

Apprentissage profond pour les systèmes de vision monoculaire 3D (Deep learning methods for monocular 3D vision systems)

Soutenance de thèse – Rémy LEROY

10 mars 2023 à 10h00

Lieu : Salle Marcel Pierre, ONERA Palaiseau

Lien zoom : <https://ec-nantes.zoom.us/j/93600414909>

Meeting ID: 936 0041 4909

Passcode: #c\$Q5LWC

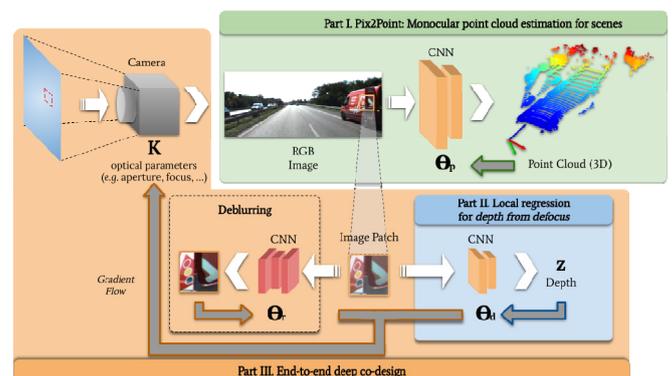
Devant le jury composé de :

Thierry CHATEAU	Université Saint-Etienne	<i>Rapporteur</i>
Loïc DENIS	Université Clermont Auvergne	<i>Rapporteur</i>
Bernadette DORIZZI	Télécom SudParis	<i>Examinatrice</i>
Christine GUILLEMOT	INRIA	<i>Examinatrice</i>
Renaud MARLET	École des Ponts – Valeo	<i>Examineur</i>
Frédéric CHAMPAGNAT	ONERA/DTIS – IVA	Directeur
Bertrand LE SAUX	ESA/ESRIN – Φ -lab	Co-directeur
Pauline TROUVÉ-PELOUX	ONERA/DTIS – IVA	Encadrante

Résumé :

Dans cette thèse, nous étudions l'apport de l'apprentissage profond pour les systèmes de vision 3D monoculaire, de l'acquisition de l'image au traitement. Nous proposons d'abord Pix2Point, une méthode d'estimation de nuage de points 3D à partir d'une seule image en utilisant des informations de contexte, et entraînée avec une fonction de coût de transport optimal. Pix2Point réalise une meilleure couverture des scènes lorsqu'il est entraîné sur des nuages de points lacunaires que les méthodes d'estimation de profondeur monoculaire, entraînées sur des cartes de profondeur lacunaires. Deuxièmement, pour exploiter les indices de profondeur provenant du capteur, nous proposons une méthode de régression de profondeur à partir d'un patch défocalisé.

Cette méthode surpasse la classification et la régression directe, sur données simulées et réelles. Enfin, nous abordons la conception d'un système de vision RVB-D, composé d'un capteur dont l'image est traitée par notre réseau de régression de profondeur basée sur la défocalisation et par un réseau de défloutage d'image. Nous proposons un cadre d'optimisation multi-tâches, conjointement aux paramètres des capteurs et des réseaux, et nous l'appliquons à l'optimisation de la mise au point d'une lentille chromatique. Le paysage d'optimisation présente plusieurs optima liés à la tâche de régression en profondeur, tandis que la tâche de défloutage semble moins sensible au paramètre de mise au point. En résumé, cette thèse propose plusieurs contributions exploitant les réseaux de neurones pour l'estimation 3D monoculaire et ouvre la voie d'une conception conjointe de systèmes RVB-D.



Mots clés : Apprentissage profond, Estimation monoculaire de la 3D, Co-conception