A Survey on 3D Point Cloud Compression Using Machine Learning Approaches

Reetu Hooda, Dr. W. David Pan and Tamseel Mahmood Syed

Department of Electrical and Computer Engineering
University of Alabama in Huntsville
Huntsville, Alabama

1st April 2022

THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Outline

- Introduction
- Autoencoders
- Fully Connected Neural Networks
- Recurrent Neural Networks
- Conclusion

THE UNIVERSITY OF
ALABAMA IN HUNTSVILLE

INTRODUCTION

- What is a Point cloud?
  - Geometry -> 3D coordinates
  - Attributes -> color information or normal vectors

Point Cloud	
Geometry	Attribute

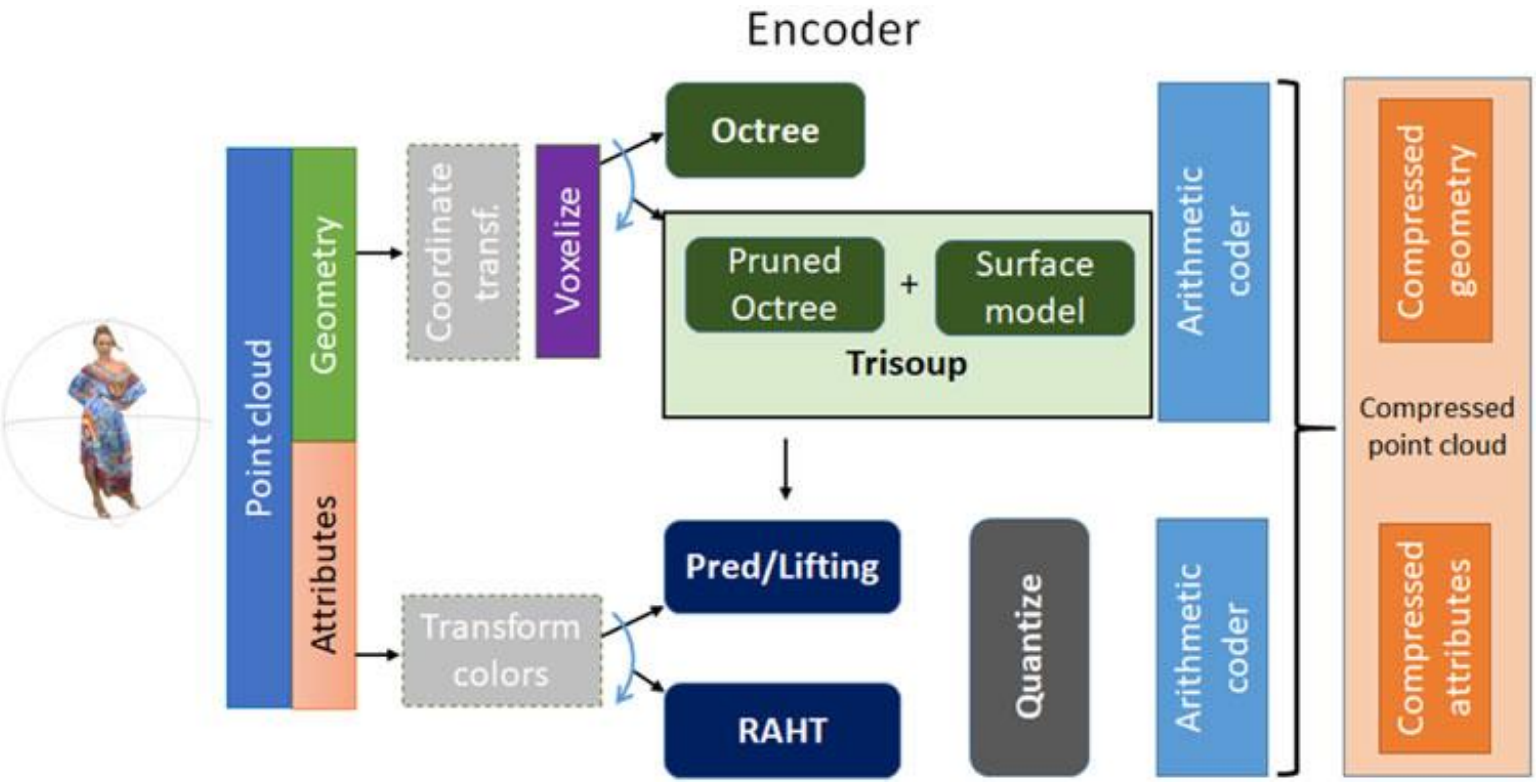
- Applications: Cultural heritage, Autonomous driving, Robotics etc.
- Need of Point Cloud Compression.
- MPEG PCC standardization:

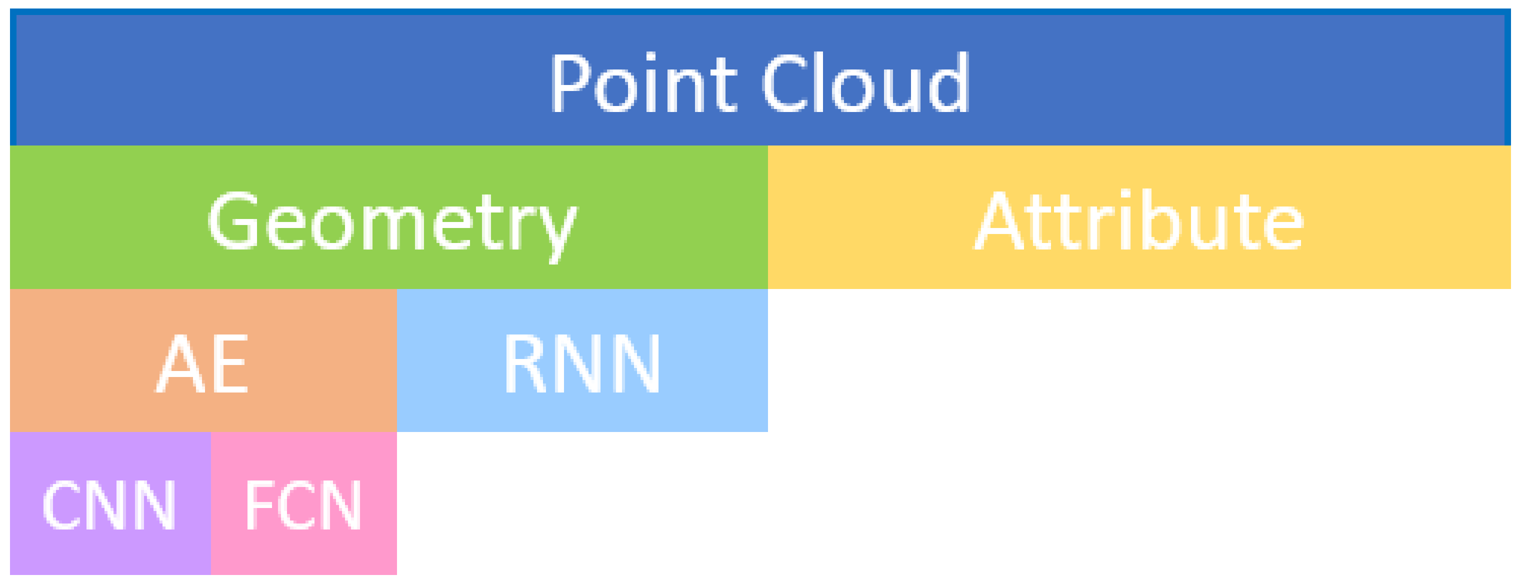
PCC

V-PCC G-PCC

OVERVIEW

- Machine learning for media applications.
- Traditional Vs Machine learning based techniques.

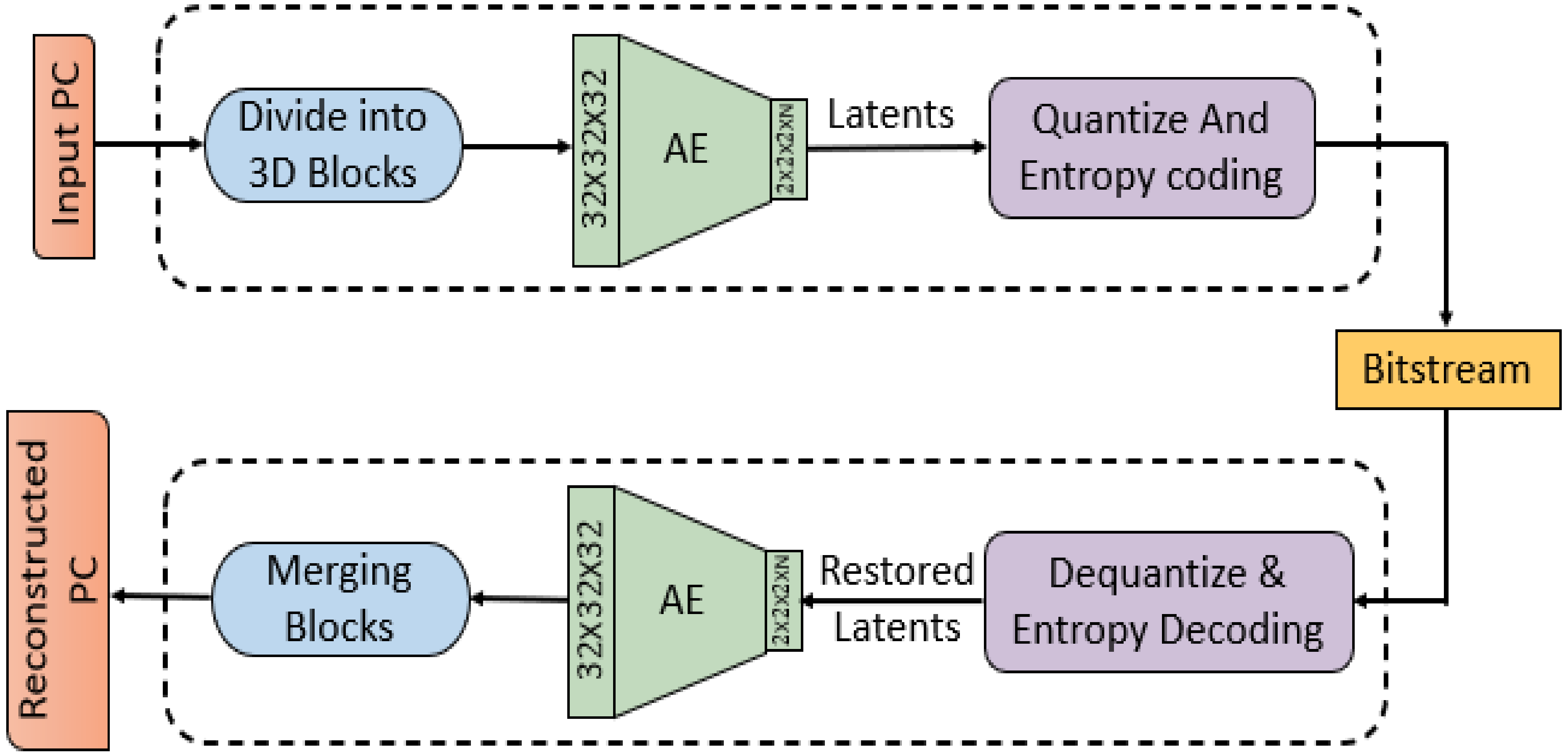




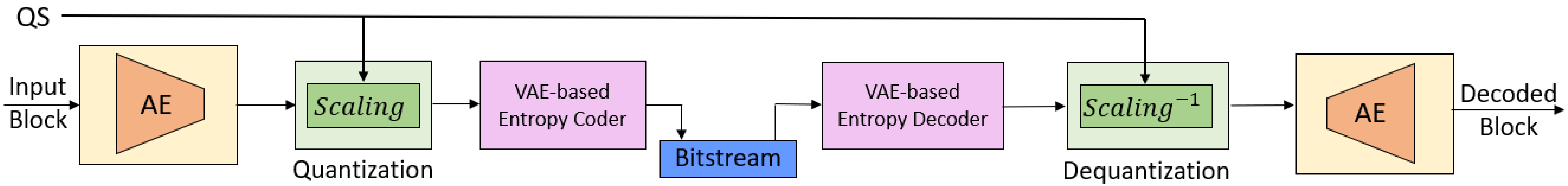
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

POINT CLOUD GEOMETRY AUTOENCODER

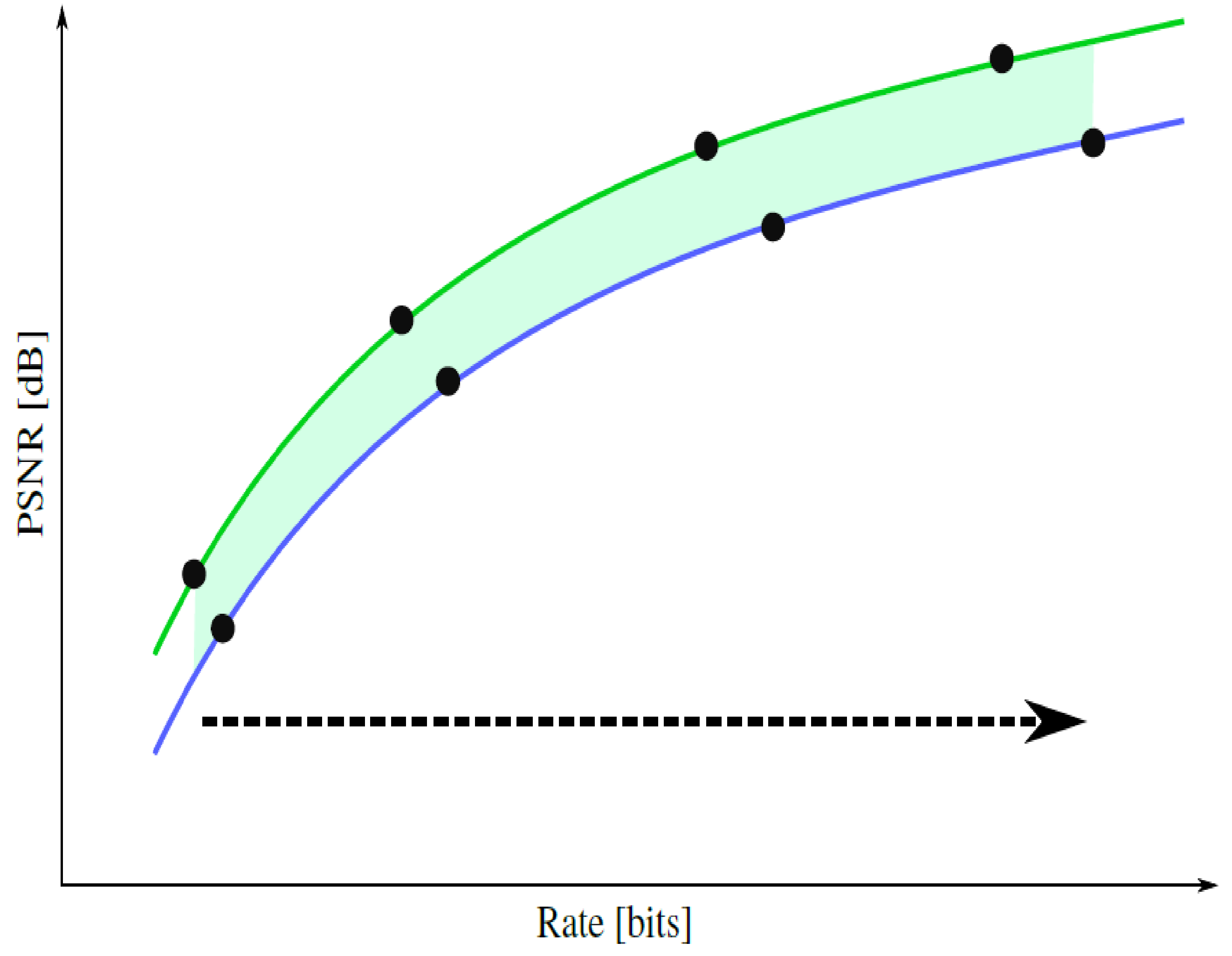
- CNNs: Most effective in extracting features.
- Characteristics found in images and videos.
- PCG-AE: Basic implementation.
- Occupancy signaled by ‘1’ and ‘0’.
- Adaptive forward and Inverse Transform.
- Four models, for N = 32, 64, 96 and 128.
- MPEG dataset, PCL as benchmark, RD performance.



RD CONTROL THROUGH IMPLICIT AND EXPLICIT QUANTIZATION

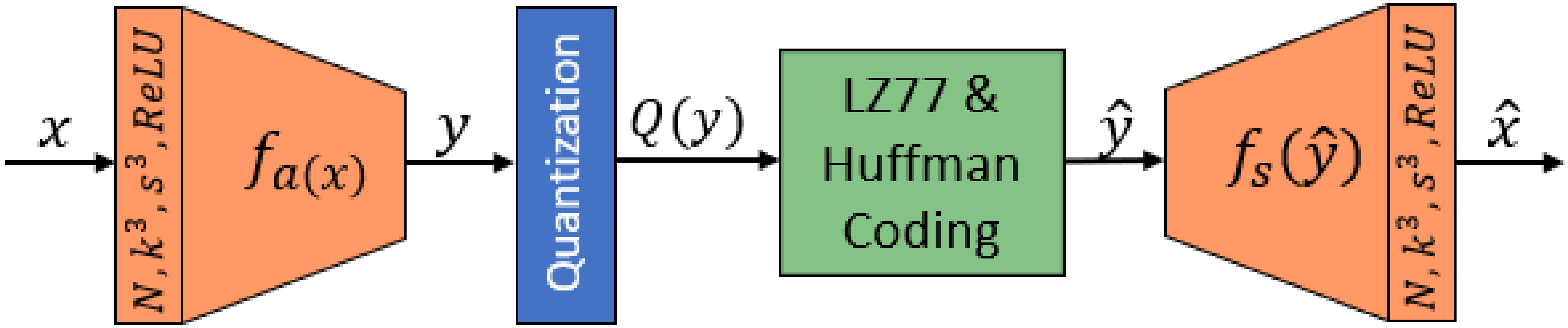


- Implicit ($Loss = dist + \lambda \times rate$) Vs Explicit (QS).
- Solves the multiple models training problem.
- Cover wide range of bitrates.
- Reduces training complexity and memory requirements.



THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

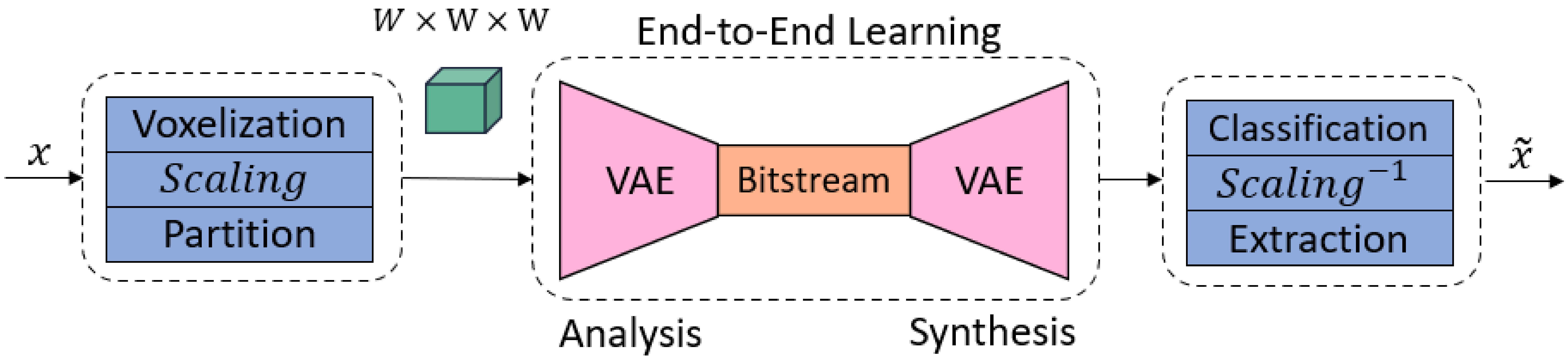
LEARNED CONVOLUTIONAL TRANSFORMS



- Directly operates on voxels and decoding-> Classification problem
- Analysis and synthesis transform.
- Number of filters, filter size and strides.
- Convolution and transpose convolution.
- Trained on ModelNet40, Tested on MVUB, MPEG anchor (51.5%).

THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

LEARNED PC GEOMETRY COMPRESSION



- Preprocessing: Voxelization is an optional step, Scaling ($\times\ s$) to reduce sparsity, Partition into non-overlapping blocks and metadata.
- VAE based encoder with Voxeption Resnet (VRN).
- Extraction: Volumetric to raw format.
- End-to-End training by changing λ.
- Training: ShapeNet, Testing: MPEG and JPEG Pleno.
- Outperforms GPCC and PCL (>50%).

THE UNIVERSITY OF
ALABAMA IN HUNTSVILLE

DEEP AE-BASED POINT CLOUD GEOMETRY COMPRESSION

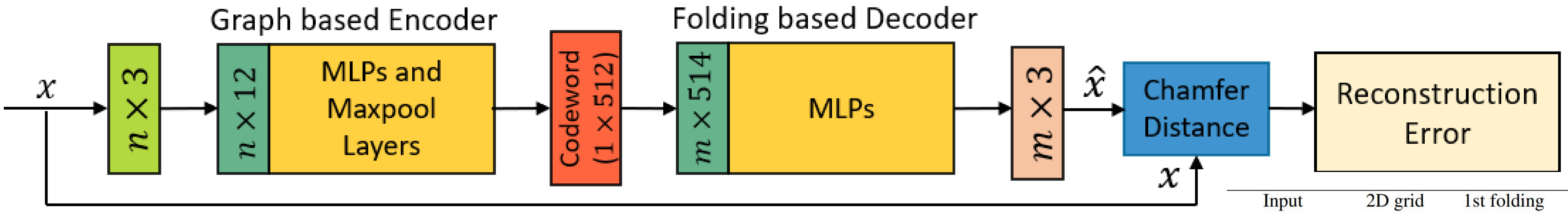
- Advantages of CNNs: Retain spatial relations and weight sharing.
- FCNN: It works directly on the geometry coordinate.
- Filters in each layer: 64, 128, 128, 256 and k (no. of input points).
- ShapeNet with 90%, 10% split from a single class.
- MPEG-GPCC: 73.17% average gain.

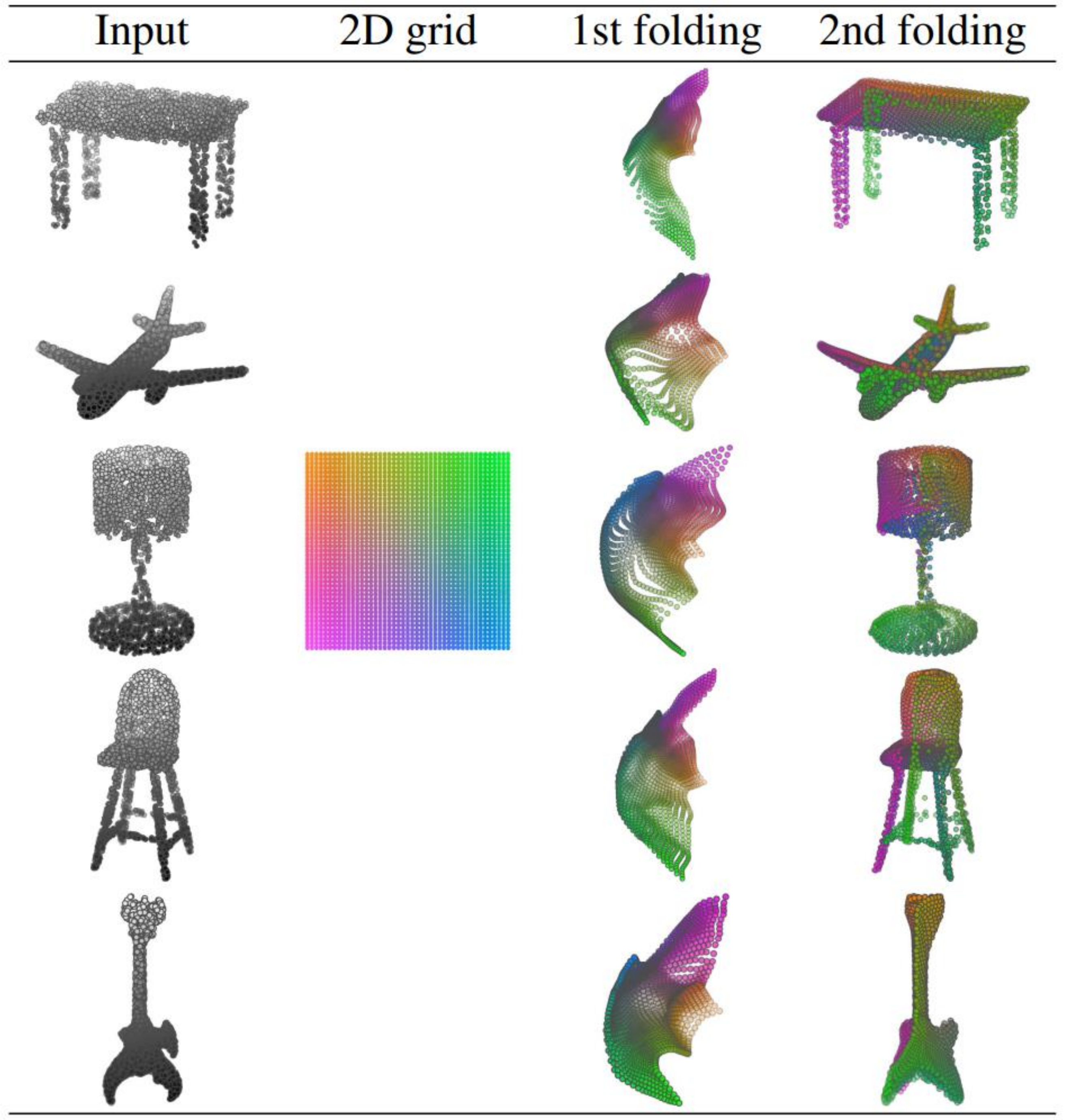
$$z = D(Q(E(S(x))))$$

$$d_{CH}(S_1, S_2) = \sum_{x \in S_1} \min_{y \in S_2} ||x - y||_2^2 + \sum_{y \in S_2} \min_{x \in S_1} ||x - y||_2^2$$

THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

FOLDING NET: DEEP GRID FORMATION

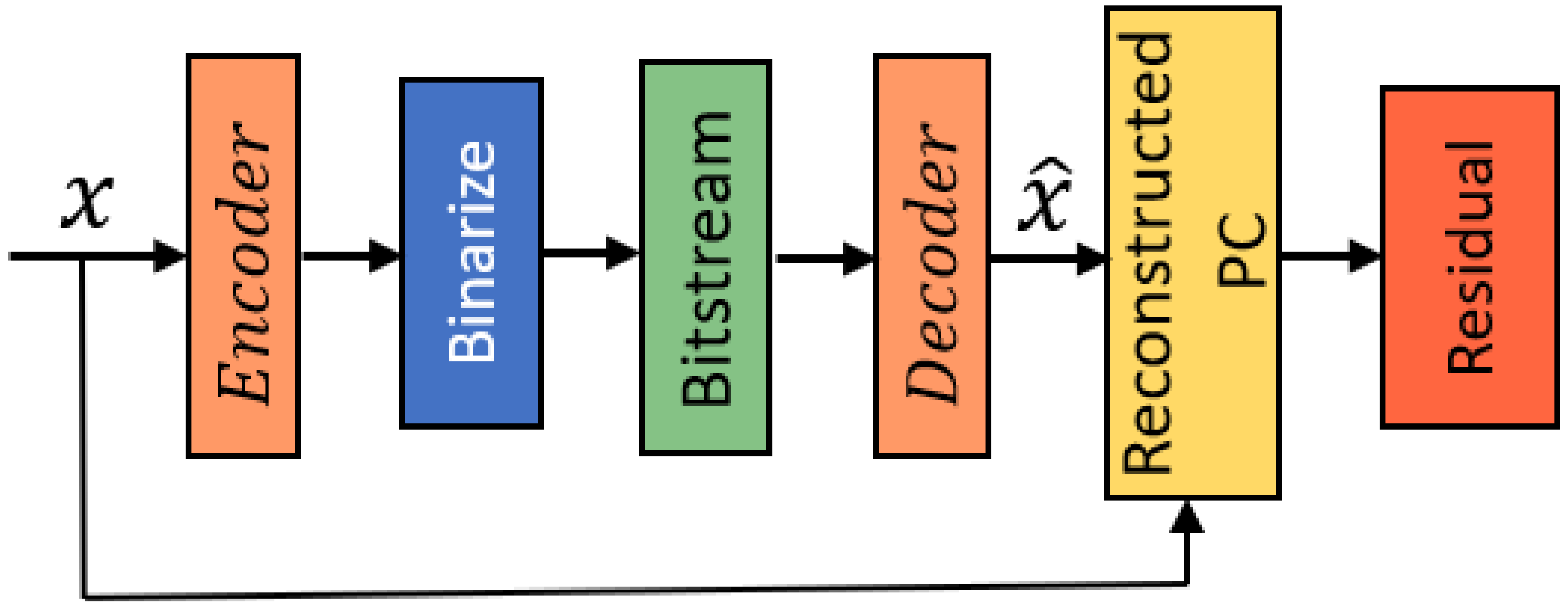




- Voxelization is performed to make CNN more meaningful.
- However, it sacrifices the original representation.
- Folding operation: 2D to 3D mapping.
- Starts with random folding.
- ShapeNet dataset and uses only 7% parameters of a FC decoder.

RECURRENT NEURAL NETWORK (RNN)

- Limited research on usage of RNN for PCC.
- Focused on data from 3D LiDAR sensors (Driving data from 11 areas of Japan)
- Challenging data to compress and hence works on raw packets.
- Raw packets to 2D matrices losslessly.
- Uses residual blocks for accurate decompression.



CONCLUSION

COMPARISON OF MACHINE LEARNING BASED METHODS FOR PCC.

	Feature	Technique	Dataset	Benchmark	Metrics	Loss Function	Optimizer
CNN	Geometry	CNN-based AE	MPEG	PCL	RD curve, D1	SGD (BCE)	Adam
CNN	Geometry	CNN-based AE	MPEG	MPEG GPCC	RD curve, D1	RD Loss	-
CNN	Geometry	3D CNN AE	Train: ModelNet40 Test: MVUB	MPEG GPCC	D1, D2	RD Loss	Adam
CNN	Geometry	3D Stacked CNN	Train: ShapeNet Test :MPEG, JPEG	MPEG GPCC, PCL MPEG VPCC	D1, D2	RD Loss WBCE	-
FCNN	Geometry	FCNN, MLPs	ShapeNet	MPEG GPCC	D1	RD Loss Chamfer Dist	Adam
FCNN	Geometry	Folding-based NN	ShapeNet ModelNet10	Fully connected decoder	-	Chamfer Dist	Adam
RNN	Geometry	RNN	TierIV	Octree, JPEG	SNNRMSE	-	-

- An emerging research area of PCC using ML/DL.
- Most existing work focuses on geometry compression.
- Most common choice is CNN-based AE with few FCNN and even fewer RNN.

THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Thank You!

Questions?