

Titre : Résoudre l'occlusion et la faible densité à distance des nuages de points pour la détection d'objets 3D par perception collaborative

Mots clés : perception collaborative, V2X, détection objets en 3D.

Résumé : La détection précise d'objets 3D est un enjeu majeur pour l'intégration sécurisée des véhicules autonomes dans le trafic routier. Le LiDAR, offrant des mesures de profondeur précises et relativement denses, est très présent dans les bases de données de référence. Cependant, ses nuages de points sont clairsemés à longue distance et soumis aux occlusions. L'état de l'art propose alors des techniques de suréchantillonnage, par fusion avec des caméras ou par distillation des connaissances, afin d'obtenir de bonnes détections. La première méthode permet de reconstruire la profondeur des pixels afin de générer des points supplémentaires tandis que la seconde vise à obtenir des nuages de points imitant ceux sans occlusion ni dispersion.

Comme ces approches utilisent des mesures obtenues par le véhicule ego à chaque pas de temps, la détection est inévitablement affectée par ces régions inobservables. Conscient des limites des méthodes à perspective unique, ces travaux de thèse s'efforcent à résoudre les problèmes d'occlusion et de rareté en exploitant des perspectives multiples. Notre approche exploite d'une part les mesures du véhicule ego par séquences: au cours de son déplacement dans le temps. D'autre part, nous proposons une perception collaborative basée sur la fusion des informations obtenues par de multiples agents connectés. Les expériences, menées sur la base de données V2X-Sim, montrent que notre approche permet au véhicule ego d'effectuer une détection précise dans des régions inobservables.

Title : Toward Solving Occlusion and Sparsity in Deep Learning-Based 3D Object Detection through Collaborative Perception .

Keywords : collaborative perception, V2X, 3D object detection, deep learning, LiDAR.

Abstract : Detecting objects at a high precision in 3D is critical for the safety of autonomous vehicles. Thanks to LiDARs' ability to provide accurate depth at a relatively high density, methods based on this modality have dominated every public benchmark. Point clouds frequently suffer from occlusion and sparsity due to their reliance on bouncing laser beams off surfaces. Prior works address these challenges by upsampling point clouds via fusion with RGB cameras or their representations via knowledge distillation. The former infers depth for pixels that are not measured by LiDAR to generate pseudo-LiDAR points. The latter makes the representation of regular point clouds imitate occlusion-and-sparsity-free ones. Because these approaches are designed to use point clouds obtained only by the ego vehicle at a single timestep, they are affected by occlusion and

and sparsity. Aware of the limitations of single-perspective methods, this thesis strives to resolve occlusion and sparsity by leveraging point clouds obtained from multiple perspectives to the fullest. The core of our approach is made of two components that are respectively built on the utilization of point cloud sequences and collaborative perception via V2X communication. Experiments in the V2X-Sim dataset show that our approach enables the ego vehicle to make accurate detection in unobservable regions.