



7ème édition de la journée sur la

## «Co-conception : capteurs hybrides et algorithmes pour des systèmes innovants»

**18 Mai 2021**

*Journée des GDR ISIS, MIA, et Ondes*

Cette journée se tiendra en distanciel. Pour des raisons techniques liées au nombre de connexions simultanées, **l'inscription est gratuite, mais obligatoire** sur le site du GdR Isis. Les informations concernant la visioconférence seront envoyées aux inscrits quelques jours avant la réunion.

### PROGRAMME

**Durée des exposés : 15 minutes + 5 minutes de questions**

**9h - 9h30 : Accueil et Introduction**

**9h30 - 10h35 : Qualité image**

**9h30 - 10h15 : Présentation invitée : Rafael Falcón - Co-design on smartphone cameras: Image quality challenges and solutions of miniaturized optical systems**

Rafael Falcón, Image Quality Project Leader, DxOMark Image Labs, Boulogne, France

**10h15 - 10h35 : HDROmni : Extension optique de gamme dynamique en vision panoramique pour la robotique**

Julien Ducrocq<sup>1</sup>, Guillaume Caron <sup>1,2</sup> et El Mustapha Mouaddib<sup>1</sup>

1. Université de Picardie Jules Verne, laboratoire Modélisation, Information et Systèmes, France

2. CNRS-AIST Joint Robotics Laboratory, IRL, Tsukuba, Japan

**10h35 - 10h45 Pause**

**10h45 - 11h05 : Entraînement auto-supervisé pour le débruitage aveugle de vidéo**

Valéry Dewil, Jérémy Anger, Axel Davy, Thibaud Ehret, Gabriele Facciolo et Pablo Arias  
Centre Borelli, Ecole Normale Supérieure Paris-Saclay, Gif- sur-Yvette, France

**11h05 - 11h25: Neuro-inspired quantisation mechanism**

Effrosyni Doutsis,<sup>1,2</sup> Marc Antonini<sup>1</sup>, Panagiotis Tsakalides<sup>2,3</sup>

1. I3S, CNRS, UCA

2. Signal Processing Lab, FORTH/ICS, Greece,

3. University of Crete, Greece

**11h25 - 11h45: Self-supervised multi-image super-resolution for push-frame satellite images**

Ngoc-Long Nguyen, Jérémy Anger, Axel Davy, Pablo Arias et Gabriele Facciolo  
Centre Borelli, Ecole Normale Supérieure Paris-Saclay, Gif- sur-Yvette, France

**11h45 - 11h55 Pause et Discussion**

**11h55 - 12h05 : Optique non conventionnelle et réseau de neurones**

**11h55 - 12h15 : Algorithmes d'apprentissage pour la segmentation d'images hyperspectrales compressées**

Maud Biquard<sup>1</sup>, Antoine Rouxel<sup>1,2</sup>, Simon Lacroix<sup>1</sup>, Antoine Monmayrant<sup>1</sup>

1. LAAS-CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse, France

2. Airbus Defence and Space, 31 Rue des Cosmonautes, 31400 Toulouse, France

**12h15 - 12h35 : Apport des réseaux profonds à extracteur complexe sur données de microscopie augmentée pour le diagnostic du paludisme**

Houda Hassini<sup>1</sup>, Lyes Bouchama<sup>2,1</sup>, Arthur Baroni<sup>1</sup>, Bernadette Dorizzi<sup>1</sup>, Yaneck Gottesman<sup>1</sup>, Jacques Klossa<sup>2</sup>, Marc Thellier<sup>3</sup>

1. Telecom SudParis

2. TRIBVN

3. Centre de Référence du Paludisme de la Pitié Salpêtrière

**12h35 - 13h30 : Pause Déjeuner**

**13h30 - 14h30 : Discussion**

**14h30 - 15h10 : Déconvolution**

**14h30 - 14h50 : Comparaison de modèles en déconvolution d'image : probabilité de modèles et échantillonnage stochastique**

Benjamin Harroué<sup>1,2</sup>, Jean-François Giovannelli<sup>1</sup> et Marcelo Pereyra<sup>2</sup>

1. Laboratoire IMS (Univ. Bordeaux – CNRS – BINP), Talence, France

2. MACS, Heriot-Watt University, Edinburgh, United Kingdom

**14h50 - 15h10 : Comment la déconvolution par zones influence-t-elle les performances de systèmes conventionnels et de systèmes co-conçus ?**

Alice Fontbonne, Hervé Sauer, François Goudail

Université Paris-Saclay, Institut d'Optique Graduate School, CNRS, Laboratoire Charles Fabry, 91127 Palaiseau, France

**15h10 - 15h20 : Pause**

**15h20 - 17h45 : Modélisation et optimisation conjointe**

**15h20 - 15h40 : Conception optique ex-nihilo à partir de la méthode « Simultaneous Multiple Surfaces »**

Thibaut Mayeur<sup>1,2</sup>, Jean-Baptiste Volatier<sup>1</sup>

ONERA Palaiseau, France

Lynred, Palaiseau, France

**15h40 - 16h00 : Quelle influence peut avoir le modèle d'imagerie sur l'optimisation de masques de phase augmentant la profondeur de champ en microscopie de localisation ?**

Olivier Lévêque, Caroline Kulcsár et François Goudail  
Université Paris-Saclay, Institut d'Optique Graduate School, CNRS, Laboratoire Charles Fabry, 91127, Palaiseau, France.

**16h - 16h20 : Pause**

**16h20 - 16h40 : Amélioration des performances du SLAM RGB-D en tenant compte des spécifications du capteur pour la localisation indoor.**

Imad El Bouazzaoui, Sergio Rodriguez Florez, Abdelhafid El Ouardi  
SATIE-CNRS UMR 8029, Paris-Saclay University, France.

**16h40 - 17h00 : Co-conception optique/réseau de neurones à partir d'un modèle optique par tracé de rayon différentiel**

Aymeric Halé, Pauline Trouvé-Peloux, Jean-Baptiste Volatier  
ONERA, Palaiseau, France

**17h00 - 17h45 : Présentation invitée : Gordon Wetzstein - End to end optimization of optics and algorithm**

Gordon Wetzstein, Assistant Professor of Electrical Engineering and, by courtesy, Computer Science at Stanford University

# Présentations invitées

## **Rafael Falcón : Co-design on smartphone cameras: Image quality challenges and solutions of miniaturized optical systems**

### **Abstract:**

With the introduction of the smartphone in 2007, digital photography got, literally and figuratively, on hands of millions of users. Camera quality being one of the main points customers look for when buying smartphones, it has become one of the largest drives for research and development in the embedded camera industry. On this work we study how codesign has been used by the largest camera manufacturers to tackle the challenges of miniaturization, power efficiency, computational efficiency, and cost. We cover technologies anchored in co-design, such as color shading and GRGB imbalance-correction, HDR, PDAF, OIS plus EIS, super-resolution, zoom with fix focal lens, ZSL, among others. These improvements in optical design and image processing have helped the smartphone companies to gain up to 7.3 stops of image quality, compared to similar camera performances in 2005, while keeping a relatively simple point-and-click system in the eyes of the user.

### **Co-authors:**

Rafael Falcon, Benoit Pochon, and Frederic Guichard  
DxOMark Image Labs, Boulogne, France

## **Gordon Wetzstein : End to end optimization of optics and algorithm**

### **Abstract:**

Neural networks and other advanced image processing algorithms excel in a wide variety of computer vision and imaging applications, but their high performance also comes at a high computational cost, and their success is sometimes limited. In this talk, we explore hybrid optical-digital strategies to computational imaging that outsource parts of the algorithm into the optical domain. Using such a co-design of optics and image processing, we can design application-domain-specific cameras or compute parts of a convolutional neural network in optics. Optical computing happens at the speed of light and without any memory or power requirements, thereby opening new directions for intelligent imaging systems.

### **Biography:**

Gordon Wetzstein is an Assistant Professor of Electrical Engineering and, by courtesy, Computer Science at Stanford University. He is the leader of the Stanford Computational Imaging Lab and a faculty co-director of the Stanford Center for Image Systems Engineering. At the intersection of computer graphics and vision, computational optics, and applied vision science, Prof. Wetzstein's research has a wide range of applications in next-generation imaging, display, wearable computing, and microscopy systems. Prior to joining Stanford in 2014, Prof. Wetzstein was a Research Scientist at MIT. He received a Ph.D. in Computer Science from the University of British Columbia in 2011 and graduated with Honors from the Bauhaus in Weimar, Germany prior to that. He is the recipient of an NSF CAREER Award, an Alfred P. Sloan Fellowship, an ACM SIGGRAPH Significant New Researcher Award, a Presidential Early Career Award for Scientists and Engineers (PECASE), an SPIE Early Career Achievement Award, a Terman Fellowship, an Okawa Research Grant, the Electronic Imaging Scientist of the Year 2017 Award, an Alain Fournier Ph.D. Dissertation Award, and a Laval Virtual Award, as well receiving Best Paper and Demo Awards at ICCP 2011, 2014, and 2016, and at ICIP 2016.

# Résumés des contributions

## **HDR Omni : Extension optique de gamme dynamique en vision panoramique pour la robotique**

Julien Ducrocq<sup>1</sup>, Guillaume Caron <sup>1,2</sup> et El Mustapha Mouaddib<sup>1</sup>

1. Université de Picardie Jules Verne, laboratoire Modélisation, Information et Systèmes, France

2. CNRS-AIST Joint Robotics Laboratory, IRL, Tsukuba, Japan

Cette communication introduit une nouvelle méthode de conception qui permet d'étendre optiquement la gamme dynamique d'une caméra panoramique pour la robotique. La large gamme de radiance d'une scène est capturée en temps réel par une combinaison de miroirs paraboliques, une lentille télécentrique et des filtres à densité neutre [1]. Le premier prototype de caméra panoramique à large gamme dynamique (HDR), nommé HDR Omni, est constitué de quatre miroirs et de trois filtres à densité neutre. Il augmente 100 fois la gamme dynamique de la caméra utilisée. Le prototype capture ainsi quatre images d'expositions différentes simultanément, chacune avec une faible gamme dynamique (LDR). Ces images sont ensuite fusionnées en une image HDR. Plusieurs évaluations comparent la nouvelle caméra avec des caméras à dynamique classique dans plusieurs scènes difficiles, c'est-à-dire avec des zones sous-exposées et sur-exposées. Elles montrent qu'en moyenne, la qualité de la gamme dynamique de l'image double. En effet, aucune des quatre images LDR capturées simultanément par HDR Omni ne parvient à capturer l'intégralité de la scène. L'image HDR, en revanche, capture toutes les régions de la scène, quel que soit leur éclairage. Des évaluations supplémentaires montrent que HDR Omni permet aussi d'accroître le nombre d'appariements de points caractéristiques. Enfin, HDR Omni engendre aussi un domaine de convergence plus étendu et une meilleure précision à convergence de l'asservissement visuel direct (basé intensités de pixels) d'un robot mobile à l'extérieur.

[1] J. Ducrocq, G. Caron, et E. Mouaddib. HDR Omni : Optical extension of dynamic range for panoramic robot vision. IEEE Robotics and Automation Letters, 2021, vol. 6, no 2, p. 3561-3568.

## **Entraînement auto-supervisé pour le débruitage aveugle de vidéo**

Valéry Dewil, Jérémy Anger, Axel Davy, Thibaud Ehret, Gabriele Facciolo et Pablo Arias  
Centre Borelli, Ecole Normale Supérieure Paris-Saclay, Gif-sur-Yvette, France

Les réseaux de neurones ont atteint l'état de l'art en matière de débruitage : DnCNN [6] reste une référence en matière de débruitage d'images, tout comme FastDVDnet [4] pour le débruitage de vidéos. Ces réseaux de neurones sont particulièrement efficaces sur un type et niveau de bruit spécifique et pour laquelle ils ont été entraînés. Cependant, leur performance décroît très rapidement lorsqu'ils sont évalués sur des bruits de distribution différents. Bien qu'ils puissent en général être réentraînés efficacement de manière supervisée pour débruiter n'importe quelle autre distribution de bruit, de tels réentraînements sont extrêmement coûteux et nécessitent une large base de données avec des paires bruitées/débruitées.

Récemment, Ehret et al ont introduit Frame2Frame(F2F) [2]. Cette méthode permet de réapprendre les poids d'un réseau de débruitage initialement entraîné pour un type de bruit spécifique, afin qu'il s'adapte à un nouveau type de bruit (ou un autre niveau de bruit). La force de cette méthode de "fine-tuning" est qu'elle est auto-supervisée : elle ne nécessite pas d'autres données que la vidéo dégradée elle-même. Cet apprentissage est basé sur l'entraînement noise2noise [3]. Néanmoins cette méthode a le défaut d'être appliquée à un débruiteur d'images (DnCNN), conduisant à une forte inconsistance temporelle. En nous basant sur F2F, nous proposons MF2F (Multi Frame2Frame), en l'adaptant plus particulièrement à FastDVDnet qui atteint l'actuel état de l'art, bien que notre méthode puisse aussi être appliquée à d'autres réseaux multi-images. Nos résultats montrent que partant de poids pré-entraînés pour du bruit gaussien et fournis par les auteurs de [4], les réseaux réentraînés à partir d'une unique séquence vidéo atteignent les performances du réseau initial entraîné de manière supervisée sur une large base de données. Et ce pour une vaste gamme de types de bruit et pour des niveaux de bruit différents.

En outre, nos réseaux réentraînés dépassent parfois les performances des réseaux entraînés par supervision.

Par ailleurs, des expériences sur des sauts de bruit au cours de séquences vidéo ont montré que l'apprentissage en ligne permet de s'adapter à un bruit non constant temporellement au cours de la vidéo.

Le réseau FastDVDnet prend en entrée une carte de niveau de bruit. Diverses expériences ont été menées sur cette carte de bruit. Partant d'une carte de bruit fixée, notre méthode peut adapter la carte de bruit en gardant les poids du réseau fixes. Nous avons mis au point différentes stratégies pour traiter des cas particuliers de débruitage : carte uniforme pour du bruit additif gaussien homoscédastique, apprentissage d'une carte de bruit spatialement variée pour du bruit hétéroscédastique, apprentissage d'une carte constante par niveau pour les bruits dépendant du signal (particulièrement adapté au bruit de Poisson).

Les expériences précédentes ont montré l'efficacité de notre méthode sur du bruit simulé (bruits synthétiques). Plus récemment, nous avons appliqué notre méthode pour le cas de bruit réel. Le récent réseau RViDeNet proposé par [5] atteint des performances état-de-l'art en matière de débruitage de séquences vidéo brutes (raw). Pourtant, il souffre de nombreux artefacts. Les auteurs de [5] ont eux aussi proposé un "fine-tuning" de leur réseau pour s'adapter à des séquences réelles, mais ce "fine-tuning" est réalisé en amont et les poids sont fixés lors de l'inférence sur une nouvelle séquence. A l'inverse, notre méthode s'adapte au bruit particulièrement présent dans la séquence à débruiter. Notre méthode donne des résultats sensiblement meilleurs que ceux obtenus avec RViDeNet (plus de détails, plus de netteté, une meilleure reconstruction en général), et ce pour diverses valeurs d'ISO. Par ailleurs, sur un bruit réel simulé selon [1] à partir d'images propres, et pour cinq différents ISO échelonnés de 1600 (peu bruité) à 25600 (très bruité), notre méthode atteint un PSNR en moyenne sur un dataset de 0.8dB au dessus de RViDeNet. Ces récentes expériences montrent l'intérêt de notre méthode en matière de débruitage de bruit réel.

## Références

[1] Tim Brooks, Ben Mildenhall, Tianfan Xue, Jiawen Chen, Dillon Sharlet, and Jonathan T Barron. Un-processing images for learned raw denoising. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 11036–11045, June 2019.

[2] Thibaud Ehret, Axel Davy, Jean-Michel Morel, Gabriele Facciolo, and Pablo Arias. Model-blind video denoising via frame-to-frame training. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[3] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2Noise: Learning image restoration without clean data. In Proceedings of the 35th International Conference on Machine Learning, volume 80 of Journal of Machine Learning Research, pages 2965–2974. PMLR, 10–15 Jul 2018.

[4] Matias Tassano, Julie Delon, and Thomas Veit. Fastdvdnet: Towards real-time deep video denoising without flow estimation. In The IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pages 1354–1363, June 2020.

[5] Huanjing Yue, Cong Cao, Lei Liao, Ronghe Chu, and Jingyu Yang. Supervised raw video denoising with a benchmark dataset on dynamic scenes. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2301–2310, June 2020.

[6] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. IEEE TIP, 26(7):3142–3155, 7 2017.

## Neuro-inspired quantisation mechanism

Effrosyni Doutsis<sup>1,2</sup>, Marc Antonini<sup>1</sup>, Panagiotis Tsakalides<sup>2,3</sup>

1. I3S, CNRS, UCA

2. Signal Processing Lab, FORTH/ICS, Greece

3. University of Crete, Greece

The 3D ultra-high-resolution world that is captured by the visual system is sensed, processed and transferred through a dense network of tiny cells, called neurons. An understanding of neuronal communication has the

potential to open new horizons for the development of ground-breaking image and video compression systems. A recently proposed neuro-inspired compression system promises to change the framework of the current state-of-the-art compression algorithms. In this presentation we will discuss the properties of this neuro-inspired compression. The presenter will show some interesting comparisons to the state-of-the-art while she will highlight the potential extensions such that this ground-breaking algorithm will be able to process higher dimensionality data and be integrated into machine learning based architectures.

## **Self-supervised multi-image super-resolution for push-frame satellite images**

Ngoc-Long Nguyen, Jérémy Anger, Axel Davy, Pablo Arias et Gabriele Facciolo  
Centre Borelli, Ecole Normale Supérieure Paris-Saclay, Gif-sur-Yvette, France

Recent constellations of optical satellites are adopting multi-image super-resolution (MISR) from bursts of push-frame images as a way to increase the resolution and reduce the noise of their products while maintaining a lower cost of operation. Most MISR techniques are currently based on the aggregation of samples from registered low resolution images. A promising research trend aimed at incorporating natural image priors in MISR consists in using data-driven neural networks. However, due to the unavailability of ground truth high resolution data, these networks cannot be trained on real satellite images. In this paper, we present a framework for training MISR algorithms from bursts of satellite images without requiring high resolution groundtruth. This is achieved by adapting the recently proposed frame-to-frame framework to process bursts of satellite images. In addition we propose an architecture based on feature aggregation that allows to fuse a variable number of frames and is capable of handling degenerate samplings while also reducing noise. On synthetic datasets, the proposed self-supervision strategy attains results on par with those obtained with a supervised training. We applied our framework to real SkySat satellite image bursts leading to results that are more resolved and less noisy than the L1B product from Planet.

## **Algorithmes d'apprentissage pour la segmentation d'images hyperspectrales compressées**

Maud Biguard<sup>1</sup>, Antoine Rouxel<sup>1,2</sup>, Simon Lacroix<sup>1</sup>, Antoine Monmayrant<sup>1</sup>

1. LAAS-CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse, France

2. Airbus Defence and Space, 31 Rue des Cosmonautes, 31400 Toulouse, France

La conception de systèmes de type "Coded Aperture Snapshot Spectral Imager" (CASSI, [1]) nécessite le développement d'algorithmes adaptés au processus d'acquisition. Contrairement aux systèmes traditionnels, ces instruments ne sont pas conçus pour mesurer des cubes hyperspectraux entiers mais des échantillons disparates et superposés de ce cube. Ainsi les informations mesurées sont difficilement directement exploitables par les algorithmes classiques d'analyse de données hyperspectrales. Nous proposons d'exploiter ces informations à l'aide de réseaux de neurones, pour segmenter une scène directement à partir des mesures compressées, c'est-à-dire sans passer par une phase de reconstruction. La méthode proposée est adaptée aux fonctionnalités du système "double-disperseur CASSI" [2] et considère le cas d'une segmentation à partir de deux acquisitions seulement : une image compressée et une image panchromatique. L'algorithme peut se décomposer en deux parties : traitement du cube de filtrage et segmentation sur la base d'un ensemble de classes. L'entraînement du réseau de neurones est effectué sur la globalité de la structure.

Le formalisme du "cube de filtrage" permet d'exprimer en 3 dimensions le filtrage spectral effectué par le modulateur spatial de lumière situé au sein du système optique. Via une architecture peu profonde de réseau de neurones, le traitement réduit une des dimensions de ce cube grâce à l'apprentissage de ses caractéristiques intrinsèques. La partie segmentation prend en entrée le cube de filtrage pré-traitée, l'image compressée et l'image panchromatique. L'architecture de cette partie de l'algorithme est dérivée de celle de DSSNet [3], utilisée pour la segmentation de cubes hyperspectraux "classiques". Cette architecture présente un bon compromis entre précision et rapidité de calcul. Elle se décompose elle-même en deux sous-parties :

mélange spectro-spatial et classification. D'abord, les informations entre pixels voisins sont mutualisées par des opérations de convolution. Ensuite, des opérations dans la dimension spectrale viennent associer chaque pixel de l'image à une classe.

L'approche proposée permet d'effectuer une segmentation sur la base d'images hyperspectrales compressées. Le nombre d'acquisitions, et donc la quantité de données, sont considérablement réduits en comparaison avec l'imagerie hyperspectrale classique. L'utilisation de réseaux peu profonds permet de réduire le temps d'inférence et ainsi d'obtenir des systèmes fonctionnant en temps réel.

[1] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied Optics*, vol. 47, pp. B44–B51, Apr. 2008.

[2] M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Optics Express*, vol. 15, p. 14013, oct 2007.

[3] B. Pan, X. Xu, Z. Shi, N. Zhang, H. Luo, and X. Lan, "DSSNet : A simple dilated semantic segmentation network for hyperspectral imagery classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, pp. 1968–1972, nov 2020.

## **Apport des réseaux profonds à extracteur complexe sur données de microscopie augmentée pour le diagnostic du paludisme**

Houda Hassini<sup>1</sup>, Lyes Bouchama<sup>2,1</sup>, Arthur Baroni<sup>1</sup>, Bernadette Dorizzi<sup>1</sup>, Yaneck Gottesman<sup>1</sup>, Jacques Klossa<sup>2</sup>, Marc Thellier<sup>3</sup>

1. Telecom SudParis

2. TRIBVN

3. Centre de Référence du Paludisme de la Pitié Salpêtrière

Ce travail s'inscrit dans le contexte d'une détection automatisée d'hématies infectées par Plasmodium. Dans notre cas, il s'agit de développer une approche de détection automatique sur frottis aussi sensible que la référence actuelle, la goutte épaisse, de façon à proposer un diagnostic complet du paludisme sur un simple frottis sanguin.

Dans le but d'imager rapidement plusieurs milliers d'hématies, par champ microscopique un objectif grand champ de faible grossissement (typiquement 10x) est utilisé. Afin de compenser la résolution insuffisante pour la détection de jeunes parasites (de taille inférieure au micron), nous avons développé une approche de microscopie d'holographie à ouverture synthétique qui se distingue sur 2 plans : une ouverture synthétique augmentée pour une super-résolution et une bimodalité Intensité et Phase permettant une représentation complexe de l'échantillon sanguin numérisé.

Au cours de cette présentation, nous discuterons de l'apport de la bi-modalité pour la sensibilité de la détection des hématies parasitées. Notre travail est centré sur le réseau Faster-RCNN pour la détection et la classification automatique des hématies.

Nous questionnons l'apport de la bi-modalité à travers des analyses comparatives obtenues sur différentes formes d'implémentation de ces réseaux : le Faster-RCNN sur données réelles et exploitant la modalité d'intensité seule ou d'intensité et de phase. Nous introduisons également un réseau Faster-RCNN modifié permettant l'exploitation de données nativement complexes d'intensité et de phase. Les filtres de convolution réelle sont alors remplacés par des filtres de convolution complexe plus adaptés à la nature physique des images. Nous montrerons l'amélioration des performances atteintes dans le cas de l'exploitation avec représentation complexe du couple intensité/phase.

Le travail s'inscrit dans une démarche de co-conception de la technique d'imagerie de microscopie et de traitement algorithmique des données d'imagerie complexes. Il est engagé dans le cadre d'une collaboration entre Telecom SudParis (équipe d'optique et de science des données), la société TRIBVN, acteur majeur



depuis près de 20 ans dans le secteur de la santé numérique (Digital Pathology) et le Centre National de Référence du Paludisme de la Pitié Salpêtrière.

## **Comparaison de modèles en déconvolution d'image : probabilité de modèles et échantillonnage stochastique**

Benjamin Harroué<sup>1,2</sup>, [Jean-François Giovannelli](#)<sup>1</sup> et Marcelo Pereyra<sup>2</sup>

1. Laboratoire IMS (Univ. Bordeaux – CNRS – BINP), Talence, France

2. MACS, Heriot-Watt University, Edinburgh, United Kingdom

La question de la reconstruction ou de la restauration (pour des images, signaux ou sons, vidéos, objets hyperspectraux, . . .) est un sujet d'intérêt dans des domaines variés (e.g., photographie numérique, astronomie, médecine, contrôle, . . .) et pour diverses modalités (e.g., microscopie, scanner et rayons X, radar, optique). Dans ce contexte, la conception de nouveaux instruments et des traitements associés repose entre autres choses sur la connaissance avancée des systèmes imageurs. La constitution de cette connaissance est souvent compliquée par les imperfections elles-mêmes de ces systèmes et le caractère mal-posé des problèmes qui y sont attachés. La prise en compte d'information d'origines diverses et surtout la gestion cohérente de ces informations constitue alors un aspect clé.

Nous nous intéressons à la comparaison de modèle pour les instruments mais également pour l'entrée et les erreurs, dans un cadre bayésien. Dans le contexte simple du filtrage de Wiener pour la déconvolution, on compare des modèles instrument et des densités spectrales d'entrée et d'erreur, dans un schéma où les hyperparamètres (e.g., niveaux de bruit et de signal) sont inconnus.

En quelques mots [1–3], notant  $\mathbf{y}$  et  $\mathbf{x}$  les images observée et inconnue, et  $\gamma$  les hyperparamètres, le principe est de calculer les probabilités  $p(M = m | \mathbf{y})$  pour  $m = 1, \dots, M$  de chacun des  $M$  modèles candidats  $M$ . Leur calcul repose sur la vraisemblance  $p(\mathbf{y} | M)$ , appelée évidence. Il s'agit d'une vraisemblance marginale par rapport à  $\mathbf{x}$  et  $\gamma$  dont le calcul présente en général une difficulté importante. Dans le cas présent, l'objet s'intègre explicitement mais pas les hyperparamètres. Nous nous appuyons alors sur l'idée de Chib [4] qui permet d'écrire l'évidence à partir de  $p(\gamma | \mathbf{y}, M)$ , la densité a posteriori marginale des hyperparamètres. Son calcul nécessite une autre marginalisation mais qui se résout en l'écrivant comme l'espérance de la densité  $p(\gamma | \mathbf{x}, \mathbf{y}, M)$  sous la distribution  $p(\mathbf{x} | \mathbf{y}, M)$ . Cette espérance s'approche par une moyenne empirique, les tirages de  $p(\mathbf{x} | \mathbf{y}, M)$  étant obtenus comme sous-produits d'un échantillonneur de Gibbs [2, 5] pour  $p(\mathbf{x}, \gamma | \mathbf{y}, M)$ .

Nous détaillons les différentes étapes indiquées ci-dessus : reformulation de l'évidence, mise en forme et calcul numérique, notamment les étapes de l'échantillonneur de Gibbs. Nous montrons des résultats numériques, sur données simulées, concernant la sélection au travers de matrices de confusion. Pour terminer nous donnons des résultats en terme de qualité de la déconvolution.

[1] J. Ding, V. Tarokh, and Y. Yang, "Model selection techniques : An overview," IEEE Signal Proc. Mag. vol. 35, pp. 16–34, nov. 2018.

[2] C. P. Robert, The Bayesian Choice. From decision-theoretic foundations to computational implementation. Springer Texts in Statistics, New York, USA : Springer Verlag, 2007.

[3] B. Harroué, J.-F. Giovannelli, and M. Pereyra, "Sélection de modèles en restauration d'image. Approche bayésienne dans le cas gaussien," in Actes 27 e coll. GRETSI, (Lille, France), août 2019.

[4] B. P. Carlin and S. Chib, "Bayesian model choice via Markov Chain Monte Carlo methods," J. R. Statist. Soc. B, vol. 57, pp. 473–484, 1995.

[5] F. Orieux, J.-F. Giovannelli, and T. Rodet, "Bayesian estimation of regularization and point spread function parameters for Wiener–Hunt deconvolution," J. Opt. Soc. Amer., vol. 27, pp. 1593–1607, juil. 2010.

## **Comment la déconvolution par zones influence-t-elle les performances de systèmes conventionnels et de systèmes co-conçus ?**

[Alice Fontbonne](#), Hervé Sauer, François Goudail

Aujourd'hui, la plupart des systèmes d'imagerie comprennent à la fois une partie optique et une partie traitement d'image afin d'améliorer la qualité finale de l'image. Il semble donc naturel de les optimiser simultanément pour obtenir le meilleur résultat possible. Un tel système optique, dit « système co-conçu » est fait pour fonctionner avec l'algorithme de traitement avec lequel il a été optimisé. Cependant, notre équipe a déjà montré qu'il était possible de modifier l'algorithme de traitement a posteriori pour obtenir de meilleures performances d'imagerie que celles prévues initialement (1). Nous étudions la possibilité d'améliorer les performances dans le champ d'optiques conçues de manière conventionnelle et d'optiques conçues conjointement sous CodeV en utilisant des critères alternatifs (2). Pour cela, une adaptation du traitement aux performances locales est envisagée : il s'agit de la déconvolution par zones. Comment la déconvolution par zones influence-t-elle les performances des systèmes d'imageries ?

Nous répondons à cette question par le biais de simulations pour établir les différences entre systèmes conventionnels et systèmes co-conçus. En particulier, nous étudions la difficulté de déconvolution intrinsèque à chaque zone de traitement et la baisse de performance liée à l'inadéquation du traitement de la zone à la PSF locale. Cette étude permet de vérifier qu'un nombre limité de zones est suffisant pour garantir des performances optimales.

1. Comparison of linear and nonlinear deconvolution algorithms for co-optimization of depth-of-field enhancing binary phase masks. Lévêque, Olivier, Kulcsár, Caroline and Goudail, François. OSA Continuum, Vol. 4, p. 589 (2021).
2. Joint digital-optical design of complex lenses using a surrogate image quality criterion adapted to commercial optical design software. Burcklen, Marie-Anne, et al. Applied Optics, Vol. 57, p. 9005 (2018).

## **Conception optique ex-nihilo à partir de la méthode « Simultaneous Multiple Surfaces »**

Thibaut Mayeur<sup>1,2</sup>, Jean-Baptiste Volatier<sup>1</sup>  
ONERA Palaiseau, France  
Lynred, Palaiseau, France

La conception optique repose actuellement en grande partie sur l'expérience du designer. En effet l'expérience d'un designer permet de mieux transcrire les spécifications du système en une fonction de mérite optique. Suite à cela il faut trouver numériquement le système optique qui permet de minimiser cette fonction de mérite. Pour trouver ce minimum l'algorithme couramment utilisé est celui de Levenberg-Marquardt. Cet algorithme réalise une optimisation locale, ce qui implique que le système optique de départ est déterminant pour l'obtention d'un bon résultat final. C'est pour cela que l'expérience du designer est importante pour trouver les meilleurs points de départ ou alors pour sortir plus facilement des différents minima locaux qui ne satisferaient pas les spécifications optiques. Ces deux points sont clefs pour obtenir un système optique optimum. C'est pour ces raisons qu'il est intéressant d'étudier des méthodes de conception ex-nihilo (à partir de rien) qui permettraient à tout designer d'arriver au système optique optimum répondant à leur cahier des charges. De plus, les systèmes optiques récents utilisent de plus en plus des surfaces optiques sans axe ni centre de symétrie (surfaces freeform), car leur utilisation permet de concevoir des systèmes plus compacts à qualité optique équivalente. Mais la conception de ces systèmes plus performants se fait au prix d'une optimisation plus complexe due au grand nombre de degrés de libertés supplémentaires. Ceci renforce l'impact de l'expérience du designer sur le résultat final obtenu.

Ma présentation portera sur une méthode de construction directe des surfaces optiques qui permet de contourner ces problèmes. Cette méthode est la méthode « Simultaneous Multiple Surfaces ». Elle permet de déduire directement les surfaces optiques d'un système à partir de quelques paramètres (champs, chemin optique, courbures des surfaces au voisinage de l'axe optique, focale...). Cette méthode permet de se rapprocher d'une conception optique ex-nihilo, en réduisant l'impact de l'expérience du designer sur le résultat final. Elle nécessite tout de même une optimisation finale car elle repose sur la correction parfaite d'autant de points du champ que le système a de surfaces et non pas une correction uniforme du champ. Ce qui n'est généralement pas le résultat voulu par un designer. L'optimisation finale permet de résoudre cette dernière difficulté.

## **Quelle influence peut avoir le modèle d'imagerie sur l'optimisation de masques**

## de phase augmentant la profondeur de champ en microscopie de localisation ?

Olivier Lévêque, Caroline Kulcsár et François Goudail  
Université Paris-Saclay, Institut d'Optique Graduate School, CNRS, Laboratoire Charles Fabry, 91127, Palaiseau, France.

La profondeur d'imagerie en microscopie PALM, de l'anglais photo-activated localization microscopy, est fortement limitée par la profondeur de champ (PdC) naturelle de l'objectif utilisé. Celle-ci est très courte car l'ouverture numérique de l'objectif doit être grande pour maximiser la résolution latérale des images acquises. Plusieurs approches ont été étudiées pour étendre cette PdC, l'une d'entre elle consiste à placer un masque de phase dans le diaphragme de l'objectif [1, 2]. Nous avons récemment proposé des masques de phase binaires annulaires, faciles à fabriquer, basés sur cette approche. Après une optimisation des masques et un traitement numérique approprié des images acquises, nous avons montré que la PdC en microscopie PALM pouvait être améliorée d'un facteur trois [3, 4].

Pour optimiser ces masques et caractériser leurs performances, la défocalisation est généralement modélisée par un simple terme de phase quadratique dans le plan pupillaire [5]. Cependant, ce modèle ne prend pas en compte deux caractéristiques essentielles des configurations de microscopie de localisation : une ouverture numérique très élevée et un écart possible entre les indices de réfraction nominaux du liquide d'immersion et de l'échantillon biologique étudié. Il existe des modèles scalaires ou vectoriels bien plus précis qui prennent mieux en compte ces caractéristiques d'imagerie. Une question importante se pose donc : quelle influence peut avoir le modèle d'imagerie sur l'optimisation des masques de phase dans cette application ?

Nous montrons tout d'abord qu'avec le modèle de formation d'image de Gibson & Lanni (GL) [6], plus réaliste qu'un simple terme de phase quadratique, l'extension de PdC se trouve simplement réduite par un facteur d'échelle dépendant de l'ouverture numérique. Nous montrons ensuite que si ce facteur d'échelle est correctement pris en compte, les masques optimisés avec le modèle quadratique restent optimaux dans le cadre du modèle GL. Ce résultat est important car il permet l'utilisation de masques génériques optimisés à l'aide du modèle quadratique approché sans requérir une connaissance précise de la configuration du microscope utilisé.

[1] S. R. P. Pavani, M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun, and W. E. Moerner, "Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function," *Proceedings of the National Academy of Sciences*, vol. 106, no. 9, pp. 2995–2999, 2009.

[2] Y. Shechtman, S. J. Sahl, A. S. Backer, and W. E. Moerner, "Optimal Point Spread Function Design for 3D Imaging," *Phys. Rev. Lett.*, vol. 113, p. 133902, Sep 2014.

[3] O. Lévêque, C. Kulcsár, H. Sauer, A. Lee, P. Bon, L. Cognet, and F. Goudail, "Can phase masks extend depth-of-field in localization microscopy ?," in *Unconventional Optical Imaging II* (C. Fournier, M. P. Georges, and G. Popescu, eds.), vol. 11351, pp. 50 – 58, International Society for Optics and Photonics, SPIE, 2020.

[4] O. Lévêque, C. Kulcsár, A. Lee, H. Sauer, A. Aleksanyan, P. Bon, L. Cognet, and F. Goudail, "Co-designed annular binary phase masks for depth-of-field extension in single-molecule localization microscopy," *Opt. Express*, vol. 28, pp. 32426–32446, Oct 2020.

[5] J. W. Goodman, *Introduction to Fourier optics*. W.H. Freeman & Company, 4th édition. ed., 2017.

[6] S. F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," *J. Opt. Soc. Am. A*, vol. 9, pp. 154–166, Jan 1992.

## Amélioration des performances du SLAM RGB-D en tenant compte des spécifications du capteur pour la localisation indoor.

Imad El Bouazzaoui, Sergio Rodriguez Florez, Abdelhafid El Ouardi  
SATIE-CNRS UMR 8029, Paris-Saclay University, France.

Plusieurs travaux se sont focalisés sur la localisation et la cartographie simultanées (SLAM), un sujet étudié depuis plus d'une décennie pour répondre aux exigences des robots en matière de navigation dans des environnements inconnus. Le SLAM est une fonctionnalité de perception essentielle dans plusieurs applications, notamment en robotique et dans les véhicules autonomes. Les caméras RGB-D font partie des capteurs couramment utilisés dans les algorithmes SLAM récents. Elles fournissent une image RGB et la carte de profondeur associée, ce qui permet de résoudre la dérive d'échelle avec moins de complexité et de créer une représentation dense de l'environnement 3D. De nombreux algorithmes SLAM RGB-D ont été étudiés et évalués sur des jeux de données disponibles en ligne sans tenir compte des spécifications des capteurs ou des modes d'acquisition d'images qui pourraient améliorer ou diminuer la précision de la localisation. Dans ce

travail, nous étudions la localisation en indoor, en tenant en compte les spécifications du capteur. Dans ce contexte, notre contribution est une étude expérimentale approfondie pour mettre en évidence l'impact des modes d'acquisition des capteurs sur la précision de la localisation, et un protocole d'optimisation paramétrique pour une localisation précise dans un environnement donné. De plus, nous appliquons le protocole proposé pour optimiser un paramètre lié à la profondeur de l'algorithme SLAM. L'étude est basée sur un ensemble de données disponibles en ligne dans un environnement interne avec un capteur de profondeur. L'analyse des résultats de la reconstruction est fondée sur l'étude de différentes métriques impliquant des erreurs de translation et de rotation. Ces erreurs métriques sont comparées à celles obtenues avec un algorithme SLAM basé sur la vision stéréo.

## **Co-conception optique/réseau de neurones à partir d'un modèle optique par tracé de rayon différentiel**

Aymeric Halé, Pauline Trouvé-Peloux, Jean-Baptiste Volatier  
ONERA, Palaiseau, France

Le développement récent d'algorithmes basés sur des réseaux de neurones pose à présent la question de la conception conjointe d'un système utilisant ce type de traitement, en particulier pour des tâches complexes telles que la détection d'objet, la classification, l'interprétation de scènes etc. Les images étant analysées par une machine, les besoins sont alors forcément différents que ceux d'un humain et sont difficiles à anticiper lors de la conception du capteur. Par ailleurs, les outils de conception optique, s'ils permettent aujourd'hui d'intégrer des critères d'optimisation autre que purement optique, ne sont pas compatibles avec l'optimisation des réseaux de neurones.

On trouve dans l'état de l'art plusieurs travaux portant sur l'optimisation d'élément optique conjointement avec un réseau de neurones (Elmalem 2019, Chang2019 Metzler2020), l'optique est optimisée pour augmenter la dynamique de l'image. L'approche commune à ces travaux consiste à définir la première couche du réseau comme une couche de convolution dont le noyau correspond à la réponse impulsionnelle de l'optique. Les paramètres optiques dont dépend cette réponse impulsionnelle sont optimisés en même temps que les paramètres des autres couches du réseau de neurones. Cependant, dans toutes ces références, les éléments optiques sont modélisés à l'aide de modèles optiques simplifiés et seulement un nombre réduit de paramètres optiques sont optimisés.

Nous proposons d'explorer la co-conception d'une optique et d'un réseau de neurones avec des modèles optiques plus réalistes, ayant les mêmes degrés de libertés que ceux utilisés en conception optique classique (rayons de courbures des verres, matériaux, distance entre lentilles, etc) afin de concevoir conjointement tous les éléments d'un objectif et le traitement des images produites. Pour cela, nous proposons d'utiliser un modèle plus complet de l'optique, reposant sur un tracé de rayon différentiel (Volatier2017). Ce modèle décrit une optique avec les mêmes paramètres qu'un logiciel de conception optique et fournit la réponse impulsionnelle et son Jacobien par rapport à l'ensemble des paramètres optiques. Ce modèle est ainsi compatible avec le formalisme de descente de gradient qui sert à l'optimisation d'un réseau de neurones.

Dans cette présentation, nous décrivons la méthode de co-conception que nous proposons et sa validation sur des exemples simples de co-conception. Nous discutons ensuite des enjeux de cette méthode et des perspectives de travail.