EPSC Abstracts
Vol. 5, EPSC2010-590, 2010
European Planetary Science Congress 2010

Multi-Image Unsupervised Spectral Analysis

Albrecht Schmidt (1), Erwan Tréguier (1), Frédéric Schmidt (2), Maël Guiheneuf (1), Saïd Moussaoui (3), Nicolas Dobigeon (4), Camille Pelloquin (1)
(1) European Space Agency, ESAC, Spain, (2) IDES, UMR 8148, Université Paris Sud, CNRS, Orsay, France (3) IRCCYN, UMR CNRS 6597, Ecole Centrale Nantes, France (4) IRIT/INPT-ENSEEIHT, Toulouse, France
albrecht.schmidt@esa.int

Abstract

Large data sets delivered by imaging spectrometers are interesting in many ways in the Planetary Sciences. Due to the size of the data, which often prohibits conventional exploratory data analysis, unsupervised analysis methods could be a way of gathering interesting information contained in the data. In this work, we investigate some of the opportunities and limitations of unsupervised analysis based on non-negative matrix approximation [2] in planetary settings. Since typically there often is no ground truth to compare to, unsupervised rather than supervised methods allow to extract new information from data sets. Often, the practicability of these methods suffered from low performance, which made large-scale analyses almost prohibitively expensive. New research and implementation strategies [1] for non-negative matrix factorisation make it possible to extract sources and relative abundances for typical planetary data sets with reasonable resources.

In this work, we try to give an impression of some of the trade-offs and opportunities involved. Non-negative matrix factorisation is a technique which has enjoyed considerable research and been used in many application areas, from document clustering to spectral analysis.

By considering P pixels of an hyperspectral image acquired at L frequency bands, the observed spectra are gathered in a $P \times L$ data matrix $\mathbf{X}$. Each row of this matrix contains a measured spectrum at a pixel with spatial index $p = 1, \dots, P$. According to the linear mixing model, the p^{th} spectrum, $1 \leq p \leq P$, can be expressed as a linear combination of $r_i, 1 \leq r_i \leq R$, pure spectra of the surface components. Using matrix notations, this linear spectral mixing model can be written as

$$\boldsymbol{X} \approx \boldsymbol{A} \cdot \boldsymbol{S}, \tag{1}$$

where non-negative matrices $A \in M^{P \times R}$, and $S \in M^{R \times L}$ approximate $X \in M^{P \times L}$ in the sense that $\frac{1}{2}||AS - X||^2$ is minimised, where $M^{\cdot \times \cdot}$ is the space of matrices of respective dimensions with non-negative entries. The rows of matrix $\boldsymbol{S}$ now contain the pure spectra of the R components, and each element A_{pr} of matrix $\boldsymbol{A}$ corresponds to the abundance of the r^{th} component in pixel with spatial index p. For a qualitative and quantitative description of the observed scene composition, the estimation problem consists of finding matrices $\boldsymbol{S}$ and $\boldsymbol{A}$ which allow to explain the data matrix and, at the same time, have a coherent physical interpretation. This approach casts the hyperspectral unmixing as a source separation problem under a linear instantaneous mixing model. Source separation is a statistical multivariate data processing problem whose aim is to recover unknown signals (called sources) from noisy and mixed observations of these sources [3].

If we want to factor not only single images but collections of images, even more questions of consistency and interpretation come up. The spectra of same physical sources often differ in scale and, to a lesser degree, in shape across images. For large data sets it is thus necessary to present and summarise the results of the calculations on the individual images in a way that is both transparent and comprehensive. This could be done by normalisation steps followed by clustering based on a similarity measure for the sources.

We compare different approaches in terms of resource consumption and viability of result presentation while keeping in mind the physical interpretation and meaningfulness in the Planetary Sciences. It turns out that there are datasets and algorithms which return efficiently results which are meaningful in the sense that abundances, sources and matching of sources are consistent with community opinion for well-known examples such as ices on the Martian South Pole, as shown in the image. Technically, we like to stress that no prerequisites are necessary for the applicability of the algorithms other than the approximate linearity of combinations of the source spectra. This implies that there are application scenarios in which the matrix

approximation algorithms are applicable and deliver sensible results. For the subsequent analysis of the extracted source spectra and their grouping, standard clustering techniques are applied and evaluated. We note that, while the first approximation is done entirely automatically, it might be sensible to involve some human judgement in the second phase, for example, in the form of a supervised learning for the classification of the spectra so that expert opinion can be honoured.

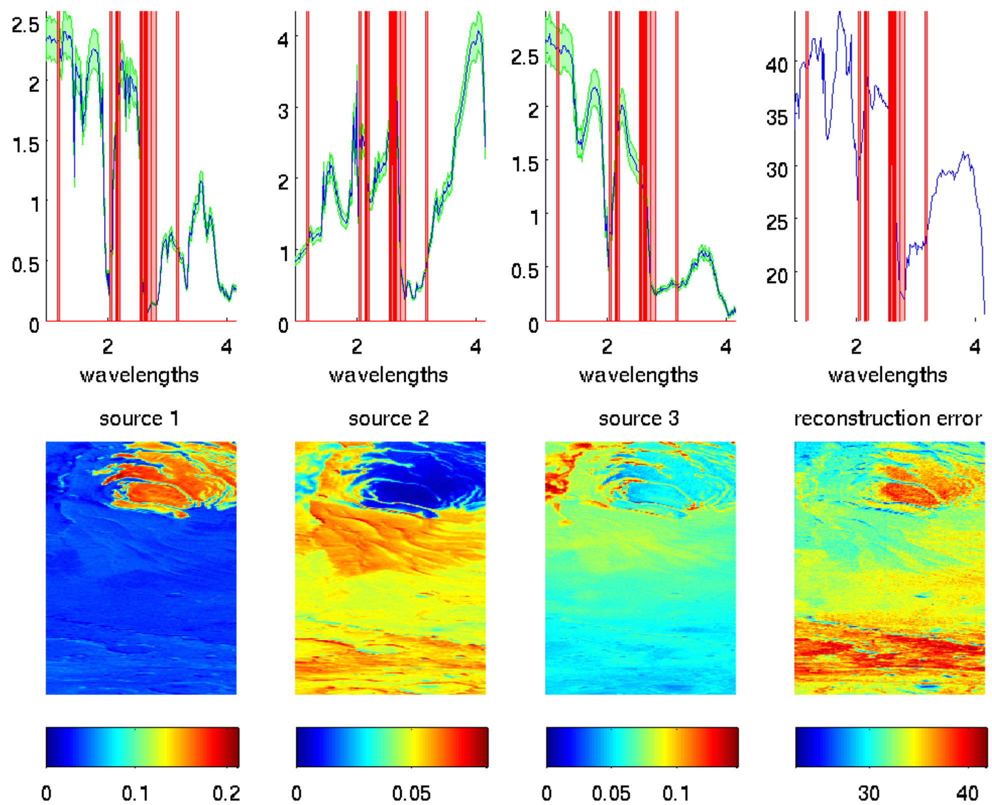
The focus of the presentation is on implementation techniques and comparative analysis of the performance and limitations of the algorithms. Opportunities for applications and engineering of systems will be discussed.

References

[1] Schmidt, F., Schmidt, A., Tréguier, E., Guiheneuf, M., Moussaoui, S., Dobigeon, N.: Implementation strategies for hyperspectral unmixing using Bayesian source separation, accepted at IEEE Transactions on Geoscience and Remote Sensing

[2] Lee, D.D., Seung, H.S.: Algorithms for non-negative matrix factorization, Advances in neural information processing systems, 2001

[3] Cichocki, A., Amari, S.-I.: Adaptive blind signal and image processing – Learning algorithms and applications, Wiley Interscience, 2002



Acknowledgements

We thank J.P. Bibring and the OMEGA Team for providing the OMEGA dataset. We are also grateful to S. Douté and B. Schmitt for their ice spectral library. We acknowledge support from the Faculty of the European Space Astronomy Centre (ESAC).