HANDLING SPECTRAL VARIABILITY WITH ALTERNATING ANGLE MINIMIZATION

Rob Heylen, Paul Gader
CISE, University of Florida
E301 CSE Building, P.O. Box 116120
Gainesville, FL 32611, USA

Paul Scheunders
iMinds-Visionlab, University of Antwerp
Universiteitsplein 1
2610 Wilrijk, Belgium

ABSTRACT
We present an algorithm for solving the spectral mixture analysis problem in situations that show spectral variability, represented by spectral libraries for each class of interest. The aim is to select a single optimal endmember candidate from each library, on a per-pixel basis. The algorithm exploits the high dimensionality of spectral data, and is based on an alternating spectral angle minimization scheme. The resulting algorithm is computationally very efficient, and yields excellent results compared to other techniques, such as multiple endmember spectral mixture analysis.

1. INTRODUCTION

A well-known problem in hyperspectral image processing is the possible presence of a large spectral variability [1, 2] within classes of interest, such as certain minerals or vegetation types. This poses a problem for spectral mixture analysis [3], as each class can no longer be represented by a single endmember. Since the properties of interest are the class abundances rather than sample or endmember abundances, traditional unmixing techniques can no longer be used in a straightforward way in cases with high variability.

Spectral variability is often captured by collecting a large set of representative spectra into a spectral library for each class. A popular technique for unmixing a target pixel with respect to these class libraries is multiple endmember spectral mixture analysis (MESMA) [4], which selects a single optimal member from each library to perform the unmixing. In practice, all possible combinations of endmembers are iteratively evaluated until a stopping criterion is met. Because this approach has a combinatorial complexity, MESMA can only be applied in cases with a low number of relatively small libraries, or with a very relaxed stopping criterion.

In its simplest form, MESMA solves the unconstrained least-squares unmixing [5] problem, and calculates the resulting reconstruction error (RE), for every possible set of endmembers. The set of endmembers, also named model, which respects the constraints on the abundances and has the smallest RE, is then retained as the solution. This process is executed for every pixel in the image, further aggravating the computational complexity. Usually a threshold error is introduced, which stops the search after a RE below this level is found. This approach leads to inconsistent results however, since typically several models will fall below this threshold.

In this paper, we present an alternative technique that allows one to quickly identify models that yield small REs, and that does not suffer from the combinatorial scaling one observes in MESMA. The technique is based on the observation that the high-dimensional convex hull spanned by any hyperspectral data set has a volume that goes to zero, and typically each target pixel will lie outside the convex hull of the library spectra. This allows the construction of the alternating angle minimization (AAM) algorithm, where at each iteration the optimal member of each library is sequentially selected. This algorithm opens up the possibility to unmix hyperspectral images with respect to large class libraries, yielding abundance maps that describe fractional class memberships. Iterative optimization schemes have been used before in hyperspectral image processing, e.g., to generate sparse unmixing results for avoiding the over-fitting problem [6].

2. THE ITERATIVE SCHEME

In spectral mixture analysis, one usually assumes the linear mixing model, which states that every pixel $x$ in a spectral space $\mathbb{R}^d$ can be composed of a linear combination of $p$ endmembers in an endmember matrix $E = (e_1, \ldots, e_p)$. Typically one assumes that the corresponding abundances have to be positive and sum to one. From a geometrical point of view, positivity will restrict solutions to the positive cone spanned by the endmembers in spectral space, and the sum to one constraint forces solutions to lie on the hyperplane through the endmembers. Both constraints together hence only allow solutions in the simplex $S(E)$ spanned by the endmembers:

$$y \in S(E) \iff \exists \{a_i \geq 0\}_{i=1}^p : \left\{ \begin{array}{l} \sum_i a_i = 1 \\ y = \sum_i a_i e_i \end{array} \right.$$  \hspace{1cm} (1)

If one considers a fixed endmember set and a single target pixel, then the RE corresponds to the Euclidean distance from this target to the closest point inside the simplex. By defining
Fig. 1: Two libraries (black and red points) in 2 dimensions, and a target to unmix (green point). The optimal solution is indicated with a gray line. The RE is the perpendicular distance to this line. Left: The target and some library spectra are inside the convex hull, and the problem is hard to solve. Right: No point is inside the convex hull of any set. The optimal model is easy to determine.

\[ P_E(x) \] as the projection operator onto the simplex spanned by the points in \( E \)

\[ y = P_E(x) \iff \forall z \in S(E) : \| z - x \| \geq \| y - x \| \quad (2) \]

the RE is given by \( \| x - P_E(x) \| \).

However, in spectral mixture analysis, there is a set of candidates for each endmember, and one is interested in finding the set, or model, that minimizes this RE. More formally, consider a target \( x \), and \( p \) spectral libraries \( \{ L_i \}_{i=1}^p \), with \( L_i = \{ e_i^1, \ldots, e_i^{N_i} \} \), and \( N_i \) the number of spectra in library \( L_i \). We want to find the index set \( \{ I_i \}_{i=1}^p \), \( I_i \in \{ 1, \ldots, N_i \} \) that identifies the model with the smallest RE:

\[ I = \arg \min_{\{ I_1, \ldots, I_p \}} \| x - P_E(x) \|_2, \quad E = \left( e_1^1, \ldots, e_p^p \right) \quad (3) \]

Without making any assumptions, one is forced to calculate the RE for every one of the \( \prod_{i=1}^p N_i \) index sets, which is a combinatorial problem. However, by using the curse of dimensionality to our advantage, it is possible to create a much faster geometrical scheme that will yield good solutions.

Since the spectral dimension \( d \) is very large, the volume of any polygon spanned by any set of points will be very small due to the curse of dimensionality. This means that the probability of finding a point inside the convex hull of any subset of points will be very small as well [7]. Any target pixel \( x \) one wants to unmix within the convex hull spanned by all library spectra. See Fig. 1 for a simple example in two dimensions. In practice, for the AVIRIS Cuprite and Indian Pines data sets, it is possible to explicitly construct a plane separating each pixel from all other pixels, showing that every pixel will indeed lie on the convex hull surface.

This observation can be exploited to create an alternating optimization scheme. Suppose one has randomly selected a model as an initial solution, corresponding to a randomly selected endmember from each library. If one disregards the positivity constraint on the abundances, the RE corresponds to the orthogonal distance to the endmember plane. If we are allowed to replace the first endmember with another one from the first library, the ideal endmember at this point is the one that minimizes the dihedral angle with the target, with respect to the pivot induced by all other endmembers in the model (See also Fig. 2). Since these angles are easy to determine, the best candidate from the first library can be quickly determined. We can then iteratively cycle through all endmembers, and repeat this process until a (local) optimum has been reached, or a maximum number of iterations has been performed. Since this process disregards the positivity constraint, fully-constrained least-squares unmixing is performed once the endmembers have been determined. This process will decrease the (unconstrained) RE at each iteration, but it is not guaranteed that the global minimum will be reached.

Note with \( O_F(x) \) the orthogonal projection of \( x \) onto the plane through the points in the set \( F \). The AAM algorithm is then given in Algorithm 1. The input of the algorithm is the set of endmember libraries. These libraries do not need to be mutually exclusive, and a single library may be chosen for every endmember, as long as care is taken to avoid repetition of endmembers. In the present version of the AAM algorithm, the number of iterations is fixed for every target pixel.

Algorithm 1: AAM algorithm

Select a number of iterations \( K \).

Select a random index set \( \{ I_i \}_{i=1}^p \), \( I_i \in \{ 1, \ldots, N_i \} \).

for iteration \( \in [1, \ldots, K] \) do

for \( i \in [1, \ldots, p] \) do

\[ F = \{ e_1^1, \ldots, e_p^p \} \]

\[ F_i = F / e_i^i \]

\[ G_i = F_i \cup x \]

for \( n \in [1, \ldots, N_i] \) do

\[ p_n = \arcsin \left( \frac{\| e_n^i - O_{F_i}(e_n^i) \|}{\| e_n^i \|} \right) \]

if \( (x - O_{F_i}(x)) \cdot (e_n^i - O_{F_i}(e_n^i)) < 0 \) then

\[ p_n = \pi - p_n \]

\[ I_i = \arg \min_{\{ p_n \}_{n=1}^N} \]

end if

end for

end for

end for

end for
Fig. 3: Histogram of the logarithm of the REs obtained by fully-constrained unmixing of a single point with respect to all 161300 possible 3-endmember combinations. The lowest value is -2.98, while the AAM algorithm returned -2.78.

3. EXPERIMENTS

3.1. Comparison with optimal solution

We have executed the AAM on 1000 randomly selected pixels from the AVIRIS Cuprite scene, and have unmixed them with respect to three libraries, each containing the same 100 randomly selected spectra from the USGS spectral database. This yields 161300 possible unique models. By fully-constrained unmixing of all possible combinations with the fast SPU algorithm [8], the optimal solution in terms of RE could be determined. To assess the AAM algorithm, we compare its RE with the optimal solution. The number of iterations $K$ was chosen 10.

If we unmix a pixel with respect to all possible models, a typical RE distribution is given in Fig. 3. The result obtained by the AAM algorithm in this case did not correspond to the optimal model, but the difference in RE is minimal. We repeated the same experiment for all 1000 pixels, and plotted a histogram of the REs obtained by the AAM algorithm and the true minimal errors in Fig. 4. Overall, the models found by AAM will have slightly higher REs, but will be of similar magnitude for any target.

Next, we computed the ratio between the errors obtained by both approaches. A histogram of this ratio is plotted in Fig. 5, and shows that in 460 out of 1000 cases, the optimal model is correctly identified. Whenever another model is found, the associated error stays comparable in magnitude to the optimal value. Furthermore, since the AAM algorithm uses a random initialization, processing the same point multiple times can improve the model RE. The main advantage of the AAM algorithm is the improved runtime: 1.02 seconds versus 1506 seconds for the enumeration approach in these experiments.

3.2. Real-world application: Indian Pines

As a real-world application of the AAM algorithm, we have processed the well-known Indian Pines data set, displayed in Fig. 6, in following way: From each of the 16 known classes present in this data set, we randomly selected 5% of the pixels to construct 16 libraries. Next, we have unmixed the entire data set with respect to these 16 libraries with the AAM algorithm with $K = 10$, and have plotted the abundance maps. The total number of candidate models was $6.15 \cdot 10^{19}$, so it is infeasible to iterate all of them.

With this approach, the abundance maps can be considered as fractional membership maps with respect to spectral classes represented by sample libraries. Some examples of such abundance maps are given in Fig. 7. In Fig. 7 (a), the abundance and classification maps of three classes of soy vegetation in different physical configurations are displayed. From these maps, it is clear that these classes show spectral mixing, and that the best reconstructions can be obtained by

Fig. 4: Histogram of the REs obtained by the optimal method and the AAM algorithm, over 1000 points of the Cuprite data set unmixed with respect to 3 libraries containing the same 100 randomly selected endmembers from the USGS database.

Fig. 5: Histogram of the error ratio.

Fig. 6: Approximate RGB image of the Indian Pines data set.
mixing members from two or three classes. This provides a possible solution for the well-known problem of classification and labeling ambiguity: Classes are not always mutually exclusive, and labeling errors might be present. A similar conclusion can be drawn for the grass classes in Fig. 7 (b). While in general, the abundance and classification maps correspond nicely, areas with low grass content can be observed, corresponding to mixing with other classes, such as the woods and the buildings-grass-trees-drives classes.

From a computational point of view, the AAM algorithm takes 33 minutes to process the Indian Pines experiment on a single-core laptop processor. This is larger than the time required for traditional unmixing, but much smaller than the classical MESMA approach would take on the same data set. The algorithm can be sped up by inclusion of steps that detect if convergence has been reached, and by parallelization.

4. CONCLUSIONS

We have introduced a new algorithm for spectral mixture analysis based on iterative angle minimization. This alternating angle minimization algorithm is capable of dealing with endmember libraries instead of a single set of endmembers, selecting the optimal endmembers from each library on a per-pixel basis. This opens up the possibility to unmix with respect to classes, represented by sample libraries.

We have shown that for small problems, the results obtained by the AAM algorithm are comparable to those obtained by the MESMA algorithm. For real problems, such as class unmixing of the Indian Pines data set, the AAM algorithm gives realistic class abundance maps, and provides a possible solution for the class ambiguity problem.

Future work includes the construction of a stopping criterion, which will improve the computational performance, and a theoretical investigation on the influence and possible inclusion of the abundance positivity constraint.

5. REFERENCES


