

PROPOSITION DE SUJET DE THESE

Intitulé : Réseaux de neurones pour la segmentation sémantique de données 3D temporelles

Référence **TIS-DTIS-18-25**
(à rappeler dans toute correspondance)

Laboratoire d'accueil à l'ONERA :

Domaine : Traitement de l'information et modélisation Lieu (centre ONERA) : Palaiseau

Département : Département du traitement de l'information et des systèmes

Unité : Image Vision Apprentissage Tél. : +33180386573

Responsable ONERA : Bertrand Le Saux, Email : bertrand.le_saux@onera.fr
Alexandre Boulch alexandre.boulch@onera.fr

Directeur de thèse envisagé :

Nom : Renaud Marlet

Adresse : IMAGINE École des Pints ParisTech Champs-sur-Marne 77455

Tél. : +33 1 64 15 21 86 Email : renaud.marlet@enpc.fr

Sujet : Réseaux de neurones pour la segmentation sémantique de données 3D temporelles

Contexte

Les données 3D sont en train de remplacer les images pour la perception de l'environnement dans de nombreux cas de figure : voiture autonomes, robotique, cartographie urbaine, biométrie, etc. Elles sont généralement constituées de nuages de points (positions dans l'espace) avec une radiométrie associée (information de couleur). Avec l'émergence de capteurs abordables pour le grand public (ex. Kinect, capteur 3D de l'iPhone X) ou pour les professionnels (scanners laser) ainsi que le développement de techniques robustes de photogrammétrie (reconstruction de la 3D à partir de plusieurs photos), les nuages de points sont devenus les données d'entrée de nombreux algorithmes pour la reconstruction de surfaces, la compréhension sémantique (donner du sens à la scène observée), ou encore l'animation 3D.

Ces capteurs permettent non seulement une acquisition spatiale précise mais aussi la répétition des acquisitions sur une même scène, par exemple l'enregistrement de séquences avec la Kinect.

Grâce au développement des techniques d'apprentissage, notamment à base de réseaux de neurones convolutifs, de nombreux progrès ont été effectués en robotique basée vision tant en navigation qu'en analyse de l'environnement 3D, en ligne ou a posteriori [Boulch et al. 2017, Guerry et al. 2017]. De nouvelles applications voient aussi le jour, comme des aides à la maintenance en milieu difficile avec réalité augmentée où l'imagerie 2D ou 3D permettant la reconnaissance des objets d'intérêt doit être robuste à l'évolution de la position de la personne portant le capteur (robustesse temporelle).

Problématiques

La thèse proposée a pour but d'améliorer la compréhension des scènes de nuages de points pour des applications robotiques. Dans ce but, plusieurs problématiques seront abordées au cours de la thèse :

1. Comment rendre la segmentation sémantique de données 3D plus performante ?

La compréhension de scènes 3D à base de nuages de points, notamment via la détection d'objets et la segmentation sémantique, est une étape essentielle à la navigation et à la prise de décision pour un agent autonome. Les techniques actuelles utilisent des méthodes héritées de la segmentation d'images [Su et al. 2015] ou bien directement des traitements voxeliques [Maturana & Scherer 2015] ou non structurés [Qi et al. 2017]. Le développement d'algorithmes et d'architectures de réseaux de neurones pour la segmentation et la détection dans des données 3D, notamment de données générées par des agents autonomes constitueront la ligne directrice de la thèse. De plus, dans le cas de jeux de données acquis par odométrie ou SLAM (par exemple en robotique), des séquences d'images ayant servi à la création du nuage de points sont aussi disponibles et constituent une source d'information complémentaire.

2. Comment utiliser des informations géométriques de bas niveau pour améliorer la segmentation d'objets ?

La sémantique recherchée se situe généralement à l'échelle de l'objet (mur, porte, meubles, route, arbres...). La recherche de primitives géométriques (plans, cylindres...) peut être un moyen de régulariser la segmentation sémantique et d'apporter de l'information sur la cohérence d'un ensemble de points en tant qu'objet. Par exemple, en intérieur, le mobilier et les pièces sont constitués de plans (ex. tables, murs...) et/ou de cylindres (piliers) et/ou de parallélépipèdes (armoires). Par ailleurs, cette information géométrique permet aussi de compléter un nuage de points lorsqu'il comporte du bruit ou des trous (données manquantes).

3. Comment identifier et caractériser les changements pour des acquisitions à différentes dates ?

Enfin, les nuages de points sont généralement traités de manière statique : soit les points sont accumulés pour une analyse globale, soit la scène est étudiée acquisition par acquisition. Or avec l'émergence de capteurs à bas coût, il est aujourd'hui possible de disposer de nuages de points "temporels" comme par exemple lors d'acquisitions successives sur un chantier ou dans le cas d'un robot évoluant dans un environnement. Le problème soulevé ici est la variabilité des données, dans la scène ou due au mode d'acquisition : les points échantillonnés changent d'une acquisition à une autre, y compris lorsqu'ils échantillonnent la même surface. Pour identifier et caractériser les changements, différentes stratégies seront étudiées : directement sur les données ou après une première étape de sémantisation. Ces analyses reposeront à la fois sur la performance de la segmentation sémantique (point 1) et sur une exploitation des primitives géométriques persistantes (point 2).

Informations pratiques

Cette thèse se déroulera en étroite collaboration entre ONERA/DTIS (Alexandre Boulch, Bertrand Le Saux) et l'équipe Imagine du LIGM/ENPC (Renaud Marlet, directeur de thèse). Le doctorant travaillera sur les sites de l'ONERA à Palaiseau et de l'ENPC à Champs-sur-Marne (région parisienne).

Bibliographie succincte

[Boulch et al. 2013] Boulch, Alexandre and Houiller, Simon and Marlet, Renaud and Tournaire, Olivier, Semantizing Complex 3D Scenes using Constrained Attribute Grammars, Computer Graphics Forum 2013

[Boulch & Marlet 2016] Boulch, Alexandre and Marlet, Renaud, Deep Learning for Robust Normal Estimation in unstructured Point Clouds, Computer Graphics Forum 2016

[Boulch et al. 2017] Boulch, Alexandre and Guerry, Joris and Le Saux, Bertrand and Audebert, Nicolas, SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks, Computer and Graphics 2017

[Guerry et al. 2017] Guerry, Joris and Boulch, Alexandre and Le Saux, Bertrand and Moras, Julien and Plyer, Aurélien and Filliat, David, SnapNet-R: Consistent 3D Multi-View Semantic Labeling for Robotics, ICCV/3DRMS 2017

[Maturana & Scherer 2015] Maturana, D. and Scherer, S., VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition, IROS, 2015.

[Qi et al. 2017] Qi, Charles R and Su, Hao and Mo, Kaichun and Guibas, Leonidas J, PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation, CVPR, 2017.

[Su et al. 2015] Su, Hang and Maji, Subhransu and Kalogerakis, Evangelos and Learned-Miller, Erik, Multi-view Convolutional Neural Networks for 3D Shape Recognition, CVPR 2015

Collaborations extérieures : Équipe IMAGINE à École des Ponts ParisTech

PROFIL DU CANDIDAT

Formation :

Master Sciences ou École d'ingénieur avec très bons résultats

Spécificités souhaitées :

Connaissances approfondies en informatique et traitement de données avec une spécialisation en : apprentissage statistique, vision par ordinateur, computer graphics.

Pour postuler, merci d'envoyer CV, lettre de motivation et lettres de recommandation au directeur de thèse et aux encadrants.