Characterizing the Geometric Complexity of G-PCC Compressed Point Clouds

Annalisa Gallina, Hadi Amirpour, Sara Baldoni, Giuseppe Valenzise, Federica Battisti

Characterizing the Geometric Complexity of G-PCC Compressed Point Clouds

Annalisa Gallina[1], Hadi Amirpour[2], Sara Baldoni[1], Giuseppe Valenzise[3], and Federica Battisti[1]

[1] Department of Information Engineering, University of Padova, Italy
[2] Christian Doppler Laboratory ATHENA, Alpen-Adria-Universitat, Klagenfurt, Austria
[3] Université Paris-Saclay, CNRS, CentraleSupélec, L2S (UMR 8506), Gif-sur-Yvette, France

***Abstract*—Measuring the complexity of visual content is crucial in various applications, such as selecting sources to test processing algorithms, designing subjective studies, and efficiently determining the appropriate encoding parameters and bandwidth allocation for streaming. While spatial and temporal complexity measures exist for 2D videos, a geometric complexity measure for 3D content is still lacking. In this paper, we present the first study to characterize the geometric complexity of 3D point clouds. Inspired by existing complexity measures, we propose several compression-based definitions of geometric complexity derived from the rate-distortion curves obtained by compressing a dataset of point clouds using G-PCC. Additionally, we introduce density-based and geometry-based descriptors to predict complexity. Our initial results show that even simple density measures can accurately predict the geometric complexity of point clouds.**

***Index Terms*—Point Cloud, complexity, compression, G-PCC.**

I. Introduction

With the unprecedented diffusion of Extended Reality (XR) technologies and the availability of many market-ready Head Mounted Displays (HMDs), the transition from 2D to 3D visual content is underway [1]. The XR technology may represent the next major paradigm shift in telecommunications [2]. Consequently, volumetric video, particularly in the form of Point Clouds (PCs), is expected to be the next generation of video technology [3]. The live streaming of 3D content is a key enabler for shared immersive experiences, which are a core component of the metaverse concept [4]. The possibility of live streaming remote environments, dynamic objects and people opens up unlimited applications, including remote collaboration, training and learning, entertainment, and remote monitoring [5].

The real-time use of this new generation of video typically requires the acquisition, encoding, transmission, decoding, and rendering of PCs [6]. According to [2], XR network requirements will be highly dependent on the application. In the short term, conversational Augmented Reality (AR) is foreseen to be one of the main applications that require the transmission of volumetric data stream between source and receiver. Medium- and long-term scenarios include more complex conditions in which the AR objects should adapt dynamically to physical world changes, and user interactivity should be supported. The overall required bitrate will depend on the content complexity, resolution, and frame rate [2]. Based on this, the possibility of determining appropriate compression levels and bandwidth allocation for 3D visual content is becoming of paramount importance.

For 2D content, the possibility of evaluating video complexity is considered a crucial step for optimizing adaptive streaming [7], [8]. Traditionally, spatial information (SI) and temporal information (TI) are used to represent the spatial and temporal complexity of video sequences using the Sobel filter. Specifically, in the ITU-T P.910 recommendations [9], SI and TI features are used to select test scenes for subjective testing. Additionally, new open-source tools called Video Complexity Analyzer (VCA) [10] and Enhanced Video Complexity Analyzer (EVCA) [11] have been developed. These tools measure spatial and temporal complexity features through an energy function based on the Discrete Cosine Transform (DCT), building on methodologies from [12]. Furthermore, a novel supervised feature extraction method called DeepVCA [13] has been introduced, leveraging deep neural networks to assess the spatial and temporal complexity of video sequences. In this method, the number of encoding bits required for each frame in intra-mode and inter-mode serves as labels for spatial and temporal complexity, respectively.

However, to the best of our knowledge, similar metrics for 3D visual content are currently lacking. This work aims to propose and study complexity measures for PCs, with a specific focus on PC compression applications. Specifically, we consider the recently standardized Geometry-based Point Cloud Compression (G-PCC) codec [14]. G-PCC allows for direct encoding of content in 3D space and separates the coding of geometry (*i.e.,* point positions) from the coding of attributes (*i.e.,* color information).

As a proof of concept, this work focuses only on geometry. Specifically, we compress a dataset of PCs at different bitrates and compute the corresponding Rate-Distortion (RD) curves. From this, we derive a ground-truth measure of geometric complexity. Subsequently, we investigate how various geometric descriptors could serve as predictors of this ground-truth geometric complexity. Our preliminary experiments offer insights into which features of PCs most significantly impact geometric complexity, paving the way for the design of an effective and low-cost measure of geometric complexity.

The rest of the paper is organized as follows: Section II describes the generation of the dataset of compressed PCs used to define complexity measures; Section III describes

three definitions of ground-truth PC complexity; Section IV presents statistical and geometric measures to predict the ground-truth geometric complexity; Section V evaluates the proposed definitions of complexity, as well as the accuracy of complexity predictors; finally, Section VI concludes the paper.

II. Dataset preparation

We employ the G-PCC codec v23.0 to encode PCs at different bitrates, using the Octree coding scheme. This method is based on a recursive subdivision of the three-dimensional space into octants. In particular, each subset is further divided into eight parts only if there is at least one point inside it; otherwise, it becomes a leaf of the recursion tree. Then, the geometry of each node of the tree containing at least one point is encoded with a byte, where each bit represents the occupancy of a child node. Once the octree is created, G-PCC uses context-based arithmetic coding to efficiently entropy code the octree symbols. For the purpose of this study, we do not employ other coding modes such as trisoup, which will be considered in future work.

The PCs used in this paper are from the BASICS dataset [15]. This dataset contains 75 pristine PCs categorized into three semantic groups: (*i*) Humans & Animals, (*ii*) Inanimate Objects, and (*iii*) Buildings & Landscapes. All PCs are voxelized by 10-bit quantization prior to coding. After compressing all the PCs using G-PCC, RD curves are computed to evaluate the compression results. The distortion introduced by the codec is assessed using the PCC quality measurement software version 0.14.1 provided by MPEG [16]. This software computes both point-to-point (D1) and point-to-plane (D2) distances, allowing the estimation of the overall distortion based on the Mean Squared Error (MSE) or the corresponding Peak Signal to Noise Ratio (PSNR) value. In this analysis, PSNR related to point-to-point distance is considered. The rate is measured by calculating the average number of bits per input point, with the total bitstream size derived from the output of the encoder.

III. Definitions of ground-truth complexity

In the literature, there are multiple definitions of the coding complexity of a signal. In classical RD theory, coding complexity characterizes the RD function of a given source and depends solely on the source's statistics [17]. A more general definition is related to the concept of Kolmogorov complexity, which is the number of bits required to describe a source. Based on this concept, [18] defines the ground truth complexity for 2D images using the compression ratio, which is the ratio of the output size to the input size at a specific quality level. In this work, we consider three approaches to define the ground-truth geometric complexity of PCs, which are detailed in the following.

A. Point-based definition

We consider an approach similar to [18], *i.e.,* we evaluate the bitrate required by the G-PCC codec for a given quantization step. In G-PCC there are six configurations corresponding to different quality levels. In this work, we consider two bitrate values, $BPP_{0.125}$ and $BPP_{0.9375}$, which represent the bits per point when using scaling factors of 0.125 and 0.9375, respectively. In G-PCC, these scaling factors correspond to quantizing the point coordinates with a step size that is inversely proportional to the scaling factor. The selected factors correspond to the minimum and maximum quality levels available in G-PCC. The main limitation of this point-based approach is that it only considers *single operational points* in the RD plane and does not account for compression complexity across different quality levels.

B. Area-based definition with content-specific interval

To overcome the limitations of the previous approach and measure the *average* complexity across different bitrates, we measure the area under the RD curve for a given PC, taking inspiration from the Bjontegaard delta metrics. Specifically, two measures are considered, depending on whether the area is calculated with respect to the rate (I_{RATE}) or the distortion (I_{PSNR}). I_{RATE} and I_{PSNR} are computed on 6 points according to the GPCC standard configuration files. It is important to note that this measure still depends on the specific codec configuration used to produce the curves. Despite this dependency, the metric enables the assessment of performance across different quality levels, making it a more general measure of complexity.

C. Area-based definition with uniform interval

The integration interval for the area-based measure described above might vary significantly with contents. To overcome this variability, we propose to integrate over the *intersection* of the supports of the curves for different contents, avoiding extrapolation outside defined regions. RD curves are interpolated using Piecewise Cubic Hermite Polynomials for reliability.

The resulting measure, I_{INTERS} (for simplicity, computed using the PSNR as the integration axis), depends on the contents of the dataset through the integration interval. For large and diverse datasets, the intersection of all supports might be empty, thus special care should be considered when computing this measure. We assess and compare these definitions of ground-truth geometric complexity in Section V.

IV. Geometry complexity predictors

In this section, we examine density-based and geometry-based features of a point cloud as potential predictors of its ground-truth geometric complexity.

A. Density-based predictors

Local density plays a key role in PC coding. Density is defined as the number of points per unit volume. In denser regions of a PC, points are closer to each other. This intuitively indicates a stronger spatial dependency among points, which typically facilitates the coding task.

Based on this observation, we propose three density-based predictors of coding complexity:

- Average nearest neighbor distance (AVG_1): to compute this metric, we evaluate the Euclidean distance between a reference point in the PC and its closest neighbor. This procedure is repeated for all points (or, to reduce complexity, on a random subset of points) in the PC. The average of these distances is then computed as the predictor.
- Average k-nearest neighbor distance (AVG_k): this measure generalizes AVG_1 for $k > 1$, since AVG_1 becomes meaningless if most points have the same minimum distance, which can occur in voxelized datasets. Additionally, relying on a single point may not accurately reflect the density of the PC.
- Standard deviation of the k-nearest neighbor distances (STD_k): a high standard deviation indicates a more irregular distribution, suggesting greater difficulty in encoding. For instance, in the case of the described encoder, which uses a constant scale factor and a fixed number of octree subdivisions, a significant standard deviation implies non-uniform quality and potentially increased final distortion.

B. Geometry-based predictors

While density-based descriptors simply measure the occupation of 3D space, geometry-based descriptors provide information about the local structures and surfaces in the PC. To capture these properties, in this paper we consider the local curvature as a geometric feature.

Curvature measures how much the content deviates from being locally planar at each point. Typically, high curvature values indicate significant geometric variability, such as edges, corners, or transition regions. Conversely, lower curvature values generally correspond to flatter areas or simple surfaces. In the literature, curvature is employed in various metrics to assess the quality of 3D content. For example, in [19], a metric based on the computation of mean curvature is proposed to evaluate the quality of a compressed PC by comparing it to the original content. Curvature is an effective indicator for defining the difficulty of compressing content, making it a strong candidate for predicting the geometric complexity of a PC. Specifically, if the content is characterized by a very irregular surface with high curvature values, it is likely that the losses introduced by compression will be significant.

To estimate the curvature at a point p, we first compute the principal curvatures using Principal Component Analysis (PCA). With the two principal curvatures, κ_{min} and κ_{max}, we can then calculate the mean curvature, defined as:

$$\kappa_m = \frac{\kappa_{\text{min}} + \kappa_{\text{max}}}{2}. \tag{1}$$

Similarly, the Gaussian curvature can be computed as the product of the two principal curvatures:

$$\kappa_g = \kappa_{\text{min}} \cdot \kappa_{\text{max}}. \tag{2}$$

For this analysis, we focus on the Gaussian curvature and obtain a single measure of the PC's complexity K_g by averaging the Gaussian curvature contributions across all points.

V. Experimental validation

In this section, we first compare the different ground-truth geometric complexity definitions proposed in Section III, and then we evaluate the accuracy of complexity predictors. We employ Pearson's Correlation Coefficient as a measure of linear correlation.

A. Correlation among different definitions of complexity

These definitions provide valid alternatives for computing the ground-truth complexity of a PC. However, the possibility of using them interchangeably needs to be investigated. To this aim, at least for the employed dataset, we compute the pairwise correlations between measures, reported in Table I. This analysis suggests that the complexity assessment is relatively independent from the specific method used. Fig. 1 further supports this finding by comparing the proposed complexity measures with the integral computed with respect to distortion (I_{PSNR}) as a reference. The results closely align with the identity function, with only one offset factor changing, indicating that all measures capture nearly the same aspect of complexity.

Based on this observation, we employ I_{PSNR} as the ground-truth complexity measure in the following sections.

B. Accuracy of complexity predictors

1) Density-based predictors: Table II presents the results of the density-based complexity prediction measures, which are based on the distance between neighboring points. The average (AVG) and the standard deviation (STD) are considered for neighbor radii of 1, 5, and 10. The high correlations observed, regardless of the radius, indicate a strong ability of these descriptors to predict geometric complexity, at least with the G-PCC Octree codec. More specifically, we observe that increasing the neighbor radius from 1 to 5 leads to better correlations for both the average and standard deviation. This is intuitive, as the statistics computed on a single point ($k = 1$) are very noisy. However, when the radius grows too much ($k = 10$) the correlations decrease. This is likely because, for some sparser point clouds, the 10 nearest neighbors may be spatially distant from each other, resulting in statistics that fail to capture the true spatial relationships of the points. Finding the optimal number of neighbors depends significantly on the dataset used and should be tuned accordingly.

Although, based on the correlation values reported in Table II, the three radii appear to be promising, a more in-depth analysis reveals that considering only the nearest neighbor yields almost meaningless results. With the exception of a few very sparse contents, the majority of the point clouds in the dataset exhibit an average nearest neighbor distance exactly equal to 1. This is due to the voxelization of the original contents, and it makes this measure inappropriate for predicting the compression complexity.

For the best-performing number of neighbors (*i.e.,* $k = 5$) we provide in Fig. 2 the comparison between the average and the standard deviation of the average neighbors distance. Both metrics show a linear trend with respect to the RD complexity,

TABLE I
CORRELATION AMONG GROUND-TRUTH COMPLEXITY MEASURES.

	I_{PSNR}	I_{RATE}	$BPP_{0.125}$	$BPP_{0.9375}$	I_{INTERS}
I_{PSNR}	1.0000				
I_{RATE}	0.9547	1.0000			
$BPP_{0.125}$	0.9519	0.8183	1.0000		
$BPP_{0.9375}$	0.9714	0.9980	0.8527	1.0000	
I_{INTERS}	0.9998	0.9493	0.9571	0.9671	1.000

TABLE II
CORRELATION AMONG I_{PSNR} AND DIFFERENT DISTANCE METRICS.

Features	AVG_1	STD_1	AVG_5	STD_5	AVG_{10}	STD_{10}
CC	0.978	0.955	0.992	0.973	0.989	0.937

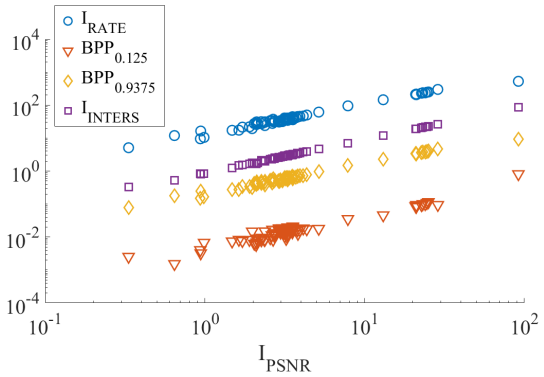

Fig. 1. Comparison among ground-truth complexity measures.

thus proving their effectiveness as predictors. Extending this analysis to additional coding configurations and learning-based codecs would be a valuable next step.

2) Geometry-based predictor: The evaluation of the mean Gaussian curvature reveals a correlation coefficient of only 0.685. This suggests that the measure does not fully capture the overall complexity of the content.

Part of this result depends on how the curvature was obtained. The software CloudCompare was used for this purpose. The procedure, as described in [20], is similar to the one previously outlined but requires fixing the PCA kernel size. If the kernel size is too small, the final metric becomes unreliable because most points cannot approximate the surface locally with too few neighbors, resulting in statistics based on a limited portion of the object. In contrast, a too large kernel size leads to a noisy estimation of the surface, making the final result once again unreliable. When dealing with a large and diverse dataset, it is challenging to define a kernel size that fits all point clouds. In this study, a kernel size of 5 was considered, which seems to be a reasonable trade-off between the needs of sparse and dense objects. However, it remains challenging to find a size that fits all types of point clouds perfectly.

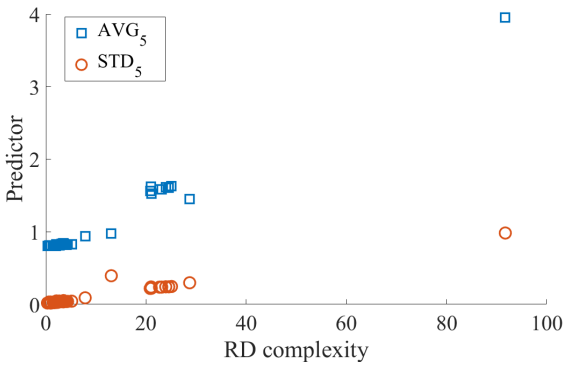

Fig. 2. Comparison between the average and the standard deviation of the average neighbors distance with $k = 5$.

VI. CONCLUSIONS AND FUTURE WORK

While the complexity of 2D videos has been extensively studied, the complexity of 3D content remains less explored. In this paper, we first investigated various definitions of ground-truth geometric complexity for PCs and proposed several predictors for estimating this complexity. We analyzed traditional coding complexity measures, such as bitrate and the area under the RD curve, along with geometry-based metrics, including curvature. Our findings suggest that, despite some variation, these complexity definitions generally align well and capture similar aspects of PC complexity. The density-based predictors, such as the average nearest neighbor distance and its standard deviation, demonstrated a strong correlation with compression-based complexity metrics, when using the G-PCC Octree codec. These predictors proved effective, though the optimal number of neighbors remains dataset-dependent and requires careful tuning. On the other hand, the curvature-based metric, specifically the average Gaussian curvature, showed a lower correlation with ground-truth complexity, indicating that it may not fully capture the overall complexity.

As future work, we will extend our study by incorporating additional codecs, color information, and a variety of datasets to evaluate the generalizability of the proposed complexity measures. This will include exploring higher precision datasets and analyzing trends among semantic groups to better understand content-specific influences. Additionally, we will assess the impact of color on complexity prediction. These steps will strengthen the robustness of our findings across different encoding schemes and data types.

ACKNOWLEDGMENTS

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) Mission 4, Component 2, Investment 1.3, CUP C93C22005250001, partnership on “Telecommunications of the Future” (PE00000001 - program “RESTART”).

REFERENCES

[1] Jeroen Van Der Hooft, Hadi Amirpour, Maria Torres Vega, Yago Sanchez, Raimund Schatz, Thomas Schierl, and Christian Timmerer,

"A Tutorial on Immersive Video Delivery: From Omnidirectional Video to Holography," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1336–1375, 2023.

[2] Fredrik Alriksson, Oskar Drugge, Anders Furuskär, Du Ho Kang, Jonas Kronander, Jose Luis Pradas, and Ying Sun, "Future network requirements for extended reality applications," online, April 2023, available at https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/future-network-requirements-for-xr-apps#:~:text=Our%20analysis%20demonstrates%20that%20XR,supporting%20today's%20mobile%20broadband%20services.

[3] Jie Li, Huiyu Wang, Zhi Liu, Pengyuan Zhou, Xianfu Chen, Qiyue Li, and Richang Hong, "Toward Optimal Real-Time Volumetric Video Streaming: A Rolling Optimization and Deep Reinforcement Learning Based Approach," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 33, no. 12, pp. 7870–7883, 2023.

[4] Leif Van Holland, Patrick Stotko, Stefan Krumpen, Reinhard Klein, and Michael Weinmann, "Efficient 3D Reconstruction, Streaming and Visualization of Static and Dynamic Scene Parts for Multi-client Live-telepresence in Large-scale Environments," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4258–4272.

[5] Giuseppe Valenzise, Martin Alain, Emin Zerman, and Cagri Ozcinar, *Immersive Video Technologies*, Elsevier, Oct. 2022.

[6] Ignacio Reimat, Yanni Mei, Evangelos Alexiou, Jack Jansen, Jie Li, Shishir Subramanyam, Irene Viola, Johan Oomen, and Pablo Cesar, "Mediascape XR: A Cultural Heritage Experience in Social VR," in *Proceedings of the 30th ACM International Conference on Multimedia*, New York, NY, USA, 2022, MM '22, p. 6955–6957, Association for Computing Machinery.

[7] Tao Jin, Mallesham Dasa, Connor Smith, Kittipat Apicharttrisorn, Srinivasan Seshan, and Anthony Rowe, "MeshReduce: Scalable and Bandwidth Efficient 3D Scene Capture," in *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, Orlando, FL, USA, Mar. 2024, pp. 20–30, IEEE.

[8] Chao Yang, Siqian Qin, Ping An, Xinpeng Huang, and Liquan Shen, "Content adaptive spatial–temporal rescaling for video coding optimization," *Expert Systems with Applications*, vol. 254, pp. 124482, Nov. 2024.

[9] ITU-T, "Subjective video quality assessment methods for multimedia applications," https://www.itu.int/rec/T-REC-P.910-201902-I/en, 2019, ITU-T P.910.

[10] Vignesh V Menon, Christian Feldmann, Hadi Amirpour, Mohammad Ghanbari, and Christian Timmerer, "VCA: Video Complexity Analyzer," in *Proceedings of the 13th ACM Multimedia Systems Conference*, June 2022, pp. 259–264.

[11] Hadi Amirpour, Mohammad Ghasempour, and Christian Timmerer, "EVCA: Enhanced Video Complexity Analyzer," in *Proceedings of the 15th ACM Multimedia Systems Conference*, 2024.

[12] Michael King, Zinovi Tauber, and Ze-Nian Li, "A New Energy Function for Segmentation and Compression," in *Multimedia and Expo, 2007 IEEE International Conference on*, Beijing, China, July 2007, pp. 1647–1650, IEEE.

[13] Hadi Amirpour, Klaus Shoeffmann, Mohammad Ghanbari, and Christian Timmerer, "DeepVCA: Deep Video Complexity Analzer," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2024.

[14] ISO/IEC, "23090-9:2023: Information technology — Coded representation of immersive media. Part 9: Geometry-based point cloud compression," Standard, ISO/IEC, Mar. 2023.

[15] Ali Ak, Emin Zerman, Maurice Quach, Aladine Chetouani, Aljosa Smolic, Giuseppe Valenzise, and Patrick Le Callet, "Basics: Broad quality assessment of static point clouds in a compression scenario," *IEEE Transactions on Multimedia*, vol. 26, pp. 6730–6742, 2024.

[16] Minh Nguyen, "MPEG-3DG mpeg-pcc-dmetric metric software," online, 2024, available at https://github.com/minhkstn/mpeg-pcc-dmetric.

[17] Nuggehally S Jayant and Peter Noll, *Digital coding of waveforms: principles and applications to speech and video*, vol. 2, Prentice-Hall Englewood Cliffs, NJ, 1984.

[18] Honghai Yu and Stefan Winkler, "Image complexity and spatial information," in *2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*, 2013, pp. 12–17.

[19] Gabriel Meynet, Julie Digne, and Guillaume Lavoué, "Pc-msdm: A quality metric for 3d point clouds," in *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*, 2019, pp. 1–3.

[20] Zvi Har'el, "Curvature of curves and surfaces – a parabolic approach," 1995.