NATIONAL INSTITUTE OF SCIENCE EDUCATION AND RESEARCH

DENOISING OF HYPERSPECTRAL IMAGE USING PARAFAC DECOMPOSITION

Submitted in Partial Fulfilment of the Requirements

for P452

by

Nelson Kshetrimayum

Under the guidance of

Prof. Guneshwar Thangjam *&*

Prof. Subhasish Basak

School of Earth and Planetary Sciences

School of Physical Sciences

2022

Abstract

In hyperspectral imagery noise are ingrained in the data taken by our instruments which arise from different reasons such as instrumental noise, thermal noise, etc. In this project, PARAFAC model has been used to denoise the HSI of the lunar surface taken by Chandrayaan-2 from the IIRS mission. In the present there are lack of well processed denoised data for the IIRS data. Here it has been demonstrated that the PARAFAC model has been able to generate much improved images of a cropped image (100×100 pixels over 250 spectral bands) of a sample data. Further it has proposed to run the model for the whole image which is would need more time because of the large data, approximately 10000×250 pixels over 150 spectral bands.

1 Introduction

1.1 Hyperspectral Image

A HYPERSPECTRAL image (HSI) is a multidimensional array also named as a tensor, and it normally consists of hundreds of spectral bands which provide much useful information in geography, agriculture, and military.

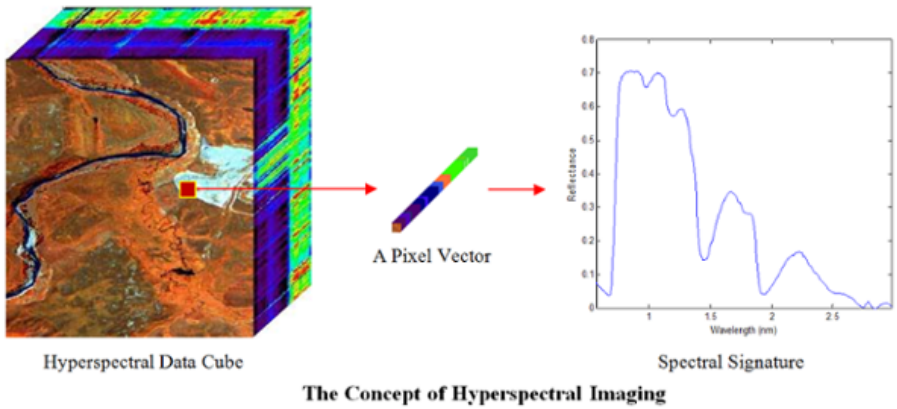

Figure 1: An image that describes the idea of HSI (Source: udayton.edu)

However, some radiometric noise, due to sensor, photon effects, and calibration error degrades classification results[1].

1.2 Noise Reduction in HSI

Noise reduction is a preprocessing step for hyperspectral imagery. Noises are ingrained in the data taken by the sensors from various causes, for example, instrumental noise, thermal noise, etc. With time the approaches of hyperspectral noise reduction techniques have evolved substantially. Various approaches have been taken in this problem. To name a few mainstream approaches [2],

1. Two dimensional and three dimensional bandwise techniques
2. 3D model-based and 3D filtering approaches
3. Spectral and spatial-spectral penalty-based approaches

4. Low-rank model-based approaches

PARAFAC model is one of the low-rank model-based approaches. Other models like Tucker3 decomposition are also there. In [1], they have shown that the PARAFAC model outperforms Tucker3.

2 PARAFAC Model

2.1 PARAFAC Decomposition

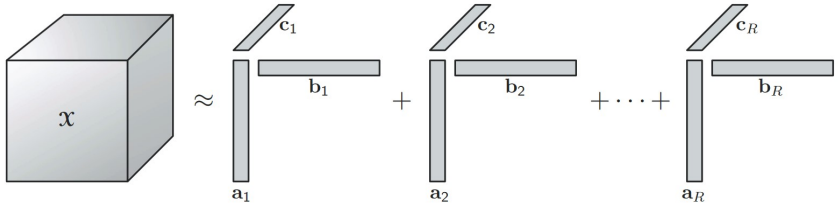

Figure 2: PARAFAC decomposition of a three-way array (Source: [3])

The PARAFAC model factorizes a tensor into a sum of rank-1 tensors [3]. A tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ can be expressed as

$$\mathcal{X} \approx \hat{\mathcal{X}} = \sum_{r=1}^{R} \mathcal{X}_r = \sum_{r=1}^{R} \lambda_r \boldsymbol{a}_r^{(1)} \circ \boldsymbol{a}_r^{(2)} \circ \boldsymbol{a}_r^{(3)} \tag{1}$$

where $\hat{\mathcal{X}}$ is the rank- R (Kruskal rank) approximation of $\mathcal{X}$; $\mathcal{X}_r \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is rank-1 tensor; $\boldsymbol{a}_r^{(1)}, \boldsymbol{a}_r^{(2)}, \boldsymbol{a}_r^{(3)} \in \mathbb{R}^{I_n}$ are normalized unite-norm vectors of the n-mode space of $\mathcal{X}$ and $\lambda_r = \left\|\boldsymbol{a}_r^{(1)}\right\| \cdot \left\|\boldsymbol{a}_r^{(2)}\right\| \cdot \left\|\boldsymbol{a}_r^{(3)}\right\|, r = 1, 2, \ldots, R$, which we will refer to as weight.

Elementwise, equation 1 is written as

$$\hat{x}_{ijk} = \sum_{r=1}^{R} \lambda_r a_{ir}^{(1)} a_{jr}^{(2)} a_{kr}^{(3)}, \quad \text{for} \quad i = 1, \ldots, I_1, \quad j = 1, \ldots, I_2, \quad k = 1, \ldots, I_3. \tag{2}$$

PARAFAC decomposition is used to compute $\hat{\mathcal{X}}$ with R components that approximate the best value of $\mathcal{X}$ by minimizing the square error $e = \|\mathcal{X} - \hat{\mathcal{X}}\|^2$. Using Khatri-Rao product, the n mode unfolding matrix of $\mathcal{X}$ is given by [3]

$$\begin{aligned}
\hat{\mathbf{X}}_1 &= \mathbf{A}^{(1)}\mathbf{\Lambda}\left(\mathbf{A}^{(3)} \odot \mathbf{A}^{(2)}\right)^T \\
\hat{\mathbf{X}}_2 &= \mathbf{A}^{(2)}\mathbf{\Lambda}\left(\mathbf{A}^{(3)} \odot \mathbf{A}^{(1)}\right)^T \\
\hat{\mathbf{X}}_3 &= \mathbf{A}^{(3)}\mathbf{\Lambda}\left(\mathbf{A}^{(2)} \odot \mathbf{A}^{(1)}\right)^T
\end{aligned} \tag{3}$$

where $\mathbf{A}^{(n)} = \left[\boldsymbol{a}_1^{(n)}, \ldots, \boldsymbol{a}_R^{(n)}\right] \quad (n = 1, 2, 3), \mathbf{\Lambda} = \text{diag}\,(\lambda_1, \ldots, , \lambda_R)$. Thus, minimizing $e = \|\mathcal{X} - \hat{\mathcal{X}}\|^2$ is transformed to find $\mathbf{A}^{(n)}$ where

$$\mathbf{A}^{(n)} = \text{argmin}(e) = \text{argmin}(\|\hat{\mathbf{X}}_n - \mathbf{X}_n\|^2) \tag{4}$$

2.2 PARAFAC ALS algorithm

To estimate $\mathbf{A}^{(n)}$, a "PARAFAC ALS algorithm" [3] is adapted.

1: Set k=0, randomly initialize $\mathbf{A}^{(n),0}$, n=1,2,3 or initialize $\mathbf{A}^{(n),0} = R$ leading left singular vectors of $\mathbf{X}_n$.
2: Let $e_0 = 0$.
3: a) Estimate $\mathbf{A}^{(n),k+1}$:
4: **for** $i = 1$ to 3 **do**
5: $\quad \mathbf{U}^{(i),k+1} = \mathbf{A}^{(p),k} \odot \mathbf{A}^{(q),k}$ ▷ for i=1,2,3; (p,q) = (3,2),(3,1),(2,1) respectively
6: $\quad \mathbf{A}^{(i),k+1} = \mathbf{X}_1\mathbf{U}^{(i),k+1}\left(\mathbf{U}^{(i),k+11^T}\mathbf{U}^{(i),k+1}\right)^{\sharp}$
7: $\quad \lambda_r^{(i),k+1} = \left\|\boldsymbol{a}_r^{(i),k+1}\right\|, \boldsymbol{a}_r^{(i),k+1} = \boldsymbol{a}_r^{(i),k+1}/\lambda_r^{(i),k+1}, \quad r = 1, \ldots, R.$
8: **end for**
9: b) Compute $\lambda_r^{k+1} = \lambda_r^{(1),k+1}\lambda_r^{(2),k+1}\lambda_r^{(3),k+1}, \quad r = 1, \ldots, R,$
10: $\quad \hat{\mathbf{X}}_3^{k+1} = \mathbf{A}^{(3),k+1}\mathbf{\Lambda}^{k+1}\mathbf{U}^{(3),k+1}T,$
11: $\quad \mathbf{\Lambda}^{k+1} = \text{diag}\left(\lambda_1^{k+1}, , \ldots, \lambda_R^{k+1}\right).$
12: c) $e_{k+1} = \left\|\mathbf{X}_3 - \hat{\mathbf{X}}_3^{k+1}\right\|^2, \quad$ if $|e_{k+1} - e_k| > \varepsilon$ and k is less than the maximum
13: $\quad$ number of iteration, $k = k + 1$ and go back to step 3(a).
14: **return** $\mathbf{A}^{(1)} = \mathbf{A}^{(1),k+1}, \mathbf{A}^{(2)} = \mathbf{A}^{(2),k+1}, \mathbf{A}^{(3)} = \mathbf{A}^{(3),k+1}$ and $\lambda_r = \lambda_r^{k+1}, r = 1, \ldots, R$

---

3 Results

The task we have in hand is to denoise the data of the lunar surface from IIRS mission. One sample of these data have approximately 10000×250 pixels on the spatial plane. This along with 250 spectral bands for each pixel, the order-3 data we have for a sample is of size around $10000 \times 250 \times 250$. Running the algorithm with data of this scale for a high rank is computationally expensive and with the systems available to us now plus considering the very short time we have worked on this project we could not generate the result of the whole data. To demonstrate that our method works, we ran with a much smaller cropped data of size $100 \times 100 \times 250$. The results are as following:

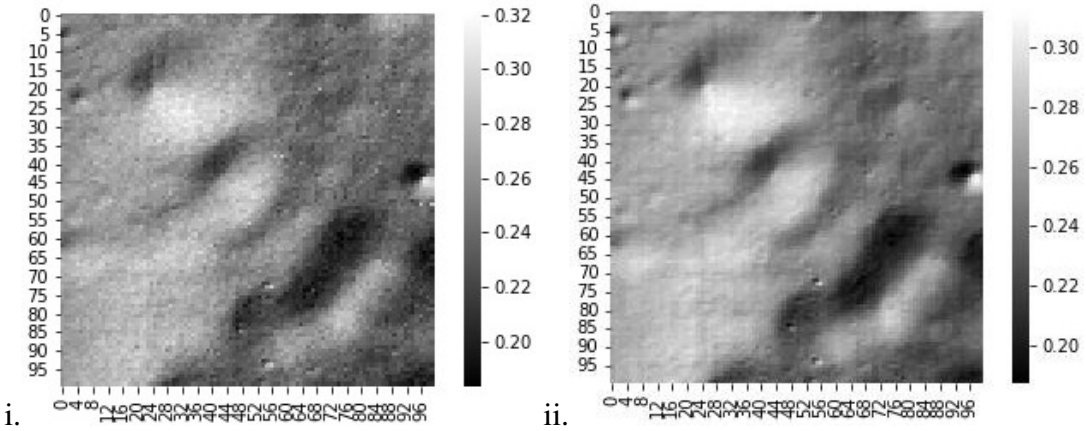

Figure 3: i. Cropped image from IIRS archive, ii. Image after denoising (213 rank)

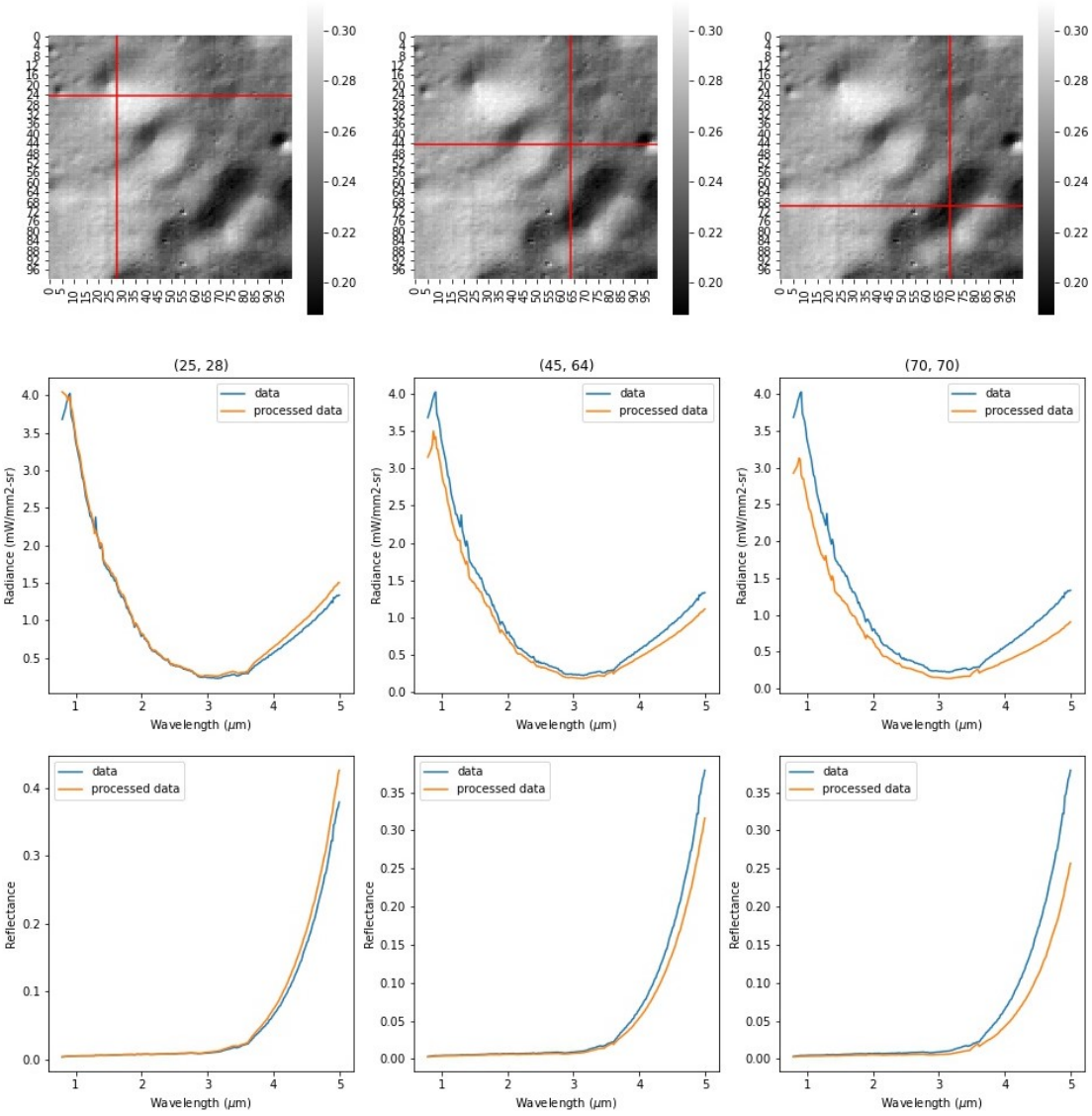

Figure 4: Radiance and reflectance plots; (i) a light spot (leftmost), (ii) medium lighted spot (center) and (iii) a dark spot (rightmost)

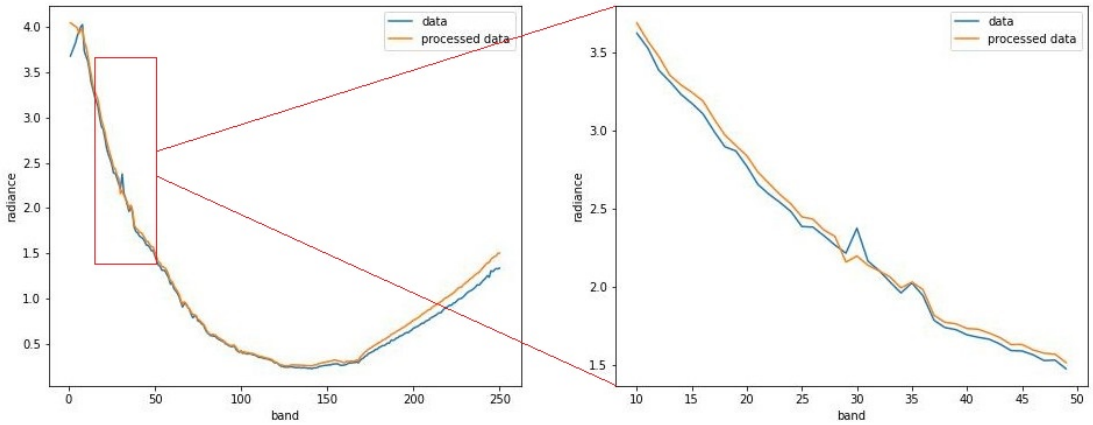

Figure 5: Enlarged image of Radiance vs band

4 Conclusion and Future works

From fig. 3 and fig. 5 it is clear that the image is smoothed up to a certain degree. Thus we have successfully demonstrated that our model is working as per our expectation and towards our objective. There are a lots of work to be done to further extend this result to the full scale image. The works that need to be done are

1. Figure out a way to find the optimal rank from the generated images to get the best image out. One way to go is use Cramer-Rao lower bound as an evaluation tool as done in [1]. Or we can find other or develop our own approach to find the optimal rank.

2. Run the algorithm for a full image and check the results.

3. If we successfully get appreciable results then we can denoise for the whole data available for from the IIRS archive and obtain a database of the obtained improved images.

References

[1] Xuefeng Liu et al. “Denoising of Hyperspectral Images Using the PARAFAC Model and Statistical Performance Analysis”. In: vol. 50. IEEE TRANSACTIONS ON GEOSCIENCE and REMOTE SENSING, Oct. 2012.

[2] Behnood Rasti et al. “Noise Reduction in Hyperspectral Imagery: Overview and Application”. In: doi:10.3390/rs10030482. MDPI, Remote Sensing, Mar. 2018.

[3] Tamara G. Kolda and Brett W. Bader. “Tensor Decompositions and Applications”. In: vol. 51, No. 3. Society for Industrial and Applied Mathematics, 2009, pp. 455–500.