

# Approches par apprentissage pour l'estimation de mouvement multiframe en vidéo

Soutenance de thèse – Pierre Godet

le 22 janvier 2021 à 10h30

Salle Marcel Pierre - ONERA Palaiseau,  
6 Chemin de la Vauve aux Granges, 91120 Palaiseau

## Devant le jury composé de :

Pascal MONASSE	Professeur École Nationale des Ponts et Chaussées	Rapporteur
Nicolas THOME	Professeur, Conservatoire National des Arts et Métiers	Rapporteur
Samia BOUCHAFA	Professeur, UEVE/Université Paris-Saclay	Examinatrice
David FILLIAT	Professeur, ENSTA	Examineur
Clément MALLET	Directeur de recherche, Institut Géographique National	Examineur
Patrick PEREZ	Directeur de recherche, Valeo.ai	Examineur
Guy LE BESNERAIS	Directeur de Recherches, ONERA	Directeur
Alexandre BOULCH	Valeo.ai	Co-encadrant

## Membres invités :

Véronique Serfaty	Direction Générale de l'Armement	
Aurélien Plyer	ONERA	Co-encadrant

## Résumé :

Ce travail porte sur l'exploitation de l'information temporelle sur une séquence de plus de deux images pour l'estimation du flot optique, défini comme le champ dense (en tout pixel) des mouvements apparents dans le repère image. Nous étudions d'une part l'utilisation d'une base de modèles temporels, appris par analyse en composantes principales à partir des données étudiées, pour modéliser la dépendance temporelle du mouvement. Cette première étude se focalise sur le contexte de la vélocimétrie par images de particules en mécanique des fluides. D'autre part, le nouvel état de l'art de l'estimation de flot optique ayant récemment été établi par des méthodes basées sur l'apprentissage profond, nous entraînons des réseaux de neurones convolutifs à estimer le flot optique en profitant de la continuité temporelle, dans le cas de séquences d'images naturelles. Nous proposons ensuite STaRFlow, un réseau de neurones convolutif exploitant une mémoire de l'information du passé au moyen d'une récurrence temporelle. Par application répétée d'une même cellule récurrente, les mêmes paramètres appris sont utilisés pour les différents instants considérés et pour les différents niveaux d'un processus multi-échelle. Cette architecture est plus légère que les réseaux concurrents tout en conférant à STaRFlow des performances à l'état de l'art. Au fil de nos travaux, nous mettons en évidence plusieurs cas où l'utilisation de l'information temporelle permet d'améliorer la qualité de l'estimation, en particulier en présence d'occultations, lorsque la qualité image est dégradée (flou, bruit), ou encore dans le cas d'objets fins.

## Mots clés :

Flot optique, Vision par ordinateur, Apprentissage profond, Apprentissage par ordinateur