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

What is the Addressed Problem

- We would like to learn the **Occupancy** values of points inside a bounding box
- We learn the features and the dynamic planes
- During Inference we reconstruct meshes with Multiresolution IsoSurface Extraction

Table of Contents

FAUST Dataset

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

- Each samples inside the training set is composed by the scan and the registration, while we have only the scan for the test set.
- The test set is composed by 200 scans, while the training has 100 scans.
- About 80 % fo the initial training set has been used for training, while the other 20 % has been used for validation.

Data Processing

Starting from the meshes given by the Dataset, some pre-processing is needed:

- We need to randomly sample **3000** points from the mesh surface and add a perturbation (Gaussian noise) to them.
- We need to randomly sample **2048** points from the bounding box that contains the original mesh.

These noisy clouds are then used to learn the features and the geometry of the object through the Encoder, while the other points are used in the decoding part of the network.

Data Augmentation

Here is some introductory text explaining the data augmentation process. This text spans the full width of the frame.

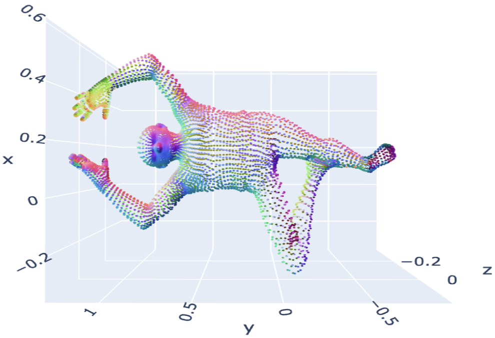


Figure: Registration cloud

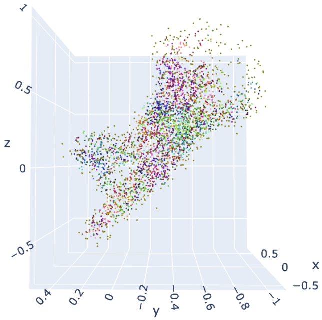


Figure: Augmented cloud

Label Generation

For each Sampled Cloud, labels have been generated with the following procedure:

- Measure the distance between the original mesh and the sampled points
- If the distance is positive then the point are inside the mesh
- If the distance is 0 the points are on the surface
- If the distance is negative outside a small threshold I am considering them outside

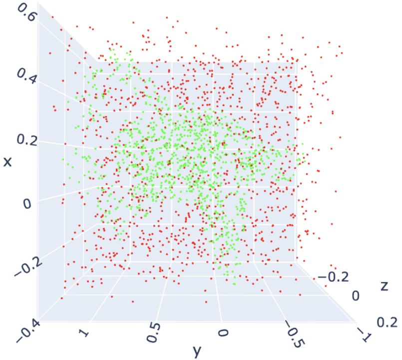
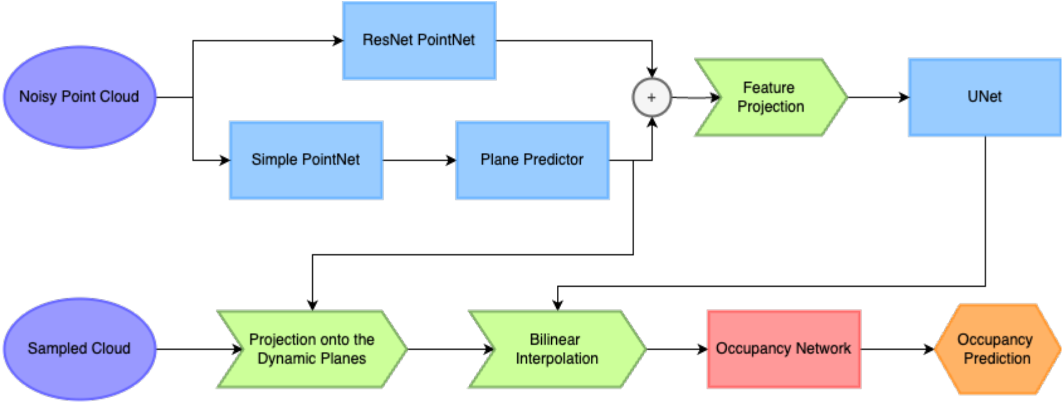


Figure: Example of the Sampled cloud, where with green we represent occupied points and with red the unoccupied ones

Table of Contents

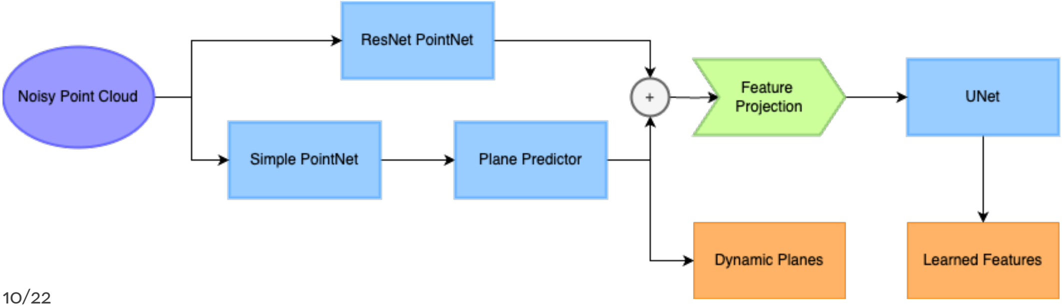
Architecture design



Encoder

The Encoding part takes in input the Noisy Cloud and it’s composed by the following steps:

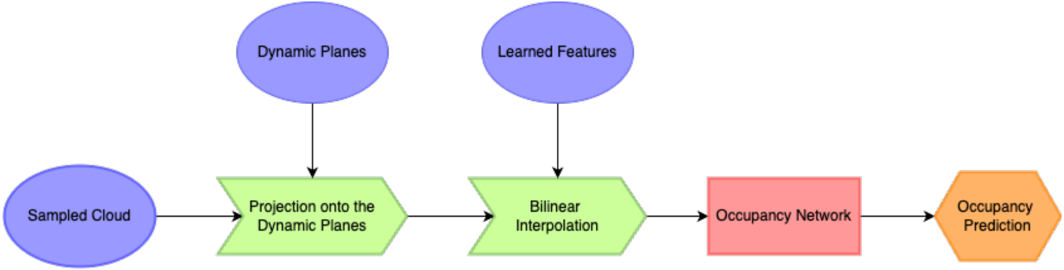
- ResNet PointNet
- Simple PointNet + Plane Predictor
- Feature summation + projection
- UNet



Decoder

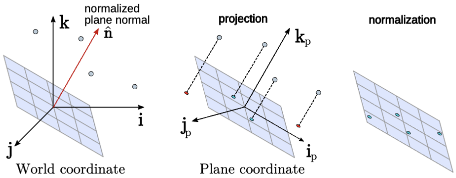
The Decoding part is composed by:

- Feature Projection and Bilinear Interpolation
- Occupancy Network



Both the Encoder and the Decoder have a projection operation involving the features and the **Dynamic Planes**.

- Change of Basis
- Orthographic Projection
- Normalization



$$\mathbf{R} = \mathbb{I} + [v]_\times + [v]_\times^2 \frac{1 - k \cdot \hat{\mathbf{n}}}{||v||^2} \tag{1}$$

Where $v = k \times \hat{\mathbf{n}}$ and $[v]_\times$ it's the skew symmetric matrix of the vector $v \in \mathbb{R}^3$.

Bilinear Interpolation

- **Definition:**
  - Bilinear interpolation is a resampling technique used to estimate new pixel values on a 2D grid.
- **Concept:**
  - It computes the output value as a weighted average of the four nearest pixel values.
  - The interpolation is performed first in one direction (e.g., horizontal) and then in the other (vertical).

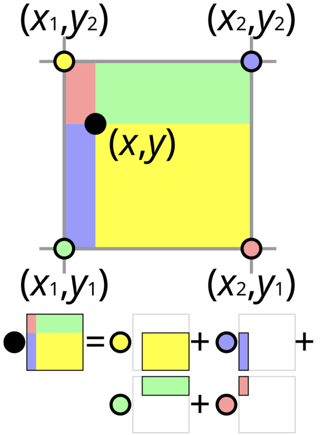


Figure: Bilinear Interpolation Visualization (credits: Wikipedia)

Table of Contents

Multiresolution IsoSurface Extraction (MISE)

- Create a grid over all the bounding box
- Evaluate the occupancy of each corner of the voxels
- Define the Active voxels as the one composed with at least one occupied corner and one not
- Subdivide each Active voxels into 8 subvoxels (2x2x2 grid) and evaluate the new points occupancy
- Repeat until the desired resolution is obtained

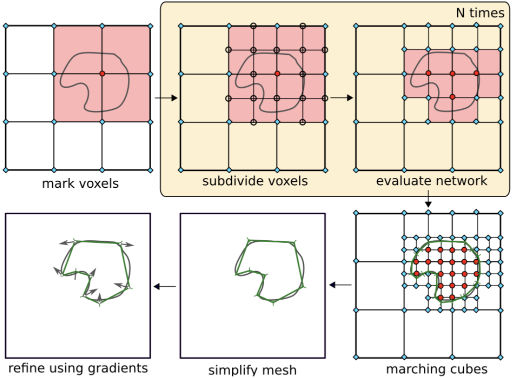


Table of Contents

5 Results

Metrics

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

Table of Contents

6 Improvements

Possible Changes and Future Improvements

Point Cloud Occupancy with Dynamic Planes

Thank you for listening