# PhD – Data Fusion and Semi-Unsupervised Learning for Hyperspectral Image Super-Resolution. Application to JWST.

**Keywords.** Inverse problems, prior learning, heterogeneous data fusion, high dimensions, images.

**Tools.** Optimization, Bayesian approach, machine learning.

**Applications.** Hyperspectral imaging, astronomy.

**Supervision.** F. Orieux (L2S – Univ. Paris-Saclay), A. Abergel (IAS – Univ. Paris-Saclay).

**Funding.** The subject is fully funded, by CNES for 50% (application necessary before April 2) and the joint laboratory LAB4S2 for 50%.

**Contact and location.** orieux@l2s.centralesupelec.fr, L2S, CentraleSupélec, Gif-sur-Yvette, and alain.abergel@ias.u-psud.fr, IAS, Orsay

## 1 Context

The objective of the subject is to develop new methods, notably inspired by machine learning approach, for fusion of heterogenous data, hyper and multispectral images in particular. It takes place in the international James Webb Space Telescope (JWST) project, see Fig. 1, the most ambitious space telescope scheduled for launch in October 2021 and which will carry four instruments with unprecedented hyperspectral and multispectral observing capabilities.

Hyperspectral and multispectral imaging are ubiquitous in many observation modalities like Earth observations (like the European Copernicus project), astronomy, medical imaging, material analysis. Thus the work carried out during the PhD should also be applicable to many other modalities.

This thesis is part of the scientific project of the joint laboratory LAB4S2 (Innovative Laboratory for Space Spectroscopy) between the IAS and the company ACRI-ST (https://www.acri- st.fr/fr). The objective of LAB4S2 is to develop fusion methods for the joint exploitation of hyperspectral and multispectral data obtained, on the one hand, by the mid-infrared instrument of the JWST (MIRI), and, on the other hand, by the instruments on board the Sentinel missions of the Copernicus project.

## 2 Methodology

### 2.1 Data model and fusion

The aim of the thesis is to develop efficient algorithms for joint processing:

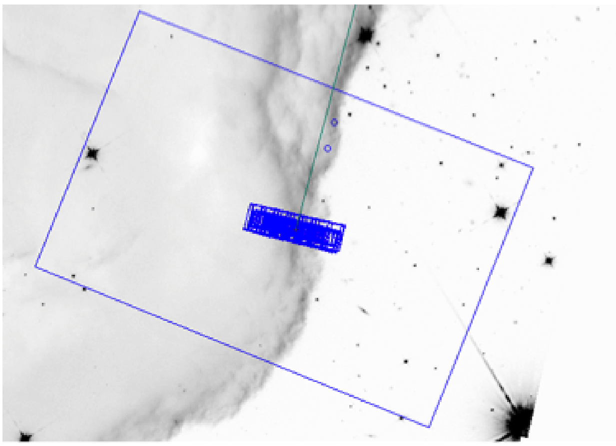

Figure 1: JWST observations of the Horsehead nebula. Large rectangle : multispectral imaging. Blue stripe : hyperspectral observations.

- of multispectral data obtained with imager over a large field of view and broadband filters ;
- with hyperspectral data with high spectral resolution but for small fields of view.

This problem, which can be likened to a data fusion problem, has been addressed in Earth observations (with so-called pansharpening methods), in a context where the effects due to the observing instruments (limited spatial resolution and insufficient spatial sampling) are negligible.

The problem can be solved by minimization of a loss function

$$J(\boldsymbol{x}) = \|\boldsymbol{y}_{img} - \boldsymbol{H}_{img}\boldsymbol{x}\|^2_{\boldsymbol{\Sigma}_{img}} + \|\boldsymbol{y}_{hyp} - \boldsymbol{H}_{hyp}\boldsymbol{x}\|^2_{\boldsymbol{\Sigma}_{hyp}}$$

with explicit data models $\boldsymbol{H}_{img}$ and $\boldsymbol{H}_{hyp}$ for multispectral (imaging) and hyperspectral (spectroscopic) instruments, with data $\boldsymbol{y}$ and unknown $\boldsymbol{x}$. This approach solves the problem of measurement heterogeneity by relying on a "virtual" instrument that combines imaging and spectroscopy.

### 2.2 Semi-Unsupervised Learning for Ill-Posed Inverse Problem

The above-mentioned method is model driven. However, the model can be inefficient is some case, when $ker(\boldsymbol{H})$ is not null for instance, or when the degradations due to the instruments, like spatial blurring, are important. A common approach is to add a regularizer $R(\boldsymbol{x})$

$$\bar{J}(\boldsymbol{x}) = J(\boldsymbol{x}) + \mu R(\boldsymbol{x}).$$

This approach can have some limitations. In particular, the prior models are often ad hoc and the compromise between the different sources of information, parametrized by $\mu$, must be tuned. During this work we will explore alternative methods based on machine learning to construct more adapted prior [2, 3]. The challenge here will be the lake of a big database necessary for supervised approach. Therefore we will explore new methods based on small databases like Few-Shot approach [3]. This approach consists of adapting trained DNN with existing databases (like ImageNet) with another database that have only a few number of examples.

## 3 Planning

- In a first time, and especially concerning the instrument models, the student will use work from the PhD of A. Hadj-Youcef (defended in 2018, see [4]) and R. Abi-Rizk (defense planned for September 2021, see [1]).
- Then it will be necessary to develop an inversion algorithm. The first approach will use a new composite criteria that must be minimized. The inversion algorithm will be tested on simulated data, and also on the first JWST flight data which should be received from mid-2022. The practical implementation of the inversion algorithm will be carried out with the technical support of engineers from the LAB4S2 joint laboratory.
- Then, the students will study learning approach (dimension reduction, invertible networks [2], Few-Shot [3], structured networks, semi-supervised . . . ) to develop new regularization approach.

## 4 Supervision

The subject will be supervised at the Laboratoire des Signaux et Systèmes (L2S) by François Orieux on the data processing and inversion aspects and the Institut d'Astrophysique Spatiale (IAS) by Alain Abergel on the instrument and astrophysical aspects.

The students will benefit from years of collaboration between the L2S and the IAS. The work will also use the expertise present at Paris-Saclay on the JWST instruments and, if necessary, at the Space Telescope Science Institute in Baltimore.

## References


[1] Ralph Abirizk, Francois Orieux, and Alain Abergel. "Non-Stationary Hyperspectral Forward Model and High-Resolution". In: *Proc. of 27th IEEE Int. Conf. on Image Processing*. 27th IEEE International Conference on Image Processing, ICIP 2020. Abu-Dhabi, United Arab Emirates, Oct. 2020, p. 5.

[2] Lynton Ardizzone et al. *Analyzing Inverse Problems with Invertible Neural Networks*. Feb. 6, 2019. arXiv: 1808.04730 [cs, stat]. URL: http://arxiv.org/abs/1808.04730 (visited on 09/24/2020).

[3] Guneet S. Dhillon et al. *A Baseline for Few-Shot Image Classification*. Oct. 21, 2020. arXiv: 1909.02729 [cs, stat]. URL: http://arxiv.org/abs/1909.02729 (visited on 02/15/2021).

[4] Claire Guilloteau et al. "Fusion of Hyperspectral and Multispectral Infrared Astronomical Images". In: *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. 2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM). June 2020, pp. 1–5. DOI: 10.1109/SAM48682.2020.9104393.

[5] Mohamed Amine Hadj-Youcef et al. "Fast Joint Multiband Reconstruction from Wideband Images Based on Low Rank Approximation". In: *IEEE Trans. Comput. Imaging* (May 28, 2020), pp. 922–933. ISSN: 2333-9403, 2334-0118, 2573-0436. DOI: 10.1109/TCI.2020.2998170.

[6] François Orieux et al. "Estimating Hyperparameters and Instrument Parameters in Regularized Inversion. Illustration for Herschel/Spire Map Making." In: *Astronomy & Astrophysics* 549.A83 (Jan. 2013). DOI: 10.1051/0004-s6361/201219950.