What's new in Point Cloud Compression?

Chao CAO, Marius PREDA and Titus ZAHARIA

ARTEMIS, Télécom SudParis, Institut Mines-Télécom, Institut Polytechnique de Paris, France

*Corresponding author: Chao CAO, ARTEMIS, Télécom SudParis, Institut Mines-Télécom, Institut Polytechnique de Paris, France.

Abstract

3D point cloud is a simple data structure representing both static and dynamic 3D objects. Though providing close-to-reality visualization, point clouds with high density raise massive demand for storage while sparse point clouds used for navigation require high precision. The emergency of compression technologies is mandatory in the form of standardization. A walkthrough for the technical approaches exploited in the two frameworks (V-PCC and G-PCC) under the MPEG standardization process is presented. Though with certain limitations, the achieved promising compression performances indicate a foreseeable evolution of the standard and a bright future for point cloud compression technologies.

Keywords: Point cloud, Compression, MPEG, V-PCC, G-PCC

Introduction

With the advancement of 3D data acquisition technologies, it is becoming easier to reconstruct point clouds from real-world objects. A 3D point cloud is defined as a set of 3D points, where each point is described by its Cartesian coordinates and various attributes such as colors or normal as shown in Figure 1.
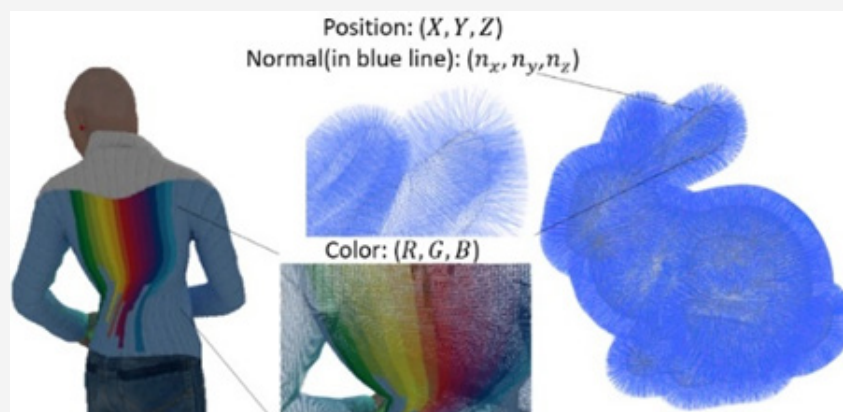

Figure 1: Example of single frame point cloud (left) from dynamic dataset Queen produced by Technicolor and Bunny (right) from Stanford dataset with normal information denoted with blue lines.

A Point Cloud is considered a relatively simple 3D representation allowing a realistic visualization and can be captured by various setups. For example, Microsoft's Kinect and Apple's PrimeSense are now being used in many interactive mobile applications to capture 3D scenes and models. Light Detection And Ranging (LiDAR) is another well-known technology for acquiring point clouds. The

realism and precision of the captured objects are obtained with the cost of huge amount of points to be store/transmitted. In order to make Point Cloud useful in applications, compression is required.

The development of compression technologies for point cloud has gained increasing attention lately in the research community to make them suitable for real-time, portable applications such as autonomous navigation [1] and Virtual and Augmented Reality [2]. A "Universe Map" for point cloud compression techniques is proposed in [3] and illustrated in Figure 2.
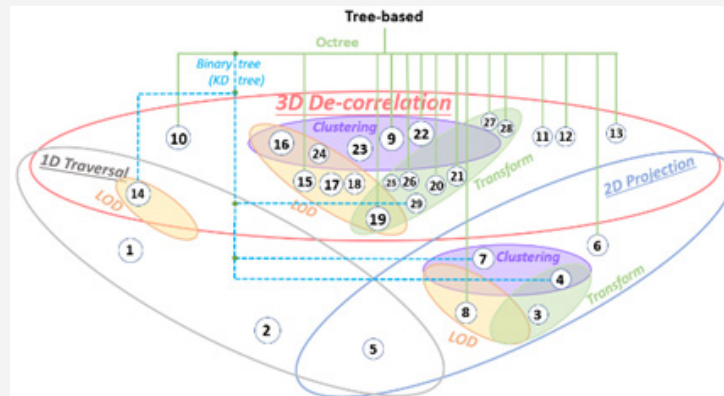

Figure 2: "Universe Map" for 3D point cloud compression methods. (The numbers refer to the paper cited in [3]).

The diversity of the applications when Point Clouds can be used requires a common representation even in the compressed domain to ensure the interoperability. Recognizing this gap, in 2014, the Moving Picture Experts Group (MPEG) initiated an exploration activity on Point Cloud Compression (PCC) [2]. As a first result, a call for proposals for PCC was issued in 2017 and 13 technical solutions were analyzed in October 2017. As an outcome, three different technologies were chosen as test models for three different types of content to explore compression technologies according to the different requirements of different 3D applications [4]. Later, two of them were merged, resulting in two standards Video-based Point Cloud Compression (V-PCC) for dense content and Geometry-based Point Cloud Compression (G-PCC) for sparse content.

V-PCC is planned for promotion as International Standard in July 2020, G-PCC in October 2020. V-PCC, as a toolbox leveraging the existing and future video compression technologies, aims at providing low complexity decoding capabilities. The current reference test model encoder shows compression rates of 125:1 while achieving very good perceptual quality [5]. The codec is further improved with minor refinements along the standardization process. However, it is expected as the first commercial product to be significantly better than the test model. Meanwhile, G-PCC is considered to provide efficient lossless and lossy compression for the deployment of autonomous driving, 3D maps, and other applications that exploit LiDAR generated and other sparse point clouds.

Overview of V-PCC

The compression process of V-PCC is shown in Figure 3. It is considered as a video-based method due to the decomposition of the 3D dynamic point clouds into a set of 2D patches, packetized into 2D canvas and encoded with traditional video codecs.
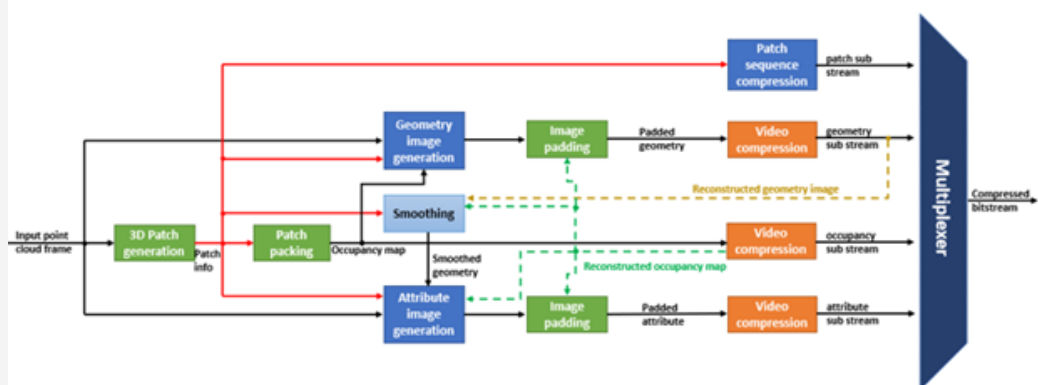

Figure 3: The V-PCC encoder diagram.

The test model implementation of V-PCC made available by MPEG decomposes the input point cloud with the help of a local planar surface approximation step. Let us note that such decomposition is not standard, any more advanced decomposition

can be used. In the test model, a normal per point is estimated using Principal Components Analysis. The points are clustered to form patches based on the closeness of the normal and the distance between points in Euclidian space. Note that kd-tree is used to separate the point cloud and find neighbors. The resulting patches are mapped from 3D to 2D grids using orthographic projection to pre-defined ten planes. A refinement step is executed to minimize the number of patches.
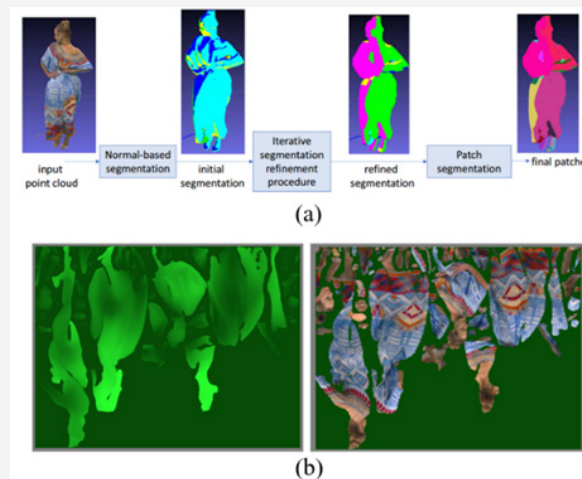

Figure 4: (a) Overview of the patch generation process. (b) Example of geometry image (left) and texture image (right) [6].

The projected patches are grouped, and they compose two types of images: one for the geometry, encoding a distance between each point and the corresponding projection plane, and another one for the attribute containing the color of each 3D point. The process of patch generation and the generated geometry and texture image after packing are illustrated in Figure 4.

A binary image, illustrating whether the pixels on the geometry and texture image are occupied is generated as well. It is used to help point cloud reconstruction and smoothing. Different packing strategies are defined in V-PCC to minimize the empty space of the image. Generally, the patches are padded in descending order of patch size. Additionally, different from the spatial packing methods, a strategy prioritizing temporal consistency of the patches can be selected to maintain the 2D packing location of the patches between frames.

Each of the three channels (geometry, color and occupancy) is encoded by existing video codec. A V-PCC bitstream is constructed with these video sub-streams and metadata are further encoded.

Two modes, All-Intra and Random-Access coding are enabled in V-PCC. Both lossy and lossless configuration are defined for different network conditions. It is shown in [6] that depending on the tested point cloud sequences, compression factors between 100:1 to 500:1 can be achieved with good to acceptable visual quality. This approach is considered as the state-of-the-art method for compressing dynamic point clouds with dense distribution.
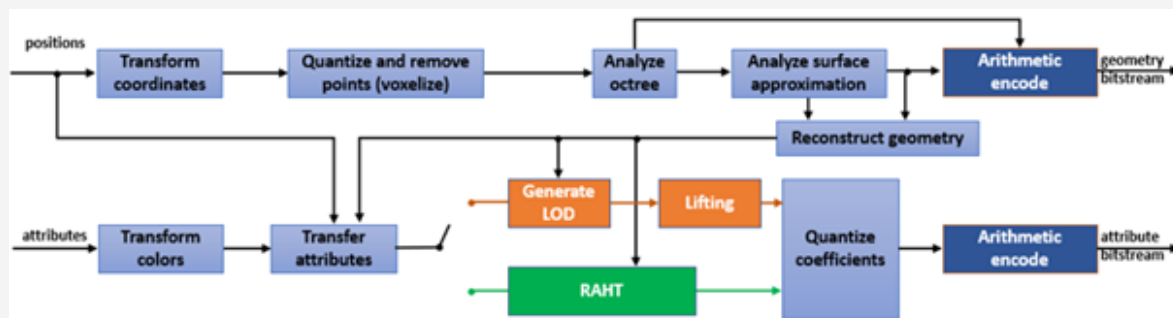
Overview of G-PCC

Figure 5: The G-PCC encoder diagram.

The G-PCC encoder diagram is shown in Figure 5. G-PCC framework integrates various methods to exploit the spatial redundancy of point clouds. The geometry positions are first coded as an octree. The attributes are coded depending on the decoded geometry.

A pre-processing step is executed to transform the positions and colors into a more compression-friendly representation. (e.g., integer values and YUV color space respectively).

Different approaches are utilized for geometry coding after the octree decomposition. In one method, a surface-based approximation method uses triangulation process to estimate the voxels on the surface within each octree leaf. Meanwhile, the octree geometry codec is set with different configurations to provide a more predictive octree using occupancy.

An attribute transfer procedure is launched to minimize the attribute distortions between the given input point cloud and the reconstructed point cloud. Then the coding method is selected between the LOD-based approach and transform-based approach depending on the property of the input point cloud. A Lifting Transform is built on top of the Predicting Transform after the LOD generation. It defines an update operator to recursively reduce the value of prediction residuals. In another mode, the colors are spatially transformed with RAHT (Region-Adaptive Hierarchical Transform) introduced in [7] and further quantized and entropy coded.

G-PCC, comparing to V-PCC, provides an approach exploiting the more native 3D representation of a point cloud and potential improvements are still possible. For more sparse distribution, G-PCC is more appropriate and decent performance is observed in [6].

Conclusion

In this paper, a brief review of point cloud compression is provided together with an overview of the V-PCC and G-PCC framework, two recent standards developed by MPEG. With the increasing popularity of augmented and virtual reality applications and the usage of 3D point cloud representation, the importance of developing efficient compression technologies is unquestionable. The technical aspects of the ongoing MPEG standardization are described under the overview of the V-PCC and G-PCC framework.

It is shown that significant compression performance has been achieved for both static and dynamic point clouds with different density. However, limitations in exploiting temporal correlations still exist for V-PCC and subjective tests for 3D visualization may be investigated in the future. Meanwhile, V-PCC will evolve with the advancement of video coding technologies while G-PCC has great potential to be improved and used in various fields.

Acknowledgement

Conflict of Interest

No conflict of interest.

References

1. X Chen, H Ma, J Wan, B Li, T Xia (2017) 30th IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2017, US.

2. K Guo (2019) The relightables: Volumetric performance capture of humans with realistic relighting. ACM Trans Graph 38(6):1-19.

3. C Cao, M Preda, T Zaharia (2019) 24th International ACM Conference on 3D Web Technology. Web3D 2019, US.

4. L Cui, R Mekuria, M Preda, ES Jang (2019) Point-Cloud Compression: Moving Picture Experts Group's New Standard in 2020 IEEE Consum Electron Mag 8(4): 17-21.

5. 3DG (2019) Liaison on MPEG-I Point Cloud Compression.

6. S Schwarz (2019) Emerging MPEG Standards for Point Cloud Compression IEEE J Emerg Sel Top Circuits Syst 9(1):133-148.

7. RL De Queiroz, PA Chou (2016) Compression of 3D Point Clouds Using a Region-Adaptive Hierarchical Transform IEEE Trans. Image Process 25(8): 3947-3956.