DENOISING CHANDRAYAAN-2 IMAGES

USING PARAFAC DECOMPOSITION

Nelson Kshetrimayum

National Institute of Science Education and Research

A HYPERSPECTRAL image (HSI) is a multidimensional array. The ones we will be dealing with are 3-dimensional arrays, or tensor of order 3.

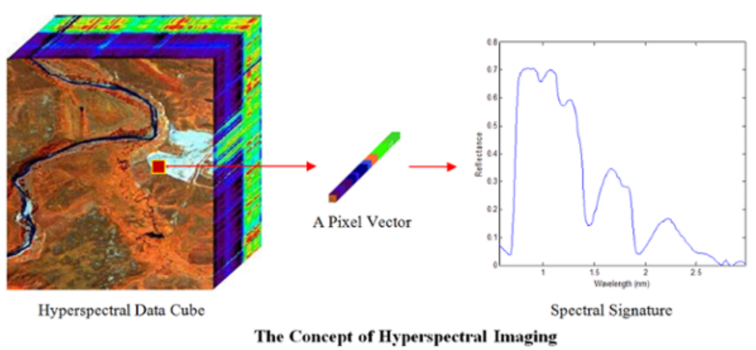


Figure: An image that describes the idea of HSI (Source: udayton.edu)

Noise Reduction in HSI

To name a few mainstream approaches ,

- Two dimensional and three dimensional bandwise techniques
- 3D model-based and 3D filtering approaches
- Spectral and spatial-spectral penalty-based approaches
- **Low-rank model-based approaches**

Tucker3 and PARAFAC are the main ones.

In our main reference paper by Xuefeng Liu et al., they have shown that the PARAFAC model outperforms Tucker3 on AVIRIS data .

PARAFAC Decomposition

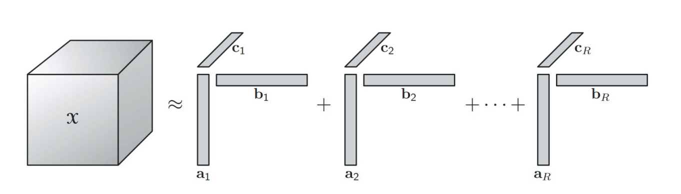


Figure: PARAFAC decomposition of a three-way array (Source: Tamara G.)

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

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

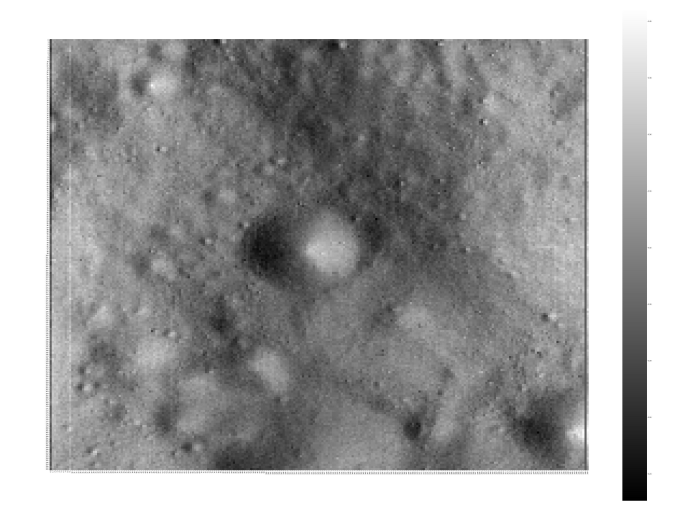
where $\hat{\mathcal{X}}$ is the rank- R (Kruskal rank) approximation of $\mathcal{X}$; $\mathcal{X}_r \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is rank-1 tensor; $\boldsymbol{a}_r^{(1)}, \boldsymbol{a}_r^{(2)}, \boldsymbol{a}_r^{(3)} \in \mathbb{R}^{I_n}$ are normalized unite-norm vectors and λ_r is the weight.

PARAFAC Decomposition

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

- Tensorly python library's PARAFAC ALS Algorithm has been used for the decomposition of the tensor.
- **Return** $A^{(1)}$, $A^{(2)}$, $A^{(3)}$ and Λ.
- The image is reconstructed by element-wise multiplication.

Runtime problem



200×250 pixels

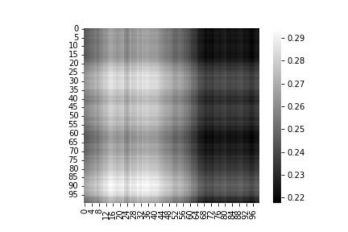
Runtime problem

The image shown is a cropped image from a sample image of 10756×250 pixels with 256 spectral bands.

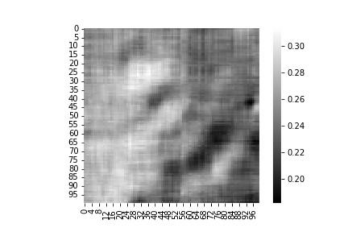
- Decomposition of this large images take a huge time and processing power
- When we run for R around 20, it takes a day.

So for the scope of this term paper, to demonstrate this method works, we have taken a cropped sample of 100×100 pixels and produce the results (in the following slides).

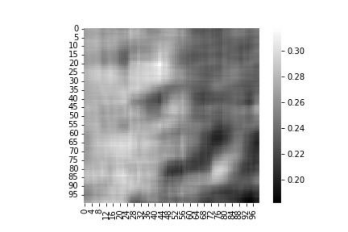
Results



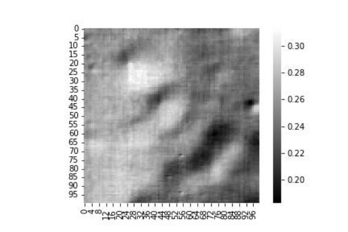
R = 1



R = 40

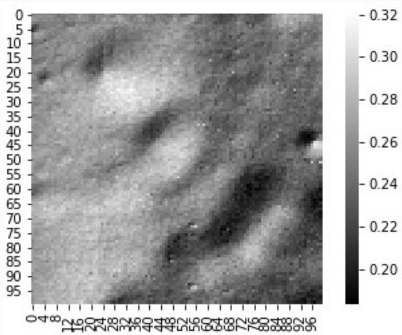


R = 10

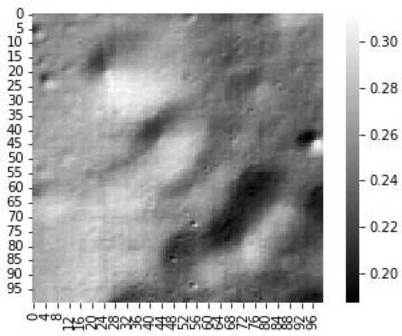


R = 100

Results



i.



ii.

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

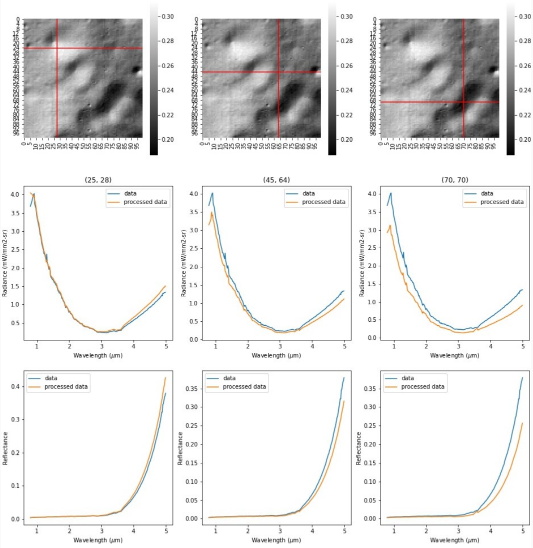


Figure: Radiance and reflectance plots; (i) a light spot (leftmost), (ii) medium lighted spot (center) and (iii) a dark spot (rightmost)

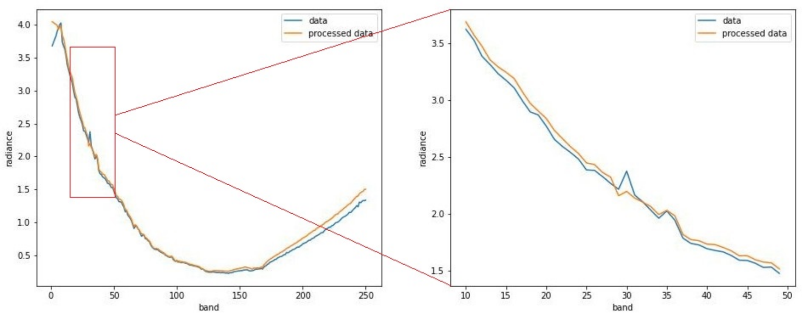


Figure: Enlarged image of Radiance vs band

Needs to be done: Statistical Performance Analysis

- We are searching of the best decomposition that gives us out data without loosing features in the images but also the best non-inclusion of the noise.
- The optimal rank will be the best tadeoff between this two.
- We need to 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 by Xuefeng Liu et al. or we can find other methods.

Summary

From the results shown 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. This has been achieve for a very smaller cropped data then the whole image. In future we need to

1. 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 [2].
2. Run the algorithm for a full image and check the results.
3. Then for all the data of IIRS mission.

References I

[1] B. R. et al.
Noise Reduction in Hyperspectral Imagery: Overview and Application.
MDPI, Remote Sensing, 3 2018.
doi:10.3390/rs10030482.

[2] X. L. et al.
***Denoising of Hyperspectral Images Using the PARAFAC Model and Statistical Performance Analysis*, volume 50.**
IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 10 2012.

[3] T. G. Kolda and B. W. Bader.
***Tensor Decompositions and Applications*, volume 51, No. 3, pages 455–500.**
Society for Industrial and Applied Mathematics, 2009.

Additional Slides: n-modes unfolding

The concept is easier to understand using an example. Let the frontal slices of $\mathcal{X} \in \mathbb{R}^{3\times4\times2}$ be

$$\mathbf{X}_1 = \begin{bmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} 13 & 16 & 19 & 22 \\ 14 & 17 & 20 & 23 \\ 15 & 18 & 21 & 24 \end{bmatrix}$$

Then the three mode- *n* unfoldings are

$$\mathbf{X}_{(1)} = \begin{bmatrix} 1 & 4 & 7 & 10 & 13 & 16 & 19 & 22 \\ 2 & 5 & 8 & 11 & 14 & 17 & 20 & 23 \\ 3 & 6 & 9 & 12 & 15 & 18 & 21 & 24 \end{bmatrix}$$

$$\mathbf{X}_{(2)} = \begin{bmatrix} 1 & 2 & 3 & 13 & 14 & 15 \\ 4 & 5 & 6 & 16 & 17 & 18 \\ 7 & 8 & 9 & 19 & 20 & 21 \\ 10 & 11 & 12 & 22 & 23 & 24 \end{bmatrix},$$

$$\mathbf{X}_{(3)} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & \cdots & 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 & 17 & \cdots & 21 & 22 & 23 & 24 \end{bmatrix}$$

Additional Slides: How PARAFAC decomposition is done

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]

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

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

$$
\mathbf{A}^{(n)} = \mathrm{argmin}(e) = \mathrm{argmin}(\|\hat{\mathbf{X}}_n - \mathbf{X}_n\|^2) \tag{3}
$$