The emergence of transformers opens a wide field of research subjects that matches well with the research axis “Vision and 3D Perception for Scene Understanding”. The goal of this postdoc is to attack some of them.

1. A first objective is to investigate the specificity of transformers for 3D. It includes the question of positional encoding, which is widely different in 3D from the 1D (text) and 2D (image) perspectives. It also includes the square complexity issues with cross- and self-attention, taking into account the sparsity of point clouds. More generally, it concerns as well the time and space complexity of point cloud processing, in the perspective of real-time, embedded software.

This study of 3D transformers will be made through downstream tasks that include object detection, semantic or panoptic segmentation, and denser depth predictions, which are key tasks for autonomous driving. We will study in particular the impact of the specific sparsity and data patterns induced by vehicle sensors. We will also consider a stream of point clouds, as available from a lidar, taking time into account in a 4D perspective.

2. Besides, we will investigate the use of transformers backbone regarding self-supervision. This approach, which is a key approach to bring supervised learning to a new level by saving the cost of annotating large datasets. We will study new pretext tasks in 3D that transformers more specifically leverage, as well as contrastive learning techniques that are tightly linked to the attention mechanism of transformers.

A PhD student is about to start on self-supervision for 3D at Valeo and another works on 2D supervision for 3D at Inria, though none of them focusing on transformers. The postdoc will be able to work jointly with either PhD student on self-supervision issues.

Moreover, we will study the transferability of learned transformer models in the perspective of domain adaptation. In particular, we will investigate the disentangling of latent space representations, working towards domain-invariant features by enforcing orthogonality of the domain features while enabling the discovery of exclusive task or domain features, through their realization via multi-head attention.

Another PhD student in Valeo is about to start working on 3D domain adaptation, although with a different perspective (using optical transport). There will nonetheless be a number of collaboration opportunities between this PhD student and the postdoc regarding adapting transformer-based features.

A last research direction concerns multi-modality, when lidar point clouds are acquired together with camera images, to leverage the similarity and complementarity of sensor information. One technical subject concerns a possible interplay between the forms of transformer attention used in 2D and the kinds of attention that are and will be developed in 3D. Another more general question is the joint self-supervision from the interaction of 2D and 3D, or from cross-task representations. Last, we will study the intertwined relation of geometry and semantics through the semantic scene completion task.

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.
Skills

Applicants should have defended or be finishing their PhD and have a strong publications record. They should have a solid background in computer vision (including 3D processing) and machine learning, particularly in deep learning, with strong PyTorch coding skills.

Applicants should apply on this platform:
- a cover letter explaining their interest and adequacy for the postdoc topic,
- their CV/resume,
- possibly, references or recommendation letters.

Apply as soon as possible. Applications are reviewed on a rolling basis. Starting date October 1st 2022 or before (firm)

Benefits package

- Subsidised catering service
- Partially-reimbursed public transport
- Flexible working hours
- Sports facilities