COMPARATIVE ANALYSIS OF DIFFERENT APPROACHES FOR HYPERSPECTRAL UNMIXING

[1]Chirag Hirpara, [2]Vankar Jyoti P.
M.E, Electronics and Communication Engineering,
Sarvajanik college of Engineering and Technology, Surat,
Email:[1]chiraghirpara0@gmail.com, [2]jyoti.ec@gmail.com

Abstract - The hyperspectral cameras used for imaging are having low spatial resolution, and thus the pixels in the captured image will be mixtures of spectra of various materials present in the scene. Thus spectral unmixing comes as an unavoidable step in hyperspectral image processing. Spectral unmixing is an important problem in hyperspectral data exploitation. It amounts at characterizing the mixed spectral signatures collected by an imaging instrument in the form of a combination of pure spectral constituents (endmembers), weighted by their correspondent abundance fractions. This paper presents a comparative study and performance analysis of different approaches for unmixing. Three geometrical algorithms for spectral unmixing, namely PPI, N-FINDER VCA and MVSA are introduced in this paper. Also the paper represents the algorithms based on statistical approaches, those are NMF, NMF with piecewise smoothness constraint and NMF with sparseness constraint.

Index Terms-Hyperspectral imaging, Spectral signature, Spectral unmixing, PPI, N-FINDER, VCA, MVSA, NMF, nsNMF, NMFSC.

I. INTRODUCTION

HYPERSPECTRAL sensor collects 3-dimensional data with two spatial dimensions and one spectral dimension as shown in Fig. 1. Spectral dimension contain the data in hundreds of very narrow contiguous bands, and this provides a good way for the identification of various materials over the observed scene captured by the sensor.

The various materials are discriminated on their unique spectral signatures as shown in Fig. 2. Hyperspectral imaging is having a wide range of applications in various fields as in agriculture, planetary remote sensing, military, environmental monitoring etc[1]. The hyperspectral imaging sensors can capture many contiguous bands which is having very high spectral resolution and this will be covering not only visible regions but also the infra red regions of electromagnetic spectrum(0.3-2.5μm)[2],[3].

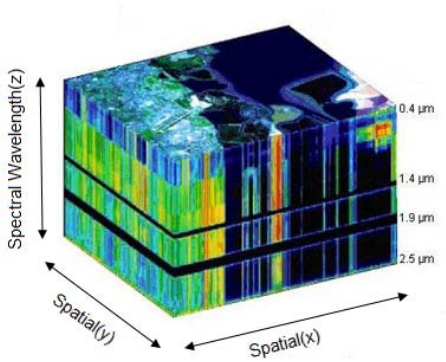

Fig. 1. An Example of a Hyperspectral Image

Advanced hyperspectral sensors like AVIRIS of NASA is now able to cover the above mentioned wave-length region using about 200 spectral channels. In the case of hyperspectral images, depending upon the spatial resolution of sensor, the individual pixels in the captured scene may comprise of more than one material.

Each pixel will be the mixture of various materials of the surface patch and thus the spectra observed will contain multiple endmembers (or spectral signatures) and thus the further analysis becomes difficult. This happens mainly because of the poor spatial resolution of the sensor used. There arises the need of hyperspectral unmixing. Hyperspectral unmixing aims at the decomposition of the observed spectra into a set of pure reference materials(endmembers)and their abundance fractions. Thus unmixing process gives both spectral signatures and corresponding abundance maps of materials present in the scene. This unmixing problem has been a subject to many investigative studies for the past many years.

Hyper Spectral unmixing is basically a blind source separation problem. Hyperspectral sense contain sources which are statistically dependent and they may combine either in a linear or nonlinear fashion. This makes spectral unmixing problem to be placed in higher level compared to other source separation problems.

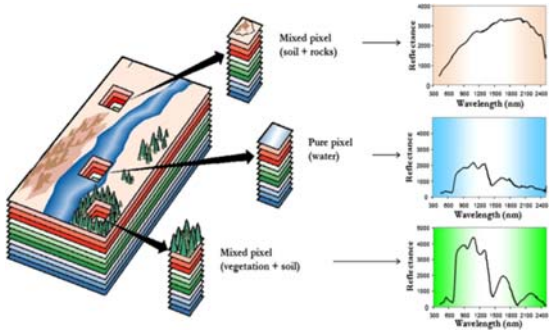

Fig. 2.Reflectance Spectrum of the pixels.

A. Mixing Models

Unmixing can be classified to linear[3] and Non linear[3].Linear models assume that the mixing scale is macroscopic, and the light which falls on the surface interacts with only one material. This type of mixing takes place due to the low spatial resolution of the sensor. Here multiple scatterings do not take place. In the case of Nonlinear mixing models the interaction between the light which is scattered by multiple materials occurs, and the mixed model becomes complicated. The interactions can be at Intimate or microscopic level. Thus non-linear unmixing becomes a difficult task. So here this paper concentrates in linear unmixing due to its simplicity, and also it's the basis of many algorithms for more than 30 years.

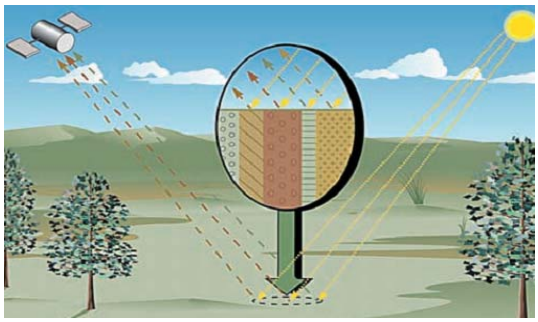

Fig. 3.Linear mixing model without any multiple scattering effects.

The linear mixing model assumes that the spectrum of a pixel in the acquired scene is a linear combination of all pure materials (endmembers) present in the scene. it is assumed that the hyperspectral sensor used for capturing the scene has L spectral bands, linear mixing model can be mathematically represented as in Fig. 4. The relation between them is given as per the follow.

$$R = X + n \quad (1)$$

Where, $X = MS$

$$R = MS + n(2)$$

Where, $S = \gamma\alpha$

$$R = M\gamma\alpha + n(3)$$

Here term **R** is the original image cube. We can represent $\mathbf{R}_{ij} \in R_L$ as an observation array at a single pixel. Matrix $\mathbf{M} \in R_{L\times P}$ is our observed spectral endmember signature matrix having each column $\mathbf{M}_p \in R_L$ and here P represents the total number of endmembers present in our image data. Matrix **S** represents the relative abundance cube. Its every column $\mathbf{S}_{ij} \in R_P$ is defined as an abundance vector related with $\mathbf{R}_{ij}$. In abundance matrix **S**each element of it denoting the relative abundance fraction of endmember present in $\mathbf{R}_{ij}$.

For simplicity, "matrix-to-vector" alignment process is used. In this method all the rows in the matrix are concatenates together to form a single vector. This process is mainly applied for the conversion of band in **R** and abundances in **S**. So our matrices will now change in dimensions. Original **R** is of i×j×L, where K=i×j, **M** is of L×P

and **S** is of i×j×P. This dimensions will then transformed to **R** is of L × K, **M** is of L×P and **S** is of P × K, respectively, where K is the number of pixels, P in the number of endmembers, L is the total spectral bands.In this modeling both M and α have to be found by unmixing.
Here basic two constraints are used.

*ANC (abundance nonnegative constraint):*It says that the each and every component present in our signature matrix **M** should always be nonnegative so it is given as eq. (4).

$$M_{IP} \geq 0,\ 1 \leq l \leq L \quad (4)$$

ASC (abundance sum to one constraint): It says that the sum of all the relative fraction of each pixel should be equal to unity as given as eq. (5).

$$\underline{S}_{ijp} \in [0,1], \sum_{p=1}^{P} \underline{S}_{ijp} = 1 (5)$$

Fig. 4. Simplified linear mixture model

II. GEOMETRICAL APPROACHES

As per the geometrical point of view the linearly mixed vectored generated by the mixing process are always stays in a simplex set of data (positive cone). It uses basically geometric properties of an image.These approaches take advantage of the analogy between mixing models and the geometric orientation of hyperspectral data in multidimensional spaces.

A. Endmember extraction algorithms overview

Geometrical approaches come as the third category of spectral unmixing algorithms. Basically it follows the fact that, under the linear mixing model spectral vectors belong to the simplex set whose vertices correspond to the endmembers. Thus by finding out the vertices it is possible to find out the endmember in the hyperspectral image.

There are two categories in this approach. Algorithms which assume the presence of pure pixels comes under the one category and algorithms which do not assume the presence of pure pixels comes under another category. MVSA (minimum volume simplex analysis) [11], MVES(minimum volume enclosing simplex), SISAL(simplex identification via split and augmented lagrangian), etc comes under the first category. In the second category to which this paper concentrates, come the following algorithms like SVMAX (successive volume maximization), AVMAX (Alternating volume maximization), ADVMM(alternating decoupled volume max-min),SDVMM(successive decoupled volume max-min),N-FINDR,VCA(vertex component analysis), IEA (iterative error analysis), PPI(pixel purity index), etc are some of the algorithms come under this section.

B. Pure pixels based geometrical algorithms

As discussed in the previous sections geometrical algorithms with pure pixel assumption assumes the presence of at least one pure pixel per endmember. These pure pixel algorithms still belong to minimum volume class.This assumption of pure pixel make these algorithms very efficient but still creates difficulty in some datasets.In this section a brief theoretical side of each of the 5 algorithms namely PPI, N-FINDER, VCA, MVSA is given.

C. N-FINDR

This is another popular algorithm used for spectral unmixing. This also works according to winter's belief [24].This is a pure pixel based algorithm and this search for the set of pixels with largest possible volume by inflating a simplex inside the given dataset. The original N-FINDER algorithm [3] is having 4 steps as follows.

1) Feature reduction-In this the dimension of data is reduced from n toP-1 by some PCA [32] or MNF[33],where P is the number of endmembers to be identified.

2)Take some randomly selected endmembers from the dataset as $\left\{E_1^{(0)}, E_2^{(0)}, E_3^{(0)} .... E_p^{(0)}\right\}$

3) At each iteration k≥0, calculate the volume by this set of endmembers as follows.

$$v\left(E_1^{(k)}, E_2^{(k)} .... E_p^{(k)}\right) = \frac{\left|\det\begin{bmatrix} 1 & 1... & 1 \\ E_1^{(k)} & E_2^{(k)}... & E_p^{(k)} \end{bmatrix}\right|}{(p-1)!} \quad (6)$$

4) Replacement- For each and every pixel the volume corresponding to it is checked by this way,if this pixel replaces one of the given endmember positions in matrix shown above. . If the replacement of pixel results in an increase in volume,the pixel replaces the endmember. This process continues until there are no endmember replacements in the given data.

D. Vertex component analysis

The pseudo-code for the VCA method is shown in following steps. Symbols $[\widehat{M}]_{:,j}$and $[\widehat{M}]_{:,i:k}$ stand for the j_{th} column of $\widehat{M}$ and for thei_{th} to k_{th} columns of $\widehat{M}$, respectively. Symbol $\widehat{M}$ stands for the estimated mixing matrix.

Step 1 Hyperspectral data cube (R) and number of endmembers (p) are known and can be an input of VCA.

Step 2 Find out value of $SNR_{th}$for given p.

Step 3 Test if the SNR is higher than SNR_{th} in order to decide whether the data is to be projected onto a subspace of dimension p or $p-1$. In the first case the projection matrix U_d is obtained by SVD from RR^T/N. In the second case the projection is obtained by PCA from $(R - \bar{r})(R - \bar{r})^T/N$ (recall that $\bar{r}$ is the sample mean of $[R]_{:,i}$, for $i = 1,...,N$).

Steps 5 and 10 ensure that the inner product between any vector $[X]_{:,j}$ and vector u is non-negative, a crucial condition for the VCA algorithm to work correctly. The chosen value of $k = \text{argmax}_{j=1...N} \|[X]_{:,j}\|$ ensures that the colatitude angle between u and any vector $[X]_{:,j}$ is between 0° and 45°, then avoiding numerical errors that otherwise would occur for angles near 90°.

Step 15 initializes the auxiliary matrix A, which stores the projection of the estimated endmembers signatures. Assume that there exists at least one pure pixel of each endmember in the input sample R. Each time the loop for is executed, a vector f orthonormal to the space spanned by the columns of the auxiliary matrix A is randomly generated and y is projected onto f . Knowing that pure endmembers occupy the vertices of a simplex, then $a \leq f^T [Y]_{:,i} \leq b$, for i = 1,…,N, where values a and b correspond only to pure pixels. The endmember signature corresponding to max ([|a| |b|]) is stored. The next time loop for is executed, f is orthogonal to the space spanned by the signatures already determined. Since f is the projection of a zero-mean Gaussian independent random vector onto the orthogonal space spanned by the columns of $[A]_{:,1:i}$, then the probability of f being null is zero. Note that the underlying reason for generating a random vector is only to get a non null projection onto the orthogonal space generated by the columns of A. Figure 4.2 shows the input samples and the chosen pixels, after the projection $v = f^T Y$. Then a second vector f orthonormal to the endmember a is generated and the second endmember is stored. Finally, steps 25 and 27 compute the columns of matrix $\widehat{M}$, which contain the estimated endmembers signatures in the L-dimensional space.

E. Minimum volume simplex analysis algorithm

Let $Y \equiv [y_1, y_2, ..., y_N] \in R^{p \times n}$ is a matrix holding in its columns the spectral vectors $Y_i \in R^p$, for $i = 1, 2, ..., n,$ of a given hyperspectral data set. Although not strictly necessary, we assume in this version of the algorithm that a dimensionality reduction step has been applied to the data set and the vectors $Y_i \in R^p$ are represented in the signal subspace spanned by the endmember spectral signatures. Under the linear mixing model, we have spectral signatures. Under the linear mixing model, we have

$$Y = MS \quad (7)$$

$$\text{S.t.: } S \geq 0, 1_p^T S = 1_n^T$$

Where, $M \equiv [m_1, m_2, ..., m_p] \in R^{p \times p}$ is the mixing matrix (m_i denotes the ith endmember signature and p is the number of endmembers), and $S \in R^{p \times n}$ is the abundance matrix containing the fractions ($[S]_{i,j}$ denotes the fraction of material m_i at pixel j). For each pixel, the fractions should be no less than zero, and sum to 1, that is, the fraction vectors belong to the probability

simplex. Therefore, the spectral vectors Y_i belong, as well, to a simplex set with vertices m_i, for $i = 1, 2, ..., p$.

Given *Y*, and inspired by the seminal work, we infer matrices *M* and *S* by fitting a minimum volume simplex to the data subject to the constraints in eq, (7). This can be achieved by finding the matrix *M* with minimum volume defined by its columns under the constraints in eq. (7). It can be formulated as the following optimization problem:

$$M^* = \arg\min \left|\det(M)\right| \quad (8)$$

$$\text{s.t.: } QY \geq 0, 1_p^T QY = 1_{N,}^T$$

Where, $Q \equiv M^{-1}$. Since $\det(Q) = 1/\det(M)$, we can replace the problem in eq. (8) with the following:

$$Q^* = \arg\max \log \left|\det(Q)\right| \quad (9)$$

$$\text{s.t.: } QY \geq 0, 1_p^T QY = 1_{N,}^T$$

Optimizations eq. (8) and (9) are nonlinear, although the constraints are linear. Problem eq. (8) is nonconvex and has many local minima. So, problem eq. (9) is nonconcave and has many local maxima. There-fore, there is no hope in finding systematically the global optima of (9). The MVSA algorithm, we introduce below aims at "good" suboptimal solutions of optimization problem (9).

Our first step is to simplify the set of constraints $1_p^T QY = 1_N^T$ by noting that every spectral vector y in the data set can be written as a linear combination of *p* linearly independent vectors taken from the data set, say $Y_p \equiv \left[y_{i1}, y_{i2}, ..., y_{ip} \right]$, where the weights add to one: i.e., $y = Y_p \beta$, where $1_p^T \beta = 1$. It turns out then, the constraint $1_p^T QY = 1_N^T$ is equivalent to $1_p^T QY_p = 1_N^T$ or else to $1_p^T Q = 1_N^T (Y_p)^{-1}$. Defining $q_m = 1_N^T (Y_p)^{-1}$, we get the equality constraint $1_p^T Q = 1_N^T (Y_p)^{-1}$. Then, the problem (9) simplifies to

$$Q^* = \arg\max \log \left|\det(Q)\right| \quad (10)$$

$$\text{s.t.: } QY \geq 0, 1_p^T Q = q_m$$

We solve the optimization problem (10) by finding the solution of the respective Kuhn-Tucker equations using sequential quadratic programming (SQP) methods. This methods belongs to the constrained Newton (or quasi-Newton) and guarantee superliner convergence by accumulating second-order information regarding

The Kuhn-Tucker equations. Each quadratic problem builds a quadratic approximation for the Lagrangianfunction associated to (10). Forthis reason, we supply the gradient and the Hessian of fin each SQP iteration.

Usually, the hyperspectral data sets are huge and, thus, the above maximization is heavy from the computational point of view. To lighten the MVSA algorithm, we initialize it with the set of end-members $M \equiv \left[m_1, m_2, ..., m_p \right]$ generated by the VCA algorithm [6]. We selected VCA because it is the fastest among the state-of-the-art.

Pure pixel based methods. Since the output of VCA is a set of *p* vectors that are in the data set, then we can discard all vectors belonging to the convex set generated by the columns of *M*. If the number of endmembers is high, it may happen that the initial simplex provided by VCA contains very few pixels inside and, therefore, most are out-side, violating the non negativity constraints and slowing down the algorithm. In such cases, we expand the initial simplex to increase the number of pixels that are in the convex hall of the identified endmembers, which speeds up the algorithm. The pseudo code for the MVSA method is shown in below. Symbols $g(Q)_{:,j}$ and $g(Q)_{i,:}$ stand for, respectively, the j_{th} column and the i_{it} line of $g(Q)$, the gradient of $f(Q)$.

III. STATISTICAL APPROACHES

The geometrical based methods give poor results whenever the spectral mixtures are highly mixed. So whenever number of bands gets increased the result of geometrical approaches get reduces.

This is also true for increment in number of pixels in the image. As we know that in geometrical approach it always needs at least one pure pixel which is not possible for every image. So we have to move on to the statistical methods, which are powerful alternative, but come with price: higher computational complexity. In this approach we don't need every time a pure pixel as well as often the full additive constraint is also verified.

A. NMF

In the unmixing process of hyperspectral data, the method of NMF can be applied for the minimization of the objective function given by eq. (11). So basic NMF problem can be defined as:

Given a nonnegative matrix $R \in R_{m\times n}$ and a positive integer $k < \min\{m,n\}$, find nonnegative matrices $M \in R_{m\times k}$ *and* $S \in R_{k\times n}$ to minimize the objective function given in eq. (11) .

$$f(M,S)=\frac{1}{2}\|R-MS\|_F^2 \qquad (11)$$

The multiplication factor **MS** is called a NMF of our original data set **R**. Here data **R**is not needed equal to the multiplication **MS.** But the multiplication **MS**is an approximate factorization term that is having the at most k. Another key characteristic of NMF is the ability of numerical methods that minimize eq. (11) to extract underlying features as basis vectors in **M**, which can then be subsequently used for identification and classification. By not allowing negative entries in **M** and **S**, NMF enables a non-subtractive combination of parts to form a whole [6], [7].

Fundamental Algorithms:

1) Multiplicative update algorithm

The basic multiplicative algorithm was introduced by Lee and Seung (2001) [5], [6]. This multiplicative update rules contained the mean squared error objective function of eq. (11) is given below [5], [6].

Here eq. (12) and eq. (13) are the basic update equation for this algorithm.

$$S \leftarrow S\frac{M^T R}{M^T MS} \qquad (11)$$

$$M \leftarrow M\frac{R\,S^T}{M\,S\,S^T} \qquad (12)$$

2) ALS Algorithm

This algorithm comes with modification in LS algorithm where one extra least square step with changed fashion is added after the first least square step. ALS algorithms were first introduced by Paatero, in 1994. But this algorithm comes with one limitation, that the problem of eq. (11) cannot convex is both the matrix M and S. It may convex in only any one of them [5], [6].

B. Constrained NMF

In this simple NMF problem our objective function shown in eq. (11) makes the solution nonunique for the factorization of R into M and S matrix so that there is necessity to add different constraints to make the solution unique.

The adaptive potential function given by the eq. (13) is used to represent the characteristic as a piecewise smoothness of spectral data.

$$g(x)=-e^{x^2/\gamma}+1 \qquad (13)$$

Here the in this function constant 1 is added to make the results of it nonnegative. And γ is a positive parameter whichcontrols the shape of this potential function.

Hyperspectral images are having high spectral resolution so that they are having strong continuity of spectral signature as compare to abundance map. So that parameter γ is defined for both spectral correlation as well as spatial correlation as $\gamma_{\mathbf{m}}$ and $\gamma_{\mathbf{s}}$ respectively. They both should have different values. The adaptive potential function $g(\mathbf{m}-\mathbf{m}_N)$ represents the piecewise smoothness of the spectral signature **m**. this can be generated by observing the spectral signature **m** and its neighbor $\mathbf{m}_N$.

The i^{th} entry in this adaptive potential function is defined as $g(\mathbf{m}_i-\mathbf{m}_{Ni})$ as given in eq. (14).

$$g(m_i - m_{N_i}) = \sum_{i' \in N_i} g(m_i - m_{i'}) \qquad (14)$$

Here Ni = {i − 1, i + 1}.

Same concept is defined for the piecewise smoothness of endmember abundances that can be given by $g(\mathbf{S} - \mathbf{S}_N)$.The entry at position (p, k) is derived by eq. (15).

$$g\left(S_{pk} - S_{N_{pk}}\right) = g\left(\underline{S}_{ijp} - \underline{S}_{N_{ijp}}\right) = \sum_{i'j' \in N_{ijp}} g(\underline{S}_{ijp} - \underline{S}_{i'j'p}) \qquad (15)$$

Here $\mathbf{S}_{pk} = \underline{S}_{ijp}$and $N_{ijp}$is the neighborhood of $\underline{S}_{ijp}$. The assumption is taken as,
N_{ijp}= {(i − 1) j, (i + 1) j, i (j − 1), i (j + 1)}.
So our objective function that is only based on Euclidian distance given in eq. (3.1) will now added with extra constraints. It will be as follow.

$$D(M,S) = Euc(M,S) + \alpha\langle g(M - M_N)\rangle + \beta\langle g(S - S_N)\rangle \qquad (16)$$

Here ⟨.⟩defines the total sum of the matrix elements, such as

$$\langle g(M - M_N)\rangle = \sum_{l,p=1}^{L,P} g(M_{lp} - M_{N_{lp}}) \qquad (17)$$

Same rule can be applied to matrix S.

So now for this objective function our updating rules are also changed as follows.

$$M \leftarrow M.* \frac{\left(RS^T + \alpha\left(M.*h(M - M_N) - g'(M - M\right.\right.}{(MSS^T + \alpha M.*h(M - M_N))} \qquad (18)$$

$$S \leftarrow S.* \frac{(M^T R + \beta(S.*h(S - S_N) - g'(S - S_N))}{M^T MS + \beta S.*h(S - S_N))} \qquad (19)$$

Here ./and .* represents element wise division and multiplication respectively and $(.)^T$ defines the transposition of the any matrix. And,

$$h(m_i - m_{N_i}) = \frac{2}{\gamma_m} \sum_{i' \in N_i} e^{\frac{-(m_i - m_{i'})^2}{\gamma_m}} \qquad (20)$$

$$g'(m_i - m_{N_i}) = \frac{2}{\gamma_m} \sum_{i' \in N_i} (m_i - m_{i'}) e^{\frac{-(m_i - m_{i'})^2}{\gamma_m}} \qquad (21)$$

1) Piecewise Smoothness Constraint (nsNMF)

The smoothness of the image is very important parameter for unmixing process. Here in this algorithm one smoothness matrix C is included into our objective function given by eq. (11) will now changed to eq. (22)

$$f(M,S) = \frac{1}{2}\|R - MCS\|_F^2 \qquad (22)$$

Here $C \in R_{p \times p}$is positive symmetric matrix which is defined by eq. (23)

$$C = (1-\theta)\, I + \frac{\theta}{P} 11^T \qquad (23)$$

Here I is the identity matrix and 1 is the vector of ones. Smoothness of C is controls by parameter θ (0< θ<1).afar adding this parameter to our objective function the algorithm of constraint NMF will be implemented. But here now update equation for matrix S given by eq. (19) is now changed to eq. (24).

$$S \leftarrow S.* \frac{(MC^T R + \beta(S.*h(S - S_N) - g'(S - S_N)))}{MC^T MCS + \beta S.*h(S - S_N))} \qquad (24)$$

The smoothness constraint may or may not give the unique solution for our objective function given in eq. (11) so that more constraints will be used to get the unique solution.

2) NMF with Sparseness Constraint (NMFSC)

This criterion is base on the relationship betweenL1 and L2 norm of the measured data. The sparseness of the image is a measured energy of any vector contained into a small number of components. This algorithm only changes the update equations for the matrix S by setting the L1 and L2 norm as well as the nonnegativity constraint.

For NMFSC the same algorithm of constraint NMF is implemented but only 1 step is added that after updating matrix by eq. (19) and then for achieving the desired sparseness project each row of S.

IV. CONCLUSION

Today spectral unmixing is a very active and interesting research topic in the remote sensing technologies. By using the spectral sensor the captured image is a mixture of various signatures of different materials and their fractional available in that scene. But in many applications we need to identify this material and their relative fractions. So this paper represents the recent developments in the unmixing field as well as the various approaches of unmixing problem.

REFERENCES

[1] G. A. Shaw and H.-H. K. Burke, "Spectral imaging for remote sensing", *Linc. Lab. J.*, vol. 14, no. 1, pp. 3–28, 2003.

[2] N. Keshava and J. F. Mustard, "Spectral unmixing", *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan 2002.

[3] Bioucas-Dias J.M., Plaza A., Dobigeon N., Parente M., Qian Du, Gader, P., Chanussot J., "Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol.5, no.2, pp.354, April 2012.

[4] M. Parente and A. Plaza, "Survey of geometric and statistical unmixing algorithms for hyperspectral images", in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process.: Evolution in Remote Sens. (WHISPERS)*, pp. 1–4, 2010.

[5] Mehul S Raval, "Hyperspectral Imaging: A Paradigm in Remote Sensing", *CSI communication*, pp.7-9, Jan 2014.

[6] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization", *Computational Statistics & Data Analysis.,* vol. 52, no. 1, pp. 155–173, Elsevier, Sep 2007.

[7] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization", in *Adv. Neural Inform. Processing System,* vol. 13. Cambridge, MA:MIT Press, pp. 556–562, 2000.

[8] SenJia, YuntaoQian, "Constrained Nonnegative Matrix Factorization for Hyperspectral Unmixing", *IEEE Transactions on Geoscience and Remote Sensing*, vol.47, no.1, pp.161, Jan 2009.

[9] A. Pascual-Montano, J. M. Carazo, K. Kochi, D. Lehmann, and R. D. Pascual-Marqui, "Nonsmooth nonnegative matrix factorization (nsNMF)", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 3, pp. 403–415, Mar 2006.

[10] Jose M. P.Nascimento, José M. Bioucas-Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data",*IEEE Transactions on Geoscience and Remote Sensing*, vol.43 (no.4), pp.898, 910, April 2005.

[11] Jun Li and Jose M. Bioucas-Dias, "Minimum Volume Simplex Analysis: A FastAlgorithmTo Unmix Hyperspectral Data", *IEEE International on Geoscience and Remote Sensing Symposium (IGARSS)*, vol.3, pp.250, 253, July 2008.

[12] USGS Digital Spectral Library: http://speclab.cr.usgs.gov/spectral-lib.htm