University of Dayton
VISION LAB
Center of Excellence

Viewpoint Transformation for Reducing Self Occlusion in Aerial Lidar

Nina Singer[1], Vijayan K. Asari[1], Theus Aspiras[1], Jonathan Schierl[1], Andy Stokes[2], Brett Keaffaber[2], Andre Van Rynbach[2], Kevin Decker[3] and David Rabb[2]

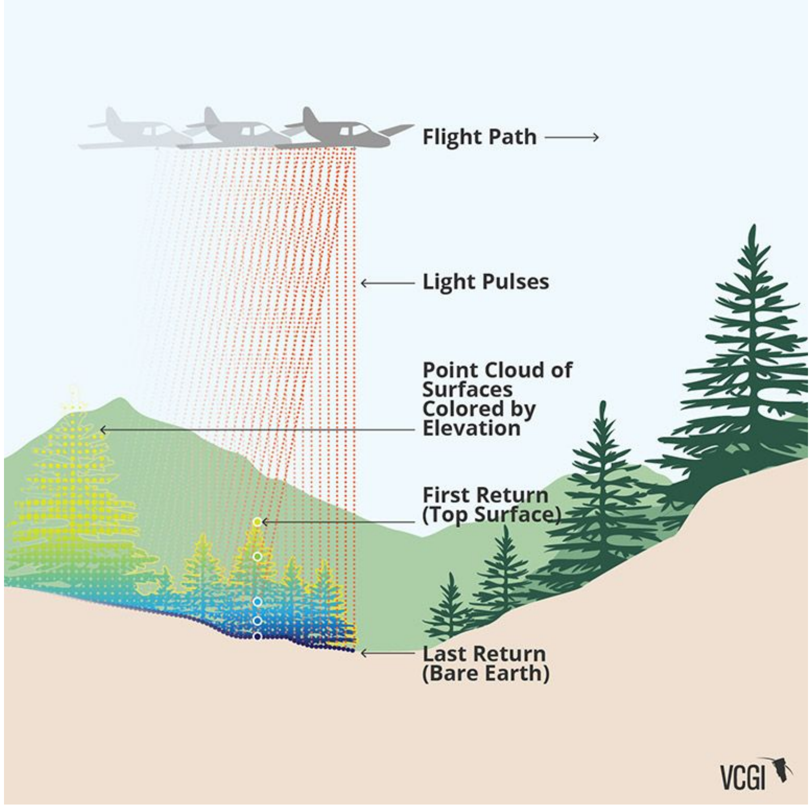
[1] University of Dayton, [2] Air Force Research Labs, [3] Defense Engineering Corp

Objective

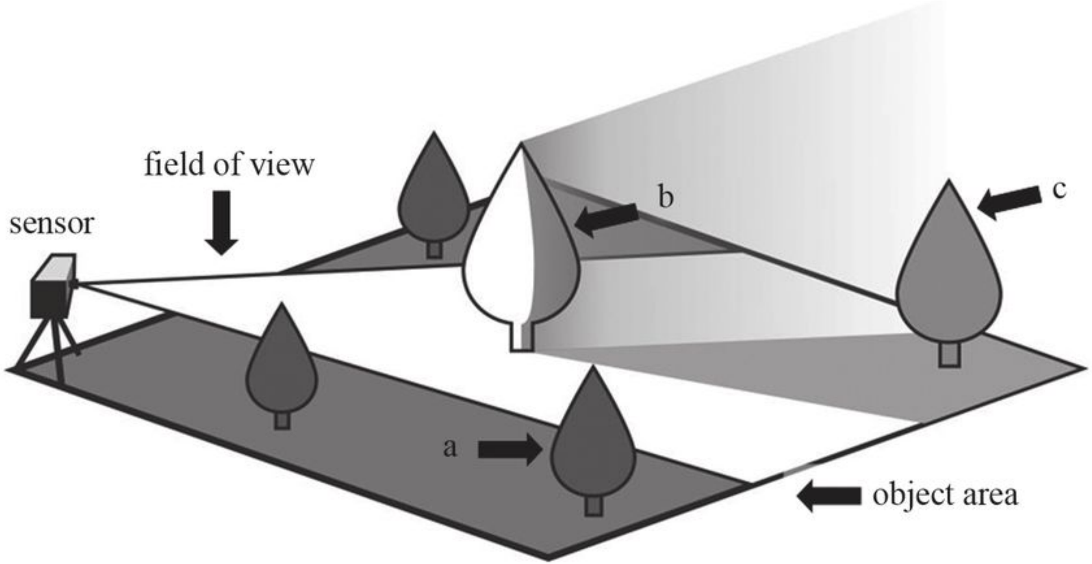
This work aims to tackle the problem of self-occlusion in single shot lidar

Problem Statement

If we have an aerial lidar point cloud with an occlusion due to sensor position, can we create a network which will generate the missing points



Occlusion Types



(a) View Occlusion

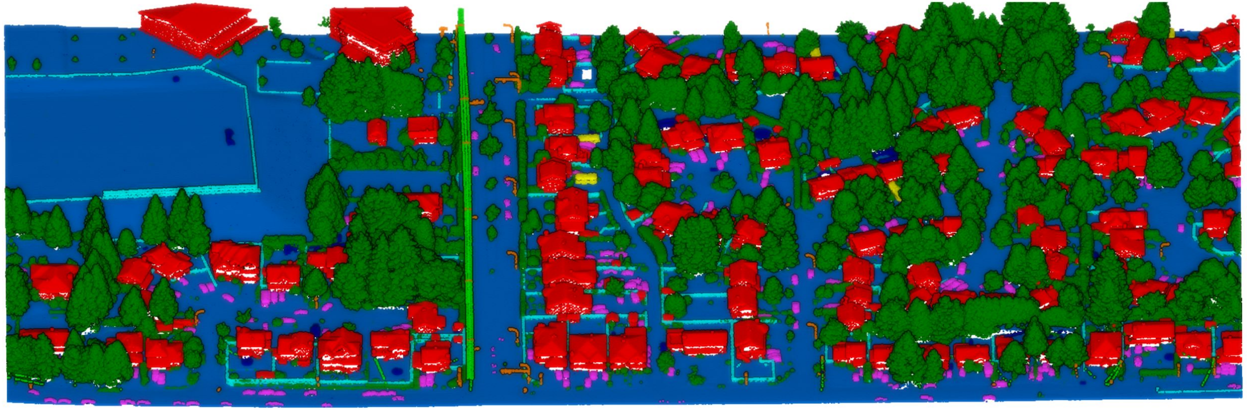
(b) Self-occlusion

(c) Ambient Occlusion

Dataset Creation: DALES Viewpoints

DALES: Dayton Annotated Laser Earth Scan

University of Dayton
VISION LAB
Center of Excellence



- Largest publicly available semantic segmentation data set for aerial LiDAR
  - 40 tiles, spanning 10 km^2 with 50 ppm
  - ~500 million expert-labeled points
  - 8 unique object categories
    - Ground, Vegetation, Poles. Fences, Power Lines, Trucks, Cars, Buildings, Unknown
  - 4 distinct scene types
    - Rural, urban, suburban, commercial

go.udayton.edu/dales3d

HELIOS ++ Overview

Survey

Default scanner settings
Fullwave settings

Leg

Platform settings
(position, speed, ...)
Scanner settings
(pulse freq., scan angle,
rotation speed, ...)

Scanner

Accuracy
Beam divergence
Deflector mechanism
Pulse Length
Fullwave settings
(overridden by Survey)

Platform

Type (static, linear, multicopter,
ground vehicle)
Mount parameters
lever arm
boresight alignment
Position noise

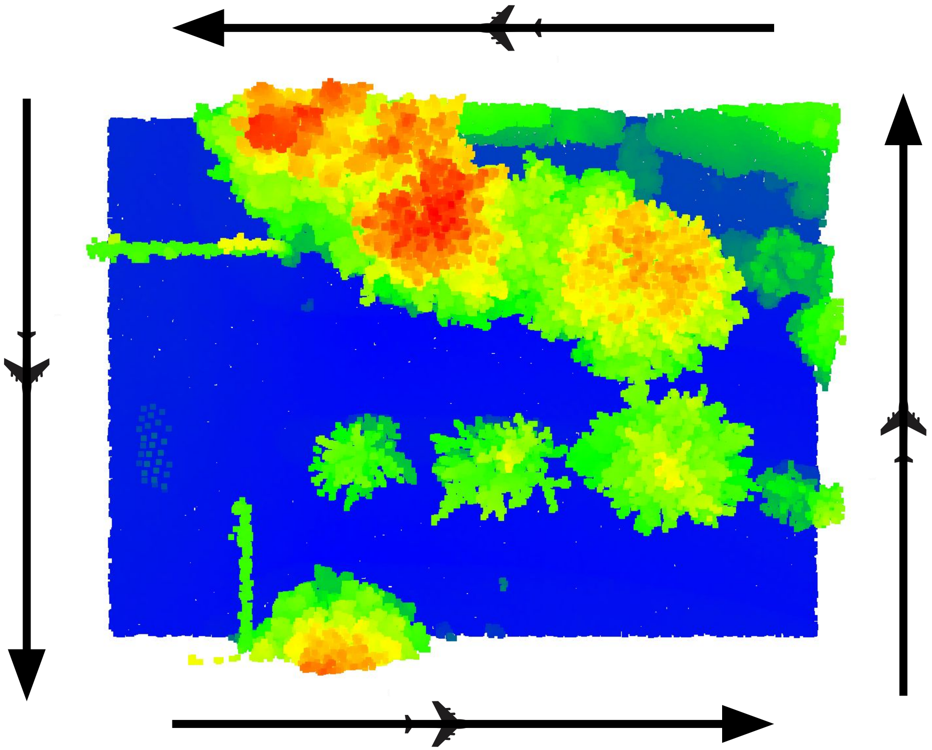
Scene

Part

Fileloader (loads a scenepart)
GeoTIFF file
Wavefront Obj
XYZ Pointcloud
vox-Voxel model
Filter (rotate, translate, scale)

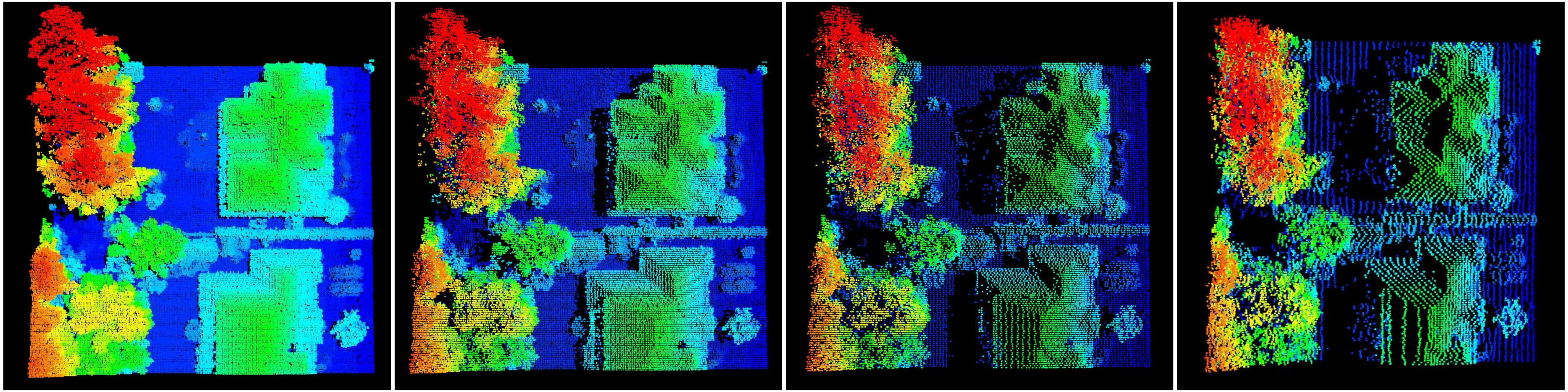
DALES Viewpoints

- Same tiles as DALES
  - 40 total tiles
    - Split into chunks, 100,000 points each
    - 4 viewpoint per chunk
    - 500 meters from each edge

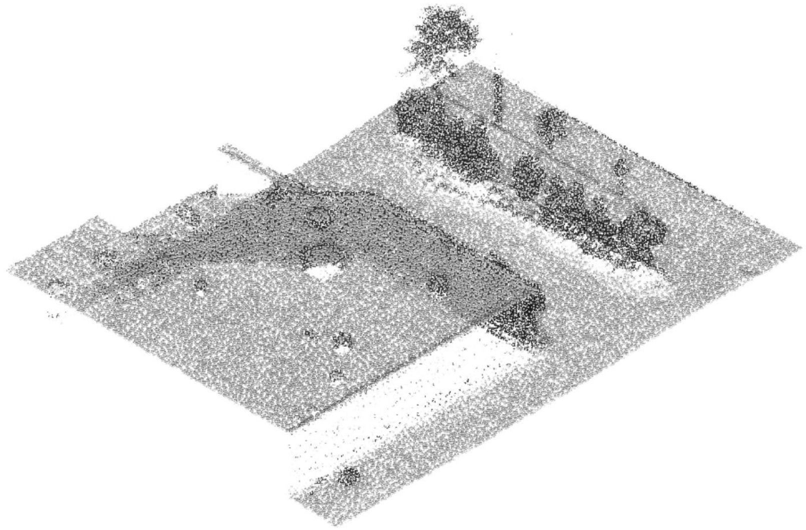


Helios ++: Creating Occlusions

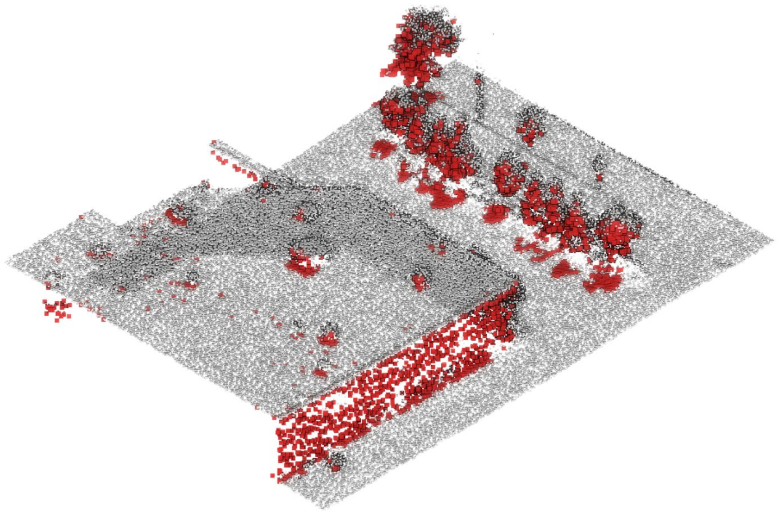
Nadir | 23° off nadir | 26.5° off nadir | 30° off nadir



DALES Viewpoints Sample



Input:
Single Shot Scene

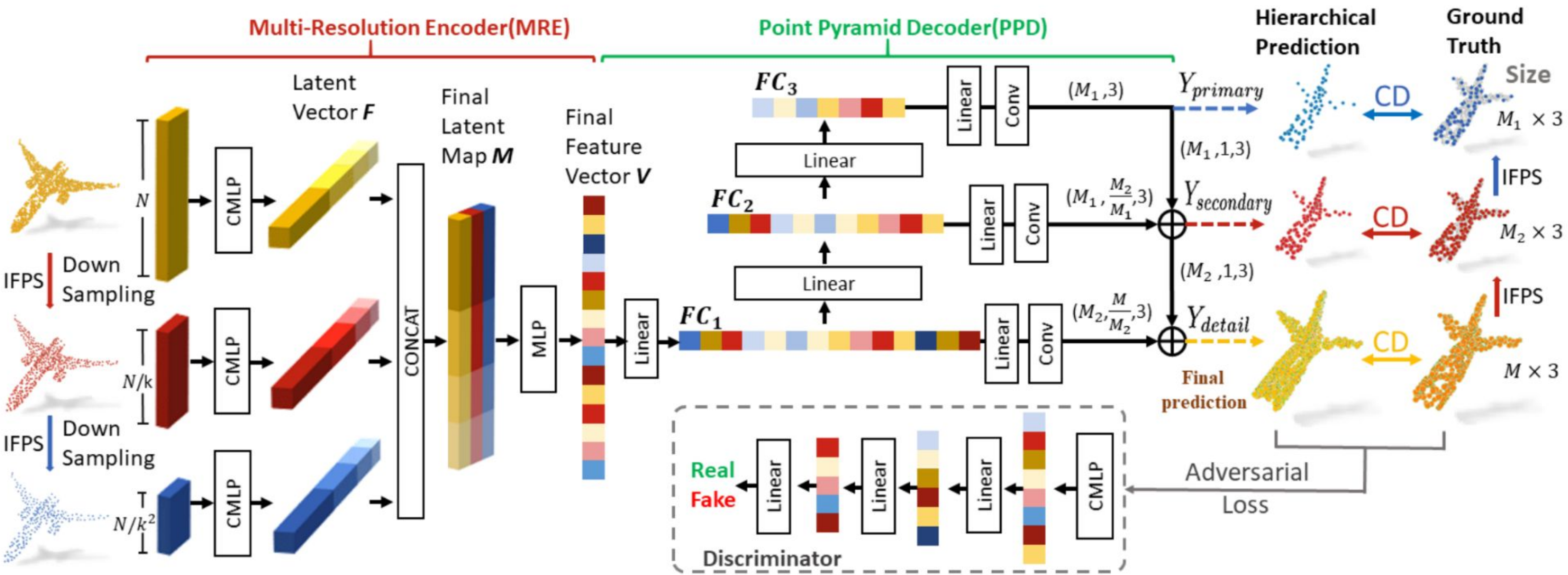


Desired Output:
Single Shot Scene w/ Occluded Points

University of Dayton VISION LAB Center of Excellence

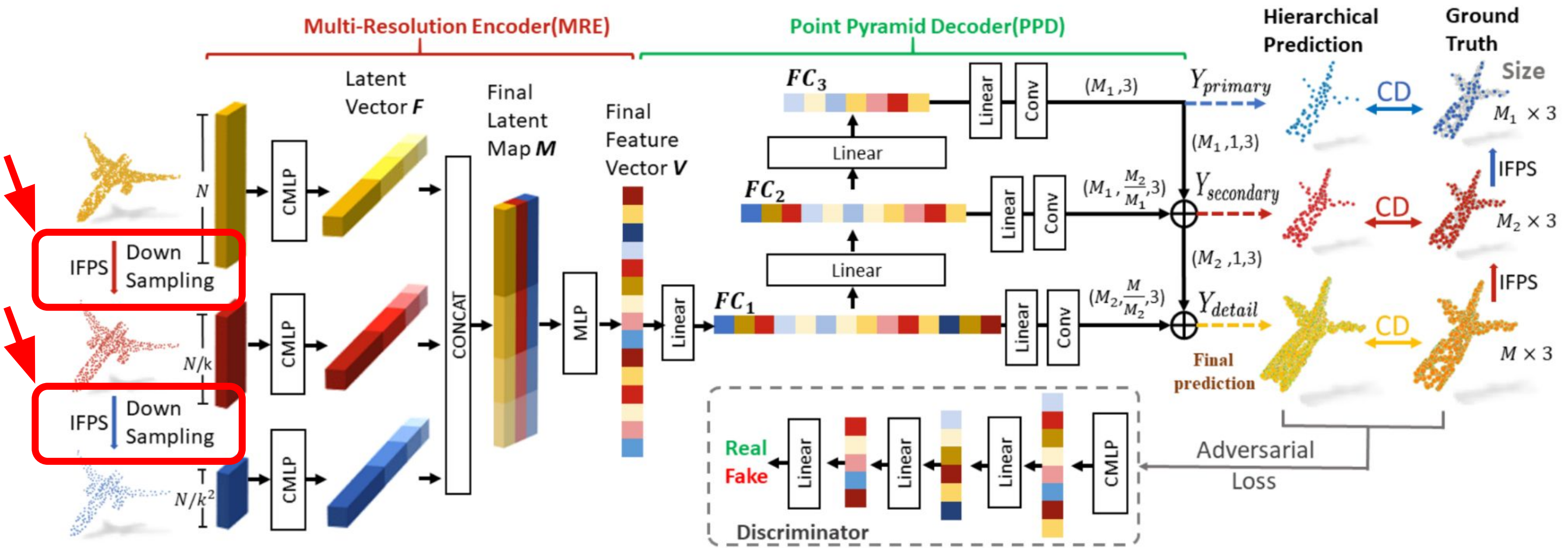
Architecture

Backbone Network: Point Fractal Network



*Huang, Zitian, et al. "Pf-net: Point fractal network for 3d point cloud completion." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.

Backbone Network: Point Fractal Network



*Huang, Zitian, et al. "Pf-net: Point fractal network for 3d point cloud completion." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.

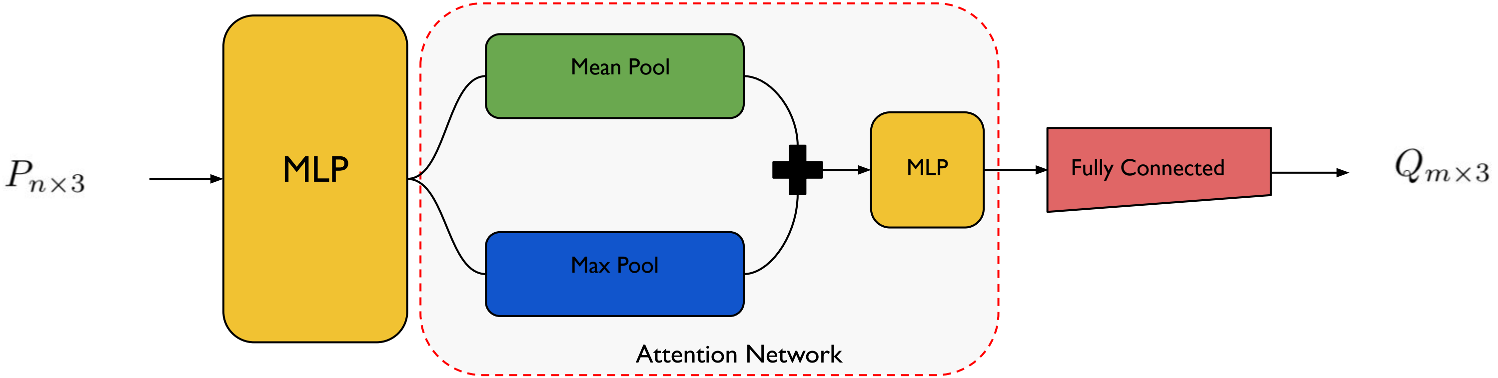
Sampling Proposal

The majority of generative point cloud algorithms are proposed for small synthetic datasets

- Aerial datasets are much larger
  - Tens of millions of points instead of thousands
  - Objects are scattered throughout the scene
    - No one point of focus

We wish to propose a sampling network which can drastically reduce the number of points, while also keeping key points

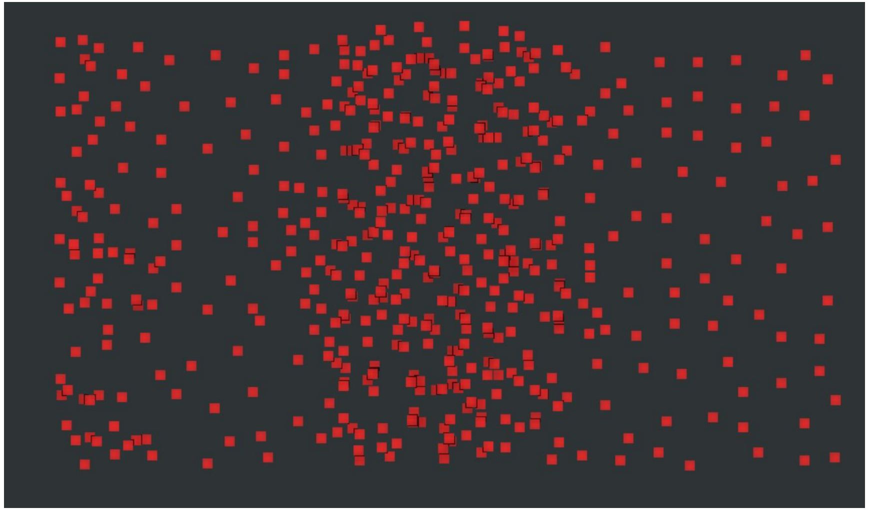
Learned Attentive Sampling



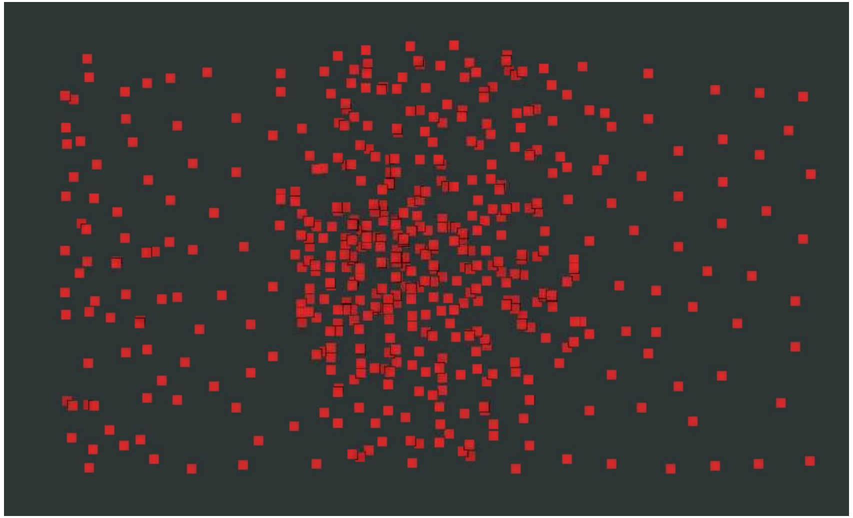
- Implement a learned attention-based sampling technique
  - Using MLP to reduce the number of points and manipulate the dimensionality
  - Propose an attention mechanism to sample key points
    - Improvement over iterative farthest point sampling techniques

Sampling Method Comparison

University *of* Dayton VISION LAB Center of Excellence

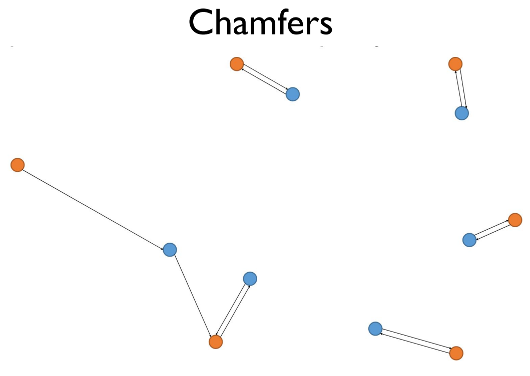


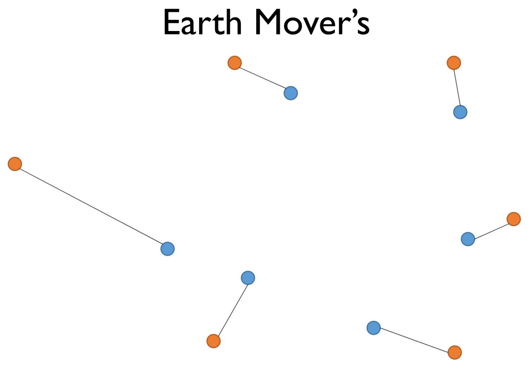
Farthest Point Sampling



Learned Sampling

Distance Metrics





We use two distance measures in our metrics

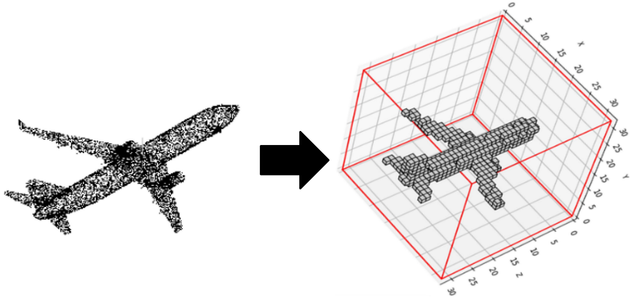
- Chamfer's: Non-bijective, fast
- Earth Mover's: Bijective, slow

Metrics

Jensen-Shannon Divergence

$$JSD(P_g||P_r) = \frac{KL(P_r||M) + KL(P_g||M)}{2}$$

$$M = \frac{P_r + P_g}{2}$$



Coverage

$$COV(A,B) = \frac{|\{argmin_{S_2 \in B} D(S_1, S_2) | S_1 \in A\}|}{|B|}$$

Minimum Matching Distance

$$MMD(A,B) = \frac{1}{|B|} \sum_{S_2 \in B} \min_{S_1 \in A} D(S_1, S_2)$$

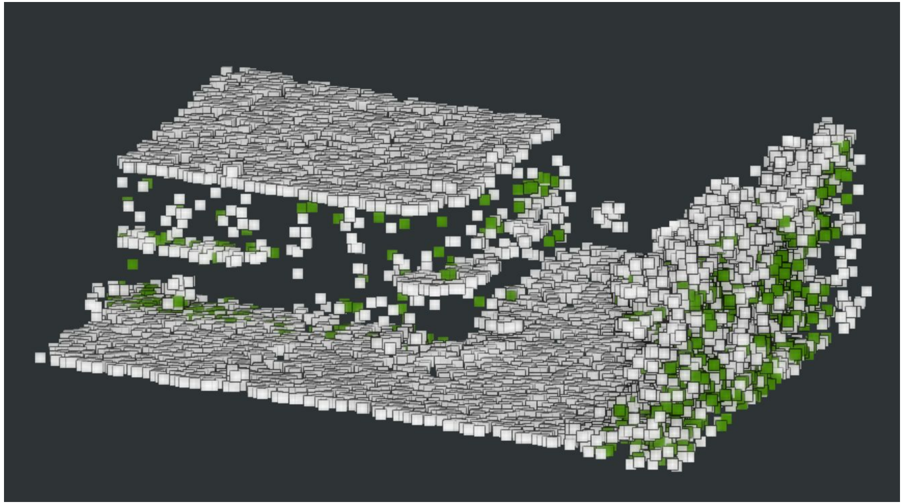
Quantitative Results

Sampling Method	↓ Jensen-Shannon Divergence
Farthest Point	0.2644
Learned Attentive	**0.0919**

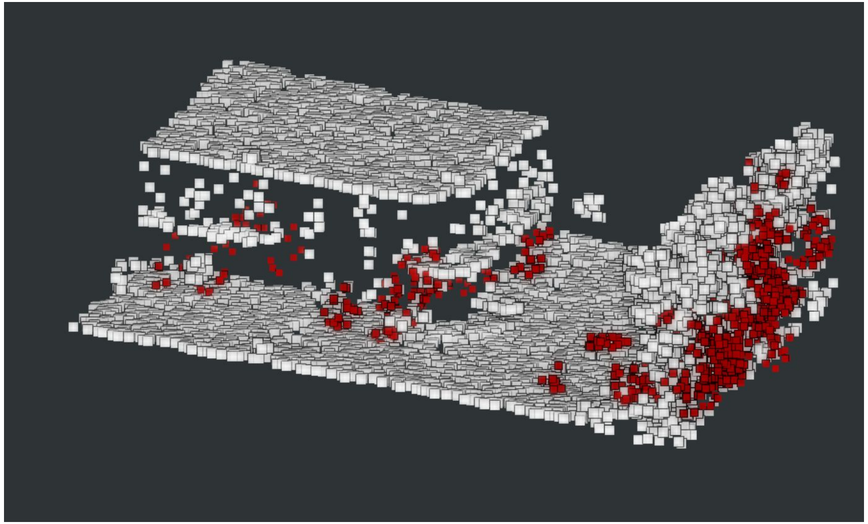
Sampling Method	↑ Coverage CD	↑ Coverage EMD
Farthest Point	**0.5077**	0.3383
Learned Attentive	0.3775	**0.4969**

Sampling Method	↓ MMD CD	↓ MMD EMD
Farthest Point	0.0308	0.3488
Learned Attentive	**0.0196**	**0.2216**

Viewpoint Transformation Results

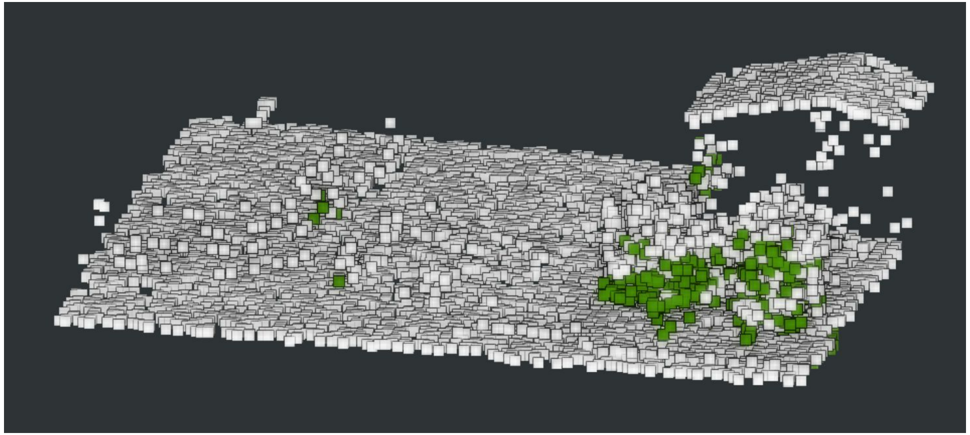


Actual Occluded Points

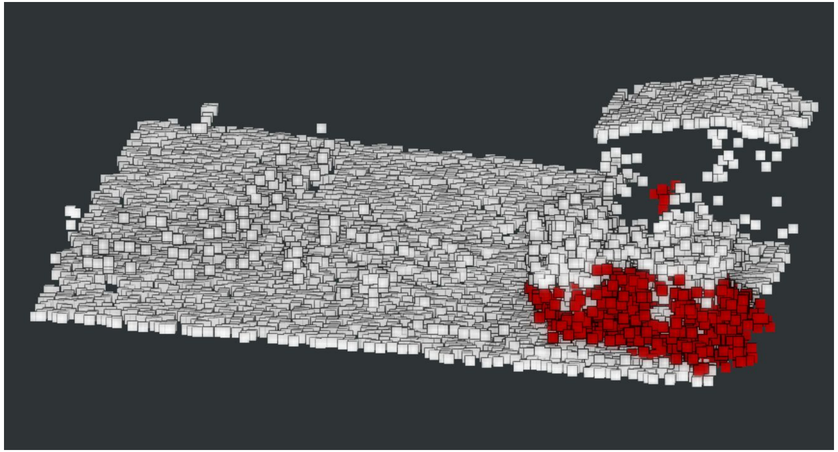


Suggested Occluded Points

Viewpoint Transformation Results



Actual Occluded Points



Suggested Occluded Points

Conclusion

- Attentive sampling is effective for our viewpoint transformation application.
- Learned sampling provides better distribution for scene based applications.

Future Work

- Need additional testing for using learned attentive sampling for other applications.
  - Segmentation, registration, classification, etc.