2012 4th Workshop on Hyperspectral Image and Signal Processing (WHISPERS)

A SPARSITY PROMOTING BILINEAR UNMIXING MODEL

Paul Gader[♯], *Dmitri Dranishnikov*[♯], *Alina Zare*[*], *Jocelyn Chanussot*[†]

[♯]Computer and Information Science and Engineering, University of Florida
[*]Electrical and Computer Engineering, University of Missouri
[+]GIPSA Lab, Grenoble Institute of Technology

ABSTRACT

An algorithm, Bilinear SPICE (BISPICE), for simultaneously estimating the number of endmembers, the endmembers, and proportions for a bilinear mixing model is derived and evaluated. BISPICE generalizes the SPICE algorithm for linear mixing. The proportion estimation steps of SPICE and BISPICE are similar. However, the endmember updates, one novel aspect of the work, are quite different. The SPICE objective function is quadratic in the endmembers. The BISPICE is a fourth degree polynomial. In SPICE, endmembers are updated simultaneously via a closed form. In BISPICE, each endmember must be updated with respect to all other endmembers. Empirically, BISPICE estimated endmembers and proportions more accurately then SPICE, even though the data fitting error was higher.

1. INTRODUCTION

Researchers have focusing intensely on the spectral unmixing problem over the past 10-15 years. A majority of this focus has been on unmixing the Linear Mixing Model (LMM) [1]. Consequently, many linear unmixing algorithms achieve excellent performance and are well-understood. In spite of receiving less attention, nonlinear unmixing algorithms have been investigated steadily throughout the period. The success in unmixing the LMM and documentation of the non-negligible effects of nonlinearities [2, 3] are sufficient to warrant an increased focus on nonlinear unmixing algorithms. Nonlinear mixing can be categorized into several subcategories [1]. Bilinear mixing is the category investigated in this paper. Several researchers have studied bilinear unmixing algorithms recently [4, 5, 6] . The foci have primarily been on devising representations of bilinear mixing to facilitate accurate estimation of proportions, or abundances [1]. Heylen et al. [7] devised a geometric algorithm for estimating endmembers in nonlinear mixing. It differs from BISPICE in that the model used to derive it was based on the generic notion of a manifold and not on the specific, physics-based functional form. In addition, Heylen et al also showed how to compute geodesic distance in a manifold defined by a bilinear model [8]. That algorithm could be very useful if incorporated with the algorithms described herein since error measures could be defined in the manifold .

New bilinear unmixing algorithms based on the computational model proposed by Nascimento and Dias (referred to here as NASDI) [4] are derived in this paper. The SPICE algorithm [9] is extended to NASDI . This Bilinear SPICE (BISPICE) algorithm provides the ability to perform full unmixing: estimating proportions, endmembers, and the number of endmembers.

2. COMPUTATIONAL MODEL AND ALGORITHM

The BISPICE algorithm estimates proportions, endmembers, and the number of endmembers for a mixture modeled by the NASDI bilinear model. It uses alternating optimization on proportions and endmembers; that is, BISPICE alternates between updating endmembers while holding proportions fixed and updating proportions while holding endmembers fixed. The proportion updating step is almost identical to SPICE. The endmember updates are quite different. In this section, the NASDI model is defined, the derivatives of the NASDI model with respect to the endmembers are calculated. The SPICE terms are then defined. Then, the BISPICE objective function is defined and the algorithm is derived.

[1]For brevity, we use proportions to refer to fractional abundances

978-1-4799-3406-5/12/$31.00©2012IEEE

2.1. The NASDI models

The notation is : $\vec{y}_n$ is the n^{th} spectrum in a hyperspectral image with N spectra, M is the number of endmembers, $\mathbf{e}_i$ is the i^{th} endmember, a_{in} is the proportion of material i reflected in pixel n and $\mathbf{b}_{ijn}$ is the proportion of the bilinear mixture of the i^{th} and j^{th} materials reflected in pixel n. The NASDI model is then given by

$$\begin{aligned} \vec{y}_n &= f(\mathbf{E}, \mathbf{a}_n, \mathbf{b}_n) := \\ &:= \sum_i^M a_{in}\mathbf{e}_i + \sum_i^M \sum_{j>i}^M b_{ijn}\mathbf{e}_i \odot \mathbf{e}_j \end{aligned} \tag{1}$$

where $\forall i, j, n; a_{in}, b_{ijn} \in [0,1]$ and

$$\sum_i^M a_{in} + \sum_i^M \sum_{j>i}^M b_{ijn} = 1 \tag{2}$$

2.2. Differentiating the NASDI model

The endmember update formulas are derived by differentiating an objective function. Thus, the derivative of the NASDI model with respect to the endmembers are required. Note that $\mathbf{e}_i \odot \mathbf{e}_j := D(\mathbf{e}_i)\mathbf{e}_j = D(\mathbf{e}_j)\mathbf{e}_i$, where $D(\mathbf{e}_i)$ is a matrix with $\mathbf{e}_i$ on the diagonal. Thus, the model for the n^{th} spectral pixel can be written as

$$f(\mathbf{E}, \mathbf{a}_n, \mathbf{b}_n) = \sum_i^M a_{in}\mathbf{e}_i + \sum_i^M \sum_{j>i}^M b_{ijn} D(\mathbf{e}_j)\mathbf{e}_i \tag{3}$$

Separating $\mathbf{e}_k$ from the other terms leads to:

$$f(\mathbf{E}, \mathbf{a}_n, \mathbf{b}_n) = g(\mathbf{a}_n, \mathbf{b}_n)\mathbf{e}_k + f(\mathbf{E}_{i \neq k}) \tag{4}$$

where

$$g(\mathbf{a}_n, \mathbf{b}_n) := \mathbf{I}a_{kn} + \sum_{i \neq k} b_{i,k} D(\mathbf{e}_i) \tag{5}$$

$$f(\mathbf{E}_{i \neq k}) = \sum_{i \neq k}^M a_{in}\mathbf{e}_i + \sum_{i \neq k}^M \sum_{j>i, j \neq k}^M b_{ijn} D(\mathbf{e}_j)\mathbf{e}_i \tag{6}$$

Note that $g(\mathbf{a}_n, \mathbf{b}_n)$ and $f(\mathbf{E}_{i \neq k})$ are not functions of $\mathbf{e}_k$. Hence, the derivative of $f(\mathbf{E}, \mathbf{a}_n, \mathbf{b}_n)$ is simply

$$\frac{\partial f(\mathbf{E}, \mathbf{a}_n, \mathbf{b}_n)}{\partial \mathbf{e}_k} = g(\mathbf{a}_n, \mathbf{b}_n) \tag{7}$$

2.3. SPICE and BISPICE

SpicE is based on the LMM. It is described in [9]; essential details are given here for completeness. Let $\mathbf{A}$ denote the $M\mathbf{x}N$ matrix whose n^{th} column is the proportion vector $\mathbf{a}_n$. SPICE is derived by differentiating an objective function. The objective function includes a sparsity promoting term, which distinguishes SPICE from the ICE algorithm upon which it is based. The objective function is

$$\begin{aligned} &J(\mathbf{E}, \mathbf{A}, \gamma; \mu, \Gamma_L) \\ &:= (1-\mu)R_L(\mathbf{E}, \mathbf{A}) + \mu V(\mathbf{E}) + S(\mathbf{A}, \gamma; \Gamma_L) \end{aligned} \tag{8}$$

where

$$R_L(\mathbf{E}, \mathbf{A}) := \sum_{n=1}^N \|(\mathbf{y}_n - \sum_i^M a_{in}\mathbf{e}_i)\|_2^2 \tag{9}$$

is the error term,

$$V(\mathbf{E}) = \frac{1}{M(M-1)} \sum_{k=1}^{M-1} \sum_{i=k+1}^M \|(\mathbf{e}_k - \mathbf{e}_i)\|_2^2 \tag{10}$$

is the sum of the squared distances between all pairs of endmembers, $\gamma = (\gamma_1, \ldots, \gamma_M)^t$, and

$$S(\mathbf{A}, \gamma; \Gamma_L) = \sum_{k=1}^M \gamma_k \sum_{i=1}^N a_{in} \tag{11}$$

is a sparsity promoting term. Holding endmembers fixed, the objective function is quadratic in the proportions (including the sparsity term). They are estimated using Quadratic Programming (QP). Holding proportions fixed, there is a closed form update formula for the endmembers [9]. The relationship between the number Γ_L and γ clarified in the SPICE/BISPICE algorithm pseudo-code later. The BISPICE objective is similar to that of SPICE except the term R_L is replaced by

$$R_B(\mathbf{E}, \mathbf{A}, \mathbf{B}) := \sum_{n=1}^N \|(\mathbf{y}_n - f(\mathbf{E}, \mathbf{a}_n, \mathbf{b}_n))\|_2^2 \tag{12}$$

where $\mathbf{B}$ is the $\frac{M(M-1)}{2}\mathbf{x}N$ matrix whose columns are the proportion vectors of the corresponding pixel with respect to the bilinear terms. There are two sparsity parameters, Γ_L and Γ_B, for linear and for bilinear.

Surprisingly, there is a closed form update formula for the endmembers, even though the bilinear model objective function R_B is a fourth degree polynomial in the endmembers as opposed to the second degree polynomial in the case of the linear mixing objective. There is a significant difference however. The updates for BISPICE require that each endmember be updated while holding the other endmembers fixed whereas in SPICE the endmember updates are simultaneous. The BISPICE endmember updates are given by

$$\mathbf{e}_k = [\mathbf{W}_k]^{-1}\left(\frac{\mu}{M(M-1)}\mathbf{S}_{\bar{k}} + \frac{1-\mu}{N}\boldsymbol{\Delta}_{\bar{k}}\right) \quad (13)$$

where

$$\mathbf{W}_k := \left(\frac{1-\mu}{N}\right)\sum_{n=1}^{N} g(\mathbf{a}_n, \mathbf{b}_n)^2 + \frac{\mu}{M}\mathbf{I} \quad (14)$$

$$\mathbf{S}_{\bar{k}} := \sum_{j \neq k}^{M} \mathbf{e}_j \quad (15)$$

$$\boldsymbol{\Delta}_{\bar{k}} := \sum_{n=1}^{N} g(\mathbf{a}_n, \mathbf{b}_n)(\mathbf{y}_n - f(\mathbf{E}_{i \neq k})) \quad (16)$$

These equations are used to update each of the endmembers as a function of all the other endmembers.

The proportion update step proceeds in the same way as SPICE, except with separate values of Γ_B for the bilinear proportions. Endmembers with low overall proportions are discarded, again similar to SPICE, with the caveat that if a non cross-term endmember is discarded all of its cross terms are discarded also. Thus, the parameters required for full unmixing of the NASDI model are estimated using the BISPICE model through alternating optimization of the endmembers and proportions. The pseudo-code is presented here with both **SPICE** and **BISPICE** to clarify the similarities and differences in the algorithms.

Algorithm: **SPICE / BISPICE**

1. *Initialize*
2. **DO** *until convergence*
   (a) *Solve for proportions using QP*
   (b) *for* $i = 1, \ldots, M$
      i. $\gamma_i = \frac{\Gamma_L}{\sum_{n=1}^{N} a_{in}}$
      ii. if BISPICE, for $j = i+1, \ldots, M$
         A. $\gamma_{ij} = \frac{\Gamma_B}{\sum_{n=1}^{N} b_{ijn}}$
      iii. *if* $\max_{n=1}^{N} a_{in} < T$, *prune* $\mathbf{e}_i$ AND
      iv. if BISPICE,$\forall j$, prune $\mathbf{e}_i \odot \mathbf{e}_j$
      v. if BISPICE,$\forall j$, if $\max_{n=1}^{N} \mathbf{e}_i \odot \mathbf{e}_j < T$, prune $\mathbf{e}_i \odot \mathbf{e}_j$
   (c) *Solve for endmembers using closed form*
      i. *if SPICE, use Eq. (7), [9]*
      ii. *if BISPICE, use Eqs.* (13)-(16)

The update of the γ_k's accelerates driving proportions to zero.

3. EXPERIMENTS

One set of controlled experiments is defined here. Additional experiments with real data are ongoing but not included due to space. The controlled experiments used data generated from spectra obtained from the ASTER library. Clay, Marble, and Sand spectra were taken from ASTER. Their wavelengths were clipped to the 0.4 - 2.5 μm range, and the resulting spectra were dimensionally reduced using window averaging to yield 3, 570-dimensional spectra (Fig. 1). These spectra were mixed using the NASDI model with proportions drawn from a uniform Dirichlet distribution. In this manner a synthetic dataset of 1000 spectra was created.

The dataset was then unmixed with the BISPICE algorithm, with initializations of $\Gamma_B = 0.01$, $\Gamma_L = 0.5$, and $\mu = 10^{-4}$, and an initial estimate of 8 endmembers (and corresponding 28 bilinear terms). For comparison we also ran ICE with 6 endmembers, and identical parameters. The evaluation measures are the MSE for the estimated and true endmembers ($\mathbf{MSE_{EM}}$) including the true cross terms, the MSE of the model and the data ($\mathbf{MSE}$), and the RMSE between the estimated and true proportions ($\mathbf{RMSE_P}$) including the proportions of the cross terms. The results of this experiment are in the following table.

Table 1. ASTER Results

Measure	ICE	BISPICE
$\mathbf{MSE}(10^{-3})$	0.116	0.338
$\mathbf{MSE_{EM}}$	1.529	0.611
$\mathbf{RMSE_P}$	0.251	0.117

978-1-4799-3406-5/12/$31.00©2012IEEE

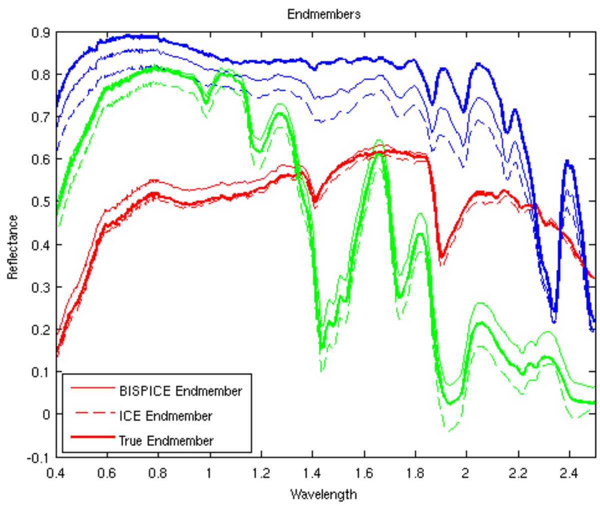

Fig. 1. This figure shows (in bold) Red - Clay, Blue - Marble, Green - Sand, with the estimations by BISPICE (solid line) and ICE (dotted line)

BISPICE correctly estimated 3 endmembers and 3 bilinear terms. Although ICE fits the data better, the ICE endmembers are not as accurate as those found by BISPICE. The MSE result is not surprising since ICE has more free parameters (6 endmembers) than BISPICE (3 linear and 3 cross terms). BISPICE also estimates number of endmembers. The endmember and proportion estimation accuracy of BISPICE is excellent.

4. CONCLUSION AND FUTURE WORK

There are many investigations to conduct. Robustness of BISPICE parameters is very important. Experiments on real data are ongoing and were not included here due to space limitations. Estimating endmembers for other bilinear models is also under investigation. BISPICE should be compared to the methods of Heylen et al. A comparison of computational models derived from reflectance theory, to those derived from generic models of manifolds would be very useful to see if either class provides better solutions or if both yield the same level of performance and interpretability.

5. REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sensing*, 2012, to appear.

[2] T. Han and D.G. Goodenough, "Investigation of nonlinearity in hyperspectral imagery using surrogate data methods," *IEEE Trans. Geoscience Remote Sensing*, vol. 46, no. 10, pp. 2840 –2847, oct. 2008.

[3] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44 –57, jan 2002.

[4] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image and Signal Processing for Remote Sensing XV*, Sept. 2009, vol. 7477, pp. 74–77.

[5] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Bilinear models for nonlinear unmixing of hyperspectral images," in *IEEE GRSS Workshop Hyperspectral Image and SIgnal Processing: Evolution in Remote Sensing (WHISPERS)*, Lisbon, Portugal, June 2011, pp. 1–4.

[6] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geoscience Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[7] R. Heylen, D. Burazerovic, and P. Scheunders, "Non-linear spectral unmixing by geodesic simplex volume maximization," *IEEE Trans. Selected Topics Signal Processing*, vol. 5, no. 3, pp. 534 –542, june 2011.

[8] R. Heylen and P. Scheunders, "Calculation of geodesic distances in nonlinear mixing models: Application to the generalized bilinear model," *IEEE Geoscience Remote Sensing Letters*, vol. PP, no. 99, pp. 1 –5, 2012.

[9] A. Zare and P. Gader, "Sparsity promoting iterated constrained endmember detection in hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 3, pp. 446 –450, july 2007.

978-1-4799-3406-5/12/$31.00©2012IEEE