COMPARISON OF VOLUME AND DISTANCE CONSTRAINT ON HYPERSPECTRAL UNMIXING

Zeheng Li[1] and Fuxiang wang[2] and Xiurui Geng[3]

[1]School of Electronic and Information Engineering, Beihang University,
Beijing Laboratory for General Aviation Technology,
Beijing Key Laboratory for Network-based Cooperative Air Traffic
Management, Beijing, 100191, P.R.China
[2]School of Electronic and Information Engineering, Beihang University,
Beijing Laboratory for General Aviation Technology,
Beijing Key Laboratory for Network-based Cooperative Air Traffic
Management, Beijing, 100191, P.R.China

ABSTRACT

Algorithms based on minimum volume constraint or sum of squared distances constraint is widely used in Hyperspectral image unmixing. However, there are few works about performing comparison between these two algorithms. In this paper, comparison analysis between two algorithms is presented to evaluate the performance of two constraints under different situations. Comparison is implemented from the following three aspects: flatness of simplex, initialization effects and robustness to noise. The analysis can provide a guideline on which constraint should be adopted under certain specific tasks.

KEYWORDS

Hyperspectral unmixing, volume constraint, distance constraint, relative capacity

1. INTRODUCTION

Characterized as extremely high spectral resolution and numerous narrow continuous bands, hyperspectral remote sensing has raised extensive concerns [1] [18]. Hyperspectral images usually consist of mixed pixels due to limited spatial resolution of sensors. Thus, hyperspectral unmixing whose purpose is to decompose the mixed pixels into material signature (endmembers) and the corresponding abundance fractions, has become a challenging task.

Hyperspectral data unmixing is commonly based on linear mixed model (LMM) [2]. LMM hypothesizes that each pixel vector can be represented as the product of the endmember matrix and abundance vector. The abundance vector need to satisfy nonnegative and sum-to-one constraints. Based on LMM, there are mainly two groups of methods solving hyperspectral unmixing problem. The algorithms in the first group require existing pure signatures in hyperspectral image, such as Pixel Purity Index (PPI) [3], N-FINDR [4], Vertex Component Analysis (VCA) [5], iterative error analysis (IEA) [6]. The algorithms in the second group can process image without requirement of pure signatures, such as single individual evolutionary strategy (SIE) [7], nonnegative matrix factorization (NMF) [8], minimum volume constraint NMF

(MVC-NMF) [10], iterative constrained endmember method (ICE) [11], robust nonnegative matrix factorization [19], minimum volume simplex analysis(MVSA) [20].

2. RELATED WORK

Since there is no requirement on pure signatures for the algorithms in the second group, they have been widely used in hyperspectral unmixing [9] [10] [11]. Among them, NMF with minimum volume constraint (VC) [10] considers simplex volume composed of the unknown endmembers during endmember extraction. VC adopts the commonly used formula to measure the 'volume' of simplex enclosed by endmembers. Though VC can efficiently restrain simplex volume, it often involves massive computation. On the other hand, ICE [11] imposes sum of squared distances constraint (SSD) on the original objective function, which can generally achieve satisfied result.

Though algorithms based on VC and SSD have been proposed for several years, there hasn't been any further analysis and comparison of these two methods. Thus, this paper mainly gives a detailed comparison of these two methods under various situations. The comparison analysis aims to give an instruction on how to choose the constraint under certain situation. To achieve this, the comparison includes flatness of simplex analysis, initialization analysis and robustness to noise analysis.

3. HYPERSPECTRAL UNMIXING ALGORITHM

In this section, we'll briefly introduce linear mixing model and unmixing algorithms based on VC and SSD.

3.1. Linear Mixture Model (LMM)

LMM [2], [12] assumes that the hyperspectral data is a linear combination of endmember spectra, with the weights being proportions. Mathematically, the model is given as:

$$x = As + \varepsilon,$$
$$\text{subjected to:} \quad s_k \geq 0, k = 1, \cdots, M, \sum_{k=1}^{M} s_k = 1, \tag{1}$$

where, x is L-dimensional vector (L is the number of bands) which is one of the pixel in image, s denotes corresponding abundance, ε represents possible errors. In real data processing, abundance should satisfy two constraints called as *nonnegative* constraint and *sum-to-one* constraint, as shown in (1).

The matrix involving all pixels in image is shown as Equation(2)

$$X = AS + \Theta, \tag{2}$$

where $X = [x_1, x_2, \cdots x_N]$ represents hyperspectral data which is assumed to be composed with material signatures $A = [a_1, a_2, \cdots a_M]$ and abundance fractions $S = [s_1, s_2, \cdots s_N]$. N is the number of pixel. M is the number of endmembers. Θ is the error matrix.

3.2. Constraints on Endmembers

For real data, there are substantial local minimum problem due to non-convexity of unconstrained objection function. As shown in

Figure 1, the red polyline indicates real simplex for hyperspectral data. Meanwhile, blue and black polyline is the solution obtained with different initial value. Since there must be corresponding abundance if endmembers enclose all scatters, it is necessary to put certain constraint on endmembers, as shown in (3):

$$f(\boldsymbol{A}, \boldsymbol{S}) = \frac{1}{2} \| \boldsymbol{X} - \boldsymbol{AS} \|_F^2 + \lambda J(\boldsymbol{A}) \tag{3}$$

where, $J(\boldsymbol{A})$ is the constraint added to endmember, λ is regularization factor to tradeoff the reconstruction error and constraint.

Before minimizing (3), several preprocessing steps will be taken to remove noises and reduce the dimension of original data, which aims to reduce computation complexity. Then, appropriate optimal strategy is used to minimize (3) and update $\boldsymbol{A}$ and $\boldsymbol{S}$ iteratively: first, given endmember matrix $\boldsymbol{A}$, calculate the abundance matrix $\boldsymbol{S}$ by the optimal strategy. Then update $\boldsymbol{A}$ by fixed $\boldsymbol{S}$ in the same way. After several iterations, (3) will approach its minimum value with $\boldsymbol{A}$ and $\boldsymbol{S}$ well-decided.
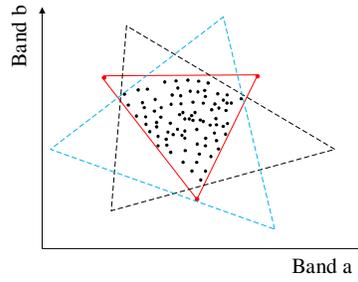
Figure 1. Endmember extraction by unconstraint algorithms. There exist many local minimum solutions due to the non-convex property. Although many solutions have relatively minor linear square error, the obtained points are still far from scattering, which cannot be regard as appropriate endmembers.
The volume and distance constraints are briefed as follows:

1. Minimum Volume Constraint

VC minimize volume [15] of simplex in its model. In VC, the expression of $J_V(\boldsymbol{A})$ is as for those.

$$J_V(\boldsymbol{A}) = \frac{1}{2(\boldsymbol{M}-1)!} \det^2 \left(\begin{bmatrix} 1_M^T \\ \boldsymbol{A} \end{bmatrix} \right) \tag{4}$$

After adding volume constraint, volume of simplex will be compressed as small as possible. Meanwhile, the hyperspectral data reconstructed by the extracted endmembers and corresponding abundance matrix can also close to real data. Therefore, we can get relatively accurate solution for endmember and abundance.

2. Sum of Squared Distance Constraint

In ICE [11], the constraint which minimizes SSD among several endmembers on hyperplane is adopted and given as equation (5). Like VC, SSD can also efficiently control the shape of simplex during the iteration by minimizing the distance between any two endmembers.

$$J_D(\boldsymbol{A}) = \sum_{k=1}^{M-1} \sum_{l=k+1}^{M} (\boldsymbol{a}_k - \boldsymbol{a}_l)^T (\boldsymbol{a}_k - \boldsymbol{a}_l) \ , \tag{5}$$

Where, M is the number of endmembers, $\boldsymbol{a}_k$ and $\boldsymbol{a}_l$ are any two endmember vectors.

In equation (3), the first term intends to decrease reconstruction error, and the second term is used to limit the overall 'volume' of simplex contrasted by endmembers. During the optimization process, we can control the tradeoff between spectral reconstruction accuracy and the distance/volume constraint $J(A)$ via λ .

Furthermore, since abundance must satisfy sum-to-one constraints, we can adjust equation (5) by adding $\mathbf{1}_M$ and $\mathbf{1}_N$ to endmember matrix A and original hyperspectral matrix X respectively. To control the influence of sum-to-one constrains, we introduce a regulation factor α to $\mathbf{1}_M$ and $\mathbf{1}_N$, as shown in Equation (6):

$$A \leftarrow \begin{bmatrix} \alpha I_M^T \\ A \end{bmatrix} \quad X \leftarrow \begin{bmatrix} \alpha I_N^T \\ X \end{bmatrix} \tag{6}$$

4. SIMPLEX PATTERN AND PARAMETER SELECTION

Since algorithm's performance fluctuated significantly with the variation of simplex pattern, analysis of simplex pattern should be significant. Two constraints mentioned above differ in most situations. It's necessary to give a comparative analysis to decide which constraint is more operative given certain simplex pattern.

4.1. Analysis of Simplex Pattern

As both VC and SSD can restrict simplex of endmember closing to original hyperspectral scattering, it is necessary to analyze the equivalence for these two constraints. We mainly do analysis in following two situations: First, for regular or quasi-regular simplex, there is a one-to-one correspondence between volume and distance for simplex, as show in Figure 2(a). In this situation, the VC and SSD have a similar performance.

However, when the simplex is not regular, namely non-regular simplex, the relationship between volume and distance of simplex is indefinite, as shown in Figure 2(b). In this case, it is necessary to analysis which one is better. Since in real hyperspectral data set, simplex for endmembers are not always regular, analysis for non-regular case is very important. For convenience, we define η as shown in (7) to measure degree of flatness in simplex.

$$\eta = \frac{\theta}{\pi} \tag{7}$$

where, θ represents the maximum generalized angle of simplex.

Because of the inequivalence in non-regular simplex, flatness analysis is given to identify the performance for each constraint. Additionally, to further differentiate VC and SSD, random initialization and anti-noise analysis are implemented to compare algorithm performance.

(a) regular simplex

(b) non-regular simplex
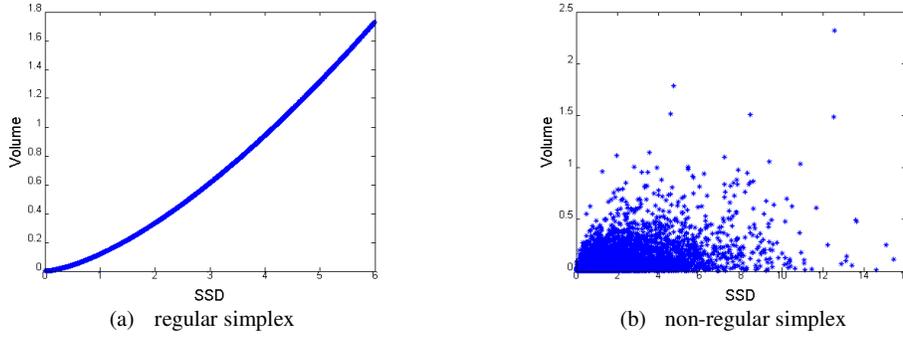
Figure 2. Relationship between volume and distance for regular/non-regular simplex. In the case of regular simplex, the relationship between two constraints is definite, while in the case of non-regular, the relationship is indefinite.

4.2. Parameter Selection

To present a fair comparison, we need to guarantee all variables except for the constraint item to be the same during unmixing process. For further details, (a) the update rule involved in these two constraints is fixed to quadratic programming method and steepest descent method respectively. (b) The regulation factor λ is carefully chosen to ensure similar ratio between constraint value and reconstruction error in each case.

To decide λ_D and λ_V for each constraint, we need to find out the relationship between volume V and sum of squared distance SSD for regular simplex. The relationship satisfies following equation:

$$\frac{V}{SSD} = c(M)d^{M-3} \tag{8}$$

where, M is the number of endmember, c is a variable only related to M, d is the distance between any two endmembers.

Thus, in order to ensure equivalence of these two constraints, we need to make:

$$\frac{\lambda_D}{\lambda_V} = \frac{1}{V/SSD} = \frac{1}{c(M)d^{M-3}} \tag{9}$$

5. EXPERIMENT

In this part, we analyse application range of these two algorithms. Then we apply these two algorithms to real hyperspectral data unmixing and compare the performance. As simplex pattern, initial value and SNR are the most important factors in hyperspectral unmixing, we mainly conduct the comparative analysis in these three aspects.

5.1. Comparison Criterion

In the process of comparing VC and SSD, a suitable comparison metric is needed to measure the unmixing performance. Since, endmember data and abundance map can be transferred to corresponding vector, we adopt angle distance(AD) which measures angular difference between two vectors as criterion, as shown in equation (10). For spectral endmember and abundance map, we refer to angle distance as spectral angle distance(SAD) and abundance angle distance(AAD) separately.

$$SAD = \cos^{-1}(\frac{\boldsymbol{x}_1^T \boldsymbol{x}_2}{\| \boldsymbol{x}_1 \| \| \boldsymbol{x}_2 \|}) \tag{10}$$

where, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are two transformed vectors. AAD is calculated in the similar way.

5.2. Analysis on Synthetic Data

For synthetic data, we pick several spectra from spectral library as endmembers. Then we create hyperspectral data by multiplying normalized endmember data with abundance map generated according to dirichlet distribution. Additionally, Gaussian noise with certain level is also added to data.

1. Flatness Analysis

In this experiment, we reconstruct hyperspectral data with two bands and three endmembers. To demonstrate unmixing capability, we increase the degree of flatness by certain value at each experiment.

Then, we implement unmixing algorithm based on VC and SSD on hyperspectral data. The initial values for both algorithms are randomly set. We compare experimental result of endmember extraction with chosen spectra from library and compute SAD. The result is shown in figure3.

According to Figure 3, we can see that SAD of VC based algorithm is becoming increasingly bigger compared to SSD with increase of degree of flatness. Whereas the resulting endmember based on VC and SSD are similar when the simplex approaches to the regular form. Therefore, we can draw the conclusion that SSD is better than VC when scattering is *flat*. In real data processing, scattering on hyperplane is usually non-regular, so algorithm based on SSD can handle most of these cases according to the result above.
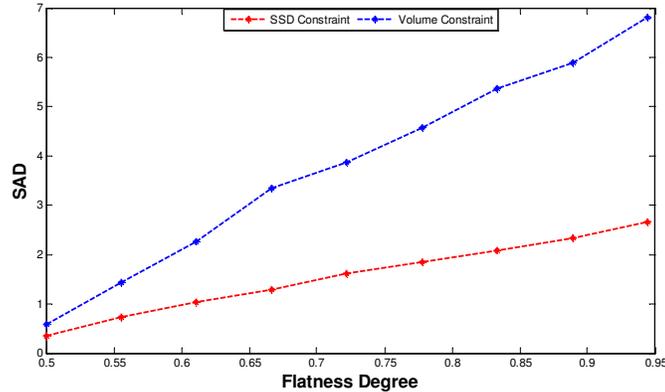
Figure 3 SAD of VC and SSD based algorithm.

2. Random Initialization Analysis

Normally, it is necessary to give an initial value. In this study, we used PPI or N-FINDR to obtain a relatively suitable initial value for following iterations. For abundance matrix, it is often initialized as random value. However, due to the complexity of real data set, these methods do not always perform well on finding well-conditioned initial values. Consequently, the result may be trapped into some local minima. Thus, random initial value analysis can give us a view that which one is more susceptible to ill initial conditions. We carry out the same random initialization on both VC and SSD algorithms, then we compare the result with original true endmember value.

We use random values following Gaussian normal distribution as initial value of endmember and abundance matrices. We conduct 50 comparison experiments with different initial values. Then we compute SAD of resulting endmember data and endmember data in spectral library, as shown in Figure 4.

We can see that endmember extraction result based on SSD is obviously closer to original endmember data. However, some of the result based on VC absolutely deviate from original endmember data. In several cases, though VC-based algorithm also can fit original hyperspectral data well, the evaluated endmember data completely differs from original endmember and consequently abundance matrix obtained by volume constraint is completely wrong.
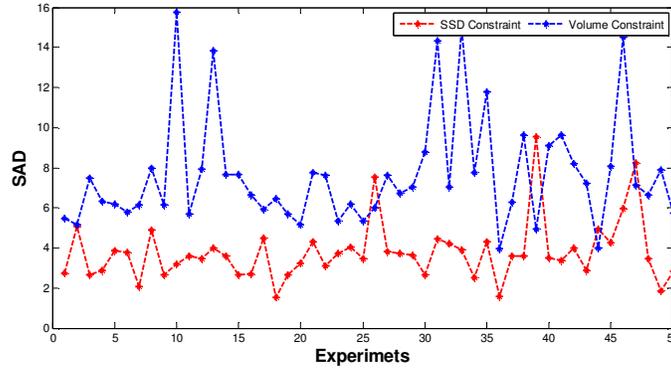
Figure 4 Solution between two constrained algorithms with random initial values. Simulated hyperspectral data consists of 5 bands and 6 endmembers.

3. Robustness Analysis for Noise

Since real data consists of much noise, original unmixing algorithm is sensitive to noise for the sake of fitting every individual data sample. However, unmixing algorithms based on VC and SSD can be applied to unmixing hyperspectral image by minimizing 'volume' of simplex. Thus, reconstructed data will not inevitably approximate all data samples. As a result, these algorithms show strong anti-noise capacity. However, as these two constraints are considered to be inequivalent in many cases, noise-sensitivity may be different with each other. Furthermore, we need to evaluate the suitable degree of SNR for VC and SSD.

We create 50×50 hyperspectral data including 5 bands and 6 endmembers. Then, we add white noise with different levels to synthetic data. The SNR is 10db, 15db, 20db, 30db. During the process of unmixing based on two constraints, we use identical iteration method with same upper limit construction error. We implement two constrained algorithms under each SNR ratio with the same initial value following Gaussian distribution during each experiment.
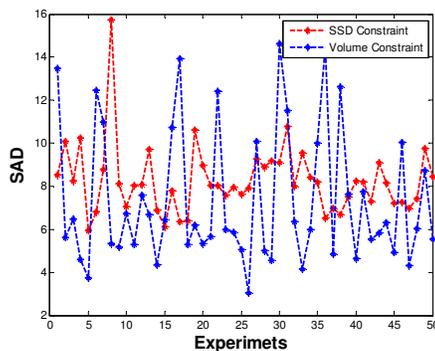
(a) SAD with 10db SNR	(b) SAD with 15db SNR
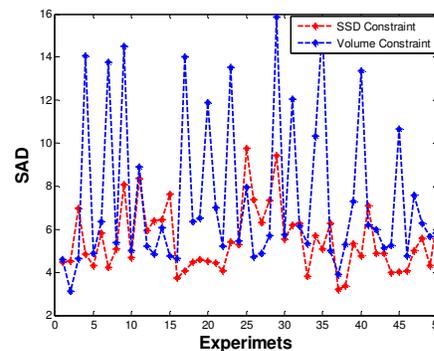
(c) SAD with 20db SNR   (d) SAD with 30db SNR
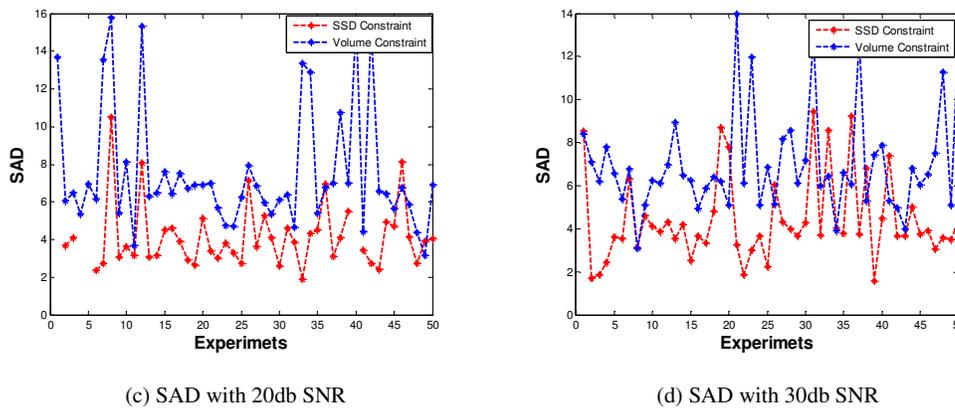
Figure 5 SAD for two constraints with different SNR

We can see from
Figure 5, SAD for VC between extracted and real endmember value change little with the decline
of SNR. On the contrary, SAD for SSD fluctuates with the SNR significantly. Thus, algorithm
based on VC is more robust than SSD in the sense of noise robustness.

5.3. Analysis on Real Data

After finishing the analysis above, we can conclude that SSD is better than VC when hyper
scattering with high degree of flatness or under ill-conditioned initial values. While VC is better
in the sense of robustness to noise. However, above-mentioned experiments are based on
synthetic data. In this experiment, we will utilize real data(AVIRIS data) to identify these two
algorithms.

The used AVIRIS data over Cuprite, Nevada totally contains 400*350 pixels and 50 bands. We
do some pre-processes to raw data to reduce computation complexity before iteration. Firstly, we
utilize principal component analysis (PCA) [16] to reduce data dimension and select principal
band numbers. Secondly, we need to find good initial endmember and abundance value to ensure
algorithms can extract real ground objects efficiently. We utilize endmember data extracted by N-
FINDR as initial endmember matrix. Since hyperspectral data can be regarded as the product of
endmember matrix and abundance matrix, we use unconstrained NMF algorithm to calculate
corresponding abundance matrix S as initial value by fixing endmember abundance A .In
addition, experiment shows that we can achieve much better results by assigning regulation factor
λ_V as 0.15 for VC and λ_D as 0.01 for SSD.

Table 1 SAD among different algorithm

	N-FINDR	Volume Constraint	SSD Constraint
Alunite	4.43	5.45	4.00
Kaolinite	3.28	5.29	5.35
Andradite	4.41	5.03	4.35
Nontronite	4.14	7.58	4.19
Muscovite	6.16	2.34	4.71
Chalcedony	3.75	6.86	3.57
Average	4.36	5.43	4.46

Then, we begin to do iteration for two constraint algorithms until it satisfies terminating
condition. we can find out best matching mineral obtained by two algorithms via comparing with
each mineral reflectance in spectral library [17].

As shown in Figure 6, it's the endmember extraction result based on SSD. Solid line and dashed line represent extraction result obtained by SSD and best matched data in spectral library respectively. The reflectance is normalized to [0, 1]. From several subgraph results, endmembers extracted by SSD are very close to the spectral of real data, like Alunite and Muscovite. The closer they are, the smaller SAD is. Table 1 represents SAD for different algorithms. For some minerals, SSD gives a better solution. For other minerals, VC performs better.
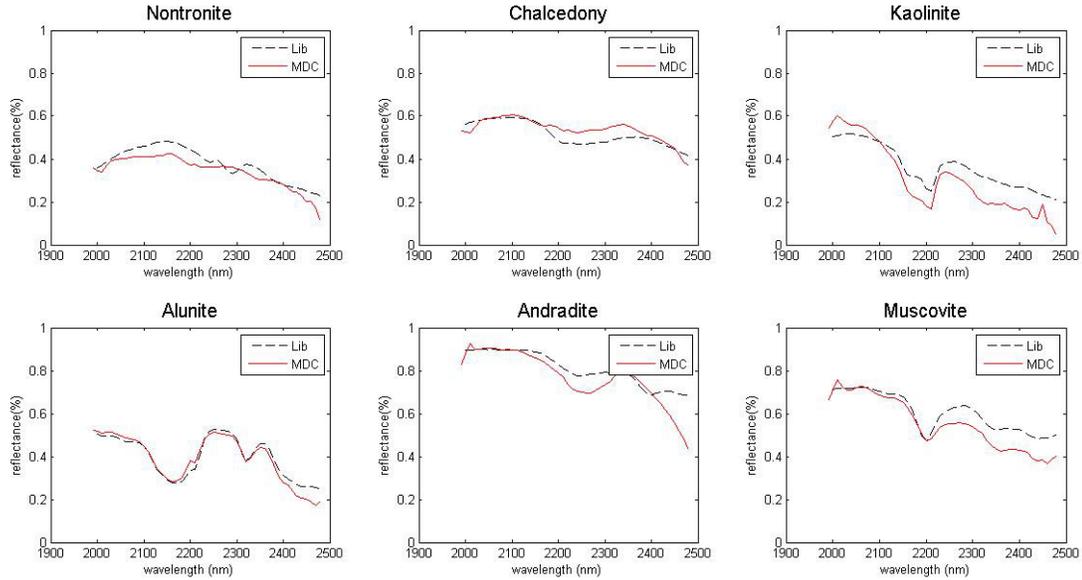
Figure 6 Extracted endmember by SSD.

6. CONCLUSION

In this paper, we analyse VC and SSD algorithm from flatness of simplex, anti-noise and initialization to discriminate these two algorithms. We aim to provide a guidance on which constraint is more suitable under some special conditions.

First, we do analysis for flatness and conduct three comparative experiments using synthetic data. For the pattern of scattering, SSD is better than VC when the scattering is flat. Whereas these two algorithms' performance resemble each other while the simplex is regular. As for initialization, endmember extracted under SSD is closer to original data in random initialization. For anti-noise performance, VC is more robust in different level of noise.

Eventually, on real data, similar solutions can be achieved for these two constrains with well-conditioned initial value. Quantitively, for some minerals, SAD of SSD is smaller, like Chalcedony in Table 1. Yet, for other minerals, like Muscovite, VC works better. Thus, VC and SSD both work similarly in hyperspectral unmixing task.

According to what mentioned above, relatively practical instruction on how to choose constraints can be attained.

REFERENCES

[1] R. N. Clark and T. L. Roush, "Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications," J. Geophys. Res., vol. 89, no. B7, pp. 6329–6340, 1984.

[2] J. J. Settle and N. A. Drake, "Linear mixing and the estimation of ground cover proportions," Int. J. Remote Sens., vol. 14, no. 6, pp. 1159–1177, 1993.

[3] J. W. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," in Proc. Summ. 4th Annu. JPL Airborne Geosci. Workshop, R. O. Green, Ed, 1994, pp. 11–14.

[4] M. E. Winter, "N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in Proc. SPIE Conf. Imaging Spectrometry V, 1999, pp. 266–275.

[5] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," IEEE Trans. Geosci. Remote Sens., vol. 43, no. 4, pp. 898–910, Apr. 2005.

[6] Sun, Lixin. "Improved iterative error analysis for endmember extraction from hyperspectral imagery." Proceedings of SPIE - The International Society for Optical Engineering 7086 (2008):70860S-70860S-10.

[7] Graña, Manuel, C. Hernandez, and J. Gallego. "A single individual evolutionary strategy for endmember search in hyperspectral images." Information Sciences 161.3(2004):181-197.

[8] D. Lee and S. Seung, "Algorithms for non-negative matrix factorization," in Advances in Neural Information Processing Systems, vol. 13. Cambridge, MA: MIT Press, 2001, pp. 556–562.

[9] V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," Lin. Alg. Appl., vol. 416, no. 1, pp. 29–47, Jul. 2006.

[10] Miao, L., & Qi, H. (2007). Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. Geoscience & Remote Sensing IEEE Transactions on, 45(3), 765-777.

[11] Berman, M, et al. "ICE: a statistical approach to identifying endmembers in hyperspectral images." Geoscience & Remote Sensing IEEE Transactions on 42.10(2008):2085-2095.

[12] D. C. Heinz, C.-I Chang, and M. L. G. Althouse, "Fully constrained least squares-based linear unmixing," in Proc. IGARSS, 1999, pp. 1401–1403.

[13] C.-I Chang and D. C. Heinz, "Constrained subpixel target detection for remotely sensed imagery," IEEE Trans. Geosci. Remote Sens., vol. 38, no. 3, pp. 1144–1159, May 2000.

[14] M. O. Smith, P. E. Johnson, and J. B. Adams, "Quantitative determination of mineral types and abundances from reflectance spectral using principal components analysis," J. Geophys. Res., vol. 90, no. B2, pp. C797–C804, Oct. 1985.

[15] G. Strang, Linear Algebra and Its Application, 3rd ed. Cambridge, MA:MIT Press, 1988

[16] C.-I Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," IEEE Trans. Geosci. Remote Sens., vol. 42, no. 3, pp. 608–619, Mar. 2004.

[17] R. N. Clark, G. A. Swayze, A. Gallagher, T. V. King, and W. M. Calvin, "The U.S. geological survey digital spectral library: Version 1: 0.2 to 3.0 µm," U.S. Geol. Surv., Washington, DC, Open File Rep. 93-592, 1993.

[18] Wang T, Lin H. Hyperspectral Remote Sensing[M]// The International Encyclopedia of Geography. John Wiley & Sons, Ltd, 2017.

[19] Févotte C, Dobigeon N. Nonlinear Hyperspectral Unmixing With Robust Nonnegative Matrix Factorization[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2015, 24(12):4810.

[20] Li J, Agathos A, Zaharie D, et al. Minimum Volume Simplex Analysis: A Fast Algorithm for Linear Hyperspectral Unmixing[J]. IEEE Transactions on Geoscience & Remote Sensing, 2015, 53(9):5067-5082.