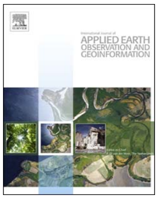
Int J Appl Earth Obs Geoinformation 64 (2018) 87–95

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

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



Research Paper

Exploring the optimal integration levels between SAR and optical data for better urban land cover mapping in the Pearl River Delta

Hongsheng Zhang[a,b,*], Ru Xu[a,b]

[a] Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong
[b] Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China

ARTICLE INFO

Keywords:
Fusion level
Optical and SAR fusion
Urban land cover
Fusion strategies

ABSTRACT

Integrating synthetic aperture radar (SAR) and optical data to improve urban land cover classification has been identified as a promising approach. However, which integration level is the most suitable remains unclear but important to many researchers and engineers. This study aimed to compare different integration levels for providing a scientific reference for a wide range of studies using optical and SAR data. SAR data from TerraSAR-X and ENVISAT ASAR in both WSM and IMP modes were used to be combined with optical data at pixel level, feature level and decision levels using four typical machine learning methods. The experimental results indicated that: 1) feature level that used both the original images and extracted features achieved a significant improvement of up to 10% compared to that using optical data alone; 2) different levels of fusion required different suitable methods depending on the data distribution and data resolution. For instance, support vector machine was the most stable at both the feature and decision levels, while random forest was suitable at the pixel level but not suitable at the decision level. 3) By examining the distribution of SAR features, some features (e.g., homogeneity) exhibited a close-to-normal distribution, explaining the improvement from the maximum likelihood method at the feature and decision levels. This indicated the benefits of using texture features from SAR data when being combined with optical data for land cover classification. Additionally, the research also shown that combining optical and SAR data does not guarantee improvement compared with using single data source for urban land cover classification, depending on the selection of appropriate fusion levels and fusion methods.

1. Introduction

The use of multi-sensor satellite images is considered a promising approach for improving urban land cover classification (Joshi et al., 2016; Pohl and van Genderen, 1998; Soergel, 2010). Synthetic Aperture Radar (SAR) is able to provide useful information about urban areas as it is sensitive to the geometric characteristics of urban land surfaces (Calabresi, 1996; Erasmi and Twele, 2009; Henderson and Xia, 1997; Joshi et al., 2016; Otukei et al., 2015; Pohl and van Genderen, 2014; Salentinig and Gamba, 2015; Soergel, 2010; Zhang et al., 2012). Numerous studies have been conducted with applications fusing SAR and optical data since the early 1990s (Calabresi, 1996; Dell'Acqua and Gamba, 2003; Erasmi and Twele, 2009; Ghassemian, 2016; Joshi et al., 2016; Lisini et al., 2006; Marceau et al., 1990; Otukei et al., 2015; Pohl and van Genderen, 2014; Salentinig and Gamba, 2015). In the latest review about data fusion in remote sensing, data fusion was defined as "the science of combining measurements, signals, or observations from different sources to obtain a result that is in some sense than what could have been achieved without this combination" (Schmitt and Zhu, 2016). Although numerous studies have addressed the fusion techniques and their applications for various multi-source satellite data, there is still no unified definition of the fusion levels in the literature (Pohl and Van Genderen, 2016). Fusion between optical and SAR data can be performed on three different levels: pixel level, feature level and decision level (Ghassemian, 2016; Lisini et al., 2006; Pohl and van Genderen, 1998; Waske and van der Linden, 2008). Generally, pixel level fusion was performed directly towards the images and it required geocoding and co-registration between the images to be fused (Baronti et al., 2011; Mitchell, 2010; Xu et al., 2014). Pixel level fusion does not require feature extraction from the images, such as spectral and textural features, even though feature points (e.g., roads or lakes) may be used during the co-registration. Especially, traditional pixel-level fusion methods (e.g., intensity-hue-saturation and principal component analysis) were reported to be unsuitable for SAR images because of speckle noise and the different imaging geometry between optical and SAR data (Zhang et al., 2010). Feature level fusion is usually performed on extracted features, such as the Grey Level Co-occurrence Matrix (GLCM) textural measures and Scale-Invariant Feature Transform (SIFT)

---

[*] Corresponding author.
E-mail address: zhangstream@gmail.com (H. Zhang).

http://dx.doi.org/10.1016/j.jag.2017.08.013
Received 21 July 2017; Received in revised form 17 August 2017; Accepted 31 August 2017

features of the objects or regions in the images (Pohl and Van Genderen, 2016; Zhong et al., 2015). Among the various features, GLCM based texture features are the most typical and commonly used. Texture features are also important for SAR data, particularly single polarization SAR data, since microwave is sensitive to the geometric properties in urban land surfaces (Dell'Acqua and Gamba, 2003; Masjedi et al., 2016). Several strategies have been proposed including layer-stacking and ensemble-learning methods, e.g., bagging, boosting, AdaBoost and Random Forest (Hall and Llinas, 1997; Rokach, 2010). The ensemble-learning methods can be conducted over different classifiers, e.g., artificial neural networks (ANN) and support vector machines (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 (Clinton et al., 2015; Waske and van der Linden, 2008). In particular, a hybrid level fusion can be designed by performing the fusion at two and more levels (Pohl and Van Genderen, 2016). However, there are some technical details that are not specified in the literature. For instance, in feature level fusion, after the features (e.g., textural features) are extracted over the neighborhood regions of pixels from the co-registered images, should the original image data be included in the fusion process or not? Furthermore, there is a lack of quantitative comparison among these different fusion levels regarding their effectiveness and applicability in fusing optical and SAR data. This study aimed at quantitatively comparing various levels of fusion between optical and SAR data for urban land cover classification and thus to provide a scientific reference for the selection of fusion levels in a wide range of applications using both optical and SAR satellite data.

2. Study area and data sets

2.1. Study sites

The Pearl River Delta (PRD) has become the fastest expanding metropolitan area in the southern part of China since the 1980s. The dramatic urbanization process has produced diverse and complex land cover over this region and resulted in a series of environmental problems, such as water and air pollution. To address these environmental issues, timely and accurate urban land cover change data are urgently required for related studies. In this study, three major cities, Guangzhou, Shenzhen and Hong Kong, were selected as study sites whose landscapes are very diverse and representative of urban areas.

Different types of optical and SAR data were selected to test the effectiveness and increase the reliability of the comparison among different fusion levels. The types of satellites or sensors, spatial resolutions and acquisition dates are listed in Table 1, showing the differences between different pairs of optical and SAR data in the three different study sites, Guangzhou, Shenzhen and Hong Kong. Table 1 demonstrates that the selected optical and SAR data were acquired at close dates to avoid possible land use or land cover changes. In particular, different types of SAR data were employed. These SAR data were obtained at different spatial resolutions at X band (TerraSAR-X) and C band (ENVISAR ASAR). The TerraSAR-X data used in this study were the StripMap product with 3 m resolution. For the ENVISAT ASAR data, two different working modes were used, including the Wide Swath Mode (WSM) at 75 m resolution and the IMage Precision mode (IMP) at 12.5 m resolution. These different types of SAR data were employed to compare the influence of different fusion levels using different combinations of optical and SAR data. Both the optical and SAR data were preprocessed with calibration. Atmospheric correction was conducted over the optical data. The SAR data were geo-coded with the corresponding digital elevation model (DEM) of the study sites. To co-register the two data sets, more than 20 pairs of tie-points were manually and carefully selected with visual interpretation over the SAR and corresponding optical images. Since optical images provided much better visual understanding, optical images were selected as the master images during the co-registration and thus the co-registered SAR images carried the same resolution as corresponding optical images. All the co-registration was completed with a Root Mean Square Error (RMSE) of less than one pixel. Additionally, all the SAR data were filtered to reduce speckle by applying the Enhanced Lee filter (Lopes et al., 1990; Xie et al., 2002). The kernel size was 3*3 in the Guangzhou case and 5*5 in the Shenzhen and Hong Kong cases. These kernel sizes were empirical values because the suitable kernel size should be related to the resolution and landscape of the study areas and in this study, the spatial resolution was 30 m in Guangzhou and 10 m in Shenzhen and Hong Kong. The geo-reference system was consistent for all the data sets after co-registration and it was a Universal Transverse Mercator (UTM) projection (Zone 50N) and Datum World Geodetic System 84 (WGS84).

Table 1
Optical and SAR data sets used for the three study sites in this research.

Study Area	Data attributes	Optical data	SAR data
Guangzhou	Satellite/Sensor	Landsat ETM+	ENVISAT ASAR (WSM)
	Resolution	30 m	75 m
	Acquisition date	31 December 2010	23 September 2010
Shenzhen	Satellite/Sensor	SPOT-5	ENVISAT ASAR (IMP)
	Resolution	10 m	12.5 m
	Acquisition date	21 November 2008	19 November 2008
Hong Kong	Satellite/Sensor	SPOT-5	TerraSAR-X
	Resolution	10 m	3 m
	Acquisition date	21 November 2008	16 November 2008

3. Methods

3.1. Framework of the research and design of the fusion levels

The framework of the proposed research was demonstrated in Fig. 1, where TSX denotes TerraSAR-X, GCPs denotes Ground Control Points (tie-points) and ULC denotes urban land cover. Before the co-registration of the optical and SAR data, preprocessing was applied to each data source, as described in Section 2. After the co-registration, different experiments were conducted with the registered optical and SAR data, as shown in Fig. 1, including feature extraction and pixel level fusion. There were three different fusion levels in this study and they are described in detail in the rest of this section.

In this study, four different levels were designed for the fusion of optical and SAR images, including the pixel level, feature level A, feature level B and the decision level. The definition of each fusion level generally followed the classical data fusion literature and latest publication (Hall and McMullen, 2005; Schmitt and Zhu, 2016). To provide a better understanding, the specific explanations of each fusion level are provided in Table 2, which can also be understood in Fig. 1. The pixel level fusion adopted in this study was also referred to as data-level fusion and observation-level fusion in the literatures (Hall and McMullen, 2005; Schmitt and Zhu, 2016). In addition, considering the special issue that the original images can be added to the extracted features in the feature level fusion, two different levels were designed, to compare the contribution of the original data in the feature level fusion in terms of the applications for urban land cover classification. Then, all the features described in Section 3.1 were employed in the two types of feature level fusion in Table 2. Moreover, decision level fusion was implemented by combining pixel level and feature level fusion results with a majority voting procedure to carry out the decision level fusion.

3.2. Feature extraction of optical and SAR data

Features of optical and SAR data were used in the feature level,