Statistical Learning for Satellite Hyperspectral Images

Hugo Jeannin

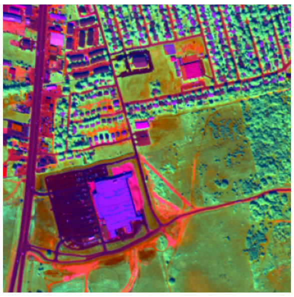
21/09/2025

Keywords: hyperspectral imaging, statistical learning, spectral unmixing, high-dimensional statistics, segmentation

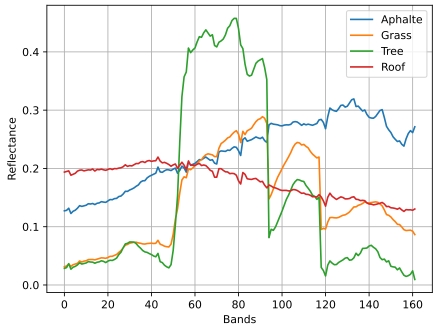
1 Context

Hyperspectral imaging records several hundred narrow, contiguous bands across the visible and infrared spectrum. This results in high-dimensional data that carry rich spectral information. However, the relatively low spatial resolution means that several distinct materials may coexist within a single pixel. These elementary materials are referred to as *endmembers*.

A fundamental problem is therefore *spectral unmixing*, which aims to recover both the spectral signatures of the endmembers and their relative proportions in each pixel. Several methods have been developed, relying on geometry [11, 12], optimization [14, 13], or machine learning [15, 16]. However, only a few studies, such as [17, 3], have been devoted to designing a consistent statistical framework that provides physically meaningful results while also quantifying the uncertainty of the solution. Some fully Bayesian methods have been proposed [10], but they often suffer from high computational cost, strong noise assumptions, reliance on dimensionality reduction, and potential identifiability issues. To overcome these limitations, we propose to leverage recent advances in random matrix theory and high-dimensional statistics [5, 1, 7, 8] to design a statistical method adapted to the high-dimensional nature of hyperspectral data. This approach exploits the covariance structure between neighboring pixels.



(a) Urban hyperspectral scene.



(b) Example of extracted endmembers.

Figure 1: Illustration of a hyperspectral image and its endmembers.

2 Previous Work and Objectives

The project aims to advance statistical learning for hyperspectral satellite images through the following objectives:

- **Estimation of the number of endmembers:** In an initial study [9], we proposed a consistent statistical method to estimate the number of endmembers in each pixel. While several approaches have been introduced in the literature [2, 3, 4], they often rely on inconsistent covariance estimators in the high-dimensional regime. Our method explicitly accounts for both the high dimensionality of hyperspectral data and the correlations between adjacent spectral bands.

- **Hyperspectral unmixing:** Building on the first objective, we aim to develop a comprehensive methodology for spectral unmixing that accounts for the high-dimensional nature of the data. The approach will rely on parametric models for the endmember spectra that satisfy regularity conditions, ensuring physical interpretability. This line of work will be developed in connection with [8].

- **Machine learning for hyperspectral data:** Finally, we seek to design machine learning algorithms specifically adapted to hyperspectral data. These algorithms will take into account high dimensionality, low-rank structures, and correlated noise to enable efficient segmentation and classification. A preliminary step in this direction has been presented in [6].

3 Information and Contact

This Ph.D. project takes place at the Laboratoire de l'Intégration du Matériau au Système (IMS), Bordeaux University, with the following team:

- **Ph.D. student:** Hugo Jeannin, IMS/Université Bordeaux **Email:** hugo.jeannin@u-bordeaux.fr
- **Thesis Director:** Prof. Pascal Vallet, IMS/Bordeaux INP **Email:** pascal.vallet@ims-bordeaux.fr
- **Co-supervisor:** Prof. Guillaume Ginolhac, LISTIC/Polytech Annecy **Email:** guillaume.ginolhac@univ-smb.fr
- **Co-supervisor:** Dr. Florent Bouchard, L2S/CentraleSupélec **Email:** florent.bouchard@cnrs.fr

References

[1] X. Mestre. Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates. *IEEE Transactions on Information Theory*, 54(11):5113–5129, 2008.

[2] J. Bioucas-Dias and J. Nascimento. Hyperspectral subspace identification. IEEE Transactions on Geoscience and Remote Sensing, 46(8):2435–2445, 2008.

[3] K. Cawse-Nicholson, S. Damelin, A. Robin, and M. Sears. Determining the intrinsic dimension of a hyperspectral image using random matrix theory. IEEE Transactions on Image Processing, 22(4):1301–1310, 2012.

[4] A. Halimi, P. Honeine, M. Kharouf, C. Richard, and J.-Y. Tourneret. Estimating the intrinsic dimension of hyperspectral images using a noise-whitened eigengap approach. IEEE Transactions on Geoscience and Remote Sensing, 54(7):3811–3821, 2016.

[5] F. Benaych-Georges and R.R. Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494–521, 2011.

[6] F. Bouchard, A. Mian, M. Tiomoko, G. Ginolhac, and F. Pascal. Random matrix theory improved Fréchet mean of symmetric positive definite matrices. In *Proceedings of the International Conference on Machine Learning*, vol. 235, pp. 4403–4415, 2024.

[7] R. Beisson, P. Vallet, A. Giremus, and G. Ginolhac. Change detection in the covariance structure of high-dimensional Gaussian low-rank models. In 2021 IEEE Statistical Signal Processing Workshop (SSP), pp. 421–425. IEEE, 2021.

[8] P. Vallet, X. Mestre, and P. Loubaton. Performance analysis of an improved MUSIC DoA estimator. *IEEE Trans. Signal Process.*, 63(23):6407–6422, 2015.

[9] H. Jeannin, F. Bouchard, P. Vallet, G. Ginolhac, A. Giremus. On the endmembers detection for hyperspectral imaging in the high-dimensional regime. 2025 European Signal Processing Conference (EUSIPCO).

[10] Dobigeon, N., Moussaoui, S., Coulon, M., Tourneret, J. Y., Hero, A. O. (2009). Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery. IEEE Transactions on Signal Processing, 57(11), 4355-4368.

[11] Nascimento, J. M. and Bioucas-Dias, J. M., "Vertex component analysis: A fast algorithm to unmix hyperspectral data," IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 4, pp. 898–910, April 2005

[12] Winter, M., "Fast autonomous spectral end-member determination in hyperspectral data," Proc. 13th Int. Conf. on Applied Geologic Remote Sensing, vol. 2, Vancouver, April 1999, pp. 337–344

[13] Miao, L. Qi, H. (2007)."Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization." IEEE Transactions on Geoscience and Remote Sensing, 45(3), 765–777

[14] Heinz, D. C. Chang, C.-I. (2001)."Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery." IEEE Transactions on Geoscience and Remote Sensing, 39(3), 529–545.

[15] Palsson, F., Sveinsson, J. R., Ulfarsson, M. O. (2018)."Hyperspectral unmixing using a neural network autoencoder."IEEE Access, 6, 25646–25656

[16] Zhang, X., Sun, Y., Wu, L., Zhang, L. (2019). "Deep spectral–spatial sparse representation for hyperspectral unmixing." IEEE Transactions on Geoscience and Remote Sensing, 57(6), 4282–4297.

[17] Terreaux, E., Ovarlez, J.-P. & Pascal, F. (2015). "Anomaly Detection and Estimation in Hyperspectral Imaging using Random Matrix Theory tools." In *Proceedings of the IEEE Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2015. DOI: 10.1109/CAMSAP.2015.7383763.