

LABO - Axe ou Equipe : Institut Pascal, ISPR, ComSEE

Directeur de thèse : Maxime Lhuillier (CR CNRS, HDR), maxime.lhuillier@uca.fr

Titre du sujet de thèse : Modèles génératifs profonds pour l'amélioration semi-supervisée de modèles photogrammétriques

Résumé du sujet de thèse :

La reconstruction 3D d'un environnement complet à partir d'images est utile dans de nombreuses applications comme la réalité virtuelle (<https://maximelhuillier.fr>) et la perception des véhicules autonomes. Plusieurs méthodes de vision par ordinateur et de photogrammétrie sont nécessaires pour résoudre ce problème. Elles incluent l'estimation de la géométrie (paramètres de caméra et nuage de points) et la reconstruction de surface. Une voie prometteuse de recherche est la conception de méthodes d'apprentissage profond qui améliorent l'étape de reconstruction de surface. Il faudrait aussi éviter les méthodes supervisées, qui nécessitent un jeu de données d'environnements générés par un scanner 3D. Il y a plusieurs raisons à cela: le prix/la disponibilité/les conditions expérimentales d'utilisation du scanner et le temps/l'effort d'acquisition. On remplace donc un jeu de données obtenu par scanner 3D par un autre jeu de données. Le jeu de données peut être synthétique et collecter des morceaux de surface que l'on sait être très probable dans les environnements. Il peut aussi inclure de grands environnements reconstruits par une méthode précédente (qui n'est pas de l'apprentissage profond), avec une minorité de corrections manuelles. Puis un réseau apprend à remplacer un morceau de surface incorrect ou improbable par un autre plus probable. Grâce à l'apprentissage, on s'attend à améliorer des méthodes précédentes de reconstruction de surface, par exemple quand les conditions expérimentales sont plus difficiles. Deux types de méthodes d'apprentissage profond peuvent potentiellement faire cela. Les méthodes non-génératives (ex: autoencoder) calculent un seul résultat, ie une surface corrigée. Elles ont des inconvénients: l'incertitude du résultat est inconnue et l'utilisateur ne peut choisir le meilleur parmi plusieurs résultats. Les méthodes génératives (ex: autoencodeur variationnel, méthode de diffusion) peuvent supprimer ces inconvénients car elles calculent plusieurs résultats, plus précisément une distribution de résultats. Dans les expériences, on se focalise sur des environnements extérieurs complètement reconstruits avec une caméra 360 (monté sur un casque ou un véhicule) qui se déplace d'au moins plusieurs centaines de mètres.

Short bibliography:

- J.Ho, A.Jain, P.Abbeel, Denoising diffusion probabilistic models, NeurIPS 2020.
- D.P.Kingma, M.Welling, Auto-encoding variational Bayes, ICLR 2014.
- M.Lhuillier, Surface reconstruction from a sparse point cloud by enforcing visibility consistency and topology constraints, CVIU 175, 2018.
- M.Lhuillier, Estimating the vertical direction in a photogrammetric 3D model, with application to visualization, CVIU 236, 2023.
- S.Peng, M.Niemeyer, L.Meschender, M.Pollefeys, A.Geiger, Convolutional occupancy networks, ECCV 2020.
- M.Prakash, A.Krull, F.Jug, Fully unsupervised diversity denoising with convolutional variational autoencoders, ICLR 2021.
- Y.Song, S.Ermon, Generative modeling by estimating gradients of the data distribution, NIPS 2019.
- R.Sulzer, L.Landrieu, R.Marlet, B.Vallet, Scalable surface reconstruction with Delaunay-graph neural networks, CGF 40(5) 2021.