 m × 10 m (Puissant, Hirsch, & Weber, 2005). Moreover, four texture measures, the homogeneity (HOM), dissimilarity (DISS), entropy (ENT), and the angular second moment (ASM), were identified as effective indicators for the texture description of different urban land cover types (Puissant et al., 2005). However, since the spatial resolution of the registered Landsat ETM+ and ASAR image is 30 m, the window size for calculating GLCM should be smaller as terrains are smaller under coarser resolution. Thus, in this study, the window size is set as 7 × 7 pixels for Shenzhen and Hong