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

Thesis supervisor: Maxime Lhuillier (CR CNRS), maxime.lhuillier@uca.fr

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] and perception of autonomous vehicles. 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), 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) can remove these drawbacks since it provides several results.

We focus on outdoor environment whose all components (buildings, ground, vegetation..) are reconstructed from a video taken by a 360 camera. This camera is helmet-held and moves at least several hundreds of meters. The 3D models of the dataset are given using a standard volumetric representation used by non-DL methods: a 3D Delaunay triangulation whose tetrahedra are labeled "empty" or "matter". The surface is defined by the set of triangular faces separating empty and matter. This representation is compact (a few millions of vertices for camera trajectory of several hundreds of meters, multiply by 7 to obtain the number of tetrahedra). But it is also non-uniform and inadequate for most DL methods. A conversion to an adequate representation for DL (e.g. signed distance, voxels, …) may be necessary. Furthermore, there are several kinds of errors: false matter, false empty, false vertex, lack of vertex. First we can focus on the labeling errors, especially the false empty errors which have a strong impact on the surface quality and which are often easy to be manually corrected.

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.*