Gaussian Splatting for Urban scenes

Research Internship at Astra-Vision Inria/valeo.ai
Machine Learning, Computer Vision

Astra-Vision

Astra is a joint team between Inria and Valeo for research on autonomous driving. Composed of Inria and valeo.ai members, Astra-Vision is dedicated to the study of 2D vision and 3D perception for robust scene understanding. The research focuses on relaxing the use of abundant data and supervision, stepping towards weak-/un-supervised vision algorithms, while providing models that are more interpretable. It primarily addresses autonomous driving but research expands to a variety of indoor and outdoor applications. The Astra-Vision group is a dynamic group with regular group reading, seminars, etc.
Most of the research is open sourced and published in all top-tier venues of computer vision: CVPR, ICCV, IJCV, TPAMI, ECCV, 3DV, WACV, etc. More information about the team can be found here: https://astra-vision.github.io/

Internships at valeo.ai

We regularly hire as interns a number of MSc students in their final year, to work on research topics. Most of these internships result in paper submissions to top-tier conferences such as CVPR, ECCV or ICCV. Some trainees go on to do a PhD thesis in the lab. Practical information:

- **Location:**
  valeo.ai, 100 rue de Courcelles, 75017 Paris.
  Inria, 2 rue Simone IFF, 75012 Paris
- **Time:** usually spring and summer (typically 6 months).
- **Applicants:** finishing their MSc with a solid background in computer vision and machine learning, particularly in deep learning with strong PyTorch coding skills.
- **Main advisors:** Raoul de Charette, Alexandre Boulch
- **Contact:** raoul.de-charette@inria.fr, alexandre.boulch@valeo.com

Subject

Being able to reconstruct a complete 360 degree 3D scene remains challenging, in particular in outdoor scenarios. A few years back, Neural Radiance Fields (NeRF) impulsed a new state-of-the-art, allowing to hallucinate novel views of complex scenes from a couple of input images. Still, they suffer at reconstructing thin structures, and preserving some fine-grained details. To address this, the recent 3D Gaussian Splatting for Radiance Fields (Kerbl *et al.*, 2023) relies on sparse, differentiable and compact encoding of the scene radiance. This has pushed quality further while being significantly faster at inference and training, but they are important remaining challenges to address large outdoor scenes.

This internship seeks to study 3D Gaussian Splatting and its application to driving scenes from short video sequences (10-20 seconds), thus enabling generation of novel (unseen) views. Applying such radiance fields to outdoor urban scenes introduces significant challenges:

- **Unbounded Outdoor Scenes**: The vastness of outdoor driving scenes introduces computational challenges. Approaches such as Block-NeRF (Tancik *et al.*, 2022; Rematas *et al.,* 2022) are promising but come at a significant loss of quality at distance.
- **Dynamic scenes**: Typical outdoor urban scenes are dynamic inducing occlusion and motion. Existing methods either focus on static scenes or utilize annotated 3D bounding boxes (Wu *et al.*, 2023). The potential integration of LiDAR supervision (Huang *et al.*, 2023), optical flow, feature tracking or disentanglement losses may offer a solution.
- **High Level of Details**: Accurately capturing details like blinkers and traffic signs is crucial for a realistic representation of driving scenes. The optimal balance between rendering time, memory and accuracy for the intended usage is to be found.
- **Lighting & Surface Interactions**: Changing lighting, alongside reflections and transparencies from surfaces like cars, windows or wet roads complicates scene reconstruction. Leveraging perceptual losses (e.g., VGG-loss) and robust losses as in RobustNeRF (Sabour *et al.*, 2023) may boost the model's resilience against factors like glare, lens flare and rain.

To address these challenges the candidate will be collaborating with researchers with a strong expertise on 3D vision (Puy *et al.*, 2023; Boulch *et al.*, 2022; etc.) and NeRF (Cao and de Charette, 2023; etc.) expertise. The candidate will be also encouraged to interact with other PhDs and interns students, and participate actively in both Valeo.ai and Astra-vision group activities, including seminars, group readings, etc. Successful works will have the opportunity to submit to a top-tier venue and might lead to a continuation as PhD.

How to apply?

If you wish to apply to one or more internships, send a mail to the above contact people with:

- a cover letter explaining your interest and adequacy for the topic,
- your CV/resume,
- transcripts of your grades from last year, as well as this year when already available.

References

Kerbl et al. "3d gaussian splatting for real-time radiance field rendering." *ACM Transactions on Graphics (ToG)*. 2023
Rematas et al. "Urban radiance fields." *Conference on Computer Vision and Pattern Recognition*. 2022.
Tancik et al. "Block-nerf: Scalable large scene neural view synthesis." *Conference on Computer Vision and Pattern Recognition*. 2022.
Sabour et al. "RobustNeRF: Ignoring Distractors with Robust Losses." *Conference on Computer Vision and Pattern Recognition*. 2023.
Huang et al. "Neural LiDAR Fields for Novel View Synthesis." *arXiv.* 2023.
Yang et al. "UniSim: A Neural Closed-Loop Sensor Simulator." *Conference on Computer Vision and Pattern Recognition*. 2023.
Puy et al. "Using a waffle iron for automotive point cloud semantic segmentation." ICCV 2023.
Cao and de Charette. "Scenerf: Self-supervised monocular 3d scene reconstruction with radiance fields". ICCV 2023
Boulch et al. "POCO: Point convolution for surface reconstruction." Conference on Computer Vision and Pattern Recognition. 2022.
Wu et al. "Mars: An instance-aware, modular and realistic simulator for autonomous driving." *arXiv.* 2023.

Other related works:
Carlson et al. "Cloner: Camera-lidar fusion for occupancy grid-aided neural representations." *Robotics and Automation Letters*. 2023.
Wang et al. "Digging into Depth Priors for Outdoor Neural Radiance Fields." *International Conference on Multimedia*. 2023.
Sharma et al. "Seeing 3d objects in a single image via self-supervised static-dynamic disentanglement." *arXiv.* 2022.
Xie et al. "S-nerf: Neural radiance fields for street views." *arXiv preprint arXiv:2303.00749* (2023).
Turki et al. "SUDS: Scalable Urban Dynamic Scenes." *Conference on Computer Vision and Pattern Recognition*. 2023.
Wang et al. "BAD-NeRF: Bundle Adjusted Deblur Neural Radiance Fields." *Conference on Computer Vision and Pattern Recognition*. 2023.
Warburg et al. "NerfBusters: Removing Ghostly Artifacts from Casually Captured NeRFs." *arXiv.* 2023.