

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

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

Titre du sujet de thèse : Modèles génératifs profonds pour la correction semi-supervisée d'environnements reconstruits par photogrammétrie

Résumé du sujet de thèse :

La reconstruction 3D d'un environnement complet à partir d'images est utile dans plusieurs applications incluant la réalité virtuelle [Lhuillier23] et la perception de 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 (AP) qui corrigent les erreurs de l'étape de reconstruction de surface avec des a priori de formes. Il faudrait aussi éviter l'AP supervisé, qui nécessite un jeu de données composé 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. Ici le jeu de données inclus de grands environnements reconstruits par une méthode précédente (qui est n'est pas AP), avec une minorité de corrections manuelles. Il peut aussi contenir des morceaux de surface, que l'on sait être très probable dans les environnements, et qui sont générés synthétiquement. Puis un réseau apprend à remplacer un morceau de surface incorrect ou improbable par un autre plus probable. Grâce à l'AP, on s'attend à améliorer des méthodes précédentes de reconstruction de surface (qui ne sont pas AP), par exemple quand les conditions expérimentales sont plus difficiles. Deux types de méthodes d'AP peuvent potentiellement faire cela. Les méthodes non-génératives (ex: auto-encoder) calculent un seul résultat, ie une surface corrigée, et ont des inconvénients: l'incertitude du résultat est inconnue et l'utilisateur ne peut pas choisir le meilleur parmi plusieurs résultats. Les méthodes génératives (ex: auto-encoder variationnel [Kingma14], méthodes de diffusion [Ho20, Song19]) peuvent supprimer ces inconvénients car elles calculent plusieurs résultats.

On se focalise ici sur des environnements extérieurs dont tous les composants (bâtiments, sol, végétation ...) sont reconstruits à partir d'un vidéo prise avec une caméra 360. Cette caméra est montée sur un casque et se déplace d'au moins plusieurs centaines de mètres. Les modèles 3D du jeu de données sont fournis sous la forme d'une représentation volumique standard des méthodes non AP: une triangulation de Delaunay 3D dont les tétraèdres sont étiquetés 'vide' ou 'matière'. La surface est alors définie par l'ensemble des faces triangulaires séparant vide et matière. Cette représentation est compacte (quelques millions de sommets pour une trajectoire de caméra de quelques centaines de mètres, en multipliant par 7 pour le nombre de tétraèdres), mais non uniforme et donc inadaptée pour la plupart des méthodes d'AP. Il faudra faire une conversion avec des représentations pour l'AP (ex: distance signée, voxels...) ou bien opérer directement sur le Delaunay 3D. De plus il y a plusieurs types d'erreurs: faux pleins, faux vides, faux sommets, manque de sommets. Dans un premier temps, on pourra se concentrer sur les erreurs d'étiquetages, en particulier les faux vides qui ont un fort impact sur la qualité de surface et qui sont souvent faciles à corriger manuellement.

References :

- [Ho20], J.Ho, A.Jain, P.Abbeel, Denoising diffusion probabilistic models, NeurIPS 2020,
- [Gangloff22], H.Gangloff, M.T.Pham, L.Courtrain, S.Lefevre, Leveraging vector-quantized variational autoencoder inner metrics for anomaly detection, ICPR 2022.
- [Kingma14], D.P.Kingma, M.Welling, Auto-encoding variational Bayes, ICLR 2014.
- [Lhuillier18], M.Lhuillier, Surface reconstruction from a sparse point cloud by enforcing visibility consistency and topology constraints, CVIU 175, 2018.
- [Lhuillier23], M.Lhuillier, Estimating the vertical direction in a photogrammetric 3D model, with application to visualization, CVIU 236, 2023. (<https://maximelhuillier.fr>)
- [Peng20], S.Peng, M.Niemeyer, L.Meschender, M.Pollefeys, A.Geiger, Convolutional occupancy networks, ECCV 2020.
- [Prakash21], M.Prakash, A.Krull, F.Jug, Fully unsupervised diversity denoising with convolutional

variational autoencoders, ICLR 2021.

- [Sun24], S.Sun, C.Zhao, Y.Guo, R.Wang, X.Huang, Y.V.Chen, L.Ren, *Behind the veil: enhanced indoor 3D scene reconstruction with occluded surfaces completion*, CVPR 2024.

- [Song19], Y.Song, S.Ermon, *Generative modeling by estimating gradients of the data distribution*, NIPS 2019.

- [Sulzer21], R.Sulzer, L.Landrieu, R.Marlet, B.Vallet, *Scalable surface reconstruction with Delaunay-graph neural networks*, CGF 40(5) 2021.