Article

Adaptive Grouping Distributed Compressive Sensing Reconstruction of Plant Hyperspectral Data

Ping Xu, Junfeng Liu, Jingcheng Zhang *, Lingyun Xue * and Bo Qiu

College of Life Information Science & Instrument Engineering, Hangzhou Dianzi University, HangZhou, 310018, China; xuping@hdu.edu.cn(P.X.); fengjliu@126.com(J.L.); qb86880301@126.com(B.Q.)

* Correspondence: zhangjcrs@hdu.edu.cn (J.Z.); xly@hdu.edu.cn(L.X.); Tel.: +86-137-3228-4126 (J.Z.); +86-133-8860-3657 (L.X.)

Abstract: With the development of hyperspectral technology, to establish an effective spectral data compressive reconstruction method that can improve data storage, transmission and maintaining spectral information is critical for quantitative remote sensing research and application in vegetation. The spectral adaptive grouping distributed compressive sensing algorithm is proposed, which enables a distributed compressed sensing reconstruction of plant hyperspectral data. The spectral characteristics of hyperspectral data are analyzed and the joint sparse model is constructed. The spectral bands are adaptively grouped and the hyperspectral data are compressed and reconstructed on the basis of grouping. The experimental results showed that, comparing with orthogonal matching pursuit (OMP) and gradient projection reconstruction (GPSR), the proposed algorithm can significantly improve the visual effect of image reconstruction in the spatial domain. The PSNR in low sampling rate (sampling rate is lower than 0.2) increases by 13.72dB than OMP and 1.66dB than GPSR. In the spectral domain, the average normalized root mean square error, the mean absolute percentage error and the mean absolute error of the proposed algorithm is 35.38%, 31.83% and 33.33% lower than GPSR respectively. Besides, the proposed algorithm can achieve relatively high reconstructed efficiency.

Keywords: hyperspectral image; spectral characteristics of plants; spectral adaptive grouping; compressive sensing

1. Introduction

The hyperspectral technology is a breaking technology in agriculture remote sensing which enables the dynamic and precise monitoring of crop types and crop growth. The hyperspectral remote sensing technology has been widely used in estimating the yield of crops, agricultural resources survey, agricultural disaster monitoring and precision agriculture[1]. Plant quantitative remote sensing technology is widely used in a variety of applications by mining spectral information and setting up spectral retrieving model. For the estimation of crop yield, Nuarrsa et al extracted rice area with the overall accuracy of 87.91% using NDVI, RVI and SAVI vegetation indexes from MODIS time series data[2]. In the investigation of agricultural resources and crop disaster monitoring, Bao et al improved the Lukina variable fertilization model to further promote the variable fertilization technology towards practical direction[3]. In study of detection and recognition of wheat stripe rust, Zhang simulated TM pixels and had relatively high precision at wheat filling period with R2=0.93[4].

With the aid of internet technology, agricultural informatization and agricultural big data have become an inevitable trend. In recent years, with successive launch of hyperspectral satellite and the development of micro hyperspectral imager of UAV, the applications of hyperspectral remote sensing became widely available. However, on the other hand, the increase of data volume brings great challenge on data transmission, analysis and storage[5]. Candes, Donoho and Tao et al proposed a new data acquisition and processing theory called Compressive Sensing(CS)[6][7][8]. Compressive sensing samples data at far below the Nyquist sampling rate by constructing

uncorrelated observation matrix, and the original data is reconstructed by reconstruction algorithm. It thus provides a new way for compressing and reconstructing data with large volume.

At present, many studies have been done on applying compressive sensing in processing high-dimensional data. Kang et al.[9] proposed a method of distributed compressive sensing to grouping the video sequences efficiently by studying the correlation of the video sequences. Ly et al. [10] pointed out that the hyperspectral data should be stochastically separated by spectral and spatial partitioning. Chen et al.[11] proposed a sparse method for hyperspectral image target detection. Wang et al proposed a pixel-based distributed compressive sensing[12], which divides the hyperspectral data into endmember extraction and abundance estimation through a linear mixture model.

To promote the application of hyperspectral remote sensing in agriculture, the compressive sensing method provides a new method for the compression and recovery of hyperspectral data. However, at present, the research of compressive sensing is mainly focused on the reconstruction of spatial image, and the spectral dimension of the hyperspectral data needs to be concerned in information reconstruction. Hyperspectral images have spectral-spatial correlations. The compressive sensing of high dimensional data using the autocorrelation nature of data to improve the data sparse representation, which is able to reduce the complexity and improve the accuracy of reconstruction. In this study, through the analysis of plant spectral characteristics, a distributed spectral adaptive grouping compressive sensing is proposed and verified.

2. Methods

2.1. Spectral adaptive distributed compressive sensing

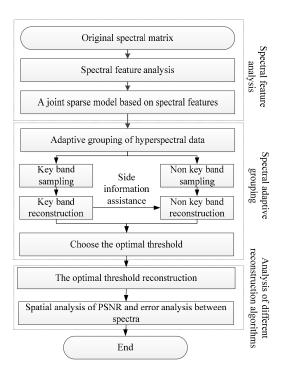


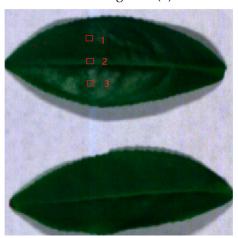
Figure 1. Flow chart of plant hyperspectral compressive sensing reconstruction

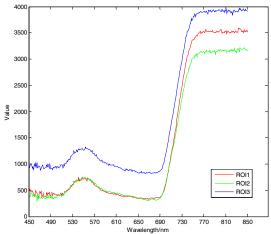
Figure 1 showed the flow chart of the proposed plant hyperspectral compressive sensing reconstruction algorithm. Firstly, the spectral characteristics of hyperspectral data are analyzed and the joint sparse model is constructed. Secondly, the spectral bands are adaptively grouped and the hyperspectral data are compressed and reconstructed on the basis of grouping, so as to determine the optimal grouping threshold. Then, to evaluate the reconstruction effect, the proposed algorithm,

OMP[13] and GPSR[14] are used to analyze the PSNR in the space and calculate errors among spectra.

2.2. Analysis of plant spectral characteristics

The data used in the experiment are hyperspectral images of tea. With wavelength 450~850nm and a total of 320 bands of data, a single pixel is defined by a 12 bits unsigned integer. The image of 661nm, 553nm and 449nm from the original data are selected as the red, green and blue channels of the false color composite RGB images as shown in Figure 2.(a). Within each leaf, 3 regions of interest (ROIs) are marked with the size of 3 pixels x 3 pixels. The averaged spectrum for each ROI is demonstrated in Figure 2.(b).





- (a) Original data space image
- (b) The averaged spectrum for each ROI

Figure 2. Spatial image of hyperspectral raw data and three regional spectral curves

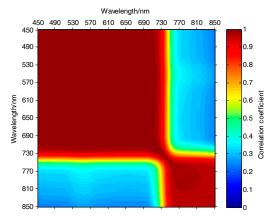


Figure 3. Correlation coefficient diagram

It can be observed from the spectral curve that despite the averaged spectral curves of different regions are different, the trend of the curves is consistent. The fluctuation of spectral reflectance values in 450~680nm are relatively small. The spectrum in 680~760nm rised rapidly. In 760~850nm, the spectral reflectance is stablized at relatively high values, but the fluctuation within the local adjacent wave bands is relatively large. In order to further study on the distribution of spectral correlation of plant hyperspectral images, the correlation matrix of the hyperspectral bands is shown in Figure 3. It can be observed from the correlation coefficient diagram that the spectral reflectance is high in 450~680nm and 760~850nm, and the absolute value of autocorrelation and cross-correlation coefficient between the bands is above 0.9. The correlation between 680~760nm is small, which is consistent with the variation of the spectral curve. Based on the qualitative analysis of the correlation characteristics of the plant spectrum, the following conclusions are obtained by analyzing the

correlation coefficient(r) of each band. 59.41% bands are extremely correlated(|r| > 0.95); 6.93% bands are highly correlated (0.8< $|r| \le 0.95$); 8.48% bands are moderately correlated (0.5< $|r| \le 0.8$); whereas 25.18% bands are lowly correlated (0.3< $|r| \le 0.5$) or not correlated ($|r| \le 0.3$).

2.3. Joint sparse model

According to the spectral characteristics of the hyperspectral images, there is a general correlation between the bands of the plant hyperspectral images and that more than half of the bands are highly correlated. The higher the correlation between the bands, less difference remained. Based on this, the present study proposed a joint sparse model of bands of plant hyperspectral images[15].

Assuming there are two bands X_i and X_{i+1} , the correlation between X_i and X_{i+1} can be calculated, where X_i represents the current band. As the result of different wavelength reflection of the same object in the different bands of hyperspectral images, X_i and X_{i+1} have the same spatial information. Meanwhile, X_i and X_{i+1} have their own unique spectral information. The X_i and X_{i+1} can be described as below:

$$X_i = X_c + X_{i_r} \tag{1}$$

$$X_{i+1} = X_c + X_{i+1_r}$$
 (2)

Where X_c is the same part of X_i and X_{i+1} , which refers to the same spatial information, X_{i_r} and X_{i+1_r} are their own unique parts, which mean the results of different reflections of different wavelengths. X_i is used as a reference to X_{i+1} , and the same spectral estimation (X_c) and the different information error coding (X_{i+1_r}) are used as the predictive value of X_{i+1} in the spectral coding. The joint sparse model can be expressed as below:

$$X_i = \Psi S_i \tag{3}$$

$$X_{i+1_r} = \Psi S_{i+1_r} \tag{4}$$

Where S_i and S_{i+1_r} are sparse representations of X_i and X_{i+1_r} , respectively, Ψ is a canonical orthogonal matrix.

2.4. Distributed compressive sensing based on spectral characteristics

In the distributed compressive sensing, hyperspectral band data are divided into a series of GOPs bands. Each GOP is consisted by several bands which contains one key band and several non-key bands. Then, the sampling rate of different sampling rates of the key band and non-key bands is realized by controlling the sampling matrix.

$$Y_i = \Phi X_i = \Phi \Psi S_i \tag{5}$$

Where Ψ is a canonical orthogonal matrix, S_i is a sparse representation of the original signal in the transform domain, $\Psi^H\Psi = \Psi\Psi^H = I$, and I is a unit matrix. The size of the Φ is $M \times N$, M << N, which is the partial block Hadamard matrix[16]. X_i is a band of hyperspectral data, and Y_i is an observed value.

At last, GPSR algorithm is used in the key band and the reconstruction of the non-key band is assisted by using the information from key band. Compressive sensing reconstruction is the process of solving equation (6). The solution of l_0 norm is a NP-Hard problem. The minimization problem of l_1 norm is equivalent to that of l_0 norm under certain conditions[17], so equation (6) can be transformed to equation (7). In this paper, GPSR and OMP are chosen to reconstruct X_i on the basis of the Y_i .

$$\hat{x} = \arg\min \|x\|_0 \quad \text{st } \Phi x = y \tag{6}$$

$$\hat{x} = \arg\min \|x\|_{1} \quad \text{st } \Phi x = y \tag{7}$$

3. Spectral adaptive threshold grouping

3.1. Spectral adaptive grouping and selection of key bands

Based on the above analysis, non-key bands can be reconstructed by key band with side information assist. Therefore, it is very important to effectively group bands and select more effective key bands. In this study, PSNR is used as a basis for adaptive grouping of hyperspectral data and selecting key bands. The steps of adaptive grouping algorithm are as follows:

Step 1: Solve all PSNRs between the first band and each of the rest bands, and those bands in the rest of the bands and those bands whose PSNRs are greater than the threshold are all selected and classified into the group of the first band.

Step 2: Set up a new set from the remaining bands and repeat Step 1 to construct a new group.

Step 3: repeat step 2 until all bands are assigned to different groups.

In the adaptive band grouping algorithm, we can see that in every grouping, the PSNR values of the first band and other bands are greater than the threshold value. Given the first band in the group has high similarity with the remaining bands, the first band of each group is determined as the key band and the other bands are non-key bands.

3.2. The results of adaptive grouping and the different sampling rate of key and non-key bands

Key bands are sampled at a high rate and all sampling rate is from 0.1bpp (bit per pixel) to 0.5bpp. In case the sampling rate of the GPSR algorithm over 0.4bpp, the PSNR value changed slowly. The sampling rate of the key band is set to 0.5bpp. The grouping results of non-key bands at different thresholds and the sampling rates of non-key bands are shown in Table 1.

It can be seen from Table 1, with the increase of the number of groupings, that the sampling rate of the non-key band decreases gradually, especially at low sampling rate. When the threshold is set to 30dB and the overall sampling rate is 0.1bpp, the number of groupings will reach 65 and the sampling rate of the non-key band will be negative.

Threshold/dB	Groups	Sampling rate/bpp					
		0.1	0.2	0.3	0.4	0.5	
20	6	0.092	0.194	0.296	0.398	0.500	
21	7	0.091	0.193	0.296	0.398	0.500	
22	7	0.091	0.193	0.296	0.398	0.500	
23	8	0.090	0.192	0.295	0.397	0.500	
24	15	0.080	0.185	0.290	0.395	0.500	
25	21	0.072	0.179	0.286	0.393	0.500	
26	29	0.060	0.170	0.280	0.390	0.500	
27	36	0.049	0.162	0.275	0.387	0.500	
28	45	0.035	0.151	0.267	0.384	0.500	
29	56	0.015	0.136	0.258	0.379	0.500	

Table 1 Sampling rates of none-key bands in different groups

3.3. Analysis of results of adaptive band grouping reconstruction

According to the sampling rate of the non-key bands calculated in Table 1, experimental results of reconstructed PSNR of key bands, non-key bands, and all bands are shown as Figure 4, Figure 5 and Figure 6.

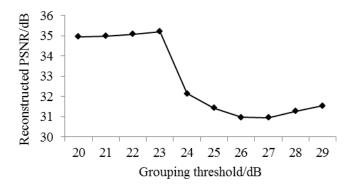


Figure 4. Average reconstructed PSNR of key bands for different groups

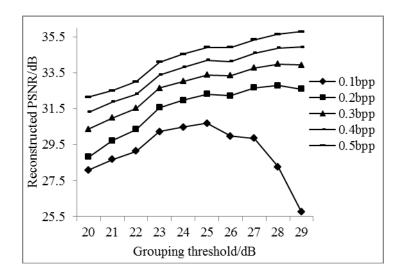


Figure 5. Average reconstructed PSNR of non-key bands for different groups

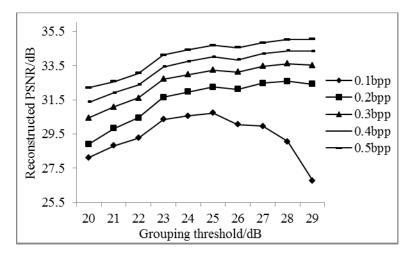


Figure 6. Average reconstructed PSNR of all bands for different groups

It can be seen from Figure 4 that the PSNR values of different grouping of key band reconstruction vary significantly, as the sampling rate of key bands is set to 0.5bpp. To explain this phenomenon, more experiments are implemented for each band at the sampling rate of 0.5bpp as shown in Figure 7.

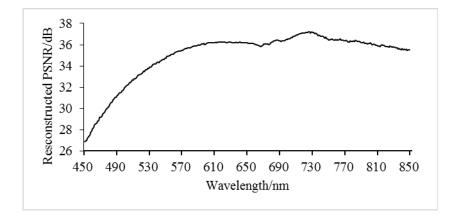


Figure 7. Reconstructed PSNR of each band at the sampling rate of 0.5bpp

In Figure 7, when the sampling rate is 0.5bpp, the reconstructed PSNR values of different bands are quite different. According to spectral characteristic curve, the difference of the spectrum with proximate reconstructed PSNR value is insignificant. The lower PSNR threshold is, more groups will emerge and the higher average PSNR value of key bands will be, and vice versa.

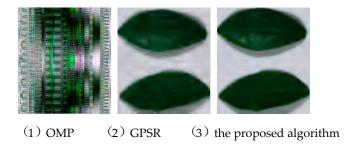
As can be seen from Figure 5 and Figure 6, when the sampling rate is larger, the reconstructed PSNR values of non-key bands and all bands are both higher. More groupings indicate less efficiency of overall reconstruction. In order to obtain good reconstructed PSNR for all sampling rate from 0.1bpp to 0.5bpp, the grouping PNSR threshold of 25dB is chosen in the following experiments.

4. Data experiment and analysis

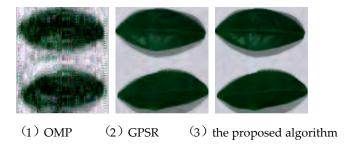
In the experiments, the software platform is Matlab R2012a and hardware platform is Lenovo notebook computer in which CPU is Intel I3-2350M clocked at 2.3GHz and memory is 6G.

4.1. Spatial domain reconstruction analysis

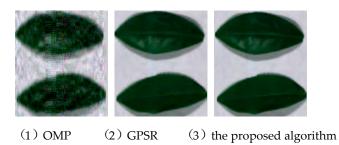
The 25dB was selected as the overall threshold for further analysis using OMP and GPSR algorithms. The reconstructed results are analyzed using the subjective evaluation and average peak signal to noise ratio in the spatial domain. To facilitate a visual comparison, 661nm, 553nm, 449nm are selected respectively as red, green and blue channels to form the synthesize RGB image. Figure 8 illustrates the experimental results for OMP, GPSR and the proposed algorithm under the sampling rate from 0.1bpp to 0.5bpp. The fidelity of reconstructed images of all algorithms are significantly related with sampling rate at low sampling rate especially. The increasing of sampling rate can significantly improve the subjective quality of reconstructed image. Better subjective quality can be obtained using the proposed algorithm comparing with the other algorithms. The reconstructed subjective quality of the proposed algorithm is very close to that of GPSR and significantly better than that of OMP at high sampling rate (great than or equal to 0.4bpp). Experimental results show that the side information assists of key bands can improve the quality of reconstruction at the low sampling rate.



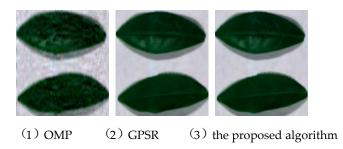
(a) reconstruction results for different algorithm at the sampling rate of 0.1bpp



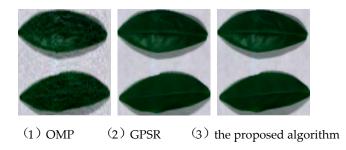
(b) reconstruction results for different algorithm at the sampling rate of 0.2bpp



(c) reconstruction results for different algorithm at the sampling rate of 0.3bpp



(d) reconstruction results for different algorithm at the sampling rate of 0.4bpp



(e) reconstruction results for different algorithm at the sampling rate of 0.5bpp

Figure 8. Comparison of reconstruction results for different algorithm at different sampling rates

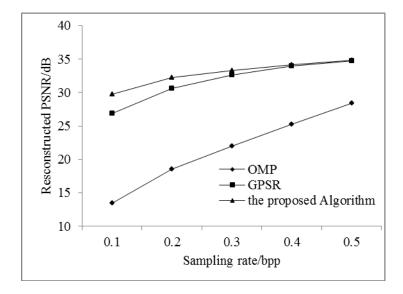
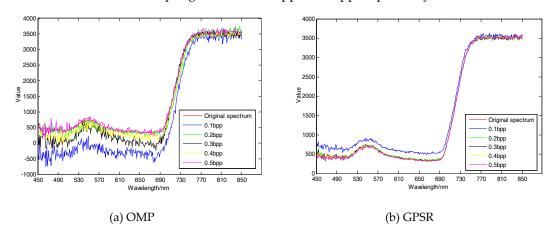


Figure 9. Comparison of reconstructed PSNR with different reconstruction methods

As can be seen from Figure 9 that the reconstructed PSNRs of the proposed method and GPSR are significantly higher than that of OMP under sampling rate from 0.1bpp to 0.5bpp. At low sampling rate from 0.1bpp to 0.3bpp, the reconstructed PSNR of the proposed algorithm is significantly higher than that of GPSR. At high sampling rate (greater than or equal to 0.4bpp), the reconstructed PSNR of the proposed algorithm approaches to that of GPSR. It is shown that the joint sparse model improves the reconstructed PSNR in the spatial domain, especially at low sampling rate.

4.2. Spectral domain reconstruction results analysis

The spectral analysis on vegetation is an important basis for monitoring of crop growing status and crop stressors in precision farming. According to regions of interest (ROIs) in Figure 1, the average reconstructed spectral curves of ROIs using OMP, GPSR and the proposed algorithm are calculated at different sampling rate from 0.1bpp to 0.5bpp respectively.



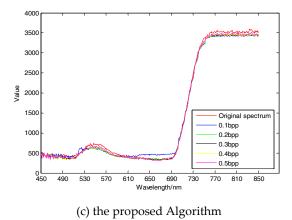
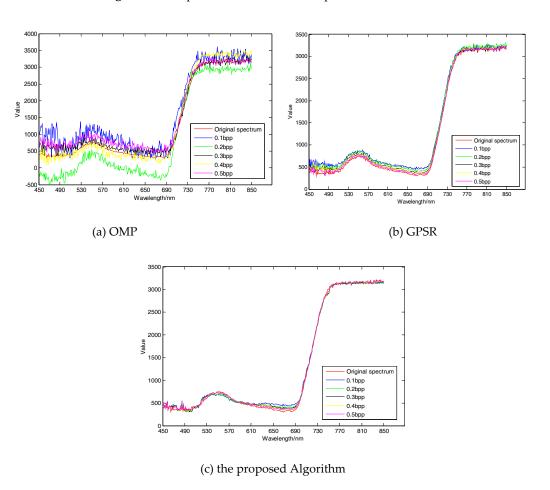


Figure 10. Comparison of reconstructed spectral curves of ROI 1



 $\textbf{Figure 11.} \ Comparison \ of \ reconstructed \ spectral \ curves \ of \ ROI \ 2$

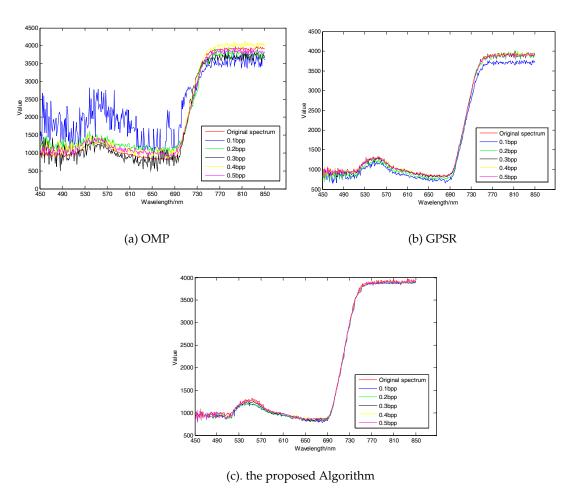


Figure 12. Comparison of reconstructed spectral curves of ROI 3

As shown in Figure 10, Figure 11 and Figure 12, different regions have different sensitivity to compressive sensing reconstruction. There is a significant correlation between the reconstructed algorithms and the sampling rates, and the reconstructed effect is improved with the increase of the sampling rate. Besides, the reconstructed spectral curves of the proposed algorithm is significantly higher than those of the others when sampling rate is less than 0.3bpp, and those of OMP is the lowest under different sampling rates. The reconstructed effectiveness of various methods is quantified by introducing the root mean square error, mean absolute error and average relative error as shown in Table 2. The root mean square error, mean absolute error and average relative error of the proposed algorithm are better than those of the others when the sampling rate is less than 0.2bpp. The root mean square error, mean absolute error and average relative error of the proposed algorithm is comparable to those of GPSR at the sampling rate of greater than or equal to 0.3bpp, and those of OMP is lower than those of the others.

Table 2 Error analysis

Error analysis of different algorithms		Sampling Rate/bpp					
Effor analysis of uniferen	interent argorithms		0.2	0.3	0.4	0.5	
	OMP	0.9683	0.5158	0.2861	0.2105	0.1598	
Mean absolute percentage error	GPSR	0.1560	0.1006	0.0796	0.0689	0.0617	
	the proposed algorithm	0.1005	0.0814	0.0728	0.0680	0.0639	

Mean absolute error	OMP	0.4599	0.2408	0.1434	0.1035	0.0787
	GPSR	0.0798	0.0529	0.0424	0.0369	0.0329
	the proposed algorithm	0.0544	0.0451	0.0401	0.0370	0.0345
Root mean square error	OMP	0.5722	0.3003	0.1803	0.1303	0.0994
	GPSR	0.1065	0.0698	0.0556	0.0481	0.0428
	the proposed algorithm	0.0710	0.0580	0.0519	0.0479	0.0447

The experimental results are shown in Figure 13. The reconstruction efficiency of the OMP algorithm is positively correlated with the sampling rate, while the reconstruction efficiency of the GPSR algorithm is negatively correlated with the sampling rate. The average reconstruction time of the proposed algorithm is relatively stable. The average reconstruction time of GPSR is much higher than that of the others. As a result, the proposed algorithm exhibited relatively high reconstructed efficiency.

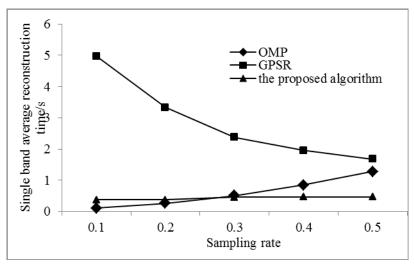


Figure 13. Comparison of average reconstructed time of all bands for different algorithms

5. Conclusions

There is a high spectral correlation for plant hyperspectral data. The joint sparse model based on spectral characteristics can not only improve the fidelity of reconstructed plant hyperspectral images, but also effectively reduce the reconstructed error in spectral domain more efficiently. Under a relatively low sampling rate (less than 0.2bpp), the PSNR of the proposed algorithm is 13.72dB higher than that of OMP algorithm, and is 1.66dB higher than that of GPSR algorithm. For the errors in the spectral domain, the average normalized root mean square error, the mean absolute percentage error and the mean absolute error of the proposed algorithm decrease by 35.38%, 31.83% and 33.33% than those of GPSR algorithm respectively. Besides, the proposed algorithm can achieve relatively high reconstructed efficiency.

More extensive research can be studied in the future. For example, new compressive sensing algorithm can be involved in denosing the plant hyperspectral data. High performance reconstructed method should be proposed to improve the reconstructed efficiency.

Acknowledgments: This work was supported by the National Nature Science Foundation of China (NSFC) under Grants Nos. 41671415 and 61205200, the National key foundation for exploring scientific instrument of

Peer-reviewed version available at Sensors 2017, 17, 1322; doi:10.3390/s17061322

13 of 13

China under Grants No 61427808 and Zhejiang public welfare Technology Application Research Project of China under Grants No 2016C32087. We are grateful to all the subjects for their participation in this study.

Author Contributions: Ping Xu conceived of and designed the study. Junfeng Liu performed the experiments and wrote the paper. Bo Qiu helped to revise the paper. Jingcheng Zhang and Lingyun Xue supervised the work

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Tang Yanlin, Huang Jingfeng. Study on Hyperspectral Remote Sensing in Agriculture. Remote Sensing Technology and Application, 2001, 16(4):248-251.
- 2. Nuarsa I W, Nishio F, Hongo C, et al. Using variance analysis of multitemporal MODIS images for rice field mapping in Bali Province, Indonesia. International Journal of Remote Sensing, 2012, 33(17): 5402-5417.
- 3. Bao Yansong, Wang Jihua, Liu Liangyun, et al. Approach to estimation of winter wheat nitrogen by using remote sensing technology on diverse scale and its application. Transactions of the Chinese Society of Agricultural Engineering, 2007, 23(2): 139-144.
- 4. Zhang Jingcheng, Li Jianyuan, Ynag Guijun, et al. Monitoring of Winter Wheat Stripe Rust Based on the Spectral Knowledge Base for TM Images . Spectroscopy and Spectral Analysis, 2010, 30(6): 1579-1585.
- 5. Zhang Haoyan,Li Zhongliang, Zou Tengfei,et al. Overview of Agriculture Big Data Research. Computer Science, 2014, 41(S2).
- 6. E.J.Candes, J.Romberg, T.Tao Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. on Information Theory, 2006, 52(2), pp. 489-509.
- 7. D.L.Donoho, Compressed sensing, IEEE Trans. on Information Theory, 2006, 52(4), pp. 1289-1306.
- 8. E.J.Candes, T.Tao, Near-optimal signal recovery from random projections: universal encoding strategies, IEEE Trans. on Information Theory, 2006,52(12), pp. 5406-5425.
- 9. Kang LW, Lu CS. Distributed compressive video sensing, IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2009:1169-1172.
- 10. Ly N H, Du Q, Fowler J E. Reconstruction From Random Projections of Hyperspectral Imagery With Spectral and Spatial Partitioning. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2013, 6(2):466-472.
- 11. Chen Y, Nasrabadi N M, Tran T D. Sparse Representation for Target Detection in Hyperspectral Imagery. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(3):629-640.
- 12. Wang Zhongliang, Feng Yan, Xiao Hua, ea al. Distributed compressive sensing imaging and reconstruction of hyperspectral imagery. Optics and Precision Engineering, 2015, 23(4):1131-1137.
- 13. Tropp J A, Gilbert A C. Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit. IEEE Transactions on Information Theory, 2007, 53(12):4655-4666.
- 14. Figueiredo M A T, Nowak R D, Wright S J. Gradient Projection for Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems. IEEE Journal of Selected Topics in Signal Processing, 2007, 1(4):586-597.
- 15. Kang L W, Lu C S. Distributed compressive video sensing. 2009, 7744: 1169-1172.
- 16. Gan L, Do T T, Tran T D. Fast compressive imaging using scrambled block Hadamard ensemble. Proc of Eusipco, 2012:1-5.
- 17. Tsaig Y, Donoho D L. Extensions of compressed sensing. Signal Processing, 2006, 86(3): 549-571.