



Contents lists available at ScienceDirect

## Information Fusion

journal homepage: [www.elsevier.com/locate/inffus](http://www.elsevier.com/locate/inffus)

# Spatio-spectral fusion of satellite images based on dictionary-pair learning



Huihui Song<sup>a</sup>, Bo Huang<sup>a,\*</sup>, Kaihua Zhang<sup>b</sup>, Hankui Zhang<sup>a</sup>

<sup>a</sup> Department of Geography and Resource Management, Chinese University of Hong Kong, Hong Kong, China

<sup>b</sup> Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China

## ARTICLE INFO

### Article history:

Received 30 July 2012

Received in revised form 30 July 2013

Accepted 19 August 2013

Available online 29 August 2013

### Keywords:

Spatio-spectral fusion

High spatial resolution

High spectral resolution

Dictionary-pair learning

Sparse non-negative matrix factorization

## ABSTRACT

This paper proposes a novel spatial and spectral fusion method for satellite multispectral and hyperspectral (or high-spectral) images based on dictionary-pair learning. By combining the spectral information from sensors with low spatial resolution but high spectral resolution (LSHS) and the spatial information from sensors with high spatial resolution but low spectral resolution (HLSL), this method aims to generate fused data with both high spatial and spectral resolution. Based on the sparse non-negative matrix factorization technique, this method first extracts spectral bases of LSHS and HLSL images by making full use of the rich spectral information in LSHS data. The spectral bases of these two categories data then formulate a dictionary-pair due to their correspondence in representing each pixel spectra of LSHS data and HLSL data, respectively. Subsequently, the LSHS image is spatial unmixed by representing the HLSL image with respect to the corresponding learned dictionary to derive its representation coefficients. Combining the spectral bases of LSHS data and the representation coefficients of HLSL data, fused data are finally derived which are characterized by the spectral resolution of LSHS data and the spatial resolution of HLSL data. The experiments are carried out by comparing the proposed method with two representative methods on both simulation data and actual satellite images, including the fusion of Landsat/ETM+ and Aqua/MODIS data and the fusion of EO-1/Hyperion and SPOT5/HRG multispectral images. By visually comparing the fusion results and quantitatively evaluating them in term of several measurement indices, it can be concluded that the proposed method is effective in preserving both the spectral information and spatial details and performs better than the comparison approaches.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

A specific feature of remote sensing images, captured from different satellites and different sensors, is a tradeoff between spatial resolution and spectral resolution. This is caused, on the one hand, by the system tradeoffs related to data volume and signal-to-noise ratio (SNR) limitations and, on the other hand, by the specific requirements of different applications for a high spatial resolution or a high spectral resolution. For example, to fulfill the high spatial resolution requirement in many land-oriented applications, sensors with spatial resolution of half meter to tens of meters are designed, including, but not limited to, ETM+(30 m) on a platform of Landsat, the sensor on QuickBird (2.4 m for multispectral bands), and the instruments on SPOT (2.5 m–10 m). Sensors like the MODerate resolution Imaging Spectroradiometer (MODIS) on-board Aqua or Terra, the MEdium Resolution Imaging Spectrometer (MERIS) on-board ENVISAT, and the Hyperion instrument on-board

the EO-1 satellite, can provide remote sensing images with high spectral resolution but with low spatial resolution. Although these remote sensing instruments have many different properties, such as the revisit period, the swath width, the purpose of the launch (commercial or science), a specific user can obtain large amounts of data from different instruments on a given study area. This promotes the development of new algorithms to obtain remote sensing data with the best resolution available by merging their complementary information.

From the application point of view, remote sensing data with high spatial resolution are beneficial for the interpretation of satellite images and the extraction of spatial details of land cover, such as in land use/cover mapping and change detection [1]; whereas remote sensing data with high spectral resolution are capable of identifying those targets that cannot be easily distinguished by human eyes, such as in geological analysis and chemical contamination analysis [2]. To merge panchromatic (with a higher spatial resolution) and multispectral (with a lower spatial resolution) images from the same or separate sensors, pansharpener algorithms have been extensively studied in the past two decades [3,4]. In this paper, we focus on the fusion of images from two

\* Corresponding author. Tel./fax: +852 39436536.

E-mail addresses: [hhsongsherry@gmail.com](mailto:hhsongsherry@gmail.com) (H. Song), [bohuang@cuhk.edu.hk](mailto:bohuang@cuhk.edu.hk) (B. Huang), [zhkhua@gmail.com](mailto:zhkhua@gmail.com) (K. Zhang), [zhanghankui@gmail.com](mailto:zhanghankui@gmail.com) (H. Zhang).

categories of sensors: one category has low spatial resolution and high spectral resolution (hereinafter abbreviated as LSHS), such as Aqua (or Terra)/MODIS and EO-1/Hyperion; while the other category has high spatial resolution but low spectral resolution (hereinafter abbreviated as HSLS), the data from which are usually termed as multispectral images, such as Landsat/ETM+ and SPOT5/HRG. The low resolution or high resolution herein is in a relative sense. Through integrating the spectral information of LSHS data and the spatial information of HSLS data, we expect to extend the applications of available satellite images, thereby meeting various demands of data users.

To address this spatial and spectral fusion problem, one category of classic method is spatial-unmixing-based algorithms [5–9]. The processing steps of these methods are: (1) geometrical co-registration between the HSLS and LSHS images; (2) the multi-spectral classification on the high spatial resolution image (HSLS) to unmix the low spatial resolution image (LSHS); and (3) determination of the class spectra via regularized linear unmixing. These methods showed good performance at fusing Landsat images with MERIS [6,7,9] images or with ASTER [8] images for land applications. However, it should be noted that whether pure spectra exist resides in the sensor's spatial resolution of the given HSLS data. These methods also have high demanding on the geometric registration accuracy of the given data (e.g., less than 0.1–0.2 of the low resolution pixel size according to Ref. [5]). The authors in [10–12] proposed to fuse multispectral and hyperspectral images in a maximum a posteriori (MAP) estimation framework by establishing an observation model between the desired image and the known image. The method reported in [10] employed a spatially varying statistical model to help exploit the correlations between multispectral and hyperspectral images. The methods developed in [11,12] made use of a stochastic mixing model of the underlying scene content to enhance the spatial resolution of hyperspectral image. The authors in [13] proposed to improve the spatial resolution of hyperspectral image by fusing with high resolution panchromatic band images based on super-resolution technique. Firstly, this method learns the spatial information from panchromatic images by sparse representation; then the high-resolution hyperspectral image is constructed based on the learned spatial structures with a spectral regularization.

Due to the low spatial resolution or the existence of homogeneous mixtures in the hyperspectral images, unmixing techniques are needed to decompose the mixed pixels observed into a set of constituents and the corresponding fractional coefficients, which denote the proportion of each constituent [14]. The first step in this spectral unmixing task is to collect a suitable set of endmembers to model the spectra of measured pixels by weighting these spectral signatures. There are two categories of endmembers according to different extraction methods: image endmember derived directly from the images and library endmember derived from known target materials in field or laboratory spectra [15]. However, employing library endmembers is risky because it is difficult to ensure that these spectra are captured under the same physical conditions as the observed data. Whereas image endmembers can avoid this problem due to the collection at the same scale as the observed data, thereby being linked to the scene features more easily [14]. A number of endmember extraction algorithms for hyperspectral data were quantitatively compared in [15]. In this paper, we employ this similar endmember extraction strategy for both HSLS and LSHS data and term the extracted spectral bases as dictionaries. For the second step of abundances estimation in hyperspectral unmixing, a popular and effective method is sparse unmixing [16]. The basic principle of this method is based on the observation that only a small portion of endmembers participated in the formation of each mixed pixel. Therefore, with the sparsity-inducing regularizers, i.e., the constraint of a few non-zero components for the

abundance vectors, the abundances can be estimated by calling for linear sparse regression techniques [16]. In this paper, we adopt this similar sparse regularization when solving the representation coefficients of HSLS and LSHS data with respect to their representation atoms (or dictionaries).

In this paper, we seek to extract a dictionary-pair for representing LSHS and HSLS data, respectively. Specifically, the representation atoms of LSHS and HSLS data are firstly extracted from the given images, respectively, and then to form the dictionary-pair. Each representation atom of the dictionary-pair herein is in correspondence. Accordingly, each pixel spectra of LSHS and HSLS data can be expressed as a linear combination of their corresponding dictionary atoms, which have the same functions as endmembers in hyperspectral unmixing. Based on this dictionary-pair, the proposed spatio-spectral fusion algorithm consists of two stages: in the first stage, the good spectral properties of the LSHS image are employed to extract the basis functions of spectra (representation atoms) and further to form the dictionary-pair by enforcing the same representation coefficients of the HSLS and the LSHS images with respect to their dictionaries; in the second stage, the good spatial properties of the HSLS image are utilized to derive the representation coefficients with respect to its dictionary. The representation coefficients herein have similar functions as abundances in hyperspectral unmixing but provide spatial location properties due to the high spatial resolution of the HSLS image. Finally, the desired high spatial and high spectral resolution image can be obtained by the multiplication of representation atoms for the LSHS image (i.e., the LSHS dictionary) and representation coefficients for the HSLS image.

The following section presents the theoretical basis of this paper. Section 3 describes the proposed method for the fusion of HSLS and LSHS data. Section 4 shows the experimental validation of the proposed algorithm through comparison with two representative algorithms on both simulated and actual satellite datasets. Finally, we conclude this paper with a discussion on the application of the proposed method and remarks about its inherent features.

## 2. Theoretical basis

As introduced in Section 1, a dictionary-pair needs to be trained from the HSLS and LSHS data. Hence, the basic principles of dictionary-pair learning will first be introduced. Taking the non-negative properties of the bases spectra and fractional abundances of HSLS and LSHS data into account, we learn the required dictionary-pair by using sparse non-negative matrix factorization method, which will be presented in the second part of this section.

### 2.1. Dictionary-pair learning

Huge amounts of information captured by human eyes or satellite sensors are superfluous in a large part caused by the related signals in real world and the oversampling of sensors. Compared to the observed signals, the underlying generating signals actually have very small dimensions. Identifying the generating signals from the relevant information is, in fact, a dimension reduction process or finding the subspace where the data lie [17]. The combination of these generating signals or bases of subspace constitutes the dictionary, each component of which is a representation atom for the observed dataset. Given the observed data, building the dictionary to provide efficient representations for the given signals, is the dictionary learning process, which has greatly promoted the development of novel dictionary learning algorithms [17–19].

For representation convenience, we will use the superscript of symbols to discriminate variables and the subscript to denote the

Download English Version:

<https://daneshyari.com/en/article/528768>

Download Persian Version:

<https://daneshyari.com/article/528768>

[Daneshyari.com](https://daneshyari.com)