LABO - Axe et Equipe : Instituit Pascal, ISPR, ComSEE

Thesis supervisor: Maxime Lhuillier (CNRS researcher, HDR), maxime.lhuillier@uca.fr
Co-supervisor :

Title of PhD subject: Deep generative models for weakly-supervised correction of environments reconstructed by photogrammetry

Summary :
The 3D reconstruction of a complete environment from images is useful for several applications including virtual reality [Lhuillier23]. Several methods of computer vision and photogrammetry are needed to solve this problem. They include the estimation of the geometry (camera parameters and cloud of 3D points) and the surface reconstruction. A promising way of research is the design of deep learning (DL) methods that correct errors of the surface reconstruction step using shape priors. We also would like to avoid supervised DL, which needs dataset of environments generated by 3D scanner. There are several reasons to do this: price/availability/experimental conditions of the scanner and time/effort of acquisition. Here the dataset includes large environments reconstructed by a previous method (which is not DL, eg [Lhuillier18]), with a minority of manual corrections. It can also includes surface segments that are known to be very probable in the environments and that are synthesized. Then a network learns to replace a wrong or improbable segment of surface by a more probable one. Thanks to DL, we expect to improve previous (non-DL) surface reconstruction methods, for example when experimental conditions are difficult. Two kinds of DL methods can potentially do this. Non-generative method (eg auto-encoder) computes only one result, ie one corrected surface. It has drawbacks: the uncertainty of the result is unknown and user cannot choose the best among several results. Generative method (eg variational auto-encoder [Kingma14], diffusion model [Ho20,Song19, rectified flow [Liu22]) can remove these drawbacks since it provides several results.

Here we mostly focus on outdoor environments whose all components (buildings, ground, vegetation, ...) are reconstructed from a video captured by a 360 camera [Lhuillier18, Lhuillier23]. This cameras is mounted on a helmet and moves along a trajectory of several hundreds of meters or more. The 3D models of the dataset are given using a standard volume representation of the non-DL methods: a 3D Delaunay triangulation whose tetrahedra are labeled "free-space" or "matter". The surface is defined by the set of the triangle faces separating free-space and matter.

- [Ho20], J.Ho, A.Jain, P.Abbeel. Denoising diffusion probabilistic models. NeurIPS 2020.
- [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)
- [Liu22], X.Liu, C.Gong, Q.Liu. Flow straight and fast: learning to generate and transfer data with rectified flow. arXiv preprint 2209.03003, 2022.
- [Song19], Y.Song, S.Ermon. Generative modeling by estimating gradients of the data distribution. NIPS 2019.

Profile and skill required:
- Master of science and/or engineering school degree, with major in IA and/or applied mathematics.
- Good proficiency in deep learning, Python, Pytorch, Linux.
- Level in French: B1, level in English: C1