

Deep Generative Modeling of LiDAR Data

July 2020

Summary

The problem in mobile robotics is the development of systems capable of fully understanding their environment. This paper provide a fully unsupervised method for both conditional and unconditional lidar generation. The author has established an evaluation framework for lidar reconstruction, allowing the comparison of methods over a spectrum of different corruption mechanisms. They have proposed a simple technique to help the model process noisy or missing data.

The data preprocessing involves converting a lidar scan containing $N(x, y, z)$ coordinates into a 2D grid. The points are clustered together from the same elevation angle into H clusters. For every cluster, the points are sorted in increasing order of azimuth angle. The grid with fixed amount of points per row the plane is divided into W bins. This creates a $H \times W$ grid, where for each cell the average coordinate (x, y, z) is stored, hence all the information is in a $H \times W \times 3$ tensor, where (x, y) channels are compressed as $d = \sqrt{x^2 + y^2}$.

Here the Deep Convolutional GANs are used for generating the images. The generator part consists of five transpose convolutions with stride two to up-sample at each layer, and ReLU activations. The discriminator uses stride two convolutions to down-sample the input, and Leaky ReLU activations. Batch normalization is used between convolution layers for easier optimization. The VAE (Variational Autoencoders) encoder setup is simply the first four layers of the discriminator, and the decoder's architecture replicates the DCGAN generator.

The model is trained on KITTI dataset. To measure how close the reconstructed output is to the original point cloud, we use the Earth-Mover's Distance. This paper got 127.0 EMD on test set reconstructions. The proposed adversarial network can generate highly realistic data, and captures both local and global features of real lidar scans.