Point Cloud Occupancy with Dynamic Planes

Computer Vision Course Project

Master's Degree in Artificial Intelligence and Robotics

Eugenio Bugli (1934824)

Academic Year 2024/2025

Sapienza
Università di Roma

Table of Contents

1 Introduction

TODO

- fix sampling
- fix decoder
- reconstruction
- metrics
- slides

What is the Addressed Problem

- In this work we are interested in performing the reconstruction of point clouds.
- The input noisy point clouds are encoded into per-point features that are projected onto multiple 2D dynamic planes.
- Then we predict the occupancy values of each point in order to find the surface of the shapes.
- The original paper applied this study to the ShapeNet Dataset.

Insert 3 images of the same sample

Table of Contents

FAUST Dataset

This Dataset is composed by high-resolution human scans of 10 different bodies in 30 different poses.

- The test set is composed by 200 scans, while the training has 100 scans.
- Each of the samples inside the training set has a corresponded ground truth alignment (registration)
- The training set has been partitioned again in order to obtain train and validation sets
- About 80 % fo the initial training set has been used for training, while the other 20 % has been used for validation

Examples from the Dataset

Insert here 1/2 images of different bodies with differente poses. Registration + Clouds

Even if I have used a different dataset, I have followed the same sampling and augmentation strategy for my input data:

- First we need randomly sample 3000 points from the cloud's surface
- Then we inject Gaussian noise with zero mean and 0.05 standard deviation

These processed clouds are then used to learn the features and the geometry of the object through the Encoder. The occupancy is then predicted over a set of 2048 points uniformly sampled over all the space, both inside and outside the shape.

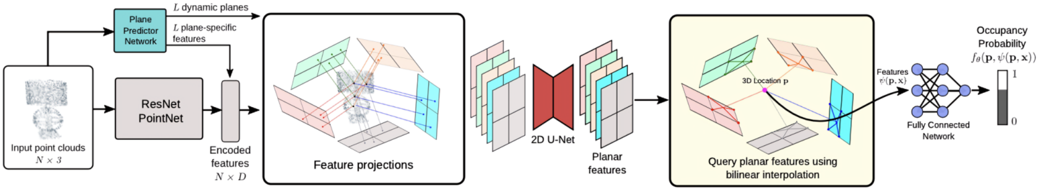
Examples of Sampled Clouds

Insert here images of the sampled point cloud in the 3 Cases + Real cloud

Table of Contents

The Architecture is characterized by an Encoder-Decoder structure:

- Encode the input clouds into 2D Feature Planes
- Decode these features into occupancy probabilities

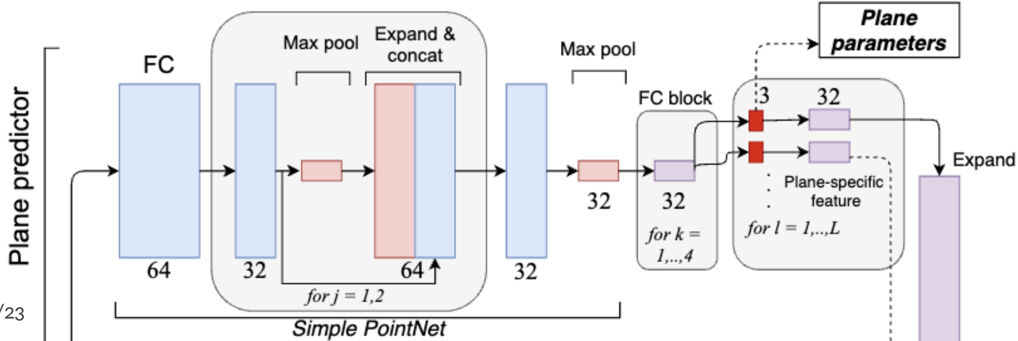


Encoder

3 Architecture

The Encoder is composed by :

- ResNet PointNet
- Plane Predictor
- UNet



Decoder

The Dec is composed by:

- Feature Projection and Bilinear Interpolation
- Occupancy Network

Table of Contents

4 Reconstruction

Reconstruction Phase

Table of Contents

5 Results

In order to evaluate the performance of our model, the following metrics have been used:

- Chamfer Distance :
  $CD(A,B) = \frac{1}{|A|}\sum_{a\in A}\min_{b\in B}\|a-b\|_2^2 + \frac{1}{|B|}\sum_{b\in B}\min_{a\in A}\|b-a\|_2^2$
- IOU : $IoU(A',B') = \frac{|A'\cap B'|}{|A'\cup B'|}$
- F-Score:

Add each formula

Loss Performances

Insert here plots

Metrics Performances

Insert here just a table with metrics, gpu usage various types of sampling

ReconstructionPerformances

Insert here just some images about reconstructions

Table of Contents

6 Improvements

Possible Changes and Future Improvements

Point Cloud Occupancy with Dynamic Planes

Eugenio Bugli

1934824