2D-3D Fusion for Road Object Detection and Tracking on Autonomous Vehicles

Ruddy THEODOSE[1], Dieumet DENIS[1], Christophe BLANC[2], Paul CHECCHIN[2]

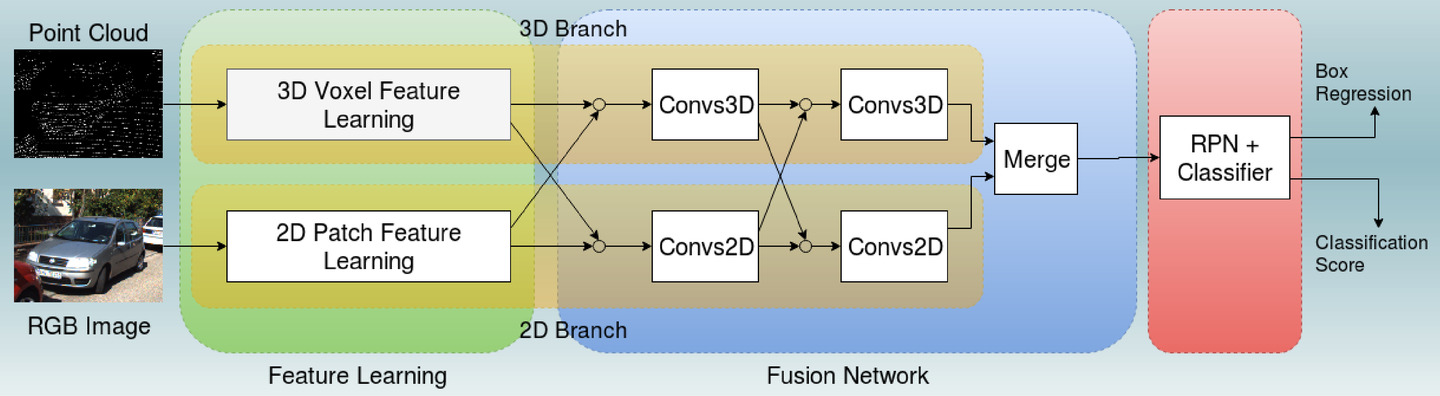
[1]Sherpa Engineering, [2]Institut Pascal

Ecole doctorale Sciences Pour l'Ingénieur

Goals

- Explore merging methods between point clouds from **automotive LiDAR** and **camera** video stream with artificial neural networks
- Develop a **obstacle detection** system using **both** modalities
- **Track** obstacles from detections and predict their trajectories

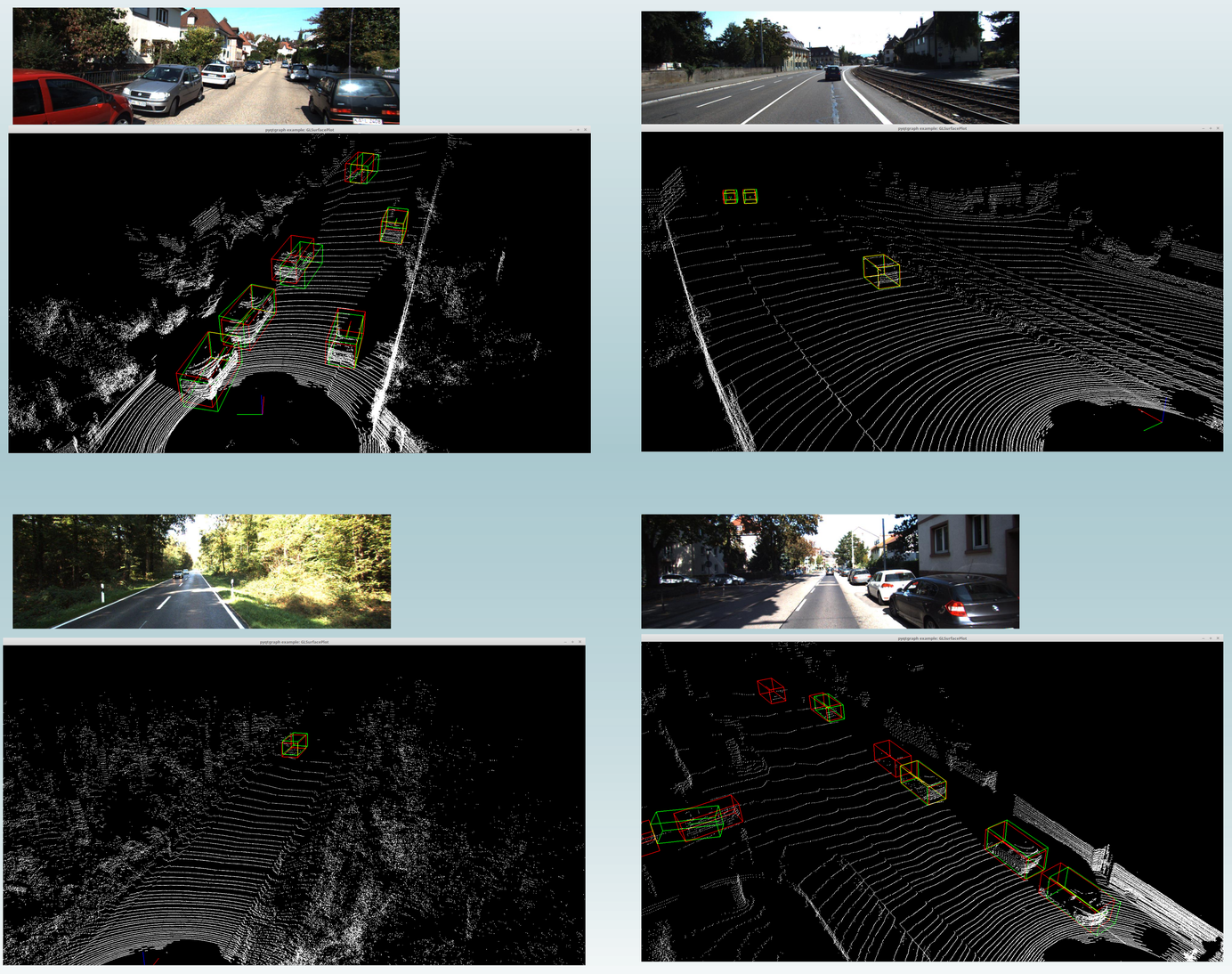
Network Architecture



Detection System

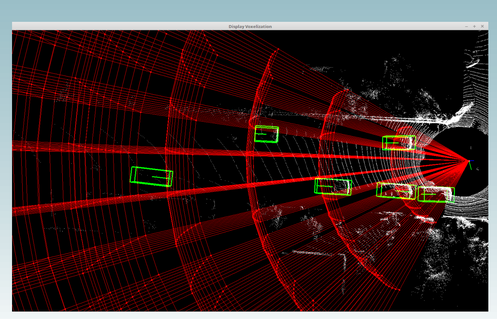
- Image patches and corresponding 3D cells learn their own representation independently from the other patches/cells
- Spherical coordinates for Voxel Feature learning
- Voxels features augmented with probabilities inspired by occupancy grid approach
- Diffusion of each branch information to the other branch

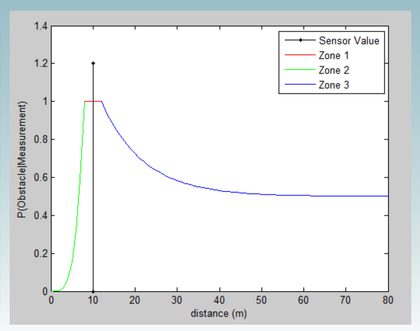
Experiments : Detection



Cell Probabilities

- For each patch, the related space region is divided into 3 zones depending on the sensor values and the distance
- Region 1 : Obstacle probably present
- Region 2 : Obstacle probably missing
- Region 3 : Missing information





Future Work

- Improve accuracy, execution time and stability on sensor loss
- Adaptation to different types of LiDAR (resolution, FOV…)
- Data augmentation : Information Densification (Depth) or Prediction (Surface Normals, Ego Motion Pose…)
- Obstacle tracking and trajectory estimation

Bibliography

- Zhou, Y., & Tuzel, O. (2017). VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection.
- Murphy, R., & Murphy, R. R. (2000). *Introduction to AI robotics*. MIT press.
- Ku, J., Mozifian, M., Lee, J., Harakeh, A., & Waslander, S. (2017). Joint 3D Proposal Generation and Object Detection from View Aggregation.
- Chen, X., Ma, H., Wan, J., Li, B., & Xia, T. (2017, July). Multi-view 3d object detection network for autonomous driving. In *IEEE CVPR* (Vol. 1, No. 2, p. 3).