

Algorithmes d'approximation de faible rang pour la super-résolution en imagerie de fluorescence

Sujet de stage de M2

Ce document contient une offre de stage financé de M2 d'une durée de 6 mois. Il sera réalisé au sein de l'équipe Imagerie Couleur du laboratoire CRISAL, à Lille. Pour toute demande, il est possible de s'adresser à clemence[dot]prevost[at]univ[dash]lille[dot]fr.

1. Contexte et motivation

L'autofluorescence définit l'émission naturelle de lumière par des tissus biologiques [7]. Il s'agit d'une intéressante propriété pour la caractérisation d'échantillons biologiques en ophtalmologie, en cytométrie en flux, ou dans l'étude des polymères [2, 6, 11]. L'exploitation de l'autofluorescence des composés présents dans le blé permet l'étude du développement du grain face à des enjeux de changement climatique et d'accroissement de la population mondiale. L'analyse des composés présents permet de suivre et de mesurer les capacités nutritionnelles et de résistance du grain à différents stades de son développement.

L'acquisition d'images de grains de blés à différents stades de développement a été conduite en juin 2024 au Synchrotron SOLEIL dans le cadre d'un partenariat avec le LASIRE (Univ. Lille), l'ENS Paris-Saclay et l'INRAE de Nantes. La source de lumière intense et hautement modulable disponible au Synchrotron a permis l'obtention d'images par des microscopes équipés de filtres spectraux. Les données résultantes sont des cubes de données possédant deux dimensions spatiales (les pixels) et une dimension spectrale (spectres d'autofluorescence).

Les acquisitions ont été réalisées sur les deux microscopes disponibles sur la ligne DISCO du synchrotron. Le microscope plein champ Télémós acquiert l'intégralité de la zone spatiale couverte par l'échantillon à haute résolution spatiale (500nm) mais avec une résolution spectrale réduite (une dizaine de bandes spectrales) sur la plage spectrale allant de 280 à 560nm. Le microspectrofluorimètre Polyphème acquiert, quant à lui, des images d'une portion d'un échantillon avec une haute résolution spectrale sur la plage de 190 à 900 nm (plusieurs centaines de bandes spectrales) mais une faible résolution spatiale (1-5 microns). Les images et spectres d'autofluorescence obtenues à partir de ces deux instruments sont illustrées aux Figures 3 et 4.

Ainsi, l'information spectrale comprise dans l'image issue de Télémós est incluse totalement dans l'information spectrale issue de Polyphème, dans sa version dégradée. A l'inverse, l'information spatiale à basse résolution issue de Polyphème couvre une partie de l'information à haute résolution contenue dans l'image provenant de Télémós. Les coordonnées des platines des microscopes, ainsi que les métadonnées provenant des logiciels d'acquisition, permettent de recalibrer spatialement l'image issue de Polyphème sur les données plein champ fournies par Télémós.

Afin d'exploiter à plein les résolutions complémentaires de chaque microscope, et d'optimiser le temps de faisceau mis à disposition des utilisateur·ices du Synchrotron, il est désirable d'obtenir des images pleinement résolues dans chaque dimension. Le but de ce stage est de développer des méthodes de super-résolution adaptées aux images d'autofluorescence acquises au Synchrotron. Les algorithmes développés devront être adaptés au grand volume des données ainsi qu'à la forte variabilité des informations présentes due à l'acquisition.

En effet, la manipulation d'échantillons vivants requiert de prendre en compte l'évolution de leurs propriétés

d'autofluorescence au cours du temps. Un échantillon primo-exposé à une source UV pourra ainsi fluorescer davantage lors de l'acquisition réalisée par le second microscope. Par ailleurs, les variétés d'intérêt présentes dans l'échantillon possèdent naturellement une forte variabilité spectrale. Si peu de composés fluorescent, et sont en pratique intéressants pour les biologistes, leur spectre d'autofluorescence varie fortement en fonction de leur localisation dans l'échantillon et du stade de croissance du grain de blé. Finalement, la source de lumière Synchrotron étant ré-alimentée périodiquement, les variations dans les conditions d'acquisition sont particulièrement importantes à prendre en compte dans la conception de solutions au problème de super-résolution.

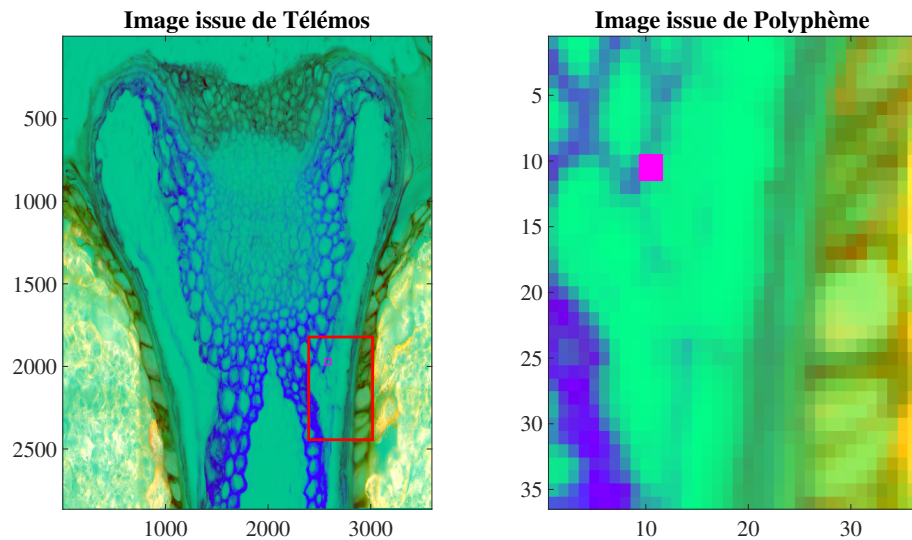


Fig. 1. Image en fausse couleur d'une image Télémós (gauche) et l'image Polyphème correspondante (droite). La portion de l'échantillon acquis par Polyphème est symbolisé en rouge sur l'image plein champ. Le pixel retenu à la Figure 4 est marqué en magenta.

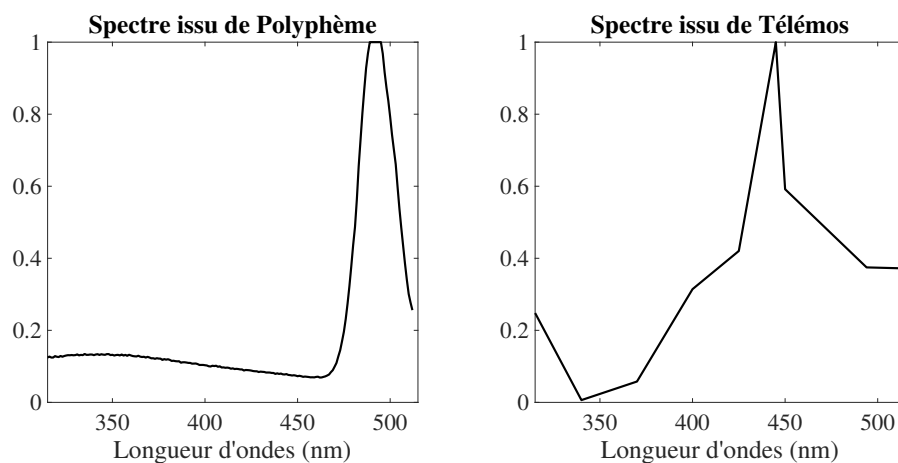


Fig. 2. Spectres d'autofluorescence obtenus à partir des images Polyphème et Télémós sur le pixel retenu en magenta à la Figure 3.

2. Approximations tensorielles de faible rang

Dans le problème de super-résolution considéré, il y a davantage d'inconnues à estimer que d'observations disponibles. On parle dans ce cas d'un problème inverse mal posé. C'est-à-dire qu'en l'absence d'hypothèses supplémentaires ou de régularisations sur les données, la solution au problème n'est pas unique. C'est un enjeu important qui compromet fortement l'interprétabilité des résultats, a fortiori pour permettre l'interprétation des images résultantes par des chercheur-euses en biologie.

Comme les images d'autofluorescence possèdent trois dimensions, elles peuvent être modélisées comme des tableaux multidimensionnels appelés tenseurs. Préserver le format tensoriel des données est un recours intéressant pour obtenir une solution unique à un problème inverse mal posé. En effet, considérer une approximation tensorielle de faible rang permet de fortement réduire le nombre d'inconnues à reconstruire et donne accès à des garanties d'unicité très flexibles pour les paramètres d'intérêt. Elle est notamment avantageuse par rapport à des modèles matriciels, qui eux, requièrent des hypothèses pour garantir une solution unique et comportent souvent davantage de paramètres à estimer [10, Chap. 1].

Les algorithmes classiques d'approximation matricielle ou tensorielle de faible rang sont généralement basés sur des mises à jour alternées de chacun des facteurs de la décomposition [4]. Les sous-problèmes relatifs à chaque facteur sont résolus alternativement en considérant les autres paramètres fixés. De nombreuses variantes de ces algorithmes existent. Elles permettent de prendre en compte des contraintes de positivité sur les données, des contaminations par des bruits non-gaussiens ou des régularisations supplémentaires [5]. Le développement d'algorithmes adaptés à la super-résolution en autofluorescence doit permettre d'exploiter à plein le potentiel méthodologique offert par le formalisme tensoriel, tout en respectant le modèle d'observation des images.

3. Objectifs du stage

3.1 Développement d'algorithmes tensoriels de super-résolution

Le premier objectif consiste en la conception d'algorithmes tensoriels de super-résolution en autofluorescence. Les méthodes développées seront basées sur la décomposition canonique polyadique (CP). Un tenseur $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ admet une décomposition CP de la forme

$$\mathcal{X} = \sum_{r=1}^R [[\mathbf{A}, \mathbf{B}, \mathbf{C}]], \quad (1)$$

où les matrices $\mathbf{A} \in \mathbb{R}^{I \times R}$, $\mathbf{B} \in \mathbb{R}^{J \times R}$ et $\mathbf{C} \in \mathbb{R}^{K \times R}$ représentent les facteurs de la décomposition. L'entier R représente le plus petit nombre de termes compris dans la somme (2); lorsqu'il est minimal, il est appelé rang de la décomposition CP. La décomposition CP présentée dans (2) jouit de propriétés d'unicité à des conditions très flexibles [3]. Elle est couramment utilisée pour résoudre des problèmes inverses en grande dimension. Dans le cadre de l'imagerie d'autofluorescence, les entiers I, J représentent le nombre de pixels dans les images tandis que l'entier K représente le nombre de bandes spectrales.

Les images à hautes résolutions reconstruites doivent être à valeurs positives afin de garantir leur interprétabilité. Cette contrainte primordiale motive la conception d'algorithmes de factorisation dits *non-négatifs*. Par ailleurs, d'autres contraintes, telles que la variation non brutale des niveaux dans la dimension spatiale et/ou spectrale, pourront être incorporées aux algorithmes pour améliorer la qualité des images re-

construites.

Le travail à faire sur ce premier objectif est le suivant :

1. Prise en main d'un algorithme de l'état de l'art basé sur une projection de gradient accélérée [9] : identification du rang optimal et des limites de l'approche;
2. Conception d'un algorithme non-négatif basé sur des mises à jour multiplicatives [8] : comparaison des performances en termes de qualité de la reconstruction, de la convergence et du temps de calcul avec l'approche par gradient projeté;
3. Incorporation de contraintes de variation non brutale spatiale et spectrale et conception d'un algorithme basé sur la méthode des multiplicateurs de Lagrange [1].

Low-rank approximation methods for super-resolution fluorescence imaging

MSc Internship

This document contains an offer for a funded MSc internship with 6-month duration. The intern will be hosted in the Color team of the CRISAL lab, in Lille. For inquiries please contact `clemence[dot]prevost[at]univ[dash]lille[dot]fr`.

1. Motivation

The property of biological tissues to naturally emit photons is called autofluorescence [7]. It is a key property of materials for many applications, e.g., the characterization of biological samples in ophthalmology, flow cytometry, and for studying polymers [2, 6, 11]. When studying wheat, leveraging the information obtained from autofluorescence allows one to study the grain's growth. This is crucial in a context of climate change and global population increase. The analysis of the chemical compounds (proteins) in the wheat allows one to measure its nutritional values as well as its resistance at diverse stages of its growth.

In June 2024, experiments were conducted at the Synchrotron SOLEIL particle accelerator within a partnership with LASIRE (Univ. Lille), ENS Paris-Saclay and INRAE (Nantes). These experiments allowed for the acquisition of wheat grain images with high resolution using microscopes, and were permitted by the intense and highly tunable light source available in the Synchrotron. The resulting observations are data cubes with two spatial dimensions (the pixels) and a spectral dimension that contains the autofluorescence spectra.

The acquisitions were conducted with the two microscopes of the DISCO UV beamline. The full-scope microscope Telemos covers the whole spatial region comprised in the sample with a high spatial resolution (500nm) but with a low spectral resolution (approx. 10 spectral bands) within the spectral range from 280nm to 560nm. Conversely, the microspectrofluorimeter Polypheme allows one for the acquisition of a small portion of a sample with a low spatial resolution (1 μ m to 5 μ m) but with a high spectral resolution spanning the range 190-900nm with hundreds of spectral bands.

Given this setup, the spectral information available with low resolution in the Telemos images are fully included in the highly-resolved Polypheme image. On the other hand, Polypheme provides low-information spatial resolution that covers a portion of the spatial information available with high resolution in Telemos. The microscope coordinates and metadata from the acquisition devices allow one for a precise alignment of the spatial information in the two observations.

In order to fully leverage the complementary resolutions of each acquisition device, and to optimize the beam time obtained by Synchrotron users, one may seek for images with high-resolution in each one of the three dimensions. This internship hence aims at designing super-resolution methods for fluorescence Synchrotron images. The proposed algorithms must be compatible with the large volume spanned by the data as well as the high variability in the acquisition process.

Indeed, the observation of living samples requires to account for the evolution of their autofluorescence properties through time. When a sample had been previously exposed to a UV light source, it may emit more photons during the second acquisition. Furthermore, the chemical varieties of interest naturally possess a high spectral variability. It means that their fluorescence spectrum greatly varies as a function of their spatial location within the sample and as a function of the grain's growth. Additionally, the Synchrotron

light source is periodically provided with new electrons. This phenomenon induces great variations in the acquisition conditions, with must be accounted for when designing super-resolution methods.

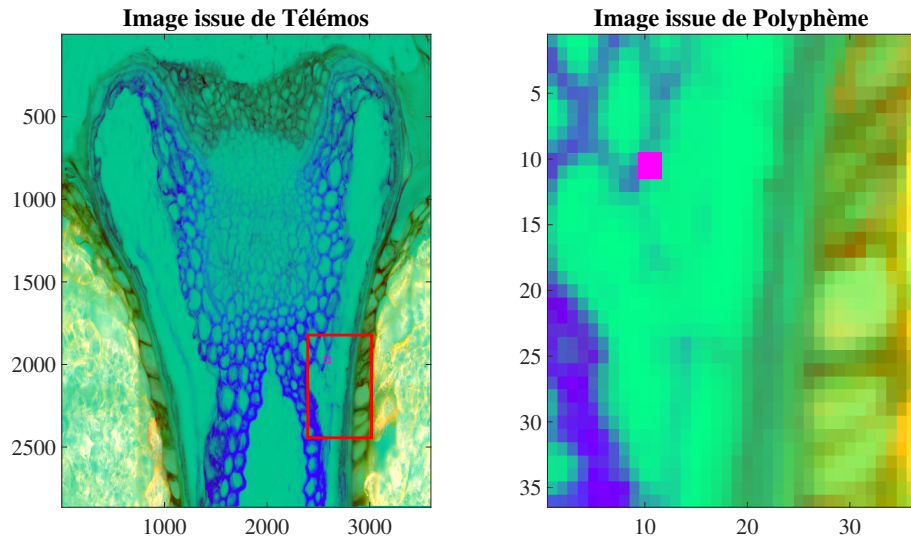


Fig. 3. False color images from Telemos (left) and corresponding Polyphème (right). The Polyphème portion is highlighted in red in Telemos. The considered pixel from Figure 4 is marked in pink.

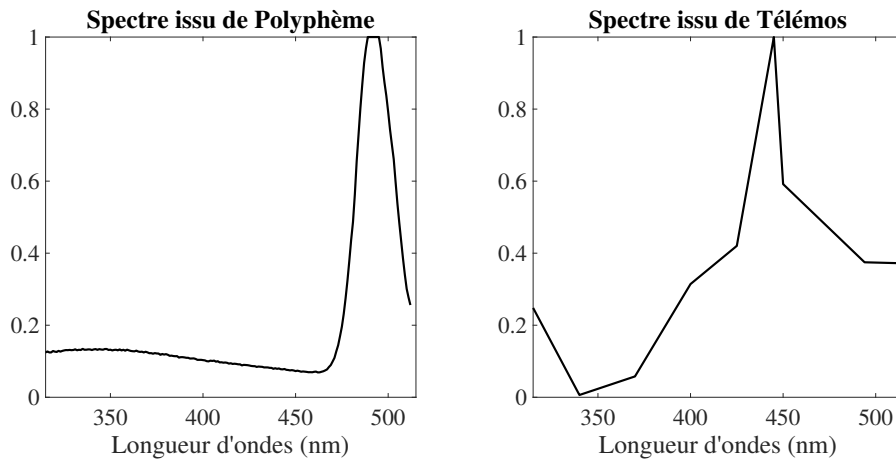


Fig. 4. Autofluorescence spectra from Polyphème and Telemos at the pink pixel from Figure 3.

2. Low-rank tensor approximations

For the considered super-resolution problem, there are more parameters to estimate than there are available observations. The problem is therefore termed as *ill-posed*. It means that its solution is not unique in the absence of additional regularizations or diversities. Uniqueness of the solution is a key issue that guarantees the interpretability of the solution for further processing by biologists.

Since the observed images have three dimensions, they are viewed as multidimensional arrays called tensors. Preserving the tensor format of the observations is a possible way to obtain a unique solution to an

ill-posed inverse problem. Indeed, low-rank tensor approximations act as a powerful form of regularization as they reduce the number of parameters to estimate. Furthermore, some tensor decompositions have interesting uniqueness guarantees under mild conditions [10, Chap. 1].

State-of-the-art (matrix or) tensor approximation methods are generally based on alternating updates of each factor in the decomposition [4]. The subproblem relative to one given factor is solved alternatively with the others while considering the other factors fixed. Such algorithms exist under many forms. Some of them allow for the addition of positivity constraints on the data, contamination by non-Gaussian noise, or other regularizations [5]. The design of super-resolution methods for autofluorescence imaging need to fully exploit the flexibility offered by the tensor formalism as well as the observation model underlying the images.

3. Goals

This internship aims at designing new tensor methods for super-resolution autofluorescence imaging. The proposed approaches will be based on the canonical polyadic (CP) decomposition. A tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ admits a CPD of the form

$$\mathcal{X} = \sum_{r=1}^R [[\mathbf{A}, \mathbf{B}, \mathbf{C}]], \quad (2)$$

where the matrices $\mathbf{A} \in \mathbb{R}^{I \times R}$, $\mathbf{B} \in \mathbb{R}^{J \times R}$ et $\mathbf{C} \in \mathbb{R}^{K \times R}$ are called the CP factors. The integer R is the number of terms in (2); when minimal, R is called the tensor rank. The CP decomposition in eqrefcp is unique under mild conditions [3], therefore it is very popular for solving high-dimensional inverse problems. In autofluorescence imaging, the integers I, J represent the number of pixels in the images whereas K stands for the number of spectral bands.

The entries of the target high-resolution images must be positive in order to be interpretable. This crucial constraint motivates for the design of non-negative factorization algorithms. Furthermore, other constraints, such as smooth variations in the spatial and/or spectral dimensions, can be incorporated to further enhance the quality of reconstructed images.

Hereafter are the tasks to be conducted in this internship.

1. Tuning of a state-of-the-art algorithm based on projected accelerated gradient [9]: identification of the optimal tensor rank and of the limits of the approach;
2. Design of a non-negative algorithm based on multiplicative updates [8]: evaluation of the reconstruction performance, convergence and computation time;
3. Incorporation of smoothness constraints and design of an algorithm based on the alternating direction method of multipliers [1].

4. References

- [1] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.
- [2] C. Chevalier, O. Leymarie, L. Sedano, B. Da Costa, C.-A. Richard, P. Maisonnasse, M. Réfrégiers, F. Jamme, and R. Le Goffic. P_{b1-f2} amyloid-like fibers correlate with proinflammatory signaling and respiratory distress in influenza-infected mice. *Journal of Biological Chemistry*, 297(1), 2021.
- [3] L. De Lathauwer. Decompositions of a higher-order tensor in block terms—part ii: Definitions and uniqueness. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1033–1066, 2008.
- [4] L. De Lathauwer and D. Nion. Decompositions of a higher-order tensor in block terms—part iii: Alternating least squares algorithms. *SIAM journal on Matrix Analysis and Applications*, 30(3):1067–1083, 2008.
- [5] N. Gillis. *Nonnegative matrix factorization*. SIAM, 2020.
- [6] F. Jamme, B. Cinquin, Y. Gohon, E. Pereiro, M. Réfrégiers, and M. Froissard. Synchrotron multimodal imaging in a whole cell reveals lipid droplet core organization. *Journal of Synchrotron Radiation*, 27(3):772–778, 2020.
- [7] F. Jamme, S. Kascakova, S. Villette, F. Allouche, S. Pallu, V. Rouam, and M. Réfrégiers. Deep uv autofluorescence microscopy for cell biology and tissue histology. *Biology of the Cell*, 105(7):277–288, 2013.
- [8] C.-J. Lin. On the convergence of multiplicative update algorithms for nonnegative matrix factorization. *IEEE Transactions on Neural Networks*, 18(6):1589–1596, 2007.
- [9] D. N. Phan, N. Gillis, et al. An inertial block majorization minimization framework for nonsmooth nonconvex optimization. *Journal of Machine Learning Research*, 24(18):1–41, 2023.
- [10] C. Prévost. *Multimodal data fusion by coupled low-rank tensor approximations*. PhD thesis, Université de Lorraine, 2021.
- [11] K. Vidot, M.-F. Devaux, C. Alvarado, S. Guyot, F. Jamme, C. Gaillard, R. Siret, and M. Lahaye. Phenolic distribution in apple epidermal and outer cortex tissue by multispectral deep-uv autofluorescence cryo-imaging. *Plant science*, 283:51–59, 2019.