

# Textured Mesh vs Coloured Point Cloud: A Subjective Study for Volumetric Video Compression

Emin Zerman\*, Cagri Ozcinar\*, Pan Gao†, and Aljosa Smolic\*

\*V-SENSE, School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland.

Email: {zermame, ozcinarc, smolica}@scss.tcd.ie

†College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China.

Email: gaopan.1005@gmail.com

**Abstract**—Volumetric video (VV) pipelines reached a high level of maturity, creating interest to use such content in interactive visualisation scenarios. VV allows real world content to be captured and represented as 3D models, which can be viewed from any chosen viewpoint and direction. Thus, VV is ideal to be used in augmented reality (AR) or virtual reality (VR) applications. Both textured polygonal meshes and point clouds are popular methods to represent VV. Even though the signal and image processing community slightly favours the point cloud due to its simpler data structure and faster acquisition, textured polygonal meshes might have other benefits such as better visual quality and easier integration with computer graphics pipelines. To better understand the difference between them, in this study, we compare these two different representation formats for a VV compression scenario utilising state-of-the-art compression techniques. For this purpose, we build a database and collect user opinion scores for subjective quality assessment of the compressed VV. The results show that meshes provide the best quality at high bitrates, while point clouds perform better for low bitrate cases. The created VV quality database will be made available online to support further scientific studies on VV quality assessment.

**Index Terms**—Volumetric video, textured mesh, point cloud, subjective quality assessment, point cloud compression, mesh compression

## I. INTRODUCTION

Volumetric video (VV) –also known as free-viewpoint video– is a new form of immersive visual media which enables viewers to look at the captured content from any direction and any viewing angle. The technical advancements made in capture and display technologies helped VV content creation and delivery pipelines reach a high level of maturity. The VV is generally captured in dedicated studios using several cameras looking inwards [1]–[5]. The captured volumetric content is ideal for augmented reality (AR) or virtual reality (VR) applications, and can be used in various applications from entertainment to education, and communication. As the interest towards VV and its applications in AR and VR grow, the visual media community continues the efforts towards

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under the Grant Number 15/RP/27760.

978-1-7281-5965-2/20/\$31.00 ©2020 IEEE

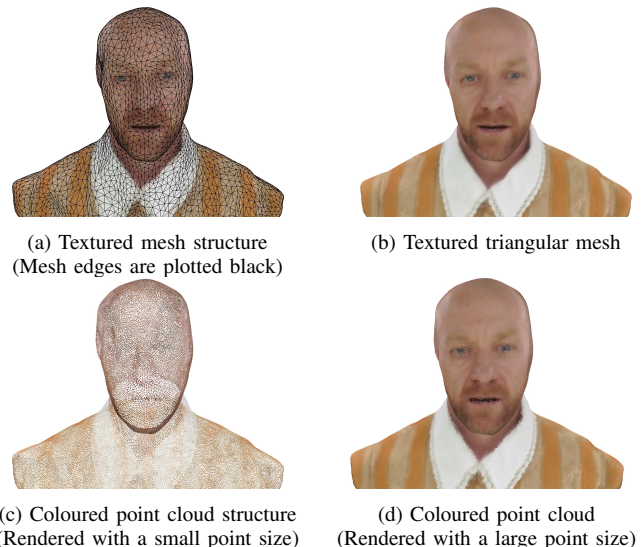


Fig. 1. Sample snapshots from the textured mesh and the point cloud of the same volumetric video.

standardisation of VV delivery. Especially, point cloud compression is a hot topic which attracts attention from researchers from academia and industry [6], [7], including JPEG [8] and MPEG [9] communities. In parallel, scientists also try to understand how perception and quality assessment (QA) change for this new immersive visual media [10]–[15]. Nevertheless, currently, there are no publicly available large QA databases for VV which can be used for understanding of QA for VV.

Textured polygonal meshes [2] and point clouds [16] are two popular formats to represent and store the captured VV. Sample figures for these formats can be seen in Fig. 1. Textured polygonal meshes consist of vertices, edges with connectivity information, and a texture atlas to colourise the 3D model. Hence, polygonal meshes need to preserve the connectivity information, and this brings some limitations during compression and transmission. Point clouds, instead, forego the connectivity information in favour of easier storage and simplicity. Moreover, since point clouds do not have connectivity information, they can be acquired and stored in any order as long as the whole of the point cloud is considered. The recent standardisation efforts [8], [9] show that the signal processing community slightly favours point clouds nowadays

due to its easier acquisition and simpler data representation. Nevertheless, textured polygonal meshes have other benefits such as its easier integration with computer graphics pipelines and a piecewise continuous surface representation.

In this paper, we aim to understand the perceptual differences between the two most commonly used VV representation formats: textured polygonal meshes and point clouds. For this purpose, we create a database including both VV representations and build a subjectively-annotated VV quality database considering the compression scenario. Besides comparing textured polygonal meshes and coloured point clouds, we also compare different compression methods. The main contributions of this paper are:

- a comparison of state-of-the-art VV compression methods: Draco [17], G-PCC [9], and V-PCC [9]
- a comparison of the effect of two different VV representation techniques (*i.e.*, textured meshes and coloured point clouds) on compression performance, and
- a QA database with subjectively annotated mean opinion score (MOS) values for VVs comprising of a total of 152 compressed VVs, generated from eight different contents and using three different compression methods.

the details of which are discussed in the following sections.

## II. RELATED WORK

Recent years have seen a peak in the interest for the creation of VV [1]–[5]. Using different number of cameras and sensors, these methods capture the real-life 3D content from the scene. For example, Alexiadis *et al.* [1] used four Microsoft Kinect sensors, each of which consists of an RGB camera and a depth scanner. A more advanced system developed by Microsoft [2] captures the scene with 106 cameras. Afterwards, the reconstructed mesh is temporally tracked and turned into a streamable bitstream. Pages *et al.* [3], instead, propose a more affordable system which uses 12 cameras for 3D mesh reconstruction. In a more recent work [5], 32 cameras are used to generate VV. Often in these studies, the reconstruction starts with point cloud generation and proceeds with meshing. Nevertheless, recent point cloud databases [16], [18] show that VV can be also represented as point clouds.

Polygonal meshes have long been used in computer graphics applications from virtual 3D representations of objects to video games. Thus, mesh compression is a widely researched topic in computer graphics [19]. Nevertheless, for most of the applications, the said 3D models were computer generated, and thus there was no need to change the connectivity (or structure) of a 3D mesh structure except for compression purposes (*e.g.*, progressive meshes). However, for VV applications, there are certain limitations stemming from 3D reconstruction. In most of the cases, the volumetric model is reconstructed for each frame of source video independently. That is, there are initially no temporal constraints imposed during the reconstruction process. This can be remedied in post-reconstruction by temporal tracking of meshes [2]. Even in this case, it is likely that these temporally consistent groups of meshes will not have the same length (*i.e.* number of frames), depending on the

nature of the motion. Therefore, representation as point clouds for storage and delivery of VV has certain advantages not requiring connectivity. Hence, point cloud compression is a hot research topic now, as many researchers [6], [7] and JPEG [8] and MPEG [9] communities are working towards it.

There are many studies conducted for the quality assessment of polygonal meshes [20]–[23] and point clouds [10]–[15]. Nevertheless, only some of these studies considered the state-of-the-art compression methods in a way which can be beneficial for a VV setting. Doumanoglou *et al.* [22] and Christaki *et al.* [23] studied different open source mesh compression algorithms and found Google’s Draco the best performing among the compared. Su *et al.* [12] used static point clouds to compare three different encoders: L-PCC, S-PCC, and V-PCC. According to the results of their subjective experiment, V-PCC was found better than other methods. Alexiou *et al.* [13] also used static point clouds to compare G-PCC and V-PCC both subjectively and objectively. These studies provide some insightful results even though they consider static 3D models. On the other hand, studies considering the QA of VVs (*i.e.*, time-sequences of either meshes or point clouds) are rather limited [14], [15]. Zerman *et al.* [14] used two VVs and considered TMC2 (*i.e.*, V-PCC), also creating an initial subjective VV quality database. Gonçalves *et al.* [15] used instead the 8i VV database [16] and compared TMC2 (*i.e.*, V-PCC) to PCC [6]. TMC2 was found to be better than PCC.

To the best of our knowledge, this paper is the first study to compare textured meshes and point clouds for a compression scenario. Even though the differences between point clouds and meshes (reconstructed from the said point clouds) were compared in only one study [24], there were no conclusions on which one is better. Instead, in this study, our findings give a clear recommendation for the VV compression scenario.

## III. SUBJECTIVE QUALITY ASSESSMENT

To build the VV quality database and collect viewer quality opinions on the comparison of textured meshes and point clouds, we conducted a subjective quality experiment. The details are discussed in the following subsections. In order to foster scientific works for VV research, the database with the collected subjective quality scores will be made publicly available on the project webpage<sup>1</sup>.

TABLE I  
CONSIDERED VV DATABASES.

Database	Point Cloud	Textured Mesh	Contents
V-SENSE	✓	✓	AxeGuy, LubnaFriends, Rafa2, Matis
8i [16]	✓	×	Longdress, Loot, Redandblack, Soldier

### A. Stimuli & Database

For the creation of the database, we selected eight different VVs in total, as shown in Table I. Four of these VVs were

<sup>1</sup>v-sense.scss.tcd.ie/research/6dof/quality-assessment-for-fvv-compression/

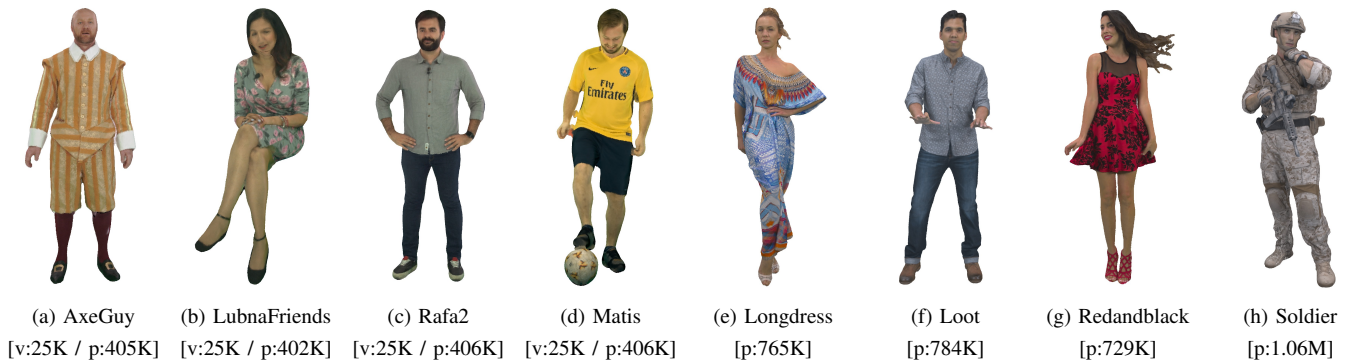


Fig. 2. Selected VVs from (a-d) V-SENSE and (e-h) 8i databases [16]. Vertex (*i.e.*, *v*) and point counts (*i.e.*, *p*) are indicated in brackets for textured mesh and point cloud representations respectively. All textured meshes have 50K faces. All VVs have 300 frames and are played with 30 fps.

generated by us (called “V-SENSE” from this point on) using the method of Pages *et al.* [3]. The remaining four of these VVs were taken from the 8i point cloud dataset [16]. While V-SENSE database had both textured meshes and point clouds, 8i database only had point clouds. Sample renders for the selected VVs are shown in Fig. 2, together with other details. The collected data contains only human bodies as this was the only content available for volumetric videos at the moment.

Anticipating that the VV will be used alongside the traditional visual media in the near future, we consider the compression scenario in this paper. Considering the findings in recent studies [22], [23], Google’s Draco encoder was selected to compress VV represented as polygonal meshes. In this case, JPEG compression was used for texture atlases. For VVs represented as point clouds, we consider the state-of-the-art compression algorithms which are being developed in MPEG standardisation efforts [9]: G-PCC and V-PCC. Draco and G-PCC were proposed for the aim of compression of static volumetric contents, and V-PCC was developed for compression of VV. Since Draco and G-PCC do not consider temporal redundancies, we included the *all-intra* option of V-PCC to compare them in a fair way. In total, we consider four different cases as shown in Table II.

The Draco encoder was downloaded from the GitHub page [17] and run on a Linux system. The texture atlases were compressed using JPEG with different Quality levels in Matlab. For the compression of point clouds, we used G-PCC (mpeg-pcc-tmc13v6.0) and V-PCC (mpeg-pcc-tmc2v5.0) encoders, following the descriptions in the MPEG document [25] for RAHT, *all-intra*, and *random-access* modes.

For the point clouds, we selected six different levels for

TABLE II  
CONSIDERED VV COMPRESSION METHODS.

Encoders	Description
<b>Draco+JPEG</b>	Draco encoder for polygonal meshes was used together with JPEG for texture atlases
<b>G-PCC (RAHT)</b>	G-PCC encoder with region-adaptive hierarchical transform (RAHT) to compress point-wise colour information
<b>V-PCC (AI)</b>	V-PCC encoder where each frame was intra coded
<b>V-PCC (RA)</b>	V-PCC encoder with random access setting

TABLE III

ENCODER PARAMETERS USED WHILE STIMULI CREATION, WHERE *QP* AND *QT* INDICATE DRACO’S QUANTISATION PARAMETERS FOR POSITION AND TEXTURE, *JPEG* INDICATES JPEG QUALITY PARAMETER, *colSt* INDICATES COLORSTEPSIZE PARAMETER OF G-PCC, *geoQP* AND *texQP* INDICATE GEOMETRY AND TEXTURE QUANTISATION PARAMETERS FOR V-PCC.

Compression Methods		Quality Levels					
		R1	R2	R3	R4	R5	R6
<b>Draco+JPEG</b>	QP	8	10	10	12	12	12
	QT	6	10	10	10	12	12
	JPEG	0	0	5	10	30	55
<b>G-PCC</b>	depth	10	10	10	10	10	10
	level	6	7	7	7	8	10
	colSt	64	32	16	8	4	1
<b>V-PCC</b>	geoQP	32	28	24	20	16	-
	texQP	42	37	32	27	22	-

G-PCC and five different levels for V-PCC as given in the MPEG’s common test conditions (CTC) for point cloud compression [25]. For the polygonal meshes, the coding parameters for six different quality levels were selected in a pilot study which was conducted with three expert viewers. Out of these six levels, five levels were selected to correspond to ACR categories [26], and another level (with a higher bandwidth requirement) was selected so that we can see the extent of the curve for the ‘Draco+JPEG’ case. The selected parameters are given in Table III. This stimuli preparation resulted in 152 distorted videos created from eight different contents.

For visualisation of the VV, a passive approach was used, which also made sure that all the participants were presented with the same stimuli. For this purpose, the compressed VVs were rendered using Blender (version 2.80) with “Point cloud visualizer” add-on<sup>2</sup>, and the rendered representations were stored as traditional videos. Using *ffmpeg*, these videos are compressed with *x264* and *-crf 15* parameter to ensure that the compression is perceptually lossless. The point sizes were selected to ensure that the stimuli would look watertight, both for textured triangular meshes and coloured point clouds. The VV was placed in the scene and camera was set to orbit the VV’s initial origin twice in the clockwise direction within video duration (*i.e.*, 10 sec).

<sup>2</sup>github.com/uhlik/bpy – Distributed under GNU General Public License

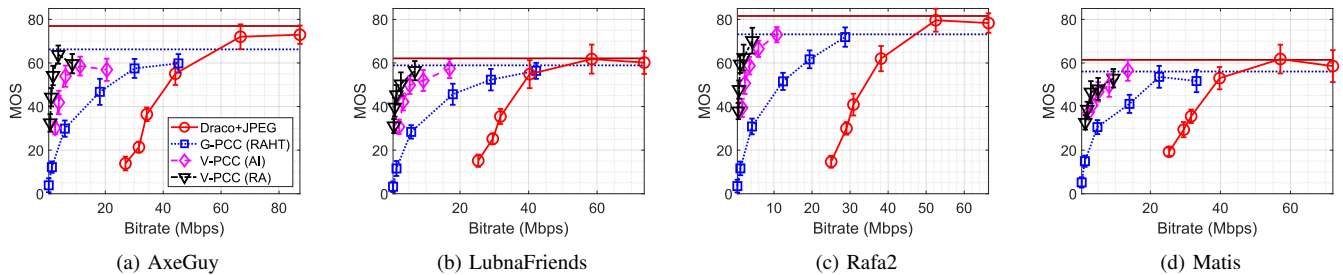


Fig. 3. Comparing textured mesh (Draco) and point cloud (G-PCC, V-PCC) compression results: MOS vs bitrate plots for the VVs in the V-SENSE database. The horizontal lines indicate MOS for reference VVs for each content: solid red line for textured mesh and dashed blue line for point cloud.

## B. Experiment Setup

As mentioned in the previous subsection, to reduce inter-subject variation, a passive approach was chosen for viewer interaction with the VV. A rendered version of each VV was shown to the participants on a 24" LCD display. The distance between the viewers and the screen was set to three times of the stimuli height, which was around 1 metre. The experiment was conducted in a dark room as recommended by ITU [26].

## C. Participants & Procedure

For this test, a call for participation was announced through e-mail lists within the university. In total 23 participants, 17 male and 6 female, volunteered to participate in the study. The average age of the participants was 30.9 (Std = 4.14).

The participants were first trained on 'Rafa' sequence [14] (different from 'Rafa2' used in this paper) with different V-PCC compression levels. Afterwards, only compressed examples from different bitrate levels were shown to the participants without providing specific instructions, to familiarise them while avoiding any bias.

Considering that both textured meshes and point clouds were present in the experiment, we wanted to conduct the test using a single stimulus methodology in order to avoid affecting the users by showing an explicit "reference" stimulus. To minimise the time required for the experiment, the absolute category rating (ACR) with hidden reference was selected as the test methodology [26]. References for both meshes and point clouds were shown separately (see Fig. 3). Each stimulus was 10 seconds long, and participants voted for each stimulus in 1.5 sec on average. The experiment was conducted in two sessions, and each session was kept under 30 minutes to avoid participant fatigue. In order to get consistent results for the comparison between meshes and point clouds, the first session consisted of the V-SENSE database with both representations, and the second session consisted of 8i dataset with point clouds. To be able to merge the subjective scores afterwards, a common set was present in both sessions. That is, 44 videos in total were present in both sessions, including both textured meshes and coloured point clouds to avoid bias.

## IV. ANALYSIS & RESULTS

In this section, we first calculate the mean opinion score (MOS) for each stimulus, and then we compare textured

meshes with point clouds. We also compare different state-of-the-art compression methods, including also a quantitative analysis based on subjective quality scores.

### A. Calculation of MOS

Although the participants were shown sample VVs in the training part prior to the experiment, we noticed that each participant had a different understanding of the scale during rating. To reduce the subject variability and keep the overall mean, we follow the procedure described by Athar *et al.* [27]. According to this, the raw opinion scores,  $s_{ij}$ , for each subject  $i$  and stimulus  $j$ , were used to calculate the z-score as follows:

$$z_{ij} = \frac{s_{ij} - \mu_i}{\sigma_i} \quad (1)$$

where  $\mu_i$  and  $\sigma_i$  are the mean and standard deviation of all the opinion scores of the subject  $i$ . The MOS was then found as follows:

$$MOS_j = \sigma_{rmos} \left( \frac{MOS_j^z - \mu_{zmos}}{\sigma_{zmos}} \right) + \mu_{rmos} \quad (2)$$

where  $MOS_j^z$  is the mean of z-scores for the stimulus  $j$ ,  $\mu$  and  $\sigma$  are the mean and standard deviation of either raw opinion scores ( $rmos$ ) or calculated z-scores ( $zmos$ ). In the remainder of this paper, we use the MOS values as computed in (2).

### B. Comparing Mesh with Point Clouds

After calculating the MOS values for all the stimuli, we plot the MOS vs bitrate figures and compare the performances of the two VV representation methods. We must note that this comparison (see Fig. 3) has only been done for the four VVs of the V-SENSE dataset, since only they are available in both textured mesh and coloured point cloud format.

It can be seen from the plots that the textured mesh has a higher maximum MOS value compared to the point cloud representation. On the other hand, point clouds seem to be a better choice for the scenarios where the transmission bandwidth and storage capacity are limited. In fact, textured meshes seem to be the best choice for a no-compression case; however, in most cases that the VV is compressed, point clouds seem to have a good balance between the perceived quality and the bitrate. The reduction in the MOS for the point cloud case can be explained by changing perceptual resolution. Since point clouds have to sample the texture of the model, this



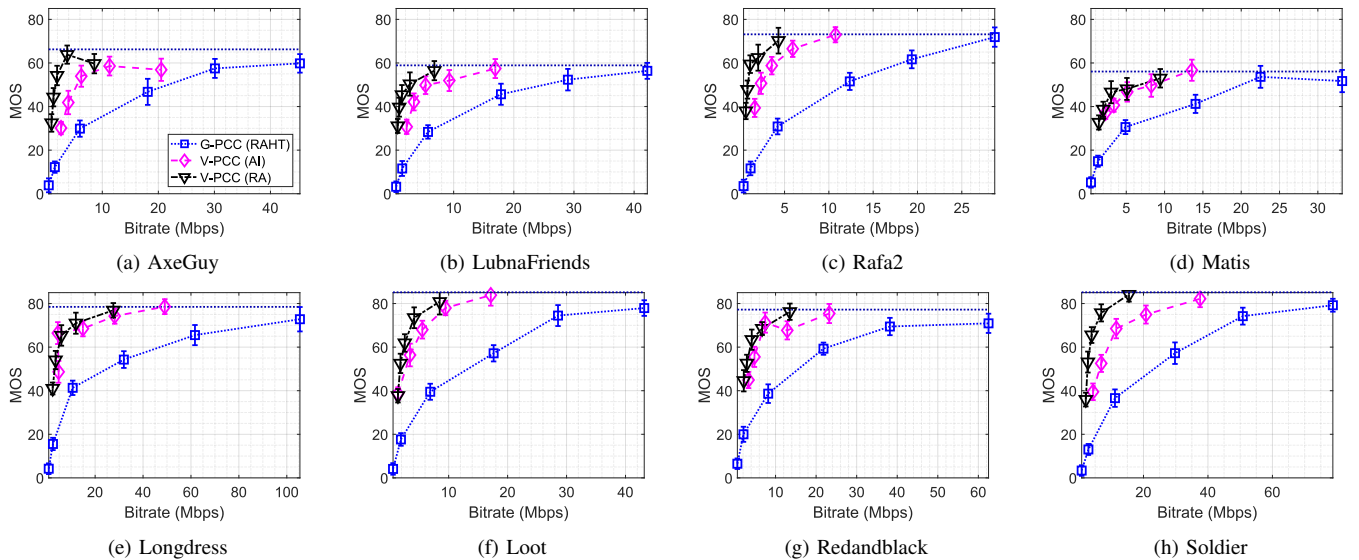


Fig. 4. Comparing different point cloud (G-PCC, V-PCC) compression results: MOS vs bitrate plots for all the VVs considered in this study. The horizontal dashed blue line indicates the MOS for the reference point cloud VV for each content.

might have the effect of reducing the apparent resolution due to sampling and point cloud sizes.

It can also be seen that two of the VVs (*LubnaFriends* and *Matis*) were found to be of lower quality than the other two VVs in the V-SENSE database. It has been noted by the providers that *LubnaFriends* and *Matis* sequences were generated with an earlier version of their technology compared to *AxeGuy* and *Rafa2*. The reason for this difference is related to the reconstruction technology. Nevertheless, we believe that including these VVs in the dataset brings value as they have different characteristics and limitations.

### C. Comparing Point Cloud Compression Methods

In addition to the comparison between meshes and point clouds, we also compare the state-of-the-art point cloud compression methods, using our database. To help readers see the results for the four VVs mentioned in the previous subsection, we plot the figures again without the “Draco+JPEG” curve. The resulting plots for all VVs are shown in Fig. 4. Examining the plots, we can observe that “G-PCC (RAHT)” is not better than V-PCC in any case. This shows that the V-PCC compression, which is based on video coding of projected point clouds, yields better results with a smaller bitrate than G-PCC, which is based on octree coding of points in 3D space. As can be expected, the “Random Access” option of V-PCC is more effective than the “All Intra” case.

Similar to the discussion in the previous subsection, the difference between maximum MOS values for each content is related to reconstruction errors. 8i content was captured in a more professional studio setting with more cameras. Additionally, another reason of this MOS difference can be the difference in point counts, i.e. resolution. The point clouds provided with the 8i database had around 800,000 points while the ones from the V-SENSE dataset had around 400,000

TABLE IV  
SUBJECTIVE BD-MOS GAIN OF THE POINT CLOUD CODING OVER THE MESH CODING. THE ANCHOR USED FOR BD-MOS CALCULATION IS THE MESH COMPRESSION SCHEME, I.E., DRACO+JPEG SCHEME, WHILE THE POINT CLOUD CODING SCHEMES ARE G-PCC (RAHT), V-PCC (AI), AND V-PCC (RA).

Sequence	Point Cloud Coding Scheme		
	G-PCC (RAHT)	V-PCC (AI)	V-PCC (RA)
AxeGuy	22.72	54.86	73.20
LubnaFriends	17.23	69.54	138.88
Rafa2	47.32	112.71	176.92
Matis	22.92	67.00	75.08
Average	27.55	76.03	115.99

points. This might also have affected the visual quality as the resolution affects the appearance significantly.

### D. Subjective BD-MOS Analysis

In this section, we quantitatively evaluate the performance of the various mesh and point cloud compression schemes. Defined similarly to the Bjontegaard Delta PSNR (BD-PSNR) [28], Bjontegaard Delta MOS (BD-MOS) was developed in [29] for measuring the quality difference between different algorithms for the same bit rate by using MOS values. As demonstrated in [29], the BD-MOS can lead to better results since it can fully take the saturation effect of the human visual system and the full nature of the artefacts into account. Therefore, we adopt BD-MOS for an additional subjective quality comparison and analysis. Table IV summarises BD-MOS gains of the various point cloud coding approaches over the mesh coding scheme, i.e., “Draco+JPEG scheme”. As can be observed, the “V-PCC (Random Access)” performs the best, which can achieve a MOS gain from 73.2 to 176.9.

Table V gives the BD-MOS gains among point cloud compression scheme, in which the “G-PCC (RAHT)” is used as the anchor. It can also be found that “V-PCC (Random Access)”

TABLE V  
SUBJECTIVE BD-MOS GAIN OF THE V-PCC SCHEME OVER THE G-PCC SCHEME. THE ANCHOR IS THE G-PCC (RAHT).

Sequence	V-PCC scheme	
	V-PCC (AI)	V-PCC (RA)
AxeGuy	17.81	36.12
LubnaFriends	18.11	30.75
Rafa2	28.97	42.45
Matis	14.35	18.53
Longdress	22.35	30.50
Loot	29.60	39.19
Redandblack	22.75	30.19
Soldier	25.47	44.64
Average	22.43	34.05

which exploits the temporal redundancy during coding yields the best subjective quality, and achieves a BD-MOS gain of 34.05 on average for all the test sequences.

## V. CONCLUSION

In this paper, we report building a subjectively annotated quality database for volumetric video (VV) in order to compare two different representations for VV considering the compression scenario: textured meshes and point clouds. Moreover, we compared different state-of-the-art point cloud and mesh compression methods with objective and subjective experiments. It can be inferred from the results that the textured meshes provide the best visualisation and are more advantageous for high-bitrate bandwidth (*e.g.*, over 50 Mbps) applications whereas the point clouds are well-suited for applications with limited bandwidth (*e.g.*, below 20 Mbps). We can also say that “V-PCC (Random Access)” was found to be more effective than all the other compression schemes considered for most of the cases. One must note that the constructed database had only human bodies and the findings of this study are only valid for human bodies.

As future work, we aim to utilise this dataset with a better perceptually suited quality assessment method, and we also aim to complement this study with a comprehensive evaluation and benchmark of the objective quality metrics for VV. The created database will be made publicly available and can be used for further scientific studies in VV compression or quality assessment.

## REFERENCES

- [1] D. S. Alexiadis, D. Zarpalas, and P. Daras, “Real-time, full 3-D reconstruction of moving foreground objects from multiple consumer depth cameras,” *IEEE Trans. Multimedia*, vol. 15, no. 2, Feb 2013.
- [2] A. Collet, M. Chuang, P. Sweeney, D. Gillett, D. Evseev, D. Calabrese, H. Hoppe, A. Kirk, and S. Sullivan, “High-quality streamable free-viewpoint video,” *ACM Trans. Graphics*, vol. 34, no. 4, Jul. 2015.
- [3] R. Pagés, K. Amplianitis, D. Monaghan, J. Ondřej, and A. Smolić, “Affordable content creation for free-viewpoint video and VR/AR applications,” *J. Visual Commun. Image Represent.*, vol. 53, 2018.
- [4] “8i,” <https://www.8i.com/>, accessed: 2020-01-25.
- [5] O. Schreer, I. Feldmann, S. Renault, M. Zepp, M. Worchel, P. Eisert, and P. Kauff, “Capture and 3D video processing of volumetric video,” in *IEEE International Conference on Image Processing (ICIP)*, Sep. 2019.
- [6] R. Mekuria, K. Blom, and P. Cesar, “Design, implementation, and evaluation of a point cloud codec for tele-immersive video,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 4, pp. 828–842, 2016.

- [7] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso, “Subjective and objective quality evaluation of compressed point clouds,” in *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2017.
- [8] S. Perry, A. Pinheiro, E. Dunic, C. da Silva, and A. Luis, “Study of subjective and objective quality evaluation of 3D point cloud data by the JPEG committee,” in *IS&T Electronic Imaging, Image Quality and System Performance XVI*, 2019, pp. 312–1–312–6.
- [9] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. A. Chou, R. A. Cohen, M. Krivokuća, S. Lasserre, Z. Li *et al.*, “Emerging MPEG standards for point cloud compression,” *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 9, no. 1, pp. 133–148, Mar 2019.
- [10] R. Mekuria, Z. Li, C. Tulvan, and P. Chou, “Evaluation criteria for PCC (Point Cloud Compression),” ISO/IEC JTC 1/SC29/WG11 Doc. N16332, 2016.
- [11] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, “Geometric distortion metrics for point cloud compression,” in *IEEE International Conference on Image Processing (ICIP)*, Sept 2017, pp. 3460–3464.
- [12] H. Su, Z. Duanmu, W. Liu, Q. Liu, and Z. Wang, “Perceptual quality assessment of 3D point clouds,” in *IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019, pp. 3182–3186.
- [13] E. Alexiou, I. Viola, T. M. Borges, T. A. Fonseca, R. L. de Queiroz, and T. Ebrahimi, “A comprehensive study of the rate-distortion performance in MPEG point cloud compression,” *APSIPA Transactions on Signal and Information Processing*, vol. 8, 2019.
- [14] E. Zerman, P. Gao, C. Ozcinar, and A. Smolic, “Subjective and objective quality assessment for volumetric video compression,” in *IS&T Electronic Imaging, Image Quality and System Performance XVI*, 2019.
- [15] M. Gonçalves, L. Agostini, D. Palomino, M. Porto, and G. Correa, “Encoding efficiency and computational cost assessment of state-of-the-art point cloud codecs,” in *IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019, pp. 3726–3730.
- [16] E. d’Eon, B. Harrison, T. Myers, and P. A. Chou, “8i voxelized full bodies, version 2 – A voxelized point cloud dataset,” ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) input document m40059/M74006, Jan 2017, Geneva, Switzerland.
- [17] “Google Draco,” <https://google.github.io/draco/>, accessed: 2020-01-25.
- [18] T. Ebner, I. Feldmann, O. Schreer, P. Kauff, and T. Unger, “HHI point cloud dataset of a boxing trainer,” ISO/IEC JTC1/SC29 Joint WG11 (MPEG) input document m42921, Jul 2018, Ljubljana, Slovenia.
- [19] A. Maglo, G. Lavoué, F. Dupont, and C. Hudelot, “3D mesh compression: Survey, comparisons, and emerging trends,” *ACM Computing Surveys (CSUR)*, vol. 47, no. 3, 2015.
- [20] J. Guo, V. Vidal, I. Cheng, A. Basu, A. Baskurt, and G. Lavoué, “Subjective and objective visual quality assessment of textured 3D meshes,” *ACM Trans. Appl. Percept. (TAP)*, vol. 14, no. 2, 2017.
- [21] K. Vanhoye, B. Sauvage, P. Kraemer, and G. Lavoué, “Visual quality assessment of 3D models: on the influence of light-material interaction,” *ACM Trans. Appl. Percept. (TAP)*, vol. 15, no. 1, 2017.
- [22] A. Doumanoglou, P. Drakoulis, N. Zioulis, D. Zarpalas, and P. Daras, “Benchmarking open-source static 3D mesh codecs for immersive media interactive live streaming,” *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 9, no. 1, pp. 190–203, March 2019.
- [23] K. Christaki, E. Christakis, P. Drakoulis, A. Doumanoglou, N. Zioulis, D. Zarpalas, and P. Daras, “Subjective visual quality assessment of immersive 3D media compressed by open-source static 3D mesh codecs,” in *25th Int. Conf. on MultiMedia Modeling (MMM)*, 2019.
- [24] E. Alexiou, M. V. Bernardo, L. A. da Silva Cruz, L. G. Dmitrovic, C. Duarte, E. Dunic, T. Ebrahimi, D. Matkovic, M. Pereira, A. Pinheiro, and A. Skodras, “Point cloud subjective evaluation methodology based on 2D rendering,” in *10th International Conference on Quality of Multimedia Experience (QoMEX)*, 2018.
- [25] S. Schwarz, P. A. Chou, and M. Budagavi, “Common test conditions for point cloud compression,” ISO/IEC JTC1/SC29 Joint WG11 (MPEG) input document N17345, Jan 2018, Gwangju, Korea.
- [26] ITU-T, “Subjective video quality assessment methods for multimedia applications,” ITU-T Recommendation P.910, Apr 2008.
- [27] S. Athar, T. Costa, K. Zeng, and Z. Wang, “Perceptual quality assessment of UHD-HDR-WCG videos,” in *IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019, pp. 1740–1744.
- [28] G. Bjontegaard, “Improvements of the BD-PSNR model,” in *ITU-T SG16/Q6, 35th VCEG Meeting, Berlin, Germany, July, 2008*.
- [29] P. Hanhart, M. Rerabek, F. De Simone, and T. Ebrahimi, “Subjective quality evaluation of the upcoming HEVC video compression standard,” in *Applications of Digital Image Processing XXXV*. SPIE, 2012.