ONERA
THE FRENCH AEROSPACE LAB

www.onera.fr

PROPOSITION DE SUJET DE THESE

Intitulé : Machine Learning data processing for Adaptive Optics assisted observations

Référence **PHY-DOTA-2024-20**
(à rappeler dans toute correspondance)

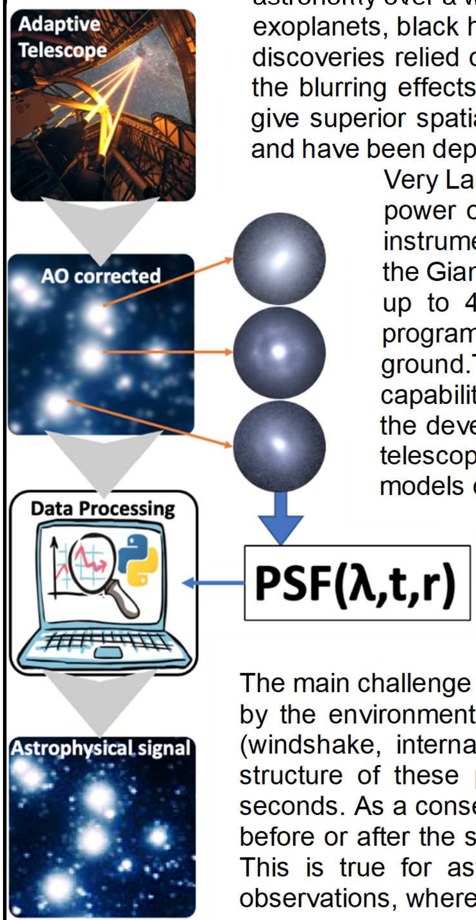
Début de la thèse : 10/2024	**Date limite de candidature :** 06/2024

Mots clés :
Astronomical Telescope, Adaptive optics, Post-processing, Machine Learning, PSF reconstruction, Astronomical Observations

Profil et compétences recherchées :
Engineering school of Optics, Master in Astronomy, Optics, Physics

Présentation du projet doctoral, contexte et objectif :

Scientific context:



High-resolution images from large ground-based telescopes have revolutionized visible and near-infrared astronomy over a wide range of astrophysical fields, including finding and characterizing exoplanets, black holes, brown dwarfs, and the earliest galaxies in the Universe. These discoveries relied on adaptive optics (AO) systems, which compensate in real-time for the blurring effects of the Earth's turbulent atmosphere (called "seeing"). AO systems give superior spatial resolution over space-based alternatives at a fraction of the cost and have been deployed on nearly all of the world's largest telescopes, as the European Very Large Telescope (VLT) and its 10m-class telescopes counterparts. The power of AO is now widely recognized and it will be built into the 1st-light instruments of ALL the next-generation giant telescopes -- the European ELT, the Giant Magellan Telescope and the Thirty Meter Telescope with diameters up to 40m. AO is also at the heart of Space Surveillance Awareness programs, providing detection and characterization capabilities from the ground.The exceptional advancement in AO technology and observational capability has, however, not been followed by a comparable advancement in the development of data analysis methods. Additionally, the increase of the telescope size introduces new effects that can be barely characterized by models of current level of complexity. This results in a difficult understanding of the AO performance, and in particular in the characterization of the AO Point Spread Function (PSF), which eventually limits the scientific analyses. **The main scientific objective of this PhD is to address this deficit and provide a set of optimized data processing tools dedicated for AO-assisted observations**.

The main challenge for the AO PSF prediction comes from the stochastic effects induced by the environment: the atmospheric turbulence but also the telescope environment (windshake, internal aberrations and vibrations). By nature, the strength and spatial structure of these perturbations are constantly evolving, on time scales faster than seconds. As a consequence, a PSF calibrator acquired with the same instrument setting before or after the science exposure can be completely uncorrelated to the actual PSF. This is true for astronomical observations, and even more a limitation for satellite observations, where the conditions can be changing rapidly.

When the PSF information cannot be properly extracted from the science image, an alternative approach is to predict the PSF shape based on auxiliary data. Different AO-PSF prediction methods have been considered over the past years, but as of today, there is no general facility providing the PSF with every AO observation. In fact, several experiments conducted using synthetic data have shown that PSF prediction can yield highly accurate predictions, however, this accuracy significantly drops when applying the methods to real telescope data.

In this PhD work, we **would like to explore a brand-new approach where the AO-PSF model is calibrated against other images made available by the system. All AO observations are usually assisted by technical cameras or wave-front sensors providing a large number of images, currently not used for PSF prediction**. By combining advanced data-fusion and learning methods, this PhD will therefore investigate the potential gain brought by these auxiliary data to predict the science PSF.

To this aim, the student will make use of 3 principal ingredients, with (1) a long-exposure AO-PSF model – called TipTop - that accurately describes the PSF shape, while being described only by less than a dozen parameter, (2) databases of on-sky PSFs and associated auxiliary data and (3) accurate instrument numerical models, a.k.a. numerical twins.

The student will combine these 3 ingredients with innovative methods based on data fusion and machine-learning algorithms, in order to propose an operational scheme for PSF prediction. The optimization of the whole data-processing chain is carried-out with data processing specialists, and the quantitative scientific impact and critical feedback on the process is evaluated with end users (astronomers, defense).

Description of the work:

The ultimate objective of the PhD program is to achieve the capability of predicting the AO-PSF solely relying on supplementary data. One initial approach could involve calibrating our analytical PSF model using actual observations and subsequently employing an inverse methodology to forecast the PSF. The key advantage of parametric PSFs lies in their ability to condense all the essential information from the physical PSF into a small number of parameters. This results in a significant simplification of the PSF prediction process, facilitating the utilization of efficient IA methods for ultimately estimating and predicting these parameters.

Nonetheless, the limited availability of on-sky samples poses a significant challenge, and the largest on-sky databases currently accessible consist of just a few hundred datasets, which is insufficient for the development of robust learning techniques. Therefore, the proposed approach involves initially training a network using realistic simulations (numerical twins) and subsequently fine-tuning the network using real on-sky data. This will be the bases of the strategy pursued in this research.

1. Numerical twins for network training

The first part of the work will be focused on developing accurate numerical twins of the AO system under study. This work will make use of existing tools (e.g. OOPAO), and adjust them to match real system conditions. This is a crucial step for (1) test the proposed techniques for PSF prediction and model calibration within a controllable simulated environment and (2) create large enough (>10.000) datasets that could be used for training learning networks. Early demonstration of the viability and accuracy of such numerical twins for PSF learning have been in our group (e.g. [Kuznetsov, Neichel, Fusco et al., AO4ELT, 2023]). The student will consolidate and extend this framework to a wider range of instruments and conditions.

2. Model adjustments from on-sky data

Once the first machine learning models are trained with synthetic data, the next step is to make use of on-sky data in order to fine-tune them. Indeed, and as shown in early experiments, a network that would only be trained on simulated data would not be accurate enough when directly used on real-data, and a model adaptation step is required. In this case however, fine-tuning requires significantly fewer samples than training a network from scratch using only on-sky samples, which is perfectly adequate in our applications when the number of on-sky samples is limited. This part of the work will make use of unique access to on-sky data from different instruments (VLT-SPHERE, VLT-AOF, Keck, LBT-SOUL, VLT-ERIS, OHP-PAPYRUS) and different environments. Access to the data is granted thanks to our collaboration network and between few hundreds to one thousand on-sky datasets are available for the instruments mentioned above. A dedicated procedure for sorting/filtering the data will be developed. Application to both astronomy and space observation is foreseen.

3. Science verification programs

The last step of the PhD will be to work with astronomers (and defense) to run science verification programs and asses the performance of the proposed method. Access to telescope time will be granted via the ESO PSF Working Group (lead Joel Vernet) and collaboration with other observatories (LBT, Keck). The goal of these science programs will be to demonstrate gains in terms of astrophysical performance, like astrometry, photometry. These science verification programs represent unprecedented and highly competitive observations. As such, they ensure a high scientific return for the work to be conducted by the PhD. But beyond the unique scientific results, they provide a critical visibility for the PSF prediction tool within the community. Therefore, when the algorithms are made publicly available through data-reduction pipelines we expect a large interest from a wide range of users. Applications to satellite observations can also be tested at this stage.

This work will be carried in close collaboration between several groups including: ONERA, LAM, INAF-Arcetri and ESO. The student will benefit from the tools, expertise and unique access to large ground-based telescopes available through this network.
The student will also work closely with the Machine Learning / Deep Learning (ML/DL) department of LAM, and the newly created LaboIA from ONERA. These groups will support the student in the development of optimized networks and machine learning strategies. This collaborative scheme, merging experts in Adaptive Optics and Artificial Intelligence already provided very successful results, and will be strengthen with this new PhD.

Collaborations envisagées :

Laboratoire d'Astrophysique de Marseille, INAF-Arcetri, ESO

Laboratoire d'accueil à l'ONERA :	**Directeur de thèse :**
Département : **DOTA**	Nom : Benoit Neichel / Thierry Fusco
Lieu (centre ONERA) : LAM (Marseille) & centre de Salon de Provence	Laboratoire : LAM / ONERA
Contact : Thierry Fusco	
Tél. : +33 6 62 48 48 36 Email : thierry.fusco@onera.fr	Email : benoit.neichel@lam.fr

Pour plus d'informations : https://www.onera.fr/rejoindre-onera/la-formation-par-la-recherche