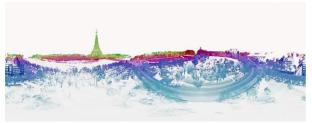## Master's internship M2 2025

LASTIG Lab – IGN / UGE

« Large-scale place recognition in 3D points clouds »

© IGN – LiDAR scan by terrestrial mobile mapping   © IGN – HD LiDAR scan by aerial mobile mapping

Keywords

Computer Vision, Photogrammetry, Machine Learning, Place Recognition, Pose Estimation, 3D Points Clouds, Big Data, LLM.

Context

Place recognition based on the visual mapping of the environment is a problem at the heart of many topical application domains, such as geolocalization for mobile mapping, digital twins update and documentation, collections annotation in digital humanities, augmented reality or fact-checking. Recognizing a location can take many forms, from the production of an annotation (e.g. "this image shows the church of Saint-Eustache in Paris") to a 6D pose that also provides information on the location of the acquisition sensor ("this is the south face of the church of Saint-Eustache photographed from 132 rue Rambuteau"). In the state of the art of computer vision, when no initial position is known, existing techniques are based on indexing and similarity search of visual content in a geolocalized image repository. Here, we study the generalization of this type of approach to 3D by considering 3D point cloud acquisition campaigns (notably LiDAR), which are becoming increasingly popular and whose richness in terms of geometry and semantics is attractive, but with a volume and diversity that are complex to handle. State of the art on place recognition for 3D points clouds exists, but it is still in its early stages and mainly focuses on the proposal, through deep learning, of points cloud descriptors with the objective of retrieving similar points clouds in a large dataset. From the seminal proposal PointNetVLAD [Uy et al, 2018], literature is enriched every year with new points cloud descriptors, which differ according to the deep backbone approach used, for better discriminance and robustness facing geometric transformations, volumetry as well as available benchmarks [Bold et al, 2019], [Liu et al, 2019], [Hui et al, 2021], [Komorowski, 2021], [Xia et al, 2021], [Hui et al, 2022], [Vidanapathirana et al, 2022]. Although there is a wide variety of 3D data acquisition equipment, most approaches focus on matching data of the same modality, typically from mobile mapping. To address this gap, [Guan et al., 2023] proposed tackling the matching of aerial and terrestrial data. It is within this context of multi-modality on a very large scale that this internship is positioned.

Subject

The internship is at the heart of the problem of indexing and retrieval in 3D point clouds, for place recognition up to the estimation of the 6D pose of the acquisition camera. The supervision team (two researchers and a PhD student) has got an expertise on points cloud descriptors for efficient retrieval in large collections of points clouds. From an input point cloud (the query), the current framework developed in the team, based on Hugging Face large language models, can retrieve the most similar sub-cloud in a geolocalized point cloud reference. These ones are associated with a geolocation that can be considered as a first answer to the place recognition problem. In practice, as with image retrieval, it is more robust to consider the *k* most similar clouds, then to analyze them more finely to judge their consistency with the query, before selecting one, or even several ones, to determine the location more precisely. This is the focus of the internship: the new recruit will have to assess the quality of the responses returned a posteriori by the 3D retrieval framework and estimate, according to criteria to be studied and implemented, which are the most similar to the query so as to possibly reorder them.

Expected output:

The expected work will be as follows:

- Understanding of current deep-based indexing and retrieval approaches, in general and for 3D point clouds;
- Study of geometric registration approaches between point clouds, starting from [Huang et al, 2021] for example;
- Implementation of one approach for registration and reranking of retrieval point clouds;
- Estimation of the 6D pose of the query from the selected cloud(s), to address the place recognition problem finely;
- Scaling the algorithm with Apache Spark framework.

As the subject of efficient indexing and retrieval in 3D point clouds, up to 6D pose estimation, is not yet well covered in the state of the art of 3D Computer Vision, particular attention will be paid to submitting a paper to an international conference or journal if the results obtained allow.

Organization

Duration: 5-6 months, starting from March/April 2025.

Workplace: LASTIG Lab, Institut National de l'Information Géographique et Forestière (IGN), Saint-Mandé (metro 1, station Saint Mandé) or Ecole Nationale des Sciences Géographiques (ENSG), Gustave Eiffel University, Champs-sur-Marne (RER A, station Noisy-Champs).

IGN (French Mapping Agency) is a Public Administrative Institution part of the French Ministry for Ecology and Sustainable Development. IGN is the national reference operator for the mapping of the territory; in particular, the agency is currently in charge of the 3D mapping program of France with LiDAR HD. The LASTIG is one of the research laboratories of IGN, attached to Ecole Nationale des Sciences Géographiques (ENSG) and Gustave Eiffel University (UGE) in Grand Paris area. It gathers about 100 researchers centered on geographical information sciences, 35 of them focusing in Image Analysis, Computer Vision, Artificial Intelligence, Photogrammetry and Remote Sensing.

Salary: internship gratification under French law (approx. 550€ net / month).

Possibility of PhD thesis afterward: yes, on the same theme.

Skills

Bac+5 in computer science, applied math or computer vision (master or engineering school); good knowledge in image or 3D data processing, as well as strong skills in Python programming. Good skills in Apache Spark, hugging Face API, LLM, PyTorch, or functional programming is a significant plus.

Submitting your candidature

Before February 15th 2025, send by e-mail to the contacts **in a single PDF file**:

- o CV
- o motivation letter
- o 2 recommendation letters, or persons to contact
- o Transcript of grades from the last two years of study
- o A list of courses followed and passed in the last two years

Contact

- o Valérie Gouet-Brunet, snior researcher, LASTIG – valerie.gouet@ign.fr
- o Laurent Caraffa, researcher, LASTIG – laurent.caraffa@ign.fr

Bibliography

[Bold et al, 2019] N. Bold, C. Zhang and T. Akashi, 3D Point Cloud Retrieval With Bidirectional Feature Match, in IEEE Access, vol. 7, p. 164194-164202, 2019.

[Geiger et al, 2013] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," vol. 32, no. 11, pp. 1231–1237, 2013.

[Guan et al, 2023] Tianrui Guan, Aswath Muthuselvam, Montana Hoover, Xijun Wang, Jing Liang, Adarsh Jagan Sathyamoorthy, Damon Conover, Dinesh Manocha, CrossLoc3D: Aerial-Ground Cross-Source 3D Place Recognition Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).

[Huang et al, 2021] Xiaoshui Huang, Guofeng Mei, Jian Zhang and Rana Abbas, A comprehensive survey on point cloud registration, Computer Science ArXiv, Published 3 March 2021.

[Hui et al, 2021] Le Hui, Hang Yang, Mingmei Cheng, Jin Xie, Jian Yang, Pyramid Point Cloud Transformer for Large-Scale Place Recognition, Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, p.p. 6098-6107.

[Hui et al, 2022] L. Hui, M. Cheng, J. Xie, J. Yang and M.-M. Cheng, Efficient 3D Point Cloud Feature Learning for Large-Scale Place Recognition, in IEEE Transactions on Image Processing, vol. 31, p. 1258-1270, 2022.

[Kim and al, 2020] G. Kim, Y. S. Park, Y. Cho, J. Jeong, and A. Kim, "MulRan: Multimodal range dataset for urban place recognition," in 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 6246–6253, IEEE.

[Komorowski, 2021] J. Komorowski, MinkLoc3D: Point Cloud Based Large-Scale Place Recognition, in 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2021 pp. 1789-1798.

[Liu et al, 2019] Zhe Liu, Shunbo Zhou, Chuanzhe Suo, Peng Yin, Wen Chen, Hesheng Wang, Haoang Li, Yunhui Liu, LPD-Net: 3D Point Cloud Learning for Large-Scale Place Recognition and Environment Analysis, in 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019 pp. 2831-2840.

[Uy et al, 2018] Mikaela Angelina Uy, Gim Hee Lee, PointNetVLAD: Deep Point Cloud Based Retrieval for Large-Scale Place Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4470-4479.

[Vidanapathirana et al, 2022] K. Vidanapathirana, M. Ramezani, P. Moghadam, S. Sridharan, and C. Fookes, "LoGG3d-net: Locally guided global descriptor learning for 3d place recognition," in 2022 International Conference on Robotics and Automation (ICRA), pp. 2215–2221.

[Xia et al, 2021] Yan Xia, Yusheng Xu, Shuang Li, Rui Wang, Juan Du, Daniel Cremers, Uwe Stilla, SOE-Net: A Self-Attention and Orientation Encoding Network for Point Cloud Based Place Recognition, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 11348-11357.