

# A MODIFIED SUNSAL-TV ALGORITHM FOR HYPERSPSCTRAL UNMIXING BASED ON SPATIAL HOMOGENEITY ANALYSIS

Wang Yuqian<sup>1</sup>, Shao Zhenfeng<sup>1</sup>, Zhang Lei<sup>1</sup>, Zhou Weixun<sup>1</sup>

<sup>1</sup>State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China.

E-mail: neo@whu.edu.cn

**Abstract.** The sparse regression framework has been introduced by many works to solve the linear spectral unmixing problem due to the knowledge that a pixel is usually mixed by less endmembers compared with the endmembers in spectral libraries or the entire hyperspectral data sets. Traditional sparse unmixing techniques focus on analyzing the spectral properties of hyperspectral imagery without incorporating spatial information. But the integration of spatial information would be beneficial to promote the performance of the linear unmixing process. An algorithm called sparse unmixing via variable splitting augmented Lagrangian and total variation (SUnSAL-TV) adds a total variation spatial regularizer besides the sparsity-inducing regularizer to the final unmixing objective function. The total variation spatial regularization is helpful to promote the fractional abundance smoothness. However, the abundance smoothness varies in the image. In this paper, the spatial smoothness is estimated through homogeneity analysis. Then the spatial regularizer is weighted for each pixel by a homogeneity index. The modified algorithm, called homogeneity analysis based SUnSAL-TV (SUnSAL-TVH), integrates the spatial information with finer modelling of spatial smoothness and is supposed insensitive to the noise and more stable. Experiments on synthetic data sets are taken and indicate the validity of our algorithm.

## 1. Introduction

Hyperspectral imagery is of wide applications in many fields due to its high spectral resolution. But the pixels of hyperspectral data may be mixed by different substances because of low spatial resolution. The mixed pixel phenomenon impedes hyperspectral data's application and therefore becomes an essential issue in hyperspectral image data analysis.

A large number of techniques are raised for hyperspectral image unmixing<sup>[1]</sup> and recently a new approach called sparse regression (SR) based unmixing<sup>[2]</sup> is proposed and developed. SR based unmixing method is semi-supervised by assuming that the observed image signatures can be expressed in the form of linear combinations of a number of pure spectral signatures known in advance. The pure spectral signatures constitute so-called endmember spectral library. The SR based unmixing method does not assume the existence of pure spectral signatures, so it is useful for some complex scene where pixels are seriously mixed. On the other hand, spatial information is involved to many spectral analysis based unmixing methods, which has been proved improving the unmixing process in many cases. For example, total variation (TV) spatial regularization is introduced to a SR based hyperspectral unmixing algorithm called SUnSAL<sup>[2]</sup> as a way to incorporate spatial smoothness of



hyperspectral image data. The novel algorithm called SUnSAL-TV has been proved to improve unmixing performance<sup>[3]</sup>. SUnSAL-TV only uses the first-order neighbourhood systems for the integration of spatial information and weights the TV terms equally in spatial space. Since the smoothness of hyperspectral image may varies in spatial space, it is natural to weight the TV terms according to the variation of smoothness.

Spatial homogeneity index (HI) quantificationally estimates the spatial smoothness and has been introduced to the incorporation of spatial information for endmember extraction process<sup>[4]</sup>. In this paper, the TV regularization of SUnSAL-TV algorithm would be carefully studied and weighted according to HI, then the SUnSAL-TV algorithm would be modified based on the spatial homogeneity analysis.

## 2. Methodology

### 2.1. Sparse unmixing model and SUnSAL-TV algorithm

The linear mixture model assumed that the spectral response of a pixel was a linear combination of all the endmembers present in the pixel. The linear model could be written as follows:

$$\mathbf{y} = \mathbf{M}\mathbf{a} + \mathbf{n} \quad (1)$$

where  $\mathbf{y}$  was a vector of dimension  $L$  (the measured spectrum of the pixel),  $\mathbf{M}$  was an  $L \times q$  matrix containing  $q$  endmembers,  $\mathbf{a}$  was a vector of dimension  $q$  (the abundances corresponding to the endmember) and  $\mathbf{n}$  was the added noise of dimension  $L$ . Sparse unmixing reformulated (1) assuming the availability of a library of spectral signatures a priori as follows:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \quad (2)$$

where  $\mathbf{A} \in R^{L \times m}$  was the library of  $m$  endmembers and  $\mathbf{x}$  was the abundance column vector of dimension  $m$ . The linear unmixing problem was expressed in linear sparse regression form and studied carefully in [2]. An equivalent form of the sparse regression linear unmixing problem is formulated as

$$\min_x \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{x} \geq 0 \quad (3)$$

where  $\lambda$  was a regularization parameter which weighted the two terms of the objective functions.

SUnSAL was an effective sparse regression algorithm to solve (3)<sup>[2]</sup>. With a TV regularizer added to SUnSAL, SUnSAL-TV took into account the relationship between each pixel and its closest horizontal or vertical neighbors and solved the following optimization problem:

$$\min_x \frac{1}{2} \|\mathbf{A}\mathbf{X} - \mathbf{Y}\|_F^2 + \lambda \|\mathbf{X}\|_{1,1} + \lambda_{TV} \text{TV}(\mathbf{X}) \quad \text{subject to} \quad \mathbf{X} \geq 0 \quad (4)$$

where

$$\text{TV}(\mathbf{X}) \equiv \sum_{\{i,j\} \in \mathcal{E}} \|\mathbf{x}_i - \mathbf{x}_j\| \quad (5)$$

was the TV regularizer and  $\mathcal{E}$  denoted the set of horizontal and vertical neighbors in the image. To solve problem (4), the TV term was replaced by two linear operators  $\mathbf{H}_h$  and  $\mathbf{H}_v$ , which calculated horizontal and vertical differences respectively as follows:

$$\begin{cases} \mathbf{H}_h(\mathbf{x}_{i,j}) = (\mathbf{x}_{i,j} - \mathbf{x}_{i,j+1}) \\ \mathbf{H}_v(\mathbf{x}_{i,j}) = (\mathbf{x}_{i,j} - \mathbf{x}_{i+1,j}) \end{cases} \quad (6)$$

where  $\mathbf{x}_{i,j}$  was an abundance vector consisting  $\mathbf{X}$  spatially located at  $(i, j)$ .  $\mathbf{H}_h$  and  $\mathbf{H}_v$  were in periodic boundaries. Set  $\mathbf{H}(\mathbf{X}) = \begin{bmatrix} \mathbf{H}_h(\mathbf{X}) \\ \mathbf{H}_v(\mathbf{X}) \end{bmatrix}$ , then  $\text{TV}(\mathbf{X}) = \|\mathbf{H}(\mathbf{X})\|_{1,1}$ . With the above definitions in place, problem (4) could be reformulated equivalently as

$$\min_x \frac{1}{2} \|\mathbf{A}\mathbf{X} - \mathbf{y}\|_F^2 + \lambda \|\mathbf{X}\|_{1,1} + \lambda_{TV} \|\mathbf{H}(\mathbf{X})\|_{1,1} + \iota_{R^+}(\mathbf{X}) \quad (7)$$

where  $\iota_{R^+}(\mathbf{X}) = \sum_{i=1}^n \iota_{R^+}(\mathbf{x}_i)$  was the indicator function. Reference [3] would supply more details for the process of SUnSAL-TV algorithm.

## 2.2. SUnSAL-TVH algorithm

The TV regularization in SUnSAL-TV algorithm considered the smooth transitions of the fractional abundance equally in the whole image, but in fact the image could not maintain spatial smoothness equally. The spatial homogeneity analysis<sup>[4][5]</sup>, which had been derived a spatial homogeneity index  $hi$ , was a good way to survey the spatial smoothness. The higher  $hi$  meant the pixel was more similar to its neighbours, so the spatial transition was smoother. In [5]  $hi$  was estimated in a local window by the weighted average of the spectral similarity between the pixel and its neighbours. Here  $hi$  was a parameter in the range of  $[0, 1]$  and was calculated as formula (8):

$$hi(i, j) = \sum_{r=i-d}^{i+d} \sum_{c=j-d}^{j+d} \omega(r-i, c-j) \times S(r-i, c-j) \quad (8)$$

where  $(i, j)$  denoted the spatial location of a pixel,  $(r, c)$  was some neighbour in the window with radius  $d$ .  $\omega$  was a weight parameter such that  $\omega$  was large when the neighbour was close to the pixel otherwise  $\omega$  was small.  $S(r-i, c-j)$  was the inner product of the normalized spectral vectors of the pixels located in  $(i, j)$  and  $(r, c)$ .  $\omega(r-i, c-j)$  and  $S(r-i, c-j)$  were computed as follows:

$$\omega(r-i, c-j) = \beta^{-1} \times e^{-\frac{(r-i)^2 + (c-j)^2}{d^2}}, \beta = \sum_{r=i-d}^{i+d} \sum_{c=j-d}^{j+d} e^{-\frac{(r-i)^2 + (c-j)^2}{d^2}} \quad (9)$$

$$S(r-i, c-j) = \frac{\langle y(r, c), y(i, j) \rangle}{\|y(r, c)\| \|y(i, j)\|} \quad (10)$$

Although homogeneity analysis owned some stability to the size of the window  $d$  it would not be too small or too large. HI of homogeneous pixels would be effected seriously by noise if  $d$  was too small. On the other hand, HI of the transitional pixels would be higher than it really was if  $d$  was too large and the windows covered many homogeneous pixels.

When  $hi$  was close to 1, the pixel was very homogeneous in a local neighbourhood, which meant the image held a smoother transition here and we could hope a strong spatial TV regularizer in the sparse unmixing. Otherwise a weak TV regularizer was supposed. To achieve above situation, an algorithm called SUnSAL-TVH was proposed. In SUnSAL-TVH, the TV regularizers were modified by HI and the linear operators  $\mathbf{H}_h$  and  $\mathbf{H}_v$  turned to be  $\hat{\mathbf{H}}_h$  and  $\hat{\mathbf{H}}_v$  as follows:

$$\begin{cases} \hat{H}_h(\mathbf{x}_{i,j}) = \hat{h}_{i,j}(\mathbf{x}_{i,j} - \mathbf{x}_{i,j+1}) \\ \hat{H}_v(\mathbf{x}_{i,j}) = \hat{h}_{i,j}(\mathbf{x}_{i,j} - \mathbf{x}_{i+1,j}) \end{cases} \quad (11)$$

where  $\hat{h}_{i,j}$  was the weight parameter located at  $(i, j)$  induced by the homogeneity analysis. Since  $hi$  may be of very small variation, we did a linear transformation that let  $\hat{h}_i = \frac{M-1}{1-m}(hi - \eta)$ , where

$m = \min_{i,j}(hi_{i,j})$  was the minimum of  $hi$ ,  $M$  was a parameter should be set and  $\eta = \frac{M \times m - 1}{M - 1}$ . The

linear transformation mapped  $hi$  from range  $[m, 1]$  to  $[1, M]$ . When  $hi_{i,j} = 1$ , the pixel at  $(i, j)$  was of just the same spectral signature with its neighbours. Hence we should give it a very strong TV regularizer, concretely  $M$  times stronger as the weakest one. If  $(i, j+1)$  or  $(i+1, j)$  located out of the image scene, we set  $\hat{H}_h(\mathbf{x}_{i,j}) = 0$  and  $\hat{H}_v(\mathbf{x}_{i,j}) = 0$ . Then the remaining procedure of SUnSAL-TVH was very similar to SUnSAL-TV.

### 3. Experiments

Above algorithms were tested by simulated data generated as formula (2). We used a selection of 498 materials from the USGS spectral library denoted splib06<sup>1</sup> which comprised spectral signatures in 224 spectral bands. Then a sublibrary  $\mathbf{A}$  was extracted that the spectral angle between any two endmembers in the sublibrary was larger than  $4.44^\circ$ . Sublibrary  $\mathbf{A}$ , owning 240 endmembers, was set to be the apriori endmember library in the sparse regression. 7 endmembers were randomly chosen from  $\mathbf{A}$ , which are present in the mixed pixels. These endmembers were distributed in a  $50 \times 50$  pixels map. The map was divided into 4 homogeneous regions and 5 transitional regions just as in Figure 1. In the homogeneous region, pixels were of the same fraction abundance. There were two homogeneous regions containing only one endmember, i.e., the pure pixel regions, another homogeneous region containing 2 endmembers with the fraction abundances 0.3 and 0.7 respectively, the last homogeneous region containing 3 endmembers with the fraction abundances 0.2, 0.3 and 0.5 respectively. The rest part of the map were transitional regions which connected homogeneous regions. The pixels in transitional regions contained all the endmembers owned by the connected homogeneous regions. Transitional region 5 was mixed more complexly, which adjoined with all the homogeneous regions and contained all the endmembers. The fraction abundances of the endmembers in the transitional region pixels were in inverse ratio to the distance of the pixel and the respectively homogeneous regions which would accord with the real scene. The simulated data was constructed to cover scenes from the most pure parts to the most mixed parts (mixed by 7 endmembers), from material stable regions (homogeneous regions) to material varied regions (transitional regions). It was also worthwhile that abundance vectors varied in transitional region 1 and 2 horizontally, transitional region 3 and 4 vertically, while in transitional region 5 horizontally and vertically. Hence the simulated data would be very proper to test SUnSAL-TV algorithm and SUnSAL-TVH algorithm which represent the difference of the pixel with its horizontal and vertical neighbours. In hyperspectral applications, the noise  $n$  in formula (2) was mostly model errors dominated by low-pass components. Hence the noise was made by low-pass filtering samples of zero-mean i.i.d. Gaussian sequences of random variables just as in reference [2] and [3]. The noise was estimated in the signal-to-noise ratio (SNR) in dB as

<sup>1</sup> <http://speclab.cr.usgs.gov/spectral.lib06>

$$SNR = 10 \log_{10} \left( \frac{E \left[ \|\mathbf{Ax}\|_2^2 \right]}{E \left[ \|\mathbf{n}\|_2^2 \right]} \right) \quad (12)$$

In the experiment, the simulated data was disturbed by noise in three levels, i.e., 20dB, 30dB and 40dB.

Homogeneous Region 1: Endmember 1	Transitional Region 1	Homogeneous Region 2: Endmember 3 and 4
Transitional Region 3	Transitional Region 5	Transitional Region 4
Homogeneous Region 3: Endmember 5, 6 and 7	Transitional Region 2	Homogeneous Region 4: Endmember 2

**Figure 1.** Homogeneous and transitional region pattern and endmember distribution map.

The size parameter  $d$  of the window in the homogeneity analysis was set to be 2, which was thought to be proper<sup>[5]</sup>. The weight parameter  $\lambda$  for the sparsity regularization and  $\lambda_{TV}$  for the TV regularization was set according to the noise level and algorithm just as the experiment analysis in [3], where the weight parameters were studied with sufficient experiment and supposed to be the best choice. The homogeneity scaling weight parameter  $M$  was set to be 10 in the experiment. The most popular performance discriminators in sparse unmixing were root mean square error (RSME) and signal-to-reconstruction error (SRE)<sup>[2][3]</sup>. RSME and SRE (measured in dB) were defined as:

$$RMSE = E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|_2 \right], SRE = 10 \log_{10} \left( \frac{E \left[ \|\mathbf{x}\|_2^2 \right]}{E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 \right]} \right) \quad (13)$$

where  $\hat{\mathbf{x}}$  was the estimated fractional abundance vector by the unmixing algorithms. The smaller RMSE meant  $\hat{\mathbf{x}}$  (the estimated fractional abundance) was closer to  $\mathbf{x}$  (the real fractional abundance), which indicated better unmixing performance. SRE gave more information regarding the power of the error in relation with the power of the signal. It was just opposite to RSME that the higher the SRE, the better the unmixing performance.

The compared experiment tested the performance of algorithms SUnSAL, SUnSAL-TV and SUnSAL-TVH with the result shown in Table 1. Two conclusions followed: (a) the performance was promoted remarkably from SUnSAL-TV to SUnSAL-TVH; (b) the performance of SUnSAL-TVH was affected by the noise as well as SUnSAL and SUnSAL-TV, but when the noise was under some acceptable level, i.e., 30dB, the affection was very small (the difference between performances in 30dB and 40dB was very small). Furthermore, we tested the algorithms performance in homogeneous regions and transitional regions respectively. Since transitional region 5 was more complex than other transitional regions, it was separately evaluated. We only showed the result in Table 2, when data was affected seriously by noise, i.e., noise was in 20dB. *RSME* was computed for the whole image in Table 1 but as an average for each pixel in Table 2, so they were in different scale. But *SRE* was in the same scale in Table 1 and 2 because it was a ratio. From Table 2, besides the above conclusions (a) and (b), we could also get (c): performance in homogeneous regions was better than transitional regions, the

more complex variation in transitional region, the worse performance. Conclusion (c) encouraged us to consider the homogeneity analysis in the process of hyperspectral image unmixing.

**Table 1.** Performance of algorithms SUnSAL, SUnSAL-TV and SUnSAL-TVH

	SUnSAL		SUnSAL-TV		SUnSAL-TVH	
	RSME	SRE	RSME	SRE	RSME	SRE
20dB	8.4275	13.5335	2.4242	24.3561	1.8980	26.4814
30dB	9.8052	12.2184	1.3551	29.4078	0.9774	32.2458
40dB	10.3867	11.7179	1.1	31.2196	0.8964	32.9972

**Table 2.** Performance of algorithms on homogeneous and transitional regions

	SUnSAL		SUnSAL-TV		SUnSAL-TVH	
	RSME	SRE	RSME	SRE	RSME	SRE
Homogeneous region	0.0927	15.0399	0.0305	25.9564	0.0228	28.7905
Transitional region 1, 2, 3, 4	0.1338	11.7692	0.04	22.9913	0.0321	25.1211
Transitional region 5	0.1643	9.5044	0.0642	17.9781	0.0634	18.1169

#### 4. Conclusion

The modified algorithm SUnSAL-TVH proposed in this paper introduced the homogeneity analysis to represent the connection between the pixel and its neighbourhood. The experiment result verified that the integration of spatial analysis would promote the sparse unmixing performance. SUnSAL-TVH owned a more accurate estimation of abundance spatial smoothness and got better performance than SUnSAL-TV. More specifically, SUnSAL-TVH was very stable to noise when the noise was at an acceptable level. We also found that the unmixing algorithms performance relied on homogeneity, so homogeneity analysis is very valuable in the consideration of integration of spatial information in hyperspectral image unmixing process.

#### Acknowledgments

This work is sponsored by National Basic Research Program of China (No.2010CB731801), National Natural Science Foundation of China (No.61172174 & No.10978003) and National Science & Technology Specific Projects of China (No.2012YQ16018505 & No.2013BAH42F03).

#### References

- [1] Bioucas-dias J M, Plaza A and Dobigeon N et al. 2012 Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches *J. IEEE J-STARS*. **5(2)** 354-79
- [2] Iordache M D and Bioucas-dias J M 2011 Sparse unmixing of hyperspectral data *J. IEEE Trans. Geosc. Remote Sensing* **49(6)** 2014-39
- [3] Iordache M D, Bioucas-dias J M, and Plaza A 2012 Total variation spatial regularization for sparse hyperspectral unmixing *J. IEEE Trans. Geosc. Remote Sensing* **50(11Part 1)** 4484-502
- [4] Martin G and Plaza A 2012 Spatial-spectral preprocessing prior to endmember identification and unmixing of remotely sensed hyperspectral data *J. IEEE J-STARS*. **5(2SI)** 380-95
- [5] Kong X B 2012 Hyperspectral unmixing based on homogeneous region analysis *D. Wuhan University*