Variability of the endmembers in spectral unmixing: recent advances
Lucas Drumetz, Jocelyn Chanussot, Christian Jutten

To cite this version:

HAL Id: hal-01358368
https://hal.archives-ouvertes.fr/hal-01358368
Submitted on 31 Aug 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
VARIABILITY OF THE ENDMEMBERS IN SPECTRAL UNMIXING: RECENT ADVANCES

L. Drumetz\textsuperscript{a}, J. Chanussot\textsuperscript{b,c}, C. Jutten\textsuperscript{a}

\textsuperscript{a}Université Grenoble Alpes, \textsuperscript{b}Grenoble Institute of Technology, Gipsa-lab, Grenoble, France
\textsuperscript{c}Faculty of Electrical and Computer Engineering, University of Iceland, Reykjavik, Iceland

ABSTRACT
Endmember variability has been identified as one of the main limitations of the usual Linear Mixing Model, conventionally used to perform spectral unmixing of hyperspectral data. The topic is currently receiving a lot of attention from the community, and many new algorithms have recently been developed to model this variability and take it into account. In this paper, we review state of the art methods dealing with this problem and classify them into three categories: the algorithms based on endmember bundles, the ones based on computational models, and the ones based on parametric physics-based models. We discuss the advantages and drawbacks of each category of methods and list some open problems and current challenges.

Index Terms— Spectral Unmixing, Endmember variability, Spectral Bundles, Extended Linear Mixing Model

1. INTRODUCTION
Spectral Unmixing (SU) is one of the most important and most studied topics in hyperspectral imaging. As a matter of fact, the fine spectral resolution provided by these images allows the accurate identification and characterization of the materials found in the observed scene. However, the spatial resolution is conversely coarser than in standard RGB or multispectral images. Thus, many pixels are mixed, meaning that the acquired spectral signature results from the combination of the spectral signatures of the different materials present in the resolution cell during the acquisition. For a given image, SU aims at extracting the signatures of the different involved materials (called “endmembers”), and at estimating their relative proportions (called “fractional abundances”) in each pixel of the image, allowing a refined understanding of the scene [1]. Over the last decades, many algorithms have been designed for this application. Most of these methods assume that the mixing process linking the signature of the pure materials and their proportions is linear. While this is very often a good approximation, nonlinear interactions are sometimes not negligible (for instance in the case of multiple reflections of the light in its path from the ground to the sensor, or indirect lighting, or intimate mixtures…). Many algorithms taking this type of phenomena into account have also been developed over the last few years [2]. However, another implicit hypothesis is often assumed when performing SU: each material is usually represented by one single spectral signature, that is used for the whole image. In practice, there often exists a non negligible intra-class variability, which can have many different causes. For example, the illumination conditions can vary along the observed scene, because of a nonflat topography or because of changes in the lighting conditions in the scene. The intrinsic variability of the materials (due to the variation of a hidden parameter, such as the concentration of chlorophyll or the hydric stress for green vegetation) can also considerably influence the spectral signature of the corresponding material, even if a pixel is pure. This issue was neglected for a long time, but has recently received a growing attention in the community, with milestone reviews on the subject [3, 4]. Taking the spectral variability into account has now become an important challenge in spectral unmixing.

In the remainder of this paper, we formalize the spectral variability problem in SU in its most general form, by considering variability inside an image (in the spatial domain). We also connect this general form with the multi-temporal and the multi-angular points of view regarding variability. Then we present and classify different existing methods. We eventually point at future research avenues.

2. A GENERAL FRAMEWORK FOR SPECTRAL VARIABILITY
In this section, we introduce the spectral variability problem in its most general form. Usually, once a mixing model (linear or not) has been chosen, the “blind” SU problem boils down to: 1) estimating the number of endmembers to consider, then 2) extracting the spectral signatures of these endmembers (usually by resorting to one of the many available algorithms [1]), before 3) estimating the corresponding relative abundances. Once the first two steps have been performed, the abundances have to be estimated from the following model:

\[ \mathbf{X} = \mathbf{f(S, A)} + \mathbf{E} \quad (1) \]

where \( \mathbf{X} \in \mathbb{R}^{L \times N} \) is the data matrix (\( L \) is the number of spectral bands, \( N \) is the number of pixels), \( \mathbf{S} \in \mathbb{R}^{L \times P} \) is the endmember matrix (containing the \( P \) spectral signatures of the \( P \) endmembers considered in its columns), \( \mathbf{A} \in \mathbb{R}^{P \times N} \) is the abundance matrix, and \( \mathbf{f} \) is a function, linear or not and possibly parametrized, of the previous two variables, which represents the actual mixing model. \( \mathbf{E} \in \mathbb{R}^{L \times N} \) is an additive noise. The abundances are usually interpreted as percentages, which constrains them to be nonnegative and to sum to one. These properties of the abundances are fundamental since they constrain the data to lie inside a simplex spanned by the endmembers, providing a clear geometric interpretation to the problem (see Fig. 1 in the case of \( P = 3 \) endmembers). In the linear case, Eq. (1) becomes

\[ \mathbf{X} = \mathbf{SA} + \mathbf{E} \quad (2) \]

In this case, once the endmembers’ signatures have been estimated, the abundances (constrained in each pixel to lie in the unit
simplex, which we denote by $A \in \text{spl}_d$ are extracted under the positivity and sum-to one constraints by:

$$\hat{A} = \arg\min_{A \in \text{spl}_d} ||X - SA||_F^2$$

Dealing with spectral variability can be seen as considering that the source matrix $S$ is not constant in space, time or, more generally, between different datasets. Mathematically, let us consider a dataset, which is partitioned into $K$ subsets indexed by $k$, and the corresponding endmember signatures:

$$\mathcal{X} = \{X_k\} \quad \text{and} \quad \mathcal{S} = \{S_k\} \quad \text{for} \quad k = 1, ..., K.$$  \hfill (4)

The index $k$ can denote a partition in the spatial domain (mostly at the pixel scale, but also possibly at the scale of larger spatial regions), or denote several datasets if we deal with a sequence of images acquired over the same scene at different time dates (temporal variability), or with different acquisition angles (angular variability), or, more generally, between distinct datasets (which are supposed to share at least one endmember).

In each subset, we have:

$$X_k = f(S_k, A_k) + E_k$$ \hfill (5)

Of course, the different subsets may not be independent: there are usually relationships and some sort of continuity between certain parameters, be it in time, space... This advocates for a joint processing of the data rather than a totally separate processing of each subset [5].

The remainder of this paper will focus (unless stated otherwise) on spectral variability in the spatial domain, but works dealing with temporal variability will also be reviewed, as well as other aspects in the perspectives. Note that most of the presented methods can be easily extended to other types of variability (temporal in particular). Likewise, most of the methods below apply primarily in a linear mixing model framework.

3. DIFFERENT CLASSES OF METHODS

The different methods dealing with spectral variability can be classified into three categories (c.f. Fig. 2 (a), where all three cases are represented in the case of a linear mixing model). The first category resorts to “spectral bundles”, that is to say each material is represented by a finite number of signatures which hopefully capture most of its variability. The second class (Fig. 2 (b) uses reference sources and uses models to characterize and constrain possible local evolutions of the endmembers, typically using statistical descriptions. The third one (Fig. 2 (c) and (d)) mostly corresponds to physics-guided models: in this case the evolution of each material is governed by a parametric model, whose parameters have to be estimated in each pixel, along with the abundances.

In all three cases, the mathematical tools used range from machine learning to Bayesian estimation and mathematical optimization.

3.1. Spectral bundles and associated methods

The key idea in this category of methods consists in extracting several instances of each endmember in order to build a dictionary, which is then used for SU, offering more than one representative spectral signature per material (Fig. 2 (a)). The simplest method to build those bundles [6] consists in extracting endmembers in randomly chosen subsets of the image. Many endmember extraction algorithms assume that there is at least one pure (abundance coefficient equal to 1) pixel per material in the image. Here, there has to be at least one pure pixel of each material in each subset. Once the extraction has been carried out, the signatures have to be grouped into classes through a clustering algorithm (e.g. k-means). Then, the dictionary is used to estimate the abundances. In order to get the total abundance of one given material, one has to sum the contributions of every representative of the corresponding class. As the dictionary may contain a lot of signatures, it can be useful to enforce different types of sparsity during the SU process, so as to select in each pixel the few signatures that actually best explain the variability. This sparsity can be global, or can be refined by taking into account the group structure of the dictionary [7].

Other works use a slightly different but related approach, in which the whole unmixing is performed locally, in spatially correlated subsets (sliding windows or segmentations of the image). In this case, the abundances are estimated locally, based on the idea that in smaller regions, variability (and some nonlinear) effects are mitigated with respect to an entire image. An example using sliding windows can be found in [8], whereas the work of [9] goes further
by proposing to build a segmentation of the image which is optimal in terms of local spectral unmixing performance. Other types of methods are more machine learning-oriented (using for instance Support Vector Machines [10], or Gaussian Processes [11]). These algorithms allow to learn the function linking the endmembers available in the dictionary (or in a supervised fashion using an a priori available dictionary). The training can be performed for instance by simulating mixed pixels with endmembers from the dictionary in different proportions [10], or simply by assigning an abundance of 1 for the material corresponding to the entries of the dictionary [11]. These methods can relatively easily model spectral variability, but they have the drawbacks of not explicitly modeling the variability, and to be very dependent on the quality of the extracted bundle.

3.2. Computational Models

For this class of methods, the mixing model takes explicitly distinct sources into account for each pixel. Most of the times, the idea is to allow the sources to vary around a known (or extracted from the data) reference. These techniques are relatively powerful since they theoretically allow to capture all kinds of variabilities, but this also harms the interpretation of the obtained results. We can cite the Bayesian approaches such as the Normal Compositional Model (NCM) [12], or the Beta Compositional Model [13], which assume that the endmembers follow a Gaussian or Beta distribution (with the advantage of being defined between 0 and 1 per construction, which is physically sound for reflectance data). These distributions are centered at reference endmembers. However, more recent methods are capable of estimating both the means and variances of the distribution (e.g. [14] which uses a generalized version of the NCM), increasing the robustness against the potential absence of pure pixels in the image. For instance, the model proposed in [14] is written as follows:

\[
\mathbf{x}_k = \sum_{p=1}^{P} a_{kp} \mathbf{s}_{kp} + \mathbf{e}_k,
\]

with \(\mathbf{e}_k \sim \mathcal{N}(0, \sigma^2_n I)\), \(\mathbf{s}_{kp} \sim \mathcal{N}(\mu_p, \text{diag}(\sigma_p))\) (6)

where \(k = 1, \ldots, N\) is a pixel index, \(\nu_k^2\) is the noise variance in the considered pixel, and \(\text{diag}(\sigma_p)\) is a diagonal covariance matrix, whose elements are variances of each endmember for the considered material, in each spectral band. Moreover, in this version of the model, the means \(\mu_p\) of the endmembers follow a Gaussian distribution themselves (the mean of this Gaussian distribution is a reference), which allows to adjust the value of the mean with respect to the basic NCM. In a Bayesian framework, the a posteriori density of the parameters is generally complex and has to be sampled using Monte-Carlo methods in order to allow the use of the Maximum A Posteriori (MAP) and Minimum Mean Squared Error (MMSE) estimators.

Even more recently, the Perturbed Linear Mixing Model (PLMM) [15] was proposed in order to explicitly model spectral variability using an additive term acting on the sources to perturb the endmembers:

\[
\mathbf{x}_k = \sum_{p=1}^{P} (\mathbf{s}_p + \mathbf{d}s_{kp})a_{kp} + \mathbf{e}_k
\]

The term \(\mathbf{d}s_{kp}\) accounts for an additive perturbation for each endmember, in each pixel, and for each wavelength. The parameters of this mixing model are then estimated through an optimization problem, in which different constraints and regularizations (spatial regularization on the abundances, proximity of the sources to a reference, limitation of the norm of the additive perturbation...) are added in order to make this Nonnegative Matrix Factorization problem better posed and more suited to the expected properties of the solution. The algorithm iteratively updates the abundances, the endmembers and the perturbations to converge to a stationary point of the objective function (the problem is globally not convex, but each subproblem is). The numerous constraints and regularizations suggest the use of the now popular Alternating Direction Method of Multipliers (ADMM) [16]. This model was recently adapted to estimate temporal variability in the case of sequences of hyperspectral images [17]. A similar model is presented in [18], but is this time derived in a Bayesian framework, with slightly different constraints. In particular, this model assumes that the additive perturbations are spectrally smooth. However, contrary to the other Bayesian algorithms presented above, the MAP estimator is here obtained through an optimization algorithm, which makes the algorithm significantly faster than standard Bayesian approaches, where the whole posterior distribution needs to be sampled. Finally, we can cite the recent approach based on tensor decomposition which has been successfully used in the case of multitemporal [19] and multangular [20] images, which allows to track the changes of the sources along the concerned modality.

3.3. Parametric Models

In this category, the sources are also allowed to vary according to a specific model, but in a more constrained way than in the previous category. The variations of the sources are now only allowed along a limited number of free parameters (typically much less than the number of spectral bands). Fig. 2 (c) (resp. Fig. 2 (d)) shows an example of this type of model, where here one (resp. two) real parameter(s) delimit(s) the accessible domain of the sources, which is in fact a parametrized manifold (of dimension 1 (resp. 2) in this example). These models are usually physics driven.

In the physics community, the Hapke model [21] is used to model the relationship between albedo and reflectance, by taking into account the geometry of the acquisition (incidence angle of the light source and viewing angle of the sensor, which vary locally with the topography of the scene), and the photometric parameters specific to the materials. In the context of macroscopic mixtures, the mixing process remains linear. By assuming that the albedo is known (or extracted by inverting the model) for a material, the model defines a (rather complex) manifold in which the endmembers can live. The model is very complex and comprises many parameters in its most general form. It is hence cumbersome to use directly in a SU context. However, as a first approximation, it can be simplified to provide the following mixture model [22], the so-called Extended Linear Mixing Model (ELMM):

\[
\mathbf{x}_k = \mathbf{S}_0 \psi_k a_k + \mathbf{e}_k
\]

where \(\psi_k\) is a diagonal matrix, comprising \(P\) scaling factors on its diagonal, which weigh each source in an identical way for all wavelengths, and \(\mathbf{S}_0\) is a reference endmember matrix. This equation explicitly models the variation of the endmembers (allowed to vary in a one dimensional manifold), where scaling factors locally absorb contrast and illumination effects in the image. [22] shows that if the scaling factor is assumed identical for each material, it can be estimated from the result of Eq. (3), by only considering the positivity constraint of the abundances. Indeed, the estimated quantity can be written as the product of an abundance coefficient and
a scaling factor (see [22] for details). The algorithm proposed in [23, 24] is specifically designed to estimate the parameters of this model through a nonconvex optimization problem with various constraints and regularization terms, and was validated using synthetic data based on the full Hapke model. In its most simple version, the optimization problem writes:

$$\hat{\mathbf{A}}, \hat{S}, \hat{\psi} = \arg \min_{\mathbf{A}, S, \psi} \frac{1}{2} \sum_{k=1}^{N} (|\mathbf{x}_k - \mathbf{S}_k \mathbf{a}_k|^2 + \lambda S_k - S_0 \psi_k |^2)$$

(9)

In the version of [24], spatial information is taken into account through spatial regularizations since abundances and geometric variability effects are known to be spatially correlated. Such regularizations also enforce a joint processing of the data rather than a pixel-wise processing ignoring spatial relationships. This algorithm is actually a hybrid approach between the last two categories of methods, since there is an explicit parametric model to explain the variation of the endmembers (along straight lines linking the origin and the reference endmembers), but the estimated sources are actually free to drift away from this exact model, in order to explain other types of variability which are not explicitly modeled (for instance intrinsic variability). Nevertheless, the scaling factors are easily physically interpretable and vary with changes in the geometry of the scene. This model (in the case of identical scaling factors for all materials) was recently adapted to address a case where the mixing model is bilinear rather than linear, showing that it is possible to simultaneously estimate parameters related to both spectral variability and nonlinear mixtures [18].

In [25], two parametric models (one is affine, the other exponential) are proposed to explain the variations in the spectrum of soil under different moisture conditions. Wavelength-dependent parameters of these two models are estimated by linear regression on in-situ measurements for two wavelengths. By assuming these parameters have been estimated for every wavelength (or at least for different wavelength ranges), the relationship between moisture and reflectance defines in each case a one-dimensional manifold (parametrized in each pixel by a real value accounting for the moisture level), which enables the local estimation of this parameter in the image.

4. DISCUSSION

In this section we summarize the previous considerations and provide a critical discussion of the different strategies.

Spectral bundles are probably the easiest way to deal with spectral variability in SU. They allow to simply represent the intra class variations of a material by a set of signatures incorporating its variability. Even when the bundles are extracted randomly in the image data, the unmixing performance usually surpasses that of the conventional linear SU. Bundles can also serve as a training set for machine learning approaches which aim at learning the relationship between the bundles and the abundances. This approach is also very sound when a spectral library is available a priori. The main drawback is that the performance of these methods highly depend on the quality of the extracted bundles. Besides, if the used endmember extraction algorithms make use of the pure pixel assumption, it has to remain valid in each used subset, which can be too strong a hypothesis when dealing with small datasets. Also, there is no explicit modeling of the variability, which can make the interpretation of the results harder.

More advanced bundle extraction algorithms, in particular using spatial information (e.g. [26] and references therein) could be used in this framework in order to improve the quality of bundle-based SU.

Computational models go further than spectral bundles because spectral variability is explicitly modeled in the mixing model used. These models usually approximate variability as a deviation w.r.t. the Linear Mixing Model, and hence can theoretically explain any type of variability. They generally only require a reference endmember to be known or extracted beforehand, and are thus relatively robust to the absence of pure pixels. Nevertheless, their flexibility can also be a drawback since, once again, the results cannot always be interpreted in a more thorough way than to detect areas where a linear mixture is not sufficient to explain the data. It is probable that other phenomena such as nonlinearities and missing endmembers could be interpreted as spectral variability within the framework of these models.

An a posteriori analysis of the extracted sources (e.g. to try to isolate “typical” sources for each material) could be a way to take advantage of the flexibility of these techniques.

The last category also resorts to parametric models to extract variability, but the key difference here is that a limited number of parameters provide rules for the spectral variation of the endmembers. Besides, these parameters are usually physics-based. Thus this type of models usually has an immediate physical interpretation. This has been recently validated in particular for illumination-based models [24]. Illumination conditions effects are correlated between materials, and hence the same model can as a first approximation be used for each of them. However, by definition, intrinsic variability is material specific, and modeling the variability of a material must be done through ad hoc models, limiting their applicability. Besides, physical modeling can lead to complex models which can be impossible to use in practice for SU, therefore requiring simplifying hypotheses.

Finally, most methods focus on considering varying endmembers in the spatial domain. However, modern hyperspectral sensors make more and more multimodal datasets available. The typical additional modalities are multitemporal and multiangular datasets. A majority of the methods listed here could be easily adapted to a different variability framework, be it temporal, angular, spatial, or any combination of those, as explained in Section 2. The works of [5],[17], and [19] are examples of algorithms for multitemporal spectral variability extraction in a joint way, rather than an independent analysis of each time frame [8].

5. CONCLUSION

We have presented a review of the most recent methods to take spectral variability into account for spectral unmixing. We have categorized these methods in three main types of approaches and described their specificities, before suggesting some research avenues for each of them. We believe that extending some of these methods to a nonlinear mixture model is possible (an encouraging step in this direction was recently made in [18]) and would definitely enrich the analysis of remote sensing images through spectral unmixing. We think that the design of models and algorithms to further refine the existing models dealing with illumination, as well as the design of new models taking into account various material specific physical phenomena is a promising research avenue for the community.
6. REFERENCES


