CLAIMS

1. A method for decoding point cloud data, comprising:

decoding, from at least a bitstream, a codeword that provides a representation of a point cloud;

decoding, from said at least a bitstream, a grid image that provides refinement information to said representation of said point cloud; and

reconstructing said point cloud based on said decoded codeword and said decoded grid image, using a neural network-based module.

2. The method of claim 1, further comprising:

accessing a component that describes an offset error along normal direction, for a point in said reconstructed point cloud; and

refining said reconstructed point cloud using said component.

3. The method of claim 2, wherein said normal is computed based on a preliminarily reconstructed point cloud.

4. The method of any one of claims 1-3, further comprising:

decoding said component from said at least a bitstream.

5. The method of any one of claims 1-4, wherein said neural network-based module includes at least a first set of layers and a second set of layers, wherein said first set of layers is responsive to said decoded codeword and said decoded grid image, and wherein said second set of layers is responsive to output of said first set of layers and said decoded codeword.

6. The method of claim 5, wherein said first set of layers corresponds to a first set of shared Multi-Layer Perceptrons (MLPs) and said second set of layers corresponds to a second set of shared MLPs.

7. The method of any one of claims 1-6, wherein said codeword is a feature vector representing said point cloud in a latent space.

8. The method of any one of claims 1-7, wherein said grid image includes a set of pixels, each pixel in said grid image includes two channels indicating a 2D position.

9. The method of any one of claims 1-8, wherein said pixel in said grid image further includes a third channel indicating an offset along a normal direction.

10. The method of any one of claims 1-9, wherein said codeword is decoded by a decoder associated with a variational autoencoder.

11. The method of any one of claims 1-10, wherein said grid image is decoded by a decoder associated with a variational autoencoder, or a 2D video decoder.

12. A method for encoding point cloud data, comprising:

generating a codeword, by a first neural network-based module, that provides a representation of an input point cloud associated with said point cloud data;

reconstructing a first point cloud, by a second neural network-based module, based on said codeword and a grid image;

adjusting said grid image to generate another grid image, by a third neural network-based module, based on said first reconstructed point cloud, said codeword, and said input point cloud; and

compressing said codeword;

compressing said another grid image; and

generating at least a bitstream including said compressed codeword and said compressed another grid image, wherein said compressed another grid image is to be used to refine said representation of said input point cloud during decoding.

13. The method of claim 12, further comprising:

identifying a nearest neighboring point, for a point in said reconstructed first point cloud, from said input point cloud;

obtaining a normal vector for said nearest neighboring point in said reconstructed first point cloud;

obtaining an error vector from said point to said nearest neighboring point; and

projecting said error vector along said normal vector to form a projected error vector for said point.

14. The method of claim 13, further comprising:

compressing a length of said projected error vector for said point in said reconstructed point cloud.

15. The method of any one of claims 12-14, wherein said second neural network-based module includes at least a first set of layers and a second set of layers, wherein said first set of layers is responsive to said decoded codeword and said decoded grid image, and wherein said second set of layers is responsive to output of said first set of layers and said decoded codeword.

16. The method of claim 15, wherein said first set of layers corresponds to a first set of shared MLPs and said second set of layers corresponds to a second set of shared MLPs.

17. The method of any one of claims 12-16, wherein said codeword is a feature vector representing said point cloud in a latent space.

18. The method of any one of claims 12-17, wherein said grid image includes a set of pixels, each pixel in said grid image includes two channels indicating a 2D position.

19. The method of any one of claims 12-18, wherein said pixel in said grid image further includes a third channel indicating an offset along a normal direction.

20. The method of any one of claims 12-19, wherein said codeword is encoded by an encoder associated with a variational autoencoder.

21. The method of any one of claims 12-20, wherein said grid image is encoded by an encoder associated with a variation autoencoder, or a 2D video encoder.

22. The method of any one of claims 12-21, further comprising:

obtaining, by said third neural network-based module, for each point P' in said first reconstructed point cloud, a corresponding point P_0 in said input point cloud; and

obtaining an error between P' and P_0.

23. The method of claim 22, wherein said third neural network-based module includes at least a third set of layers and a fourth set of layers, wherein said third set of layers is responsive to said decoded codeword, said grid image, and said error, and wherein said fourth set of layers is responsive to output of said third set of layers, said decoded codeword, said grid image, and said error.

24. The method of claim 23, wherein said third set of layers corresponds to a third set of shared MLPs and said fourth set of layers corresponds to a fourth set of shared MLPs.

25. The method of claim 23 or 24, wherein said grid image is adjusted responsive to output of said fourth set of layers to form said another grid image.

26. An apparatus, comprising one or more processors and at least one memory coupled to said one or more processors, wherein said one or more processors are configured to perform the method of any of claims 1-25.

27. A signal comprising a bitstream, formed by performing the method of any one of claims 12-25.

28. A computer readable storage medium having stored thereon instructions for encoding or decoding a point cloud according to the method of any one of claims 1-25.