https://doi.org/10.2352/EI.2022.34.17.3DIA-225

Quality analysis of point cloud coding solutions

João Prazeres, Manuela Pereira, Antonio M. G. Pinheiro
Instituto de Telecomunicações and Universidade da Beira Interior, Covilhã, Portugal

Abstract

In this paper, a subjective quality based comparison between four point clouds codecs is presented. For that, a set of six point clouds was chosen. They were coded with four different point cloud encoding solutions, notably the MPEG V-PCC and G-PCC, a deep learning coding solution RS-DLPCC and also Draco, with different bit rates. A subjective test where the distorted and reference point clouds were rotated in a video sequence side by side followed by the quality evaluation, was conducted. Then the performance of a set of four point cloud objective quality metrics of he quality, was analysed using the subjective quality evaluation results. These metrics are usually reported as providing a good representation and are often used to evaluate compression solutions. In fact, the studied metrics tend to provide a good representation for V-PCC and G-PCC, an acceptable representation for RS-DLPCC, and a bad representation for Draco. It was also concluded that V-PCC is the best codec of the studied ones. The deep learning based solution still performs worst than the two MPEG codecs.

Introduction

In the modern world, 3D data capture and transmission became a common requirement for emerging technologies. Typical 3D information representation leads to huge amounts of data. Therefore, efficient methods of data compression are needed, in order to provide efficient transmission and storage of 3D data. Recently, point cloud technology has emerged as a very popular method for 3D data representation. A point cloud is a set of Cartesian coordinates(x, y, z), with a list of attributes associated to each element, such as a RGB component, reflectance information, physical sensor information or normal vectors. Point clouds contain a large amount of information, allowing an accurate representation of an object or scene. Hence, they are a very powerful visual representation model, extremely useful in VR/AR scenarios [22], computer graphics or 3D computer vision applications, between others.

If an accurate precision point cloud of a city or building, or even of an artefact is created, the resulting file can easily have several millions of points, with several features associated to each point. Since the representation of 3D data can contain a large amount of information, several solutions for point cloud compression have been researched. MPEG provided encoding solutions, notably V-PCC (Video Point Cloud Compression) and G-PCC (Geometry Point Cloud Compression) [7]. Deep Learning technology has been considered for image and video compression. Some works had also considered that possibility for point clouds compression [5, 6, 20, 23].

In this paper, the two MPEG codecs and a deep learning based solution proposed to the JPEG Pleno Point Cloud coding call for evidence, RS-DLPCC, are considered [20]. Furthermore, the DRACO codec [26] that has gain some popularity as a royalty free coding solution for Point Clouds and meshes was also considered.

This paper aims to compare the performance of the four codecs, using subjective and objective quality evaluation models. Several works have been published considering the quality evaluation of point clouds. In [9, 10], geometry only point clouds are considered. Compression artifacts using prior encoding schemes are evaluated in [11–13]. Current efforts account for a wider range of high-performing codecs, such as the ones reported in [7,14,15]. In [7] a quality model for point clouds is established. Apart a subjective evaluation using the MPEG codecs, the work also considers a set of point cloud metrics concluding that point to point and point to plane metrics [16] are the best performing ones and provide a good representation of the subjective evaluation. A subjective quality evaluation test where GPCC, V-PCC, RS-DLPCC and Draco codecs are compared was performed. We believe this is the first study that compares these four coding solutions. Finally, a comparison between the subjective results and objective models is also performed.

Brief Description of the tested codecs

Geometry Point Cloud Compression

G-PCC (Geometry Point Cloud Compression) [3] contains two methods of point cloud compression, an octree based method, and a trisoup based, method. For this study, only the octree method was considered. The octree compression method is regulated in the codec by the positionQuantizationScale (pQS) parameter. This parameter controls the number of divisions of the octree from the root, to each leaf node leading to a regular down sampling of the input clouds. Five different rates were coded for each content, ranging from low to near perfect quality levels, with bitrates ranging from 0.09 to 12 bits per point.

Table 1: G-PCC Parameters Example

Rate	QP	pQS
R01	46	0.125
R02	40	0.25
R03	34	0.5
R04	28	0.75
R05	22	0.875

Video Point Cloud Compression

The V-PCC (Video Point Cloud Compression) [4], presents a solution which projects the point cloud in a set of planes, and then encodes the projections in the 2D domain. Those projections contain texture, depth and an occupancy map, with the textures being encoded with legacy methods and the depth being encoded with 2D video encoding methods. The occupancy map represents the pixels containing meaningful information, and is encoded with spatial quantization, combined with raster scanning and entropy encoding. The image projection sequence is encoded with the HEVC video codec. Five rates were chosen for each of the point clouds in the set, with bitrates ranging from 0.08 to 15.22 bits per point.

Table 2: V-PCC Parameters Example

Rate	Geometry QP	Texture QP	Occupancy Map
R01	36	47	4
R02	32	42	4
R03	28	37	4
R04	20	27	4
R05	16	22	2

Resolution Scalable Deep-Learning Point Cloud Compression (RS-DLPCC)

This codec uses a deep-learning approach to compress point clouds geometry [5], by using a latent representation of a point cloud, computed by an auto encoder framework. The scalability feature is made possible by interlaced block creation. The point cloud is divided into super-blocks, which are further divided by interlaced down sampling, resulting in up to eight interlaced blocks for each super-block, which are then coded separately, then enabling random-access. For each point cloud, four rates were chosen, with bitrates ranging from 0.34 to 25.88 bits per point.

This codec is likely to create some blocking artifacts due to the super blocks division.

After the geometry encoding, the color was transferred from the nearest neighbour of the original point cloud. The color for the recolored points is encoded with G-PCC, using the lossless geometry Octree coding mode, and the Predlift color encoder. The lossless Octree coding mode was chosen so that the (decoded) geometry is not changed, minimising the geometry coding effects on the color information from the G-PCC codec. This color information is then textured over the RS-DLPCC lossy decoded geometry.

Draco

Draco is a popular codec developed by Google. This codec uses KD-Tree [21] in order to efficiently organize the 3D data. Draco continuously splits the point cloud from the center, while also modifying the axes on each direction. Draco comes with four main parameters for controlling point cloud encoding. QP, which define the quantization bits for the position attributes, QT, which defines the quantization bits for the texture coordinate attribute, QN, which defines the quantization bits for the normal vector attribute, and QG, which defines the quantization bits for any generic attribute. Draco contains 32 levels of quantization (0 - 31) and 11 levels of compression. For this test, qp levels of 7, 9 and 10 were considered, which represent low, medium and high quality, respectively with the default compression level of 7. The coded point clouds resulted in bitrates ranging from 8.1 to 28.41 bpp.

Basically, Draco is a lossless codec. The parameter qp was used to control the bit rate, but basically it controls the precision of the representation. Reducing the precision (or somehow the resolution) of the point cloud representation allows the codec to be more efficient, resulting in lower bit rates. This was also the reason to consider 3 bit rates only. The resulting bit rates are much higher as qp does not reduce the number of points. It changes the point locations, reducing the representation precision, resulting at the same time lower bit rates.

Evaluation Methodology

Point Cloud Data Selection

For the comparative study, a set of six point clouds was used, containing geometry and texture information. The set consisted of a frame selected from the soldier and longdress dynamic point clouds available at [1], representing human figures. Frames 1300 and 0690 were selected for the longdress and soldier, respectively. Furthermore the point clouds rhetorician, guanyin, from EPFL dataset and point clouds romanoillamp, bumbameuboi, available at [2] were also selected. The later four point clouds represent cultural heritage artifacts. The selected point clouds are represented in figure 1. The full body point cloud redandblack (frame 1550) [1] was used for training prior to the subjective evaluation. The set was coded using V-PCC, G-PCC, RS-DLPCC and Draco with different bit rates.

Subjective evaluation

For all point clouds, a complete rotation over the vertical axis was applied. At each degree an image representing the point cloud view was extracted. These images were extracted using PCL Visualizer. The point cloud views were rendered as 12 second videos, and were displayed at 30fps and with 1920x1080 resolution. Videos were created using the FFMPEG software using a no video compression mode. To ensure no compression was applied to the extracted frames, the stream copy option in FFMPEG was used [8].

In some cases, the point size was changed to provide an improved visual representation. If holes appear in the point cloud the viewers will see the opposite part of the point cloud and that creates a very bad quality perception [10,12]. The change of point size is important to avoid this effect and to create continuous surfaces for the point cloud under observation. The point size values are represented in table 3 for each content and were obtained for the display used in this subjective test. The point clouds bumbameuboi and romanoillamp require a different solution from the remaining content. For point clouds coded with V-PCC, a modification was not required, so the default value of 1 was set. For all point clouds coded with G-PCC, the point size was set to 6 for R01 and 4 for R02. All the other rates were set to the default value. For the remaining cases the options are described in table 3 C

(a) Soldier (b) Guanyin (c) Longdress (d) Rhetorician (e) Romanoillamp (f) Bumbameuboi

Figure 1: Point Cloud testing set.

Table 3: Point size for each content

V-PCC					
Content	R01	R02	R03	R04	R05
bumbameuboi	4	4	4	4	4
Guanyin	1	1	1	1	1
Longdress	1	1	1	1	1
Rhetorician	1	1	1	1	1
Romanoillamp	2	2	2	2	2
Soldier	1	1	1	1	1
G-PCC					
Content	R01	R02	R03	R04	R05
bumbameuboi	6	4	4	4	4
Guanyin	6	4	1	1	1
Longdress	6	4	1	1	1
Rhetorician	6	4	1	1	1
Romanoillamp	3	2	2	2	2
Soldier	6	4	1	1	1
RS-DLPCC					
Content	R01	R02	R03	R04	R05
bumbameuboi	-	20	9	8	7
Guanyin	-	6	4	1	1
Longdress	-	6	4	1	1
Rhetorician	-	6	4	1	1
Romanoillamp	-	7	3	2	2
Soldier	-	6	5	1	1
Draco					
Content	R01	R02	R03	R04	R05
bumbameuboi	6	-	4	-	4
Guanyin	6	-	2	-	1
Longdress	6	-	2	-	1
Rhetorician	6	-	2		1
Romanoillamp	6	-	2	-	2
Soldier	6	-	2	-	1

For the test, a Double Stimulus Impairment Scale was used. In this method, both the reference and the coded point cloud are shown to the subject. Then the subject is asked to evaluate each point cloud pair difference in a five-level rating scale (1 - very annoying, 2 - slightly annoying, 3 - annoying, 4 - perceptible, but not annoying, 5 - imperceptible). Prior to the evaluation, a sequence of four videos was shown to the subjects to help familiarizing with the evaluation. The redandblack point cloud was selected with four different levels of degradation. This point cloud was not included in the final test sequence. Additionally, hidden reference-reference pairs were included in the test sequence, to help verifying unusual behaviour in the evaluation. The same content was never shown twice in a row. To avoid biases, half the subjects were shown videos with the reference on the right and the codec content on the left, and vice-versa. All the tests were, conducted at subjective test laboratory of Image and Video Technology Group of Universidade da Beira Interior, using a 47 inch, FULL HD LG 47LA860V, with the test environment following the specifications in [24].

Six different point clouds were selected for the subjective quality evaluation of this test, based on the experience of JPEG and MPEG evaluation test sets. Both V-PCC and G-PCC codecs had five quality levels, while RS-DLPCC had four quality levels, and Draco had three quality levels. Taking the references in account, a total of 108 scores were obtained in each session.

Table 4: Subject Information

Males	Females	Overall	Age Span	Average age
10	6	16	21-33	24.75

Subjective Evaluation Scores

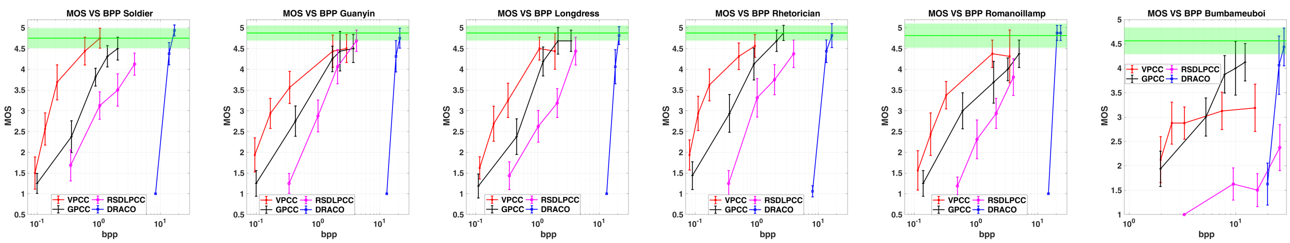
After the test, all the scores were aggregated, and the MOS for each content was computed. The bitrate, measured in bits per point (bpp), is calculated by taking the number of bits of a particular content, and dividing it by the number of points of the original content.

The MOS results are represented in figure 2. These figures also represent the Confidence Interval (considering a Gaussian distribution) The green line and the green horizontal bar represent the MOS and respective confidence interval obtained for the original point cloud (that was also in the test as hidden reference). This bar can be seen as a representation where transparent quality is reached. Although this is a simplistic approach, it is somehow representative that a given codec reaches unperceived distortions in case the MOS is inside this green horizontal bar. Moreover, can be observed: 1) The V-PCC in general provides the best quality scores, followed by G-PCC and the RS-DLPCC. Draco is the worst case, leading to much higher bit rates. The point cloud bumbameuboi is the exception to this regular behavior. This happens because this point cloud is rather sparse when compared with the others. This also reveals that further studies will be required in the future for sparse point clouds, that tend to be created with some acquisition technologies, like LIDARS.

Objective Evaluation

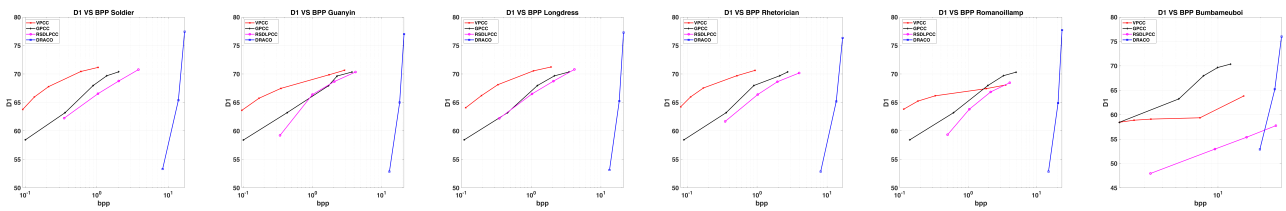
Four objective metrics were calculated:

- Point-to-point: This metric calculates the geometric distance of associated points between the reference and of the



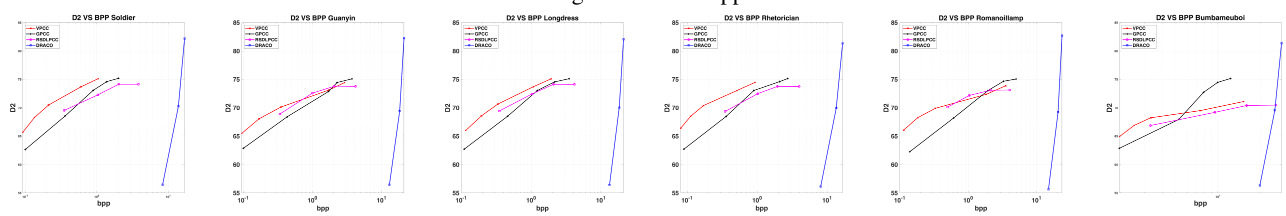
(a) Soldier (b) Guanyin (c) Longdress (d) Rhetorician (e) Romanoillamp (f) Bumbameuboi

Figure 2: MOS with Confidence intervals (assuming a Gaussian distributiion) vs bpp



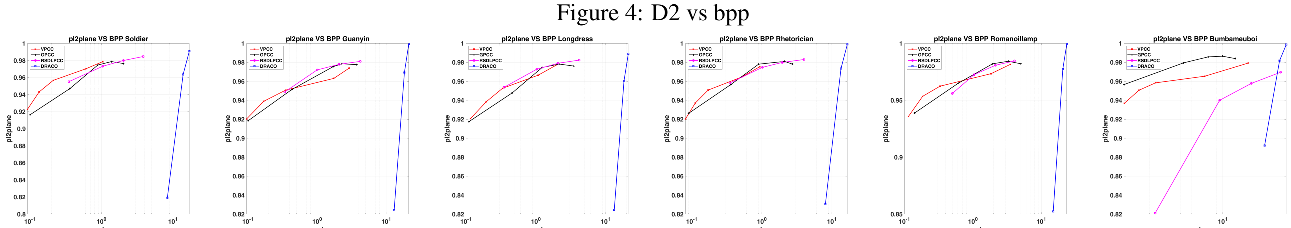
(a) Soldier (b) Guanyin (c) Longdress (d) Rhetorician (e) Romanoillamp (f) Bumbameuboi

Figure 3: D1 vs bpp



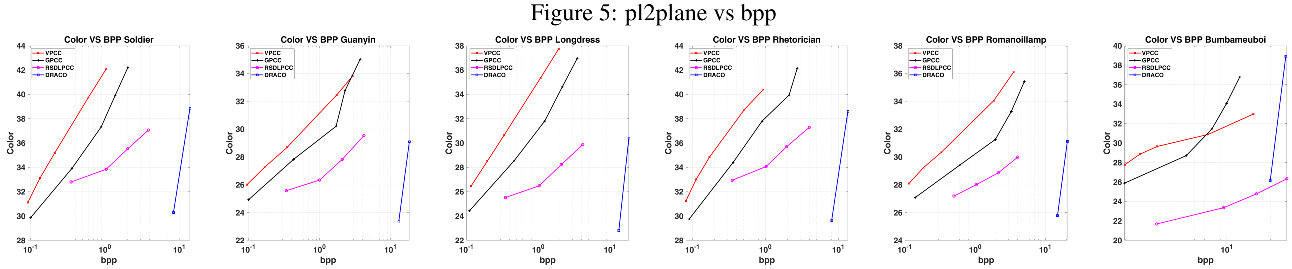
(a) Soldier (b) Guanyin (c) Longdress (d) Rhetorician (e) Romanoillamp (f) Bumbameuboi

Figure 4: D2 vs bpp



(a) Soldier (b) Guanyin (c) Longdress (d) Rhetorician (e) Romanoillamp (f) Bumbameuboi

Figure 5: pl2plane vs bpp



(a) Soldier (b) Guanyin (c) Longdress (d) Rhetorician (e) Romanoillamp (f) Bumbameuboi

Figure 6: Color vs bpp

content under evaluation. Through nearest neighbour algorithm, the corresponding point belonging to the reference point cloud of each point of the distorted point cloud is found. Individual errors are computed based on Euclidean distance [16], followed by the aggregation mechanism.

- Point-to-plane: For every point of the content under evaluation, a point in the reference cloud is identified, through the nearest neighbour algorithm. A plane is fitted into the region centered on that point that is normal to the point under consideration. This plane is computed using quadric fitting in CloudCompare [18], with a radius of 20. The individual error of the point of the content under evaluation is the dimension of the normal vector to the plane and ends on that point. The final metric is the aggregation of these individual errors [16].
- plane-to-plane: For each point in the coded content, a point is identified using the nearest neighbour algorithm. Afterwards, considering the normal vectors to the planes for the reference and coded point cloud, the angular similarity is calculated. This is computed for each point [17]. This metric requires the planes normal to the vectors of both point clouds. The planes were computed using quadric fitting in cloud compare, with a radius of 20 [18].
- color: For every point in the codec point cloud, a point is identified, belonging to the reference cloud, through nearest neighbour algorithm. An individual error is computed based on Euclidean distance, and for color attributes, the MSE is calculated for the three color components, with a RGB to YCbCr conversion being made [25].

The plots of these metrics versus bit rate are represented respectively in figures 3, 4, 5 and 6. The highest bit rate of Draco resulted in infinite values for 6 and could not be represented.

From plots 3 to 5 we can observe that the lower performance found in the subjective evaluation (2) for the deep learning solution RS-DLPCC is not visually observed as it exhibits a very close performance to the G-PCC. While the metrics studied in this work provide a good representation of the subjective quality for the MPEG codecs, they did not reveal the same representation for the deep learning solution. Moreover, these metrics are not appropriate to evaluate the Draco encoder performance. Typically, deep learning solutions lead to different types of distortions which are not properly represented by the studied metrics.

Objective Metrics Benchmarking

In order to compare the objective measures with the subjective scores, the statistical measures proposed in [19] were calculated to measure the performance of each metric. Specifically, these are the Pearson Correlation Coefficient (PCC), the Spearman Rank Order Correlation Coeficient (SROCC), the Root-Mean Squared Error (RMSE) and the Outlier Ratio (OR). The prediction of the MOS for each objective metric was computed by applying a linear (no fitting) and a logistic fitting function on the objective scores.

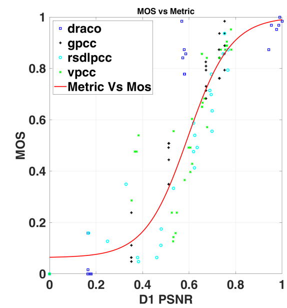
From table 5 and table 6, it can be observed that the best performing metrics were the point to point and point to plane metrics. A plot representing the relation between each metric and the MOS is represented in figure 7 including the logistic fitting curve. We can observe in this plot that while V-PCC and G-PCC data tends to be very close from the logistic curve, RS-DLPCC tends to be below the curve. This reveals that the metrics provide a good representation for the codec, but they are not suitable to compare with other codecs. Furthermore, Draco results are not appropriately represented by these metrics. This is a case were metrics could misjudge the performance of a codec.

Table 5: Linear Fitting

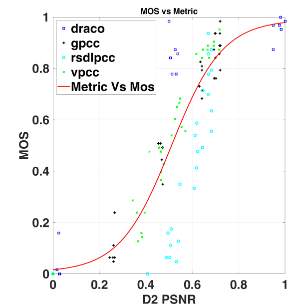
Metric	PCC	SROCC	RMSE	OR
po2point_MSE_PSNR	**0.862**	**0.884**	**0.163**	**0.735**
po2plane_MSE_PSNR	**0.814**	**0.847**	**0.187**	**0.784**
pl2plane_MSE	0.791	0.795	0.197	0.775
color_PSNR	0.488	0.679	0.280	0.833

Table 6: Logistic Fitting

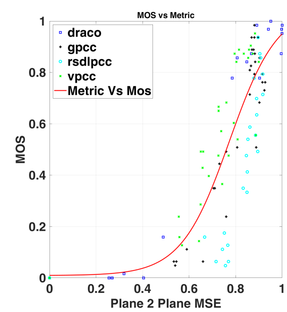
Metric	PCC	SROCC	RMSE	OR
po2point_MSE_PSNR	**0.890**	**0.884**	**0.148**	**0.618**
po2plane_MSE_PSNR	**0.851**	**0.847**	**0.169**	**0.618**
pl2plane_MSE	0.846	0.795	0.172	0.667
color_PSNR	0.670	0.679	0.240	0.719



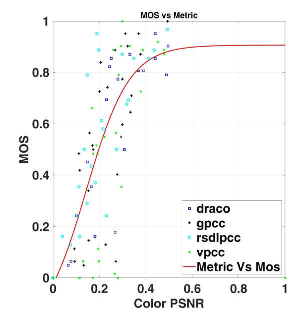
(a) D1 Logistic



(b) D2 Logistic



(c) Plane 2 Plane Logistic



(d) Color Logistic

Figure 7: Relation between metrics and MOS, and Logistic fitting curve.

Conclusions

An evaluation on the quality of V-PCC, G-PCC, RS-DLPCC and Draco codecs is presented. This paper reveals that MPEG codecs are the best performing solutions. The tested deep learning solution also provided a very good compression performance result, with space to further improvements. It was developed for geometry compression and can still improve its performance through the appropriate compression of each point associated features (RGB components in this case). Moreover, better training and better architectures can be implemented. This deep learning solution is likely to produce some blocking artifacts that we believe were the cause of the reduction of performance when compared with the MPEG codecs. Draco does not provide state of the art results for point cloud compression.

This research also revealed that the most common point cloud metrics fail to provide an accurate representation of quality when deep learning compression models are used. While the metrics show similar results for the G-PCC and the studied deep learning solution, the subjective evaluation revealed different quality. Because of that, correlations are slightly lower than 0.9 for the studied metrics. Although those are still acceptable results, they reveal that these metrics should be carefully considered when different compression technologies are used, causing different types of distortions. In the near future, a deeper analysis of the state of the art point cloud metrics will be conducted using this subjective quality evaluation data.

Another fact that was revealed is that sparse point clouds might tend to have different behaviors. In the case of bumbameuboi, which is rather sparse when compared with the others, G-PCC provided the most efficient representation and the deep learning solution provides a really bad performance.

Acknowledgemnts

This research was funded by the Portuguese FCT-Fundação para a Ciência e Tecnologia under the project UIDB/EEA/50008/2020, PLive X-0017-LX-20, and by operation Centro-01-0145-FEDER-000019 - C4 - Centro de Competencias em Cloud Computing.

The authors would like to acknowledge the authors of [20] for providing the deep learning decoded data for the reported experiment.

References

[1] JPEG Pleno Database, https://jpeg.org/plenodb/. [Online].

[2] UNIVERSITY OF SÃO PAULO POINT CLOUD DATASET http://uspaulopc.di.ubi.pt [Online]

[3] MPEG 3DG, "G-PCC Codec Description v5," ISO/IEC JTC1/SC29/WG11 N18891, Geneva, CH, October 2019.

[4] "V-PCC Test Model v8," ISO/IEC JTC1/SC29/WG11 W18884, Geneva, CH, October 2019.

[5] A. Guarda et al., "Deep Learning-based Point Cloud Geometry Coding with Resolution Scalability," IEEE International Workshop on Multimedia Signal Processing (MMSP'2020), Tampere, Finland, Sep 2020.

[6] A. Guarda, Nuno M. M. Rodrigues, F. Pereira, Adaptive Deep Learning-based Point Cloud Geometry Coding, IEEE Journal on Selected Topics in Signal Processing, Vol. 15, No. 2, pp. 415 - 430, February, 2021

[7] S. Perry et al., "Quality Evaluation Of Static Point Clouds Encoded Using MPEG Codecs," 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 3428-3432

[8] FFMPEG Documentation, https://ffmpeg.org/ffmpeg.html. [Online]

[9] E. Alexiou, E. Upenik, and T. Ebrahimi, "Towards subjective quality assessment of point cloud imaging in augmented reality," in 2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP), 2017, pp. 1–6.

[10] E. Alexiou et al., "Point cloud subjective evaluation methodology based on 2D rendering," in 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), 2018, pp. 1–6.

[11] A. Javaheri et al. "Subjective and objective quality evaluation of compressed point clouds" 2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP) pp. 1-6 2017.

[12] L. A. da Silva Cruz et al., "Point cloud quality evaluation: Towards a definition for test conditions," in 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), 2019, pp. 1–6.

[13] S. Perry et al. "Study of subjective and objective quality evaluation of 3d point cloud data by the jpeg committee" Electronic Imaging vol. 2019 no. 10 pp. 12–1-312–7 2019

[14] E. Alexiou et al., "A comprehensive study of the rate- distortion performance in mpeg point cloud compression," APSIPA Transactions on Signal and Information Processing, vol. 8, p. e27, 2019.

[15] H. Su et al., "Perceptual quality assessment of 3d point clouds," in 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 3182–3186.

[16] D.Tian, H.Ochimizu, C.Feng, R.Cohen, and A.Vetro. Geometric distortion metrics for point cloud compression. In 2017 IEEE International Conference on Image Processing (ICIP), pages 3460–3464, 2017

[17] E. Alexiou, T. Ebrahimi, "Point cloud quality assessment metric based on angular similarity", IEEE International Conference on Multimedia and Expo (ICME), July 2018.

[18] CloudCompare: 3D point cloud and mesh processing software Open Source Project, https://www.danielgm.net/cc/, 2019.

[19] ITU-T P.1401, "Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models," International Telecommunication Union, Jul. 2012.

[20] A. F. R. Guarda, N. M. M. Rodrigues and F. Pereira, "Point Cloud Coding: Adopting a Deep Learning-based Approach," 2019 Picture Coding Symposium (PCS), 2019, pp. 1-5, doi: 10.1109/PCS48520.2019.8954537.

[21] O. Devillers and P. . Gandoin, "Geometric compression for interactive transmission," in Proceedings Visualization, 2000, pp. 319–326.

[22] G. Bruder, F. Steinicke and A. Nüchter, "Poster: Immersive point cloud virtual environments," 2014 IEEE Symposium on 3D User Interfaces (3DUI), 2014, pp. 161-162, doi: 10.1109/3DUI.2014.6798870.

[23] Quach, Maurice and Valenzise, Giuseppe and Dufaux, Frederic, "Learning Convolutional Transforms for Lossy Point Cloud Geometry Compression", 2019 IEEE International Conference on Image Processing (ICIP),2019

[24] ITU-R BT.500-13, "Methodology for the subjective assessment of the quality of television pictures," International Telecommunications Union, Jan. 2012.

[25] BT709 ITU-R BT.709, Parameter values for the HDTV standards for production and international programme exchange, Jun. 2015.

[26] Google, Draco 3D Data Compression, https://github.com/google/draco. [online]