DOI: 10.58895/ksp/1000032143-16 erschienen in:
OCM 2013 - Optical Characterization of Materials - conference proceedings
DOI: 10.58895/ksp/1000032143 | https://www.ksp.kit.edu/site/books/m/10.58895/ksp/1000032143/

Evaluation of spectral unmixing using nonnegative matrix factorization on stationary hyperspectral sensor data of specifically prepared rock and mineral mixtures

Wolfgang Gross[1], Sven Borchardt[2] and Wolfgang Middelmann[1]

[1] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation, Departement Scene Analysis, Gutleuthausstraße 1, D-76275 Ettlingen
[2] University of Potsdam, Institute of Earth and Environmental Science, Am Mühlenberg 3, D-14476 Potsdam

Abstract Hyperspectral sensors are used to identify materials via spectroscopic analysis. Often, the measured spectra consist of mixed materials and depending on the problem, the mixture ratio and the pure material spectra are wanted. In this paper, linear spectral unmixing is performed using the Nonnegative Matrix Factorization to analyze its correlation to ground truth data. The results are compared to Nonnegative Least Squares unmixing using manually selected endmembers from the image. Additionally, the effect of different endmember extraction algorithms and abundance initialization methods for NMF are investigated. To test the validity of the method, several checkerboard patterns of different ground minerals/rocks with predefined mixtures were prepared. It was shown that good initialization is beneficial in terms of approximation error and correlation to ground truth.

1 Introduction

Hyperspectral sensors are used to identify materials via spectroscopic analysis. They are widely used in areas such as satellite/airborne imaging, mining and recycling. The value of hyperspectral data can be increased by having a fair knowledge about the physical composition of the recorded data. Considering that certain materials only occur in

small quantities on a sub-pixel level, typical classification approaches can be insufficient depending on the scope of work. Instead, spectral unmixing can be used to determine the individual material spectra, so called endmember and their abundances in a measured sample. As the amount of materials in an image is generally much less than the number of spectral bands, data reduction is achieved simultaneously [1].

Usually, the endmembers must be provided by the user and the abundances are computed by a Nonnegative Least Squares (NNLS) approach to approximate the original data set. However, this is only applicable when they are known in advance or can be selected manually from the data. When the data sets become larger, manual selection is increasingly difficult and automatic endmember extraction algorithms should be used. Small deviations from the actual endmember spectra, e.g. noise in the data set or variation in illumination, can lead to errors. The Nonnegative Matrix Factorization (NMF) can be used to alternately optimize endmembers and abundances to increase approximation accuracy [2].

When the focus lies on the physical interpretation of a hyperspectral data set, generally two constraints have to be introduced to the process [3]. The most important is the nonnegativity constraint for endmembers and abundances as the measured reflectance is per definition nonnegative and mutual cancellation of endmembers is impossible. The second constraint states that the abundances of one sample must sum to one. In that case, each abundance directly stands for the ratio of its associated endmember.

2 Linear spectral unmixing and initialization

The underlying optimization problem of linear spectral unmixing can be written as

$$\begin{aligned} &\min_{W,H} \|V - WH\|_F \,, \\ &\text{subject to } W, H \geq 0 \text{ per element,} \end{aligned} \tag{16.1}$$

where V is the data matrix with m bands and n samples, W is the $m \times k$ endmember matrix and H the $k \times n$ abundance matrix. $\|\cdot\|_F$ denotes the Frobenius norm. All entries of the matrices are real and nonnegative numbers. When the endmember matrix W is already known, it can be

considered as constant and optimization is done only for H. Thus, the optimization problem becomes convex and a global minimum can be computed via NNLS.

In practice, data dimensionality will always be m due to sensor noise. However, the number of materials and thus the inherent dimension of a data set is usually much smaller than m and a set of k vectors, with $k \ll m$ can approximate the data very well [1]. In this paper, we act on the assumption that k is known, as determination of the actual number of endmembers for a data set is a problem by itself. Further information on the selection of k can be found in [4].

The NMF is basically computed using the method of alternating steepest descent. As the underlying optimization problem is nonlinear due to the nonnegativity constraint, multiple (sub)optimal solutions exist and a good initialization improves the outcome [5].

The alternating multiplicative update proposed by [6] is basically a steepest descent algorithm with efficient step size calculation, to comply with the nonnegativity constraint. W and H are alternately updated. The NMF can be used when the endmembers are not known *a priori* or cannot be determined with sufficient accuracy. However, linear spectral unmixing is an ill-posed inverse problem. Because of model inaccuracies, sensor noise, external measurement conditions and variability in material spectra, it is impossible to analytically determine the solution. Depending on the initialization the steepest descent algorithm can get stuck in a local minimum due to initialization.

The endmembers can be regarded as extreme directions of the minimal convex cone containing the data cloud in m-dimensional space. All spectra within the cone can be reconstructed without residuals, using abundances as coefficients for the linear combination of endmembers. NMF iteration alternately adjusts the endmembers and their abundances to better fit the data cloud and thus reduce the approximation error.

Using the notation from (16.1), NMF can be performed as follows:

$$\begin{aligned} H^{(t+1)} &= H^{(t)} * \frac{((W^{(t)})^T V)}{(W^{(t)})^T W^{(t)} H^{(t)} + \epsilon} \\ W^{(t+1)} &= W^{(t)} * \frac{(V(H^{(t+1)})^T)}{W^{(t)} H^{(t+1)} (H^{(t+1)})^T + \epsilon} \end{aligned} \tag{16.2}$$

Here, (t) is the iteration index and $*$ and $\dot{\div}$ denotes multiplication/division per element. ϵ is a small, strictly positive value to prevent division by zero. The proof of monotony can be found in [6], further discussion on convergence in [3].

One of the most useful properties of NMF is that it usually produces a sparse representation of the data set. This means that most of the information is concentrated in a few abundances [6]. As the mixed spectra usually consist of only two or three different materials, this is a great benefit when compared to an unconstrained least squares approach where all endmembers are used per sample.

2.1 Initialization of W

Several approaches exist to extract endmember spectra directly from the data and use them to initialize the endmember matrix for linear spectral unmixing. This section contains a short explanation of the most frequently used algorithms. Further information and examples for implementation can be found in [2,5,7,8].

Random initialization: The random methods are among the fastest and easiest to code. W is initialized as a dense matrix with random numbers between 0 and 1. According to [9], these initializations have the potential to outperform every other method in terms of approximation error.

Spherical k-Means Clustering: The cluster centers are used as initialization for W. In this paper, the spectral angle is used as a metric for clustering. This is a reasonable choice, as spectra of identical mixtures should be treated identical, regardless of illumination. The algorithm, which is very similar to conventional k-Means Clustering, can be found in [2].

Pixel Purity Index (PPI): The data is projected onto random unit vectors and the most extreme samples are collected. This is repeated several times and the resulting samples are thinned out till only a set amount of spectra remains. The PPI algorithm used for endmember extraction can be found in [7].

Sequential Orthogonal Subspace Projection (SOSP): The SOSP is an analytical approach to determine the most significant vertices of the data set in m-dimensional space. Thus, an approximation to the convex hull of the data set is computed. The algorithm can be found in [8] and

performs exceptionally well on simulated data satisfying the linear mixture model. In contrast to the other methods discussed here, no random procedures are used making the result reproducible.

2.2 Initialization of H

Only two notable initialization methods exist for the abundance matrix H. The first being random initialization similar to W from Subsect. 2.1. This usually provides good results, but specifically when a sparse solution is needed another method must be chosen.

After initialization of W with an endmember extraction algorithm, the abundances in H are estimated by their fractional part, that is not accounted for by any other endmember. Each row $H_{i:}$ of H is calculated separately by

$$\begin{aligned} H_{i:} &= w_i^T P_{\mathrm{OSP}}^{-i}, \text{ where} \\ P_{\mathrm{OSP}}^{-i} &= (I - W_{-i}(W_{-i}^T W_{-i})^{-1} W_{-i}^T)V \end{aligned} \tag{16.3}$$

I is the unit matrix of suitable dimension and W_{-i} is the endmember matrix W without the i-th endmember w_i. A discussion of this method can be found in [10].

3 Experiments and discussion

Evaluation of linear spectral unmixing was performed using a 4×4 checkerboard pattern of ground minerals/rocks with known abundances per square. Experiments with three to nine classes were performed. To evaluate the effect of different endmember extraction methods at least one square was provided containing the pure material. Also, the experiments were limited to having three different materials per square at most.

For every pattern, regions of interest (ROI) were manually selected to outline the homogeneous parts of each sample. This was done to restrict the endmember extraction algorithms to spectra that comply with the linear mixture model. However, NMF was performed on the whole data set. Restricting NMF to the homogeneous regions was considered to be impractical when working with data that is not specifically prepared. The setup for three classes is shown in Fig. 16.1. The ground

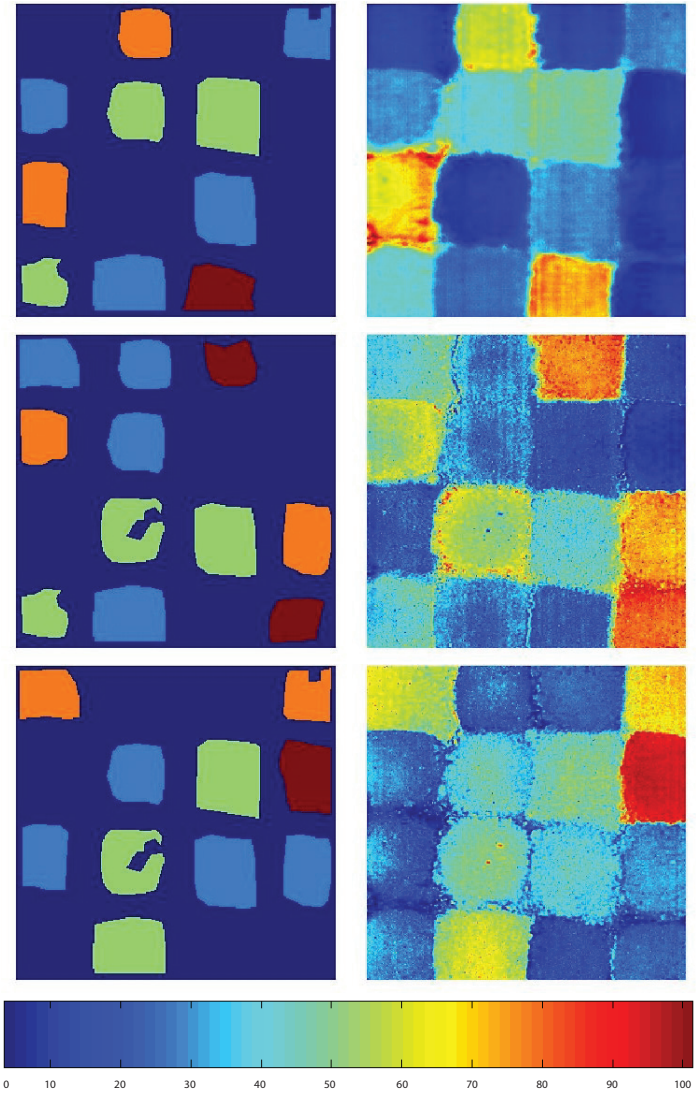

Figure 16.1: Ground truth masks (left) and corresponding abundance images (right) for basalt (top), trachyte (middle) and rhyolite (bottom); endmembers by SOSP, random initial abundances.

truth masks depict homogeneous regions of each square, light blue corresponding to a fraction of 25%, green 50%, orange 75% and red 100% of the corresponding material.

The data was recorded in a laboratory using an AISA sensor with 238 spectral bands in the wavelength range from $1 - 2.5\mu m$ and artificial illumination.

On the left side of Fig. 16.1 the ground truth images for basalt, trachyte and rhyolite are depicted from top to bottom. On their right are the corresponding abundance images. In this case, assignment was done manually as the similarities are clear. The abundances from Fig. 16.1 were calculated by initializing W with SOSP and H randomly.

To evaluate the results, the mean approximation error per sample as well as the mean correlation between the abundance images and the corresponding ground truth images were calculated. This was done for all initialization methods in section 2.1 including the random and OSP initialization for H. All methods that use random procedures were computed 10 times and the best result was saved. Only the combination of SOSP initialization for W and OSP for H can be computed analytically, always producing the same outcome. The NMF algorithm was terminated after 2000 iterations of (16.2). Without parallel processing the computation time of NNLS and NMF was comparable. The results of NMF and NNLS solutions are shown in tables 16.1 and 16.2 respectively. Depicted are the mean results over all available test sets.

Selecting the endmember candidates manually from visual judgment usually results in a good unmixing. However, manual selection is susceptible to errors especially on extensive data sets. The random initialization for both W and H gives good results in terms of correlation. However, the approximation error, especially in the case of NNLS, is worse when compared to other methods. Also, initialization methods with random procedures were evaluated 10 times due to the dependency of NMF and NNLS on a good initialization. When random procedures are involved and no ground truth is available, selecting the best among multiple solutions remains to be investigated. In the case of NNLS, random initialization of W has a very high approximation error per sample as it is treated as a fixed endmember matrix and the results depend on its similarity to the actual endmembers.

Spherical k-Means Clustering, PPI and SOSP have similar performance. When manual initialization is not possible SOSP in combination

with OSP initialization of H is preferable due to analytic computation. It was shown that SOSP always finds the correct endmembers, when the linear mixture model holds [8]. On simulated data sets a correlation of 0.9986 could be achieved while simultaneously the approximation error was the lowest among the tested methods. However, outliers and regions with nonlinear mixtures must be ignored during initialization.

Table 16.1: NMF results: Approximation error and mean correlation between abundances and corresponding ground truth images.

NMF:	H_{OSP}		H_{random}	
	approx. error	mean corr.	approx. error	mean corr.
Random	0.1850	0.8891	0.1387	0.9111
Manual initialization	0.1330	0.9012	0.1958	0.8665
k-Means Clustering	0.1247	0.8790	0.1954	0.8517
Pixel Purity Index	0.1550	0.8418	0.2417	0.9543
Sequential OSP	0.1147	0.9273	0.1942	0.9386

Table 16.2: NNLS results: Approximation error and mean correlation between abundances and corresponding ground truth images.

NNLS:	H_{OSP}		H_{random}	
	approx. error	mean corr.	approx. error	mean corr.
Random	3.6586	0.9015	3.7402	0.8900
Manual initialization	0.1541	0.9380	0.1580	0.9452
k-Means Clustering	0.1563	0.8978	0.1562	0.8955
Pixel Purity Index	0.6656	0.6475	0.6656	0.6475
Sequential OSP	0.2036	0.8239	0.2036	0.8239

4 Conclusion

Comparison of linear spectral unmixing algorithms has shown that NNLS generally gives good results provided the initial endmember matrix closely resembles the actual endmember spectra in the scene. The iterative approach of NMF is able to compensate for worse initialization by alternately updating abundances and endmembers. While the accuracy of the result, measured here by correlating the abundance im-

ages with their corresponding ground truth information, is comparable, good results are computed more consistently with NMF and the approximation error is lower. In a measurement where no ground truth data is available to quantify the result, NMF is generally more forgiving of bad initialization and the chance to get a result close to the actual physical composition in the first attempt is higher.

Considering the different initialization methods, random initialization for W and H has the potential to outperform every other method with the downside that multiple computations may be necessary. The SOSP initialization for W most consistently gave good results as the computation does not rely on random procedures. Additionally, when H is initialized with random values the NMF is able to iterate towards a result that allows easy assignment between ground truth and abundance images. It has to be noted that for this to work properly on a real data set, areas where nonlinear effects occur should be ignored while initializing W. Otherwise, endmembers might be chosen from these areas. NNLS and NMF would then be preset to prioritize the approximation of these areas over the actual linear mixtures. This can result in a lower approximation error, when a lot of nonlinear mixtures are in the scene, but in terms of comparability to the physical composition of the scene the result is worse.

Initialization of H by (16.3) already limits optimization problem (16.1) to certain solutions resulting in slightly lower correlations in our tests. This limitation can be advantageous, where nonlinear mixtures should be ignored. Also, when computation time is especially important, tests have shown that it can already be used as a crude unmixing or reduce the number of iterations needed for NMF.

The arranged test sets are valuable for further analysis as barely any hyperspectral data is available with ground truth about the mixture ratios per sample. More checkerboards were prepared where some materials were only included in mixtures. The arranged data sets can also be used to analyze nonlinear mixtures as well. Nonlinear unmixing usually needs a lot of *a priori* information about the data set, but the ground truth information is already available. This can help to improve the understanding of unmixing and the degree of model complexity that is needed for accurate results.

In future work the performance of NMF will be analyzed when no pure spectra are available for endmember extraction. Additionally, the

effects of over- or underestimating the real amount of endmembers have to be explored.

References

1. C. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 3, pp. 608–619, 2004.
2. S. Wild, J. Curry, and A. Dougherty, "Motivating non-negative matrix factorizations," *SIAM*, vol. 8, 2003.
3. M. Berry, M. Browne, A. Langville, V. Pauca, and R. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," in *Computational Statistics & Data Analysis*, vol. 52, 2007, pp. 155–173.
4. J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, 2008.
5. A. Langville, C. Meyer, and R. Albright, "Initializations for the nonnegative matrix factorization," *Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2006.
6. D. Lee and H. Seung, "Algorithms for non-negative matrix factorization," *Adv. Neural Info. Proc. Syst.*, vol. 13, pp. 556–562, 2001.
7. F. Chaudhry, C. Wu, W. Liu, C. Chang, and A. Plaza, "Pixel purity index-based endmember extraction for hyperspectral data exploitation," *Image Processing*, vol. 661, no. 2, 2006.
8. W. Gross and W. Middelmann, "Sparseness-inducing initialization for non-negative matrix factorization in hyperspectral data," *Proc. DGPF*, vol. 32, 2012.
9. S. Wild, J. Curry, and A. Dougherty, "Improving non-negative matrix factorizations through structured initialization," *Pattern Recognition*, vol. 37, no. 11, pp. 2217–2232, 2004.
10. C. Kwan, B. Ayhan, G. Chen, J. Wang, B. Ji, and C. Chang, "A novel approach for spectral unmixing, classification, and concentration estimation of chemical and biological agents," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 2, pp. 409–419, 2006.