

# Estimation simultanée de la forme d'un objet, de sa la localisation et de sa vitesse à l'aide d'une caméra plénoptique

---

**Ecole doctorale :**

SPIM (Sciences pour l'Ingénieur et Microtechniques) – ED37  
Université Bourgogne Franche-Comté (UBFC)

**Unités de recherche :**

ImViA, **VIBOT** CNRS EMR 6000 / UBFC  
<http://vibot.cnrs.fr>

en coencadrement avec

Institut **FEMTO-ST**, CNRS UMR6174 / UFC / ENSMM / UTBM  
Département **Automatique et Systèmes Micro-Mécatroniques (AS2M)**  
<https://www.femto-st.fr/en/Research-departments/AS2M/Presentation>

**Directeurs de thèse :**

Pr. Cédric Demonceaux, ImVIA, VIBOT  
Pr. Nicolas Andreff, FEMTO-ST, AS2M

**Contacts :**

[cedric.demonceaux@ubfc.fr](mailto:cedric.demonceaux@ubfc.fr)

tél : 07 64 17 16 33

<http://vibot.cnrs.fr/ceacutedric-demonceaux.html>

[nicolas.andreff@ubfc.fr](mailto:nicolas.andreff@ubfc.fr)

tél : 03 81 40 29 61

<http://members.femto-st.fr/~nicolas.andreff>

**Résumé du sujet de thèse**

Ce projet de thèse entre dans le cadre du projet iFILTER soutenu par le programme « Investissements d'Avenir », projet ISITE-BFC (contrat ANR-15-IDEX-0003) de l'Université Bourgogne Franche-Comté.

Les caméras plénoptiques, constituées d'une matrice de microlentilles placée devant le capteur photosensible, permettent de capturer en une prise de vue l'intensité lumineuse et l'orientation des rayons [1]. Ainsi, à moindre encombrement, il est possible de reconstituer la structure 3D de la scène observée et/ou de focaliser l'image à une distance choisie par l'utilisateur [2]. En particulier, dans le cadre des recherches en microrobotique intracorporelle [3], ce type de caméra pourrait permettre de réduire l'encombrement des endoscopes, de manière à laisser plus de place aux outils chirurgicaux, sans dégrader le retour visuel fourni au chirurgien.

La miniaturisation du capteur, non abordée dans cette thèse, impose de choisir un capteur à acquisition séquentielle (*rolling shutter*). Ce type de capteur souffre de distorsions optiques lorsque l'objet observé se déplace devant le capteur (voir des exemples dans [5]) mais permet également d'ouvrir la possibilité de simultanément localiser l'objet et d'estimer sa vitesse dans l'espace, si forme est connue [5]. Néanmoins, lorsque l'objet est de forme inconnue, il reste une ambiguïté sur la séparation de la forme et du mouvement [6] qu'une caméra plénoptique devrait être en mesure de résoudre.

Ainsi, dans ce travail de thèse, il s'agit de tirer parti de la particularité de ces caméras qui acquièrent l'intensité lumineuse et la direction des faisceaux de manière séquentielle. L'association d'un capteur à acquisition séquentielle et d'une matrice de microlentille n'a jusqu'ici pas été explorée d'un point de vue algorithmique. Pour ce faire, deux étapes sont primordiales pour mener à bien ce projet : le traitement des images plénoptiques et le développement d'outils géométriques spécifiques de vision par ordinateur.

En premier lieu, il s'agira donc de tenir compte de la physique des capteurs pour développer des traitements bas niveaux adaptés. A chaque prise de vue, ce capteur nous fournit une série d'images avec différentes focales. A ce jour, les auteurs proposent de traiter ce bloc d'images [10,11] ou d'analyser



l'image du plan épipolaire afférente [12] en utilisant des techniques de filtrages classiques. Dans ce travail, nous proposerons de tenir compte de la particularité d'acquisition de ces images pour développer de nouvelles méthodes robustes de filtrages adaptées au capteur. Ces filtrages adaptés permettront également de repenser les réseaux neuronaux convolutifs en s'inspirant des travaux que nous avons réalisés sur les capteurs sphériques [7], RGBD [8] et multimodaux [9].

Par la suite, le travail consistera à proposer de méthodes de vision 3D géométrique tenant compte de la matrice de microlentilles [13] et de l'acquisition séquentielle [6] dans le but d'estimer simultanément la forme de l'objet inconnu filmé, sa vitesse et sa position.

Ces travaux pourront être appliqués en microbotique intracorporelle.

### **Profil idéal**

Formation initiale à Bac+5 en mathématiques appliquées, informatique ou robotique, avec des connaissances préalables en vision par ordinateur.

Bonnes qualités de formalisation mathématique et goût pour la programmation et l'expérimentation.

Bonne capacité de dialogue.

Une bonne dose de créativité et une excellente rigueur méthodologique.

Communication écrite et orale de haut niveau en français et en anglais.

**Une attention particulière sera donnée aux questions de genre et d'équité sociale.**

### **Date de début de la thèse**

1<sup>er</sup> octobre 2021

### **Date limite de candidature**

22 mai 2021

### **Pièces à fournir :**

CV + lettre de motivation + relevé des notes des années 2019-2020 et 2020-2021

+ toute autre élément pertinent

# Simultaneous estimation of object shape, location and speed using a light-field camera

---

**Doctoral school:**

SPIM (Sciences for Engineers and Microtechnologies) – ED37  
Université Bourgogne Franche-Comté (UBFC)

**Research Institute:**

ImViA, **VIBOT** CNRS EMR 6000 / UBFC  
<http://vibot.cnrs.fr>

in cooperation with

**FEMTO-ST** Institute, CNRS UMR6174 / UFC / ENSMM / UTBM  
Automatic control and **Micro-Mechatronic Systems (AS2M)** department  
<https://www.femto-st.fr/en/Research-departments/AS2M/Presentation>

**Supervisors:**

Pr. Cédric Demonceaux, ImVIA, VIBOT  
Pr. Nicolas Andreff, FEMTO-ST, AS2M

**Contact:**

[cedric.demonceaux@ubfc.fr](mailto:cedric.demonceaux@ubfc.fr) phone : +33 764 171 633 <http://vibot.cnrs.fr/ceacutedric-demonceaux.html>  
[nicolas.andreff@ubfc.fr](mailto:nicolas.andreff@ubfc.fr) phone : +33 381 40 2 961 <http://www.femto-st.fr/~nicolas.andreff>

**Abstract**

This work is supported by the iFilter project, funded by the French “Investissements d’Avenir” program, under the ISITE-BFC project (ANR-15-IDEX-0003).

Light-field cameras, consisting of a matrix of microlenses placed in front of the photosensitive sensor, make it possible to capture the light intensity and orientation of the rays in a single shot. Thus, with a reduced space, it is possible to reconstruct the 3D structure of the observed scene and/or to focus the image at a distance chosen by the user [2]. In particular, within the framework of research in intracorporeal micro robotics [3], this type of camera could allow to reduce the size of endoscopes, so as to leave more space for surgical tools, without degrading the visual feedback provided to the surgeon.

The miniaturization of the sensor, not addressed in this thesis, requires the choice of a rolling shutter sensor. This kind of sensor suffers from optical distortions when the observed object moves in front of the sensor (see examples in [4]) but also allows to simultaneously localize the object and to estimate its speed, if the shape is known [5]. Nevertheless, when the object is of unknown shape, there remains an ambiguity about the separation of shape and motion [6] that a light-field camera should be able to resolve.

Thus, in this PhD work, the aim is to take advantage of the particularity of these cameras which acquire the light intensity and the direction of the rays in a sequential way. The association of a sequential acquisition sensor and a microlens array has not been explored so far from an algorithmic point of view. To do so, two steps are essential to complete this project: the image processing of Light-field images and the design of specific geometrical computer vision tools.

First of all, it will be necessary to take into account the physics of the sensors to develop adapted low-level image processing. At each shot, this sensor provides us with a series of images with different focal lengths. Previously, we developed techniques to deal with this block of images [10,11] or to analyze the image of the associated epipolar plane [12] by using classical filtering techniques. In this work, we propose to take into account the specificity of acquisition of these images to develop new robust filtering methods adapted to the sensor. These adapted filtering methods will also allow us to rethink convolutional neural networks by taking inspiration from the work we have done on spherical [7], RGBD [8] and multimodal [9] sensors.



Then, the work will consist in proposing methods of geometrical 3D vision taking into account the microlens grid [13] and the sequential acquisition [6] in order to simultaneously estimate the shape of the unknown object filmed, its speed and its position.

This work can be applied in intracorporeal microbotics.

**Ideal candidate profile**

M.Sc. (or equivalent) in applied mathematics, computer science or robotics.

Solid background in computer vision.

Good skills in mathematical formalization, programming and experimentation.

Good dialogue skills.

Creativity and methodological rigor.

Oral and written communication skills in English and French.

**Specific attention will be given to gender and social equity issues.**

**Starting date of the PhD work**

October 1, 2021

**Application deadline**

May 22, 2021

**Application file**

CV + motivation letter + transcript of records for academic years 2019-2020 and 2020-2021

+ any other relevant document

## References

- [1] E. H. Adelson and J. Y. A. Wang. Single Lens Stereo with Plenoptic Camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, pp. 99-106, February 1992.
- [2] C. Hahne, A. Aggoun, and V. Velisavljevic, S. Fiebig, and M. Pesch "Baseline and triangulation geometry in a standard plenoptic camera," *Int. J. of Comput. Vis. (IJCV)*, 2017.
- [3] N. Andreff et al. Micronanorobotique biomédicale. Grand Prix Scientifique de la Fondation Charles Defforey, Institut de France, 2018
- [4] L. Magerand. Calcul de pose dynamique avec les caméras CMOS utilisant une acquisition séquentielle. Thèse de Doctorat. Université Blaise Pascal. 2014
- [5] Omar Ait Aider, Nicolas Andreff, Jean-Marc Lavest, Philippe Martinet. Simultaneous Object Pose and Velocity Computation Using a Single View from a Rolling Shutter Camera. 9th European Conference on Computer Vision, ECCV'06, 2006, Graz, Austria.
- [6] Lao, Yizhen & Ait-Aider, Omar & Bartoli, Adrien. Solving Rolling Shutter 3D Vision Problems using Analogies with Non-rigidity. *International Journal of Computer Vision*. 2021.
- [7] Fernandez-Labrador, C., Facil, J. M., Perez-Yus, A., Demonceaux, C., Civera, J., & Guerrero, J. J. Corners for layout: End-to-end layout recovery from 360 images. *IEEE Robotics and Automation Letters*, 5(2), 1255-1262, 2020
- [8] Wu, Z., Allibert, G., Stolz, C., & Demonceaux, C. Depth-Adapted CNN for RGB-D cameras. In Proceedings of the Asian Conference on Computer Vision, 2020 (Oral Presentation).
- [9] Piasco, N., Sidibé, D., Gouet-Brunet, V., & Demonceaux, C. Improving image description with auxiliary modality for visual localization in challenging conditions. *International Journal of Computer Vision*, 129(1), 185-202, 2021.
- [10] Yoon, Y., Jeon, H. G., Yoo, D., Lee, J. Y., & So Kweon, I. Learning a deep convolutional network for light-field image super-resolution. In Proceedings of the IEEE international conference on computer vision workshops, 2015.
- [11] Farrugia, R. A., & Guillemot, C. Light field super-resolution using a low-rank prior and deep convolutional neural network. *IEEE transactions on pattern analysis and machine intelligence*, 42(5), 1162-1175, 2019
- [12] Wu, G., Zhao, M., Wang, L., Dai, Q., Chai, T., & Liu, Y. Light field reconstruction using deep convolutional network on EPI. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [13] Labussière, M., Teulière, C., Bernardin, F., & Ait-Aider, O. Blur Aware Calibration of Multi-Focus Plenoptic Camera. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020