PhD Proposal: Automated Semantic Analysis of Dynamic 3D Point Clouds

loic.landrieu

July 2019

Discipline: Computer Science.

Specialty: Machine Learning, Computer Vision.

Research Laboratory: STRUDEL, LaSTIG, IGN; IMAGINE, ENPC.

Workplace : IGN, St Mandé, France.

Financing: ANR READY3D.

PhD Advisor: Mathieu Aubry (ENPC).

PhD Co-advisor: Loic Landrieu (IGN).

Start Date: TBD - 2020.

Keywords: Point Clouds, Panoptic Segmentation, Dynamic 3D, Autonomous Driving.

1 Context and Objectives

Autonomous driving is set to have a profound impact throughout our entire society in the near future. However, several scientific challenges remain unanswered, such as the automated analysis of dynamic 3D point clouds from moving vehicles at a precision and speed compatible with a fully-autonomous system.

This PhD proposal, as part of the READY3D (Real-Time Analysis of Dynamic 3D Point Clouds) ANR Project, addresses the scientific and practical challenges associated with 3D perception for autonomous driving. The thesis aim is to develop a framework for the accurate semantic analysis of dynamic 3D point clouds which could operate in real time.

Another key part of this thesis will be to assist in the creation of the Dynamic3D data, an open benchmark based on mobile LiDAR data already acquired by IGN in a previous campaign [4]. This dataset will be used to evaluate the performance of submitted algorithm for the panoptic segmentation[2] task for a variety of urban dynamic scenes, with an emphasis on computational efficiency metrics.

2 Proposed Approach:

With computational efficiency in mind, the methods developed will aim to extend the superpoint graph algorithm for the large-scale semantic segmentation [3] of point clouds to the dynamic setting. This approach, based on the partition of point clouds into superpoints—groups of adjacent points likely belonging to the same object—has proven apt to process large quantity of data acquired with fixed sensors, but do not take the temporal dimension into account yet.

The structure of the acquisition of a LiDAR sensor on a moving vehicle scanning a dynamic scene is complex. However, the design such sensors induces a specific space-time configuration to the point cloud, which could be exploited for faster and more precise algorithms.

The student will tackle the four following tasks:

Task1 **Sensor Topology-Structured Adjacency Graph:** The superpoints are defined with respect to an adjacency graph of the considered point clouds, which can be time-consuming to compute. The particular structure of rotating LiDAR acquisitions could be used to build on-the-fly a graphs of meaningful spatio-temporal connections between points, as shown by [1].

Task2 **Embedding Spatio-Temporal Superpoints:** Time and space structures are intertwined for mobile LiDAR acquisitions. This raises the question of a fitting definition for spatio-temporal superpoints in this context, and a fitting method to compute their embeddings. Furthermore, the spatial information is highly redundant from scan to scan since most scanned objects are usually static. This redundancy should be exploited for added speed and precision.

Task3 **Leveraging Spatio-Temporal Context:** The semantization of dynamic superpoints would be aided by exploiting the global context of the scene at any given time. By defining fitting descriptors of the spatio-temporal relationship between adjacent superpoints, we will be able to employ powerful graph-convolution algorithm [5] for leveraging this context.

Task4 **Dynamic3D benchmark:** A key contribution of is the creation of the Dynamic3D dataset, based on the Stereopolis acquisition campaign [4]. The entire dataset is already acquired, but remained to be annotated. This task would be carried out by a private company with funds from the READY3D project, as well as setting up the submission servers.

The first three tasks can be tackled sequentially. However, some flexibility can be gained if functional baselines can be designed. The creation of Dynamic3D dataset would be helpful for testing out the efficiency of our approach on a dataset which thoroughly reflects the real-time constraints. However, since our ambition is to create a versatile approach, most of the work could be prototyped on existing dataset such as semantic KITTI, vKITTI, or the newly released Lyft Level 5 dataset.

3 Requirements

- Masters in applied mathematics or computer science with a background in machine learning or computer vision.
- A good knowledge of state-of-the-art in deep learning (graph convolutions in particular).

- A good mastery of python, and Pytorch (preferred) or Tensorflow.
- Fluency in English.
- Not required but a plus: a background in functional optimization, efficient programming, an active Github profile, some publications in ML/CV/remote sensing conferences.

For applying please send the following items in an email to `loic.landrieu@ign.fr` with the object *PHD Proposal READY3D*:

- An up-to-date CV.
- Transcripts from your most recent diploma.
- Not required but a plus: recommendation letters or referrals from previous supervisors / professors.
- Not required but a plus: report on a project you participated in / master thesis.

4 Bibliography

References

[1] Stephane Guinard and Bruno Vallet. Weighted simplicial complex reconstruction from mobile laser scanning using sensor topology. *RFIAP*, 2018.

[2] Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, and Piotr Dollár. Panoptic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9404–9413, 2019.

[3] Loic Landrieu and Martin Simonovsky. Large-scale point cloud semantic segmentation with superpoint graphs. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[4] Nicolas Paparoditis, Jean-Pierre Papelard, Bertrand Cannelle, Alexandre Devaux, Bahman Soheilian, Nicolas David, and Erwann Houzay. Stereopolis II: A multi-purpose and multi-sensor 3d mobile mapping system for street visualisation and 3d metrology. *Revue française de photogrammétrie et de télédétection*, 2012.

[5] Martin Simonovsky and Nikos Komodakis. Dynamic edge-conditioned filters in convolutional neural networks on graphs. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.