

Hyperspectral Image Unmixing for Classification and Recognition: An Overview

Mingyu Nie¹, Zhi Liu^{1*}, Hui Xu¹, Xiaoyan Xiao², Fangqi Su¹,
Jun Chang¹ and Xiaomei Li³

¹*School of Information Science and Engineering, Shandong University*

²*Department of Nephropathy, Qilu Hospital of Shandong University*

³*Department of oncology, the Second Hospital of Shandong University*

*liuzhi@sdu.edu.cn

Abstract

The limited resolution of image sensors and the complex diversity of nature, cause mixed pixel problems in hyperspectral technology. Such problems are common, and increase the complexity of hyperspectral image processing. Hyperspectral unmixing is crucial for hyperspectral image classification and recognition. In unmixing, the image signatures are represented as a linear combination of the basic materials. Unmixing is the process of decomposing a mixed pixel into constituent materials, and calculating the corresponding fractional abundance. If pure materials (end members) are present in an image, unmixing can be divided into two steps, namely, end member extraction and abundance decomposition. On the other hand, if there is no pure material, researchers have devised and investigated unsupervised and semi-supervised spectral unmixing technology. This article presents an overview of the state-of-the-art methods of hyperspectral unmixing and their extensions.

Keywords: *hyperspectral; unmixing; end member extraction; unsupervised; NMF; semi-supervised*

1. Introduction

Hyperspectral imaging [1] was discovered in the early 1980s. It employs an imaging spectrometer that can detect narrow and continuous spectra of image data, in the electromagnetic spectrum of ultraviolet, visible light and near-infrared and infrared areas (wavelength between 400 and 2400 nm). A hyperspectral image reflects the radiation and space information of targets.

* Corresponding Author

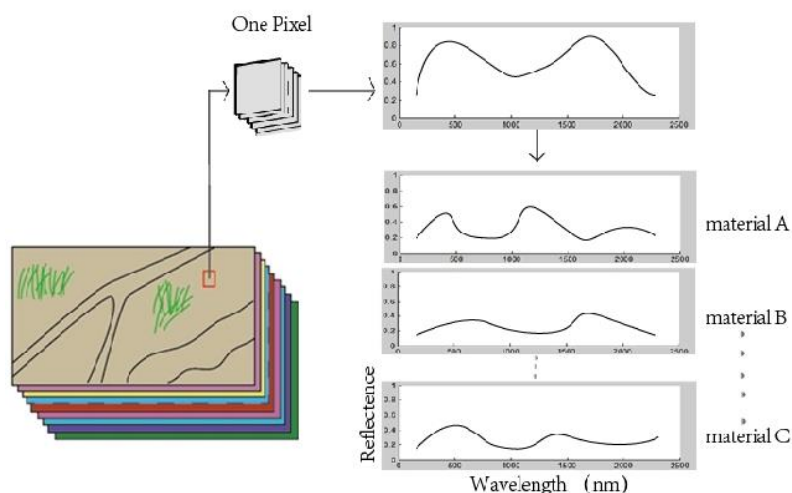


Figure 1. Graphical Illustration of a Hyperspectral Image and Unmixing

The spectrum obtained in one pixel may have mixed spectra of different materials because of the limited resolution of image sensors and the complex diversity of nature. Thus, unmixing [2] is a crucial step in hyperspectral technology that must be performed before material classification and recognition. Unmixing is implemented to decompose mixed pixel into basic components (end members), and obtain the abundance fractions. Figure 1 illustrates the process of unmixing. A pixel is a linear combination of basic materials, decomposed into material A, B, C, *etc.*, via unmixing.

The remainder of this paper is organized as follows. Section 2 describes the linear mixing mode [3] of a hyperspectral image. Section 3 introduces the traditional methods of end member extraction and abundance decomposition. Section 4 describes unsupervised spectral unmixing technology and its extensions. Section 5 introduces other trends in hyperspectral unmixing, semi-supervised unmixing, and contains sparse regression and its extensions. Section 6 presents the conclusions and future work.

2. Mixing Model for Hyperspectral Image

Hyperspectral unmixing methods can be divided into two forms, linear mixing model (LMM) and nonlinear mixing model (NMM) [4], according to the mixed materials and physical distribution space. LMM assumes that the pixels are a linear combination of end members and ignores the magnitude of the multiple scattering between different types of materials. By contrast, small mixed materials cause the input photon to have multiple scattering with more than one type of material, such as water and bread crumbs, and result in the formation of a nonlinear mixing model. A common feature of NMMs is that they all include a supplementary additive term to the standard LMM [5]. Compared with NMM, LMM has the advantages of simple modeling, definite physical meanings, and good scientific theory. Therefore, LMM was employed in this study.

In LMM, a hyperspectral image can be represented as $m \times n$ matrix X , where m is the number of spectrum and n is the number of pixels. Each row of X corresponds to a 2D image in a wavelength, and each column of X corresponds to the spectral signature of a pixel. LMM is illustrated in Figure 2. The model in practical application can be represented as

$$X \approx A \times S + E \quad (1)$$

where A is an $m \times r$ matrix that represents the signatures end-members in different wavelengths, r is the number of end-members, and S is an $r \times n$ matrix that represents

abundances, and E represents additive noise. However, abundance is non-negative (ANC), and each column element has a sum of 1 (ASC).

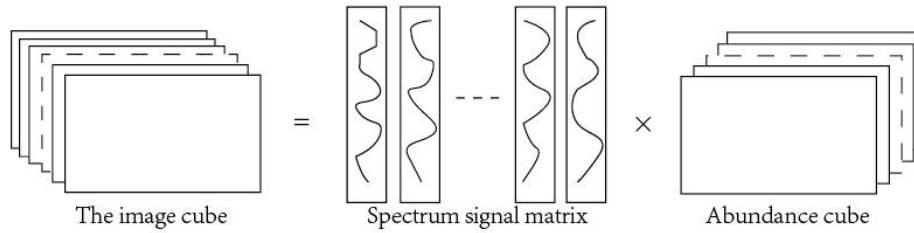


Figure 2. Linear Mixing Model

With LMM, hyperspectral unmixing can be divided into two steps (Section 3): end member extraction and abundance decomposition. Unmixing can be performed with different methods depending on the different mixed degrees of materials and whether the hypothesis (pure pixels) is established. In the presence of pure pixels, the traditional methods utilized to extract end members include pixel purity index (PPI) [6], N-FINDR [9], and vertex component analysis (VCA) [10]. When the materials are mixed thoroughly and pure pixels are not present, the traditional methods of end member extraction cannot obtain accurate results. In this case, blind source separation must be conducted through independent component analysis (ICA) [20], complexity pursuit [25], non-negative matrix factorization (NMF) [26], and their extensions. In addition to the above methods, sparse regression [34] can be utilized for hyperspectral unmixing when the spectral library has been obtained in advance.

3. Traditional Methods of End Member Extraction and Abundance Decomposition

3.1. End Member Extraction

When a hyperspectral image has high resolution, end members are assumed to be present in the image. Identifying pure signatures (contain only one material) in a hyperspectral image is important in hyperspectral data unmixing. PPI, N-FINDR, VCA, iterative error analysis (IEA) [11], and so on are the traditional methods of end member extraction.

3.1.1. PPI: PPI, developed by Boardman *et. al.*, [7], is a popular end member extraction algorithm. In PPI, a maximum noise fraction transform of the data is applied to reduce dimensionality. Then, a large number of random test vectors are generated through the data set called “skewers”. All data vectors are projected onto $skewer_j$, j is the serial number of “skewers”; if pixel r is projected to the extreme positions of $skewer_j$, $I_{s_j}(r) = 1$. The PPI score of a sample vector r can then be defined as

$$N_{PPI}(r) = \sum_{j=1} I_{s_j}(r) \quad (2)$$

Threshold value t must be defined for the PPI score. If sample vector r has $N_{PPI}(r) \geq t$, then r is an end member of the image.

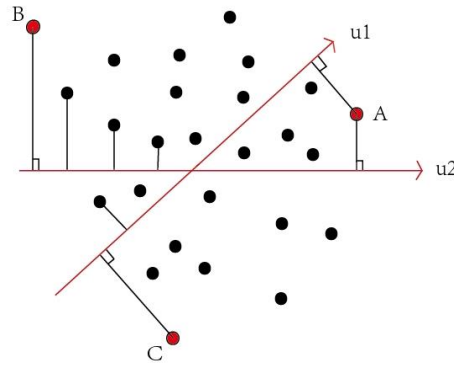


Figure 3. Simple Explanation of PPI

In PPI, a large $N_{PPI}(r)$ value means high probability that r is a pure pixel. As shown in Figure 3, u_1 and u_2 are random test vectors, pixel A is projected to the extreme positions of u_1 and u_2 , pixel B is projected to the extreme position of u_2 , and pixel C is projected to the extreme position of u_1 . As a result, A has a higher probability of being a pure pixel than B and C.

PPI has several drawbacks. This algorithm is not an iterative process and thus complicates the calculation process. Moreover, skewers are randomly generated in PPI and may produce different end members.

Chein-I Chang and Antonio Plaza developed fast iterative PPI (FIPPI) [8] to optimize PPI. In FIPPI, the number of end members is estimated by virtual dimensionality. Appropriate skewers are produced instead of using random vectors.

3.1.2. N-FINDR: In 1999, Michael E. Winter proposed N-FINDR, a method in which the image data are considered to form a convex set in high-dimensional space. The end members in this method are the vertices of the convex. Two bands and their end members as Figure 4 to illustrate the relationships of the geometric.

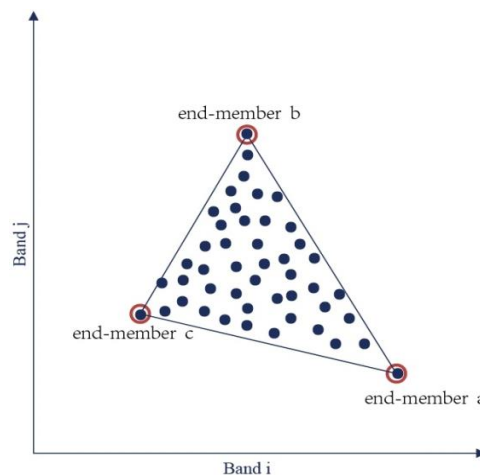


Figure 4. Simple Explanation of N-FINDR

As shown in Figure 4, the simplex formed by end members is larger than that formed by other pixel combinations. The algorithm identifies the largest volume by inflating the simplex. A random set of pixels is selected as the end members, and the volume is defined as follows: E is the augmented matrix of the end members by adding a ones row,

$$E = \begin{bmatrix} 1 & 1 & \dots & 1 \\ - & - & & - \\ e_1 & e_2 & \dots & e_l \end{bmatrix}$$

Where e_i is an m-dimension column vector of temporary end member i and l is the number of end-members. The volume (V) of the simplex is defined as

$$V(E) = \frac{1}{(l-1)!} \text{abs}(|E|) \quad (3)$$

For each pixel in the image and each end member, the spectrum of the pixel replaces the temporary end member and the volume is recalculated. If it increases, the spectrum of the pixel replaces the temporary end member. This procedure is repeated until all pixels are calculated.

End member extraction can be performed with other methods aside from the classic algorithms mentioned above.

VCA finds the orthogonal vector repeatedly, projects pixels to the orthogonal vector, and then calculates the projection distance to extract end members. The point with the largest coordinate is selected as the initial end member, a unit vector that is orthogonal with the simplex and the initial end member is identified, all pixels are projected onto the unit vector, and the pixel point with the maximum projection distance is extracted as the end member. These steps are repeated until all end members are detected by circulation.

IEA repeatedly employs linear mixed solution with constraints and selects the end members with the minimum error after linear mixed solution.

Sequential maximum angle convex cone [12] changes the orthogonal projection of VCA to oblique projection. Extraction is conducted through iteration. Each iteration extracts one end member, calculates the coefficient of the end member in all pixels, and adjusts the proportion coefficient of previously extracted end members in all pixels. At the same time, projection transformation is implemented aside from the influence of the end member from each pixel. The iteration is continued until all end members are extracted.

End member extraction can be performed with other methods aside from those mentioned above. Such methods include simplex growing algorithm [13], alternating volume maximization, and successive volume maximization [14].

3.2. Abundance Decomposition

Abundance is calculated after obtaining the end members of the image. The most commonly used methods are least squares (LS) [15], projection pursuit, and orthogonal subspace projection.

LS minimizes the sum of the squares of the error to identify the best match function. Unknown data can be easily obtained by LS, and the sum of the squares of the error between actual and calculated data is at the minimum. LS can be divided into four forms based on the degree of abundance constraints: unconstrained least squares, sum-to-one constrained least squares, no negatively constrained least squares, and fully constrained least squares [15].

The basic idea of projection pursuit [16] is to project high-dimensional data to low dimensions, with a projection index (PI) as a measure to determine the projection that can reflect the structure or characteristic of the original high-dimensional data. The projection data are then analyzed to determine the characteristics of the original data. Variance or information divergence serves as the PI.

Other methods of abundance decomposition of mixed pixels include orthogonal subspace projection [17], maximum likelihood [18], spectral signature matching [19], spectral angle mapper, and simplex volume.

4. Unsupervised Spectral Unmixing

Section 3 assumes that pure pixels (contain only one end member) exist in the image. However, thorough hyperspectral mixing may result in the absence of pure pixels. Traditional methods of end member extraction cannot obtain accurate results. In this case, unmixing is considered a blind source separation (BBS) [20] problem called unsupervised spectral unmixing.

Unsupervised spectral unmixing technology based on the information of end members has not been explored yet. Unsupervised signal processing method is employed to obtain end members and abundance according to the spectral model of mixed pixels and end member information, such as constraint conditions. Unsupervised spectral unmixing has become a research hotspot in recent years because it can overcome the limitations of the traditional method and provide a new train for hyperspectral unmixing. The most common methods of unsupervised spectral unmixing include ICA, complexity pursuit, and NMF. The following sections provide a brief introduction to ICA and complexity pursuit as well as a detailed introduction to NMF and its variants.

4.1. Hyperspectral Unmixing based on ICA

ICA, which was introduced by Comon, approximates the mixed sources by assuming that the sources are statistically independent of one another. Observation signal X establishes the objective function in accordance with the principle of statistical independence. It is decomposed into several independent components through an optimization algorithm. In Formula (4), observation signal X is known, and mixed matrix A and sources S are unknown. The hypothesis that sources S are independent and separate as much as possible is adopted.

$$X = A \times S . \quad (4)$$

The results of separation are independent of one another. Linear transformation matrix W is obtained, and X is linearly transformed. Y is an estimate of S as shown below.

$$Y = W \times X = W \times A \times S . \quad (5)$$

ICA involves different estimation methods of obtaining Y and transformation matrix W . These methods include maximum of non-Gaussian [21], minimum of mutual information [22], and maximum likelihood function estimation [23]. Different estimation methods have different objective functions.

In hyperspectral unmixing, spectrum abundance is the source signal. End member extraction is broken down into three steps: preprocessing, ICA, and post-processing [24]. However, the abundance of the end members has sum-to-one constraint. Thus, the distribution does not satisfy the premise of independence and thus reduces the solution mixing precision of ICA.

4.2. Spectral and Spatial Complexity for Hyperspectral Unmixing

In ICA, the spectrum abundances are assumed to be independent and stationary. However, this assumption does not conform to the actual situation. A previous study proposed a complexity-based BSS algorithm [25]. Compared with ICA, the complexity-based BSS algorithm reduces the complexity of the extracted signal as much as possible. Spatial complexity accounts for the spatial autocorrelation of each abundance, whereas spectrum complexity accounts for spectral smoothness. Spatial complexity combined with spectrum complexity based on blind source separation is called SSCBSS.

In SSCBSS, complexity can be formulated by predictability. If the value of a signal is easily predicted by the previous values of signals, the complexity of the signal is low.

Large predictability corresponds to low complexity. The definition of predictability for extracted signal z is

$$F_{CP} = \ln \frac{\sum_{i=1}^n (\bar{z} - z_i)^2}{\sum_{i=1}^n (z_i - z_i)^2}, \quad (6)$$

$$z_i = \lambda \tilde{z}_{i-1} + (1 - \lambda) z_{i-1}. \quad (7)$$

where \bar{z} is the mean value of z , \tilde{z}_i is the short-term moving average to predict z_i , and λ is the predictive rate. Maximizing F_{CP} is the objective function. In SSCBSS, spatial complexity based on Markov random field (MRF) is combined with spectrum complexity to constitute the objective function.

4.3. NMF and its Extensions

Aside from the abovementioned methods of unmixing based on BSS, NMF and its extensions have also been studied recently. NMF, which was proposed by Lee and Seung, is a multiplicative update algorithm. It is often applied to solve the following problem.

Given an $m \times n$ non-negative matrix V , the purpose is to find two non-negative matrices W and H , such that

$$V = W \times H. \quad (8)$$

In hyperspectral unmixing, V represents a set of m -dimensional vectors; each 2D image corresponds to a wavelength as a row of V , and each column of V corresponds to the spectral signature of a pixel [27]. V is approximately factorized into an $m \times r$ matrix W and an $r \times n$ matrix H , where r is the number of end members in the hyperspectral image, W represents the signatures of end members at different wavelengths, and H represents the abundances.

To solve the matrix decomposition indicated above, the Euclidean distance between V and WH is calculated to structure the cost function.

$$\|V - WH\|^2 = \sum_{ij} (V_{ij} - (WH)_{ij})^2, \quad \text{such that } W, H \geq 0. \quad (9)$$

In this cost function, a small Euclidean distance means a good approximate effect. To solve the problem above, the following multiplicative update rules are implemented.

$$H_{\alpha\mu} \leftarrow H_{\alpha\mu} \frac{(W^T V)_{\alpha\mu}}{(W^T W H)_{\alpha\mu}}, \quad W_{i\alpha} \leftarrow W_{i\alpha} \frac{(V H^T)_{i\alpha}}{(W H H^T)_{i\alpha}}.$$

However, NMF has several drawbacks. NMF is an NP-hard nonlinear optimization problem with many local minimum values. The optimal solution is generally non-unique, and only some local minima are identified. Many algorithms based on NMF, such as MVC-NMF [28], PSNMFSC [29], and NMU [31], have been proposed to optimize NMF.

4.3.1. Minimum Volume Constrained Non-Negative Matrix Factorization (MVC-NMF): Considering the drawbacks of NMF, Miao and Qi presented a novel method called MVC-NMF to optimize the NMF algorithm. In this method, the constraints of spectral data are non-negative, and the simplex volume composed of end members is the minimum in all possible simplexes that surround the image data. Thus, the cost function becomes a constrained optimization problem as follows:

$$\begin{aligned} \text{minimize } f(W, H) &= \frac{1}{2} \|V - WH\|^2 + \lambda J(V), \\ \text{subject to } W &\geq 0 \quad H \geq 0 \quad 1_c^T H = 1_N^T. \end{aligned} \quad (10)$$

where $J(V)$ is the simplex volume determined by the estimated end members. Parameter λ is utilized to control the tradeoff of the volume constraint.

4.3.2. PSNMFSC and $L_{1/2}$ Sparsity Constrained Non-Negative Matrix Factorization:

To achieve better estimates, Sen Jia and Yuntao Qian presented constrained non-negative matrix factorization for hyperspectral unmixing in 2009. They added two features of hyperspectral data, namely, the piecewise smoothness of spectral data and the sparseness of abundances, to NMF.

The smoothness constraint can be described by discontinuity adaptive MRF model, which can be expressed as

$$g(x) = -e^{-\frac{x^2}{\gamma}} + 1. \quad (11)$$

Examination of any row w of end member matrix W and its neighborhood w_N shows that the function $g(w - w_N)$ represents the piecewise smoothness of end member w . Likewise, the piecewise smoothness of abundances can be calculated by $g(h - h_N)$, where h is one column of H . Then, the objective function can be expressed as

$$D(W, H) = \frac{1}{2} \|V - WH\|^2 + \alpha \langle g(h - h_N) \rangle + \beta \langle g(w - w_N) \rangle. \quad (12)$$

By adding sparseness constraint to the abundances and NMF with sparseness constraints (NMFSC), the sparseness criteria can be calculated as

$$\text{sparseness}(x) = \frac{\sqrt{n} - (\sum |x_i|) / \sqrt{\sum x_i^2}}{\sqrt{n} - 1}. \quad (13)$$

where n is the dimensionality of x .

In addition to PSNMFSC, $L_{1/2}$ sparsity-constrained NMF was established by Sen Jia and Yuntao Qian in 2011 by modifying NMF and incorporating the $L_{1/2}$ sparsity constraint to NMF [30]. The researchers analyzed L_q ($0 < q < 1$) regularizer and found that is a better choice.

4.3.3. NMU: Non-negative matrix factorization based on underapproximations (NMU) [32] is another technique based on NMF. NMU allows for unmixing in a recursive manner, such as PCA, but preserves non-negativity, such as NMF. Sparse NMU adds a sparse constraint of the abundance matrix. Thus, unmixing with sparse NMU [33] is more efficient than that with NMU.

In sparse NMU, NMU is solved at each step of the recursion as

$$\min_{w \geq 0, h \geq 0} \|V - \sigma w h^T\|^2, \text{ such that } w h^T \leq V. \quad (14)$$

In the formula (14), w represents the signatures of the first extracted end member at different wavelengths while h represents its abundances in the image. Non-negative

residual matrix $R = V - wh^T \geq 0$ is obtained, and in R , the signatures of the first extracted end member are identically zero. Then, the same procedure can be applied to R and extract the second end member. The recursion will stop until all end members be extracted.

The sparse constraint is then added into the NMU model. The abundance matrix is sparse, each end member is typically present in a relatively small number of pixels, and each pixel contains only a small number of materials. Thus, h should be sparse. The recursion can be changed as

$$\min_{w \geq 0, h \geq 0} \|V - \sigma wh^T\|^2 + \mu \|h\|_0, \text{ such that } wh^T \leq V. \quad (15)$$

To obtain approximate solutions to the problem, Lagrangian dual variable Λ is introduced. The Lagrangian dual function is expressed as

$$L(\Lambda) = \min_{w \geq 0, h \geq 0} \|V - \sigma wh^T\|^2 + \mu \|h\|_0 - 2 \langle \Lambda, V - \sigma wh^T \rangle. \quad (16)$$

The Lagrangian dual function is a non-smooth convex optimization problem that can be solved using the sub-gradient scheme. Then, w and h can be update as:

$$h \leftarrow \arg \min_{h \geq 0, \|h\|_2=1} \|(V - \Lambda) - \sigma wh^T\|_F^2 + \mu \|h\|_1, \quad (17)$$

$$w \leftarrow \arg \min_{w \geq 0} \|(V - \Lambda) - \sigma wh^T\|_F^2. \quad (18)$$

Our experimental result contains real hyperspectral data set. The HYDICE Urban hyperspectral image contains 210 spectral bands, and the data have a dimension of $307 \times 307 \times 210$. The data contain six types of materials, namely, road, dirt, trees, roofs, grass, and metal. Figure 5 presents the Urban Dataset and the materials' reflectance in different bands from [32]. The result in Figure 6 is obtained after applying sparse NMU to the dataset.

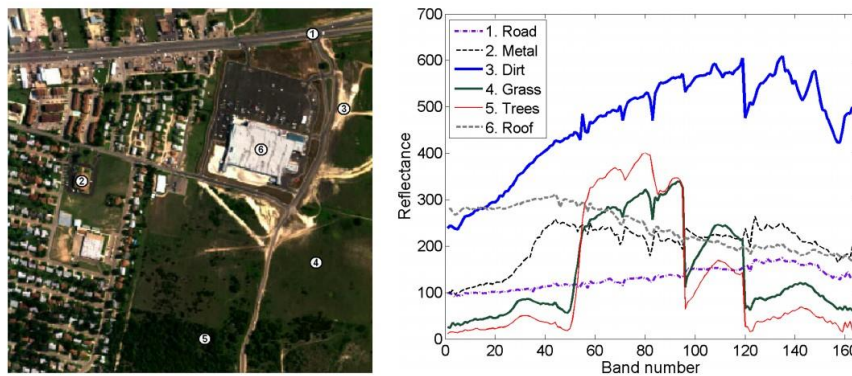


Figure 5. Data set and materials' Reflectance in Different Bands[32]



Figure 6. Result-based NMU for Hyperspectral Unmixing

5. Semi-Supervised Spectral Unmixing

Semi-supervised spectral unmixing [34-35] is different from unsupervised, unmixed amounts to identify the optimal subset from a large spectral library collected on the ground by a field spectral radiometer; the process is called sparse unmixing. This method is called semi-supervised spectral unmixing technology because the spectral library is assumed to be known priori. Sparse unmixing avoids two conventional limitations of the aforementioned classic unmixing approaches: the absence of pure pixels in the image data and the difficult task of estimating the number of end members.

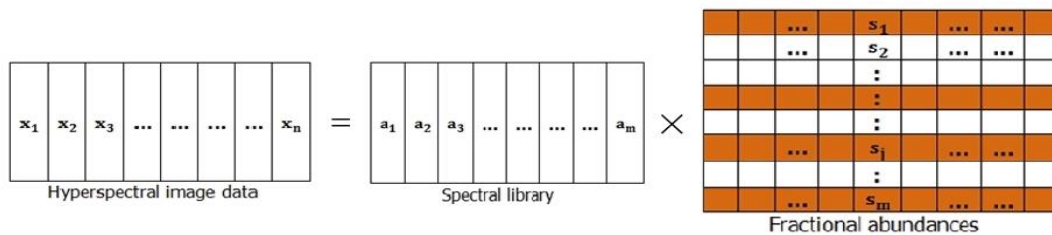


Figure 7. Graphical Illustration of the Performance of the Proposed Collaborative Regularizer

In Figure 7, $A = [a_1 \ a_2 \ \dots \ a_m]$ represents a spectral library with m spectral materials. Sparse linear regression can be implemented to determine the optimal subset. The spectral abundances can be obtained by minimizing an objective function that often includes a quadratic data term and l_1 norm sparsity. The fractional abundances of the matrix in orange color correspond to the end members (called active members) in A with a large probability in the image.

The objective function is defined as

$$\min_x \|AS - X\|_F^2 + \lambda \sum_{k=1}^m \|S^k\|_q. \quad (19)$$

where $\|X\|_F$ is the Frobenius norm, S^k denotes the k -th row of S , and q represents the l_q norm of S^k . Equation (19) is composed of two parts. The first part measures the error of unmixing, and the second part measures the sparsity of the fractional abundances. λ is the relative weight of the sparsity. The ANC and ASC

constraints of abundance are usually applied to the fractional abundances to optimize the algorithm.

To obtain the optimal value of the equation, sparse unmixing is often employed via variable splitting-augmented Lagrangian (SUnSAL). The equation then becomes

$$\min_x \left\{ \frac{1}{2} \|AS - X\|_2^2 + \lambda \sum_{k=1}^m \|S^k\|_1 \right\}, \quad s.t. \quad S^k \geq 0, 1^T S^k = 1. \quad (20)$$

However, the actual spectral library contains many similar signatures. Similar signatures have high mutual coherence, which reduces the level of unmixing via sparse regression. Several suitable regularization terms are applied to the unmixed solution to address the limitation of mutual coherence.

Iordache *et. al.*, proposed a total variation spatial regularizer [37] was applied to sparse unmixing, which represents the spatial contextual information of the image data. The study presented a new algorithm called SUnSAL-TV. Iordache *et. al.*, proposed modified SUnSAL and introduced CLSUnSAL [38]. The article opined that only a small number of end members constitute the image by a linear combination in unmixing. At this point, the pixels share the same active set of end members. Hence, CLSUnSAL enforces joint sparsity among all pixels, whereas SUnSAL employs pixel-wise independent regressions. Iordache *et. al.*, proposed introduced a novel algorithm called multiple signal classification and collaborative sparse regression[39]. Unmixing was divided into two steps. The first step applied the theory of multiple signal classification to unmixing to identify a subset of the spectral library, which contains the end member signatures. The second step is similar to CLSUnSAL, such that it also applies collaborative sparse regression.

6. Conclusions and Future Work

Hyperspectral unmixing continues to be a popular topic in remote sensing image processing. Basing on a linear mixed model, we summarized the traditional methods of end member extraction and abundance decomposition, unsupervised spectral unmixing, and semi-supervised spectral unmixing. Each method has its advantages and disadvantages. To address the disadvantages, we recommend the use of the non-negative and continuity for signal spectrum and distribution as well as the sparse feature of the end members' distribution.

Table 1. The Advantages and Disadvantages for Each Method

	Methods	Advantages	Disadvantages
Traditional methods of end member extraction	PPI	High precision, definite physical meanings, good scientific theory	End members must be present in the image
	N-FINDR		
	VCA		
	IEA		
Unsupervised spectral unmixing	ICA	Simple modeling	Abundance does not satisfy independence
	SSCBSS	Definite physical meanings	Large amount of calculation
	NMF	Definite physical meanings	Optimal solution is non-unique
	NMF's extensions	Optimize the NMF, optimal solution is stable	Computing complex
Semi-supervised spectral unmixing	Sparse unmixing	Easy to understand and high precision	Spectral library must be known priori

Although hyperspectral technology has been existing for only 30 years, it has already garnered great achievements. Future studies should focus on exploring new solutions of unmixed algorithm and improving the robustness and accuracy of traditional methods. Moreover, hyperspectral technology should be applied to ecological systems, cities, resources, medical sciences, and so on.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (No. 81100520, and No. 61105020), the Science and Technology Development Plans of Shandong Province (No. 2012GGE27073 and 2014GSF118142) and the Science and Technology Development Plans of Jinan City.

References

- [1] A. F. H. Goetz, G. Vane, J. E. Solomon and B. N. Rock, *Science*, vol. 228, no. 4704, (1985).
- [2] N. Keshava and J. F. Mustard, *IEEE Signal Processing Magazine*, vol. 19, no. 1, (2002).
- [3] C. I. Chang, *IEEE Transactions on Information Theory*, vol. 46, no. 5, (2000).
- [4] S. Liangrocapart and M. Petrou, *Proc. SPIE, Image and Signal Processing for Remote Sensing IV*, 3500, (1998).
- [5] C. Fevotte and N. Dobigeon, *arXiv*, (2014), pp. 1401.5649.
- [6] J. W. Boardman, *Proc. Ann. JPL Airborne Geosci. Workshop*, vol. 1, (1993).
- [7] J. W. Boardman, F. A. Kruse and R. O. Green, *Proc. JPL Airborne Earth Sci. Workshop*, (1995).
- [8] C. I. Chang and A. Plaza, *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 1, (2006).
- [9] M. E. Winter, *Proc. SPIE, Imaging Spectrometry V*, 3753, (1999).
- [10] J. M. P. Nascimento and J. M. Bioucas-Dias, *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, (2005).
- [11] R. A. Neville, K. Staenz, T. Szeredi, J. Lefebvre, and P. Hauff, Editors, *Fourth International Airborne Remote Sensing Conference and Exhibition /21st Canadian Symp, Remote Sens*, Ottawa, Canada, (1999) June 21-24.
- [12] J. H. Gruninger, A. J. Ratkowski and M. L. Hoke, *Proc. SPIE, Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery X.*, 5425, (2004).
- [13] T. H. Chan, W. K. Ma, A.M. Ambikapathi and C. Y. Chi, *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, (2011).
- [14] A. M. Ambikapathi, T. H. Chan, W. K. Ma and C. Y. Chi, *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2nd Workshop on, (2010) June 14-16.
- [15] D. C. Heinz and C-I. Chang, *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, (2001).
- [16] J. H. Friedman and J. W. Tukey, *IEEE Transactions on Computers*, vol. 23, (1974).
- [17] J. C. Harsanyi and C-I. Chang, *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, (1994).
- [18] J. J. Settle, *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 4, (1996).
- [19] A. S. Mazer, M. Martin, M. Lee and J. E. Solomon, *ELSEVIER, Remote Sensing of Environment.*, vol. 24, no. 1, (1988).
- [20] P. Comon, *ELSEVIER, Signal Process*, vol. 36, no. 3, (1994).
- [21] The Fast ICA MATLAB package. <http://www.cis.hut.fi/projects/ica/fastica>.
- [22] A. Hyvarinen and E. Oja, *ELSEVIER, Neural Networks*, vol. 13, (2000), pp. 4-5.
- [23] D. T. Pham and P. Garat, *IEEE Trans. Signal Processing*, vol. 45, no. 7, (1997).
- [24] J.D. Bayliss, J. A. Gualtieri and R.F. Cromp, *Proc. SPIE, 26th AIPR Workshop: Exploiting New Image Sources and Sensors*, vol. 3240, (1997).
- [25] S. Jia and Y. Qian, *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, (2007).
- [26] D. D. Lee and H. S. Seung, *Nature*, 401, (1999).
- [27] V. P. Pauca, J. Piper and R. J. Plemmons, *ELSEVIER, Linear Algebra and its Applications.*, vol. 416, no. 1, (2006).
- [28] L. Miao and H. Qi, *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, (2007).
- [29] S. Jia and Y. Qian, *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, (2009).
- [30] Y. Qian, S. Jia, J. Zhou and A. Robles-Kelly, *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, (2011).
- [31] N. Gillis and F. Glineur, *ELSEVIER, Pattern Recognition.*, vol. 43, no. 4, (2010).
- [32] N. Gillis and R. J. Plemmons, *SPIE, Optical Engineering.*, vol. 50, no. 2, (2011).
- [33] N. Gillis and R. J. Plemmons, *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 3rd Workshop, (2011) June 6-9.
- [34] B. Natarajan, *SIAMJ. Comput.*, vol. 24, no. 2, (1995).
- [35] P. Sprechmann, I. Ramirez, G. Sapiro and Y. C. Eldar, *IEEE Trans. Signal Process.*, vol. 59, no. 9, (2011).
- [36] J. M. Bioucas-Dias and M.A.T. Figueiredo, *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2nd Workshop on, (2010) June 14-16.

- [37] M.D. Iordache, J. M. Bioucas-Dias and A. Plaza, IEEE Trans. Geosci. Remote Sens., vol. 50, no. 11, (2012).
- [38] M.D. Iordache, J.M. Bioucas-Dias and A. Plaza, IEEE Trans. Geosci. Remote Sens., vol. 52, no. 1, (2014).
- [39] M.D. Iordache, J.M Bioucas-Dias, A. Plaza and B. Somers, IEEE Trans. Geosci. Remote Sens., vol. 52, no. 7, (2014).

Authors



Mingyu Nie, she is currently working toward the M.S. degree at School of Information Science and Engineering, Shandong University, Jinan, China. Her research interests now include hyperspectral unmixing, image processing and pattern recognition.



Zhi Liu, he is an associate professor at Shandong University. In 2005-2006, he worked at the Hong Kong Polytechnic University as a research assistant. He received his Ph.D. degrees in image processing and pattern recognition from the Shanghai Jiao Tong University in 2008. In 2010-2011, he worked as a postdoctor of Technical University of Denmark. His current research interests include pattern recognition, image processing, and machine learning.



Hui Xu, a engineer at Electronic Design Automation Research Institute, School of Information Science and Engineering of Shandong University. Her research interest include integrated circuit design and EDA.



Xiaoyan Xiao, she is MD and Ph.D working on clinical medicine.



Fangqi Su, he received his B.S. degree from Taishan University, Taian, Shandong, China, in 2013. He is currently working toward the M.S. degree at School of Information Science and Engineering, Shandong University. His current research interests include face recognition, finger vein recognition and machine learning.



Jun Chang, he received the B.S., M.S. and PhD degrees in Shandong University in 1988, 1991, 2006 respectively. He was visiting fellow in school of Electrical Engineering and Telecommunications in University of New South Wales during 2004 and 2005. He is currently professor with School of Information Science and Engineering of Shandong University. His research interest include optical fiber sensor and fiber laser.



Xiaomei Li, hi Xiao-Mei is an attending doctor at the Cancer Center of the Second Hospital, Shandong University. She received the Master's degree in 2004 from Shandong University, and is currently pursuing the M.D. degree at Shandong University, directed by Professor Zheng Cheng-Yun, a famous hematologist as well as an oncologist. Her research interests include cancer biotherapy, hemotopathology and multimedia analysis and retrieval.