Computers and Electronics in Agriculture 145 (2018) 379–388

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Greedy compressive sensing and reconstruction of vegetation spectra for plant physiological and biochemical parameters inversion

Xu Ping, Chen Bingqiang, Zhang Jingcheng*, Xue Lingyun*, Zhu Lei, Qiu Bo

Hangzhou Dianzi University, College of Life Information Science & Instrument Engineering, Hangzhou, China

A R T I C L E I N F O

Keywords:
Spectral analysis
Remote sensing
Compressive sensing
Physiological and biochemical parameters
Spectral index

A B S T R A C T

In order to compress and reconstruct the plant spectral data effectively, greedy compressive sensing methods are introduced to improve the storage and transmission efficiency of data while maintaining interpretation capacity of spectral information for physiological and biochemical parameters of plant. The key physiological and biochemical parameters such as water content, carotenoid content and chlorophyll content were chosen to test the retrieving efficacy of the greedy compressive sensing algorithms of MP, ROMP, OMP, StOMP and CoSaMP at different sampling rates. The performances of all algorithms are thoroughly evaluated on three different levels: the spectral level, the feature level and the model level. The experimental results show that the errors of these three levels of all methods are regularly changed with the sampling rate. For the spectral level, the error of spectral reconstruction of all algorithms stabilized within 2% when the sampling rate is higher than 0.15bpp. For the feature level, the spectral indices of all algorithms exhibited a decreasing trend with the increase of the sampling rate. When the sampling rate is higher than 0.2bpp, the errors of the reconstructed spectral indices of all methods are lower than 2%. For the model level, the reconstructed normalized root mean square error of chlorophyll and carotenoids of all methods are reduced to 8% and 12% respectively at the sampling rate of 0.25bpp. For the error of water content, the reconstruction errors of MP and ROMP are about 24% and those of OMP, StOMP and CoSaMP were lower than 15% at the sampling rate of 0.4bpp. Therefore, these greedy compressive sensing algorithms not only reduce the data volume of plant spectra significantly, but also maintain plant critical spectral characteristics.

1. Introduction

Currently, Remote sensing technology plays an important role in agricultural resource survey, agricultural resource monitoring, biomass estimation and agricultural disaster prediction (Lu, 1998). In various forms of remote sensing data, given the hyperspectral technology can obtain spectral information on the continuous wavelength, it can provide abundant information which facilitate the spectral analysis and information extraction (Tang and Huang, 2001). This advantage enables the hyperspectral technology presenting great potential in plant physiological and biochemical parameters inversion (Ruan and Niu, 2004), stress detection (Zhang et al., 2011), etc.

With the rapid development of vegetation hyperspectral technology, the massive data in the process of acquisition, transmission and analysis also challenge the traditional data acquisition and compression technology (Zhang et al., 2014). (Candès et al., 2006; Donoho, 2006; Candes and Tao, 2007) proposed a new data acquisition and processing theory namely compressive sensing (CS). CS samples data at far below the Nyquist sampling rate by constructing an uncorrelated observation matrix, and the original data is reconstructed by a reconstruction algorithm. Due to its great advantage in data processing, it has aroused the concern of related fields including wireless sensor network (WSN) and Internet of Things (IoT) since its development (Liu et al., 2015; Liu and Wang, 2017; Zheng et al., 2015, 2017).

The reconstruction algorithm of CS can be divided into three categories: convex optimization algorithm, greedy algorithm and combinatorial algorithm. Convex optimization algorithm can achieve high reconstruction precision with high computation complexity. Combinatorial optimization algorithm uses the structured method to sample the original signals with high speed, but its computation complexity is very high and the sampling matrix is too complicated to be used in real acquisition system. Greedy algorithm can not only reconstruct signals with high fidelity, but also its calculation speed is significantly higher than that of the others. Therefore, greedy algorithm has greater potential to be used in the massive plant hyperspectral data. The current greedy reconstruction algorithm mainly includes Matching

* Corresponding author.
E-mail addresses: zhangjcrs@hdu.edu.cn (J. Zhang), xly@hdu.edu.cn (L. Xue).

https://doi.org/10.1016/j.compag.2018.01.005
Received 1 August 2017; Received in revised form 4 January 2018; Accepted 6 January 2018

Pursuit (MP) (Mallat and Zhang, 1993), Orthogonal Match Pursuit (OMP) (Tropp and Gilbert, 2007), Stagewise Orthogonal Matching Pursuit (StOMP) (Donoho et al., 2006), Regularized Orthogonal Matching Pursuit (ROMP) (Needell and Vershynin, 2010) and Compressive Sampling Orthogonal Matching Pursuit (CoSOMP) (Needell and Tropp, 2008), etc.

To preserve the image texture, edge and other information, sparse representation is constructed by different sparse basis to improve the reconstruction accuracy (Lian and Chen, 2010). For hyperspectral data with the characteristics of its unification of map and high spectral-spatial correlation, Duarte proposed the Kronecker product matrices as a sparse basis (Duarte and Baraniuk, 2012). Coluccia improved the reconstruction accuracy according to inter-spectral correlation (Coluccia et al., 2013). Ly proposed random projection reconstruction of hyperspectral imagery with spectral and spatial partitioning (Ly et al., 2013). Chen proposed a sparse representation for target detection in hyperspectral imagery (Chen et al., 2011). Eason raised a total variation regularization via continuation to recover compressed hyperspectral images (Eason and Andrews, 2015). High-resolution computational spectral imaging of remote sensing based on coded sensing had been used (Shi et al., 2011). A pixel-based distributed compression sensing that rely on the endmember and abundance of hyperspectral data has been developed (Wang et al., 2015).

It should be noted that compressive sensing technology is often closely related to a specific field and application characteristics. In the application of vegetation, a common and important task is to use hyperspectral data to invert plant physiological and biochemical parameters, that is, to establish the spectral inversion model of plant physiological and biochemical parameters (Wang, 2008). These studies usually consider vegetation index as a characteristic variable in addition to the spectral reflectance (Liang et al., 2010). The leaf-based pigment index was used to invert the contents of chlorophyll and carotenoid to obtain better performance than the normalized ratio pigment index (Wang et al., 2009). Li found that the ratio of spectral reflectance at 1600 nm and 820 nm can eliminate the influence of external factors such as environmental background and canopy structure in the inversion of vegetation water content, improve the inversion accuracy, and reflect the temporal and spatial changes of water content of vegetation (Li et al., 2009). Yang adopted the radiative transfer model and the neural network method to invert the plant leaf area index (Yang et al., 2011).

To sum up, the application of hyperspectral remote sensing of vegetation is mainly focused on the spectral analysis. However, the existing compressive sensing methods are mainly for spatial analysis whereas the spectral analysis is less involved. Therefore, in order to premote the research and application of quantitative remote sensing of vegetation, there is an urgent need to extend the compressive sensing technology to the spectral domain, and establish compressive sensing reconstruction methods which can maintain the key characteristics of vegetation spectrum. Moreover, the reconstruction effect of the original spectral signals and inversion models of vegetation physiological and biochemical parameters of the compressive sensing technology should be evaluated.

This study aims to solve the above mentioned issues. Firstly, inversion models of the calculated spectral index and the original vegetation physiological and biochemical parameters were obtained using the partial least squares regression (PLSR) method in this study. Secondly, the spectral reflectance of the plants were sampled and reconstructed by five kinds of greedy reconstructed algorithms at different sampling rate. Finally, the reconstructed spectral data were analyzed at the spectral level, the feature level and the model level.

Table 1
Statistical parameters in PROSPECT model.

Input parameter	Unit	Mean ± s.d.	Min	Max
Blade structure parameters (*N*)	/	1.53 ± 0.22	1.01	2.24
Chlorophyll concentration (*Cab*)	$\mu g \cdot cm^{-2}$	34.58 ± 17.13	0.36	96.58
Carotenoid concentration (*Car*)	$\mu g \cdot cm^{-2}$	8.66 ± 3.73	0.04	24.41
Equivalent water thickness (*Cw*)	$g \cdot cm^{-2}$	0.015 ± 0.006	0.004	0.036
Leaf dry matter content (*Cm*)	$g \cdot cm^{-2}$	0.007 9 ± 0.003 3	0.000 8	0.019 5

2. Materials and methods

2.1. Data description

In this study, the experimental data were obtained by radiative transfer model of leaf. The PROSPECT model (Jacquemoud and Baret, 1990) is the most widely used leaf reflectance simulation model in the world. The leaf reflectance is simulated according to the leaf physiological and biochemical parameters and scattering parameters in the wavelength range of 400–2500 nm. The PROSPECT model was used to simulate the following physiological and biochemical parameters, including the following physiological and biochemical parameters: leaf structure parameter N, chlorophyll concentration Cab (g/cm^2), carotenoid concentration Car (g/cm^2), equivalent water thickness Cw (g/cm^2) and leaf dry matter content Cm (g/cm^2). According to the multivariate normal distribution of physiological and biochemical parameters of common vegetation (Féret et al., 2011; Cheng et al., 2012), the random 2500 groups data are generated. The mean, standard deviation, maximum and minimum values of the data are shown in Table 1.

2.2. Experimental flow chart

The experimental process includes random measurement of the sampling data, reconstruction, physiological and biochemical parameters inversion, error analysis and result comparison as shown in Fig. 1. Firstly, the original hyperspectral data are transformed based on sparse representation, and then are sampled randomly. Secondly, the reconstructed algorithms of MP, OMP, ROMP, StOMP and CoSaMP are used to reconstruct the hyperspectral data accurately. Then, the original data is used to construct the inversion models between the spectral index and the physiological and biochemical parameters. Finally, the performance of different reconstruction algorithms are evaluated and compared.

2.3. Spectral index of plant physiological and biochemical parameters

As important physiological and biochemical parameters, water content, carotenoid content and chlorophyll content play an important role in the growth of vegetation (Plummer et al., 1995). Based on the comprehensive investigation of classical vegetation indices for retrieving physiological and biochemical parameters of plants, four spectral indices are selected for each physiological and biochemical parameter (Table 2).

2.4. Construction of partial least squares regression model for plant physiological and biochemical parameters

Considering high inter-spectral correlation, partial least squares regression (PLSR) (Zhang et al., 2005) is used to construct inversion model. Taking the PLSR model of the water content as an example, the count of sample is 2500 and each sample includes the band range of