**3GPP TSG SA WG4 Meeting 133-eS4-251318**

**online, , 18 - 25 July 2025**

**Source: Qualcomm Incorporated**

**Title: Pseudo-CR on [FS\_Beyond2D] Gaussian Splats in glTF**

**Spec: 3GPP TR 26.956v1.0.0**

**Agenda item: 9.6**

**Document for: Decision**

**1. Introduction**

This provides some details on Gaussian splats.

**2. Reason for Change**

A more standardized format for splats is mission.

**3. Conclusions**

Add a standardized format based on glTF.

**4. Proposal**

It is proposed to agree the following changes to 3GPP TR 26.956v1.0.0 at SA4-133-e.

\* \* \* First Change \* \* \* \*

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] Allied Market Research, “3D Technology Market Size, Share, Competitive Landscape and Trend Analysis Report by Product, Application : Global Opportunity Analysis and Industry Forecast, 2021-2030.”, [www.alliedmarketresearch.com/3d-technology-market.](http://www.alliedmarketresearch.com/3d-technology-market.)

[3] Mordor Intelligence, “Mobile 3D Market Size & Share Analysis - Growth Trends & Forecasts (2024 - 2029).”, <https://www.mordorintelligence.com/industry-reports/mobile-3d-market.>

[4] Grand View Research, “Immersive Technology Market Size, Share & Trends Analysis Report By Component (Hardware, Software, Services), By Technology, By Application, By Industry, By Region, And Segment Forecasts, 2023 - 2030.”, [https://www.grandviewresearch.com/industry-analysis/immersive-technology-market-report.](https://www.mordorintelligence.com/industry-reports/mobile-3d-market.)

[26955] 3GPP TR 26.955: "Video codec characteristics for 5G-based services and applications".

[26118] 3GPP TS 26.118: "Virtual Reality (VR) profiles for streaming applications".

[26119] 3GPP TS 26.119: "Media Capabilities for Augmented Reality".

[26143] 3GPP TS 26.143: "Messaging Media Profiles".

[26511] 3GPP TS 26.511: "5G Media Streaming (5GMS); Profiles, codecs and formats".

[26966] 3GPP TR 26.966: " Evaluation of new HEVC coding tools".

[26265] 3GPP TS 26.265: "Media Delivery: Video Capabilities and Operating Points".

[S1] Apple HEVC Stereo Video - Interoperability Profile (Beta), Version 0.9, June 21, 2023, <https://developer.apple.com/av-foundation/HEVC-Stereo-Video-Profile.pdf>

[S2] A. Quested and B. Zegel, "3D-TV production standards - first report of the ITU-R Rapporteurs", EBU Technical Review, 2011 Q2, <https://tech.ebu.ch/publications/trev_2011-Q2_3dtv_quested>

[S3] Mike Swanson, "Spatial Video", March 7 2024, https://blog.mikeswanson.com/spatial-video/

[S4] Video Contour Map Payload, Version 0.9, June 21, 2023, <https://developer.apple.com/av-foundation/Video-Contour-Map-Metadata.pdf>

[S5] ITU-T H.273 (09/23), Coding-independent code points for video signal type identification

[S6] M. Satya, "3D Image Reconstruction From Multi-View Stereo", <https://medium.com/@satya15july_11937/3d-image-reconstruction-from-multi-view-stereo-782e6912435b>, March, 2023.

[S7] S. Khan and S. S. Channappayya, "Estimating Depth-Salient Edges and Its Application to Stereoscopic Image Quality Assessment," IEEE Transactions on Image Processing, vol. 27, no. 12, pp. 5892–5903, 2018, doi: 10.1109/TIP.2018.2860279.

[D1] Greg Turk, The Polygon File Format, Stanford University, 1994.

[D2] Volumetric Format Association VFA, <https://www.volumetricformat.org/>

[D3] "The OpenGL Graphics System: A Specification" (PDF). 4.0 (Core Profile). March 11, 2010.

[D4] V-PCC, Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC), ISO/IEC 23090-5 2nd Ed, Nov 2023.

[D5] G-PCC, Geometry-based point cloud compression, ISO/IEC 23090-9, Mar 2023

[D6] Draco Bitstream Specification, <https://google.github.io/draco/spec/>

[D7] MPEG 115, Use cases for Point Cloud Compression, https://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w16331\_Use\_Cases\_for\_Point\_Cloud\_Compression\_(PCC)\_0.docx

[M1] OpenCV, <https://opencv.org>.

[M2] Colmap, <https://colmap.github.io/index.html>.

[M3] AliceVision Photogrammetric Computer Vision Framework, <https://alicevision.org>.

[M4] Open Multiple View Geometry (openMVG), <https://github.com/openMVG/openMVG>.

[M5] Immersive Video Depth Estimation (IVDE), <https://gitlab.com/mpeg-i-visual/ivde>.

[M6] Test model for MPEG immersive video, <https://gitlab.com/mpeg-i-visual/tmiv>.

[M7] Reference view synthesizer, <https://gitlab.com/mpeg-i-visual/rvs>.

[M8] Open Realtime Depth Image Based Renderer (OpenDIBR), <https://github.com/IDLabMedia/open-dibr>.

[M9] A. Dziembowski, D. Mieloch, J. Stankowski and A. Grzelka, "IV-PSNR – the objective quality metric for immersive video applications," in IEEE Transactions on Circuits and Systems for Video Technology, doi: [10.1109/TCSVT.2022.3179575](https://doi.org/10.1109/TCSVT.2022.3179575), software: <https://gitlab.com/mpeg-i-visual/ivpsnr>

[M10] Quality Metrics for Immersive Video (QMIV), <https://gitlab.com/mpeg-i-visual/ivpsnr>.

[M11] Gerhard Tech, Ying Chen, Karsten Müller, Jens-Rainer Ohm, Anthony Vetro, Ye-Kui Wang, Overview of the Multiview and 3D Extensions of High Efficiency Video Coding, IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 1, January 2016.

[DM-1] MPEG 136, CfP for Dynamic Mesh Coding, <https://www.mpeg.org/wp-content/uploads/mpeg_meetings/136_OnLine/w20972.zip>

[DM-2] MPEG 134, Use cases for Mesh Coding, <https://www.mpeg.org/wp-content/uploads/mpeg_meetings/134_OnLine/w20364.zip>

[DM-3] Y. Choi, J. -B. Jeong, S. Lee and E. -S. Ryu, "Overview of the Video-based Dynamic Mesh Coding (V-DMC) Standard Work," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 578-581, doi: 10.1109/ICTC55196.2022.9952734.

[DM-4] Information technology - Coding of audio-visual objects - Part 16: Animation Framework eXtension (AFX), ISO/IEC 14496-16.

[DM-5] Mammou, K., Kim, J., Tourapis, A. M., Podborski, D., & Flynn, D. (2022, September). Video and subdivision based mesh coding. In 2022 10th European Workshop on Visual Information Processing (EUVIP) (pp. 1-6). IEEE.

[DM-6] HS, Yang. and X. de Foy, "RTP Payload for V-DMC", Work in Progress, Internet-Draft, draft-hsyang-avtcore-rtp-vdmc-00, 18 October 2024, <https://www.ietf.org/id/draft-hsyang-avtcore-rtp-vdmc-00.html>.

[DM-7] ISO/IEC 12113:2022, Information technology — Runtime 3D asset delivery format — Khronos glTF™ 2.0, International Organization for Standardization, 2022.

[DM-8] Dynamic Mesh Documentation - Unigine Developer. Available at: https://developer.unigine.com/en/docs/latest/objects/objects/mesh\_dynamic/ (Accessed: 19 February 2025).

[DM-9] OWLII Dynamic Human Textured Mesh Sequence Dataset MPEG Point Cloud Compression. Available at: https://mpeg-pcc.org/index.php/pcc-content-database/owlii-dynamic-human-textured-mesh-sequence-dataset/ (Accessed: 19 February 2025).

[DM-10] Pagés, Rafael & Amplianitis, Konstantinos & Ondrej, Jan & Zerman, Emin & Smolic, Aljosa. (2022). Volograms & V-SENSE Volumetric Video Dataset. 10.13140/RG.2.2.24235.31529/1.

[DM-11] Q. Yang, J. Jung, T. Deschamps, X. Xu and S. Liu, "TDMD: A Database for Dynamic Color Mesh Quality Assessment Study," in IEEE Transactions on Visualization and Computer Graphics, doi: 10.1109/TVCG.2024.3451526.

[DM-12] M. Corsini, E. D. Gelasca, T. Ebrahimi, and M. Barni, “Water marked 3-d mesh quality assessment,” IEEE Trans. Multimedia, vol. 9, no. 2, pp. 247–256, 2007.

[DM-13] FTorkhani, K. Wang, and J.-M. Chassery, “Perceptual quality assessment of 3d dynamic meshes: Subjective and objective studies,” Signal Processing: Image Communication, vol. 31, pp. 185–204, 2015.

[DM-14] B. ITU-R RECOMMENDATION, “Methodology for the subjective assessment of the quality of television pictures,” International Telecommunication Union, 2002.

[DM-15] P. ITU-T RECOMMENDATION, “Subjective video quality assessment methods for multimedia applications,” International Telecommunication Union, 1999.

[DM-16] Y. Nehm´ e, F. Dupont, J.-P. Farrugia, P. Le Callet, and G. Lavou´ e, “Visual quality of 3d meshes with diffuse colors in virtual reality: Subjective and objective evaluation,” IEEE Trans. Visualization and Computer Graphics, vol. 27, no. 3, pp. 2202–2219, 2020.

[DM-17] ITU-T, “Subjective test method for interactive virtual reality applications,” https://www.itu.int/ITU-T/workprog/wp item.aspx?isn= 17817.

[DM-18] MPEG, mpeg-pcc-mmetric, <https://github.com/MPEGGroup/mpeg-pcc-mmetric>

[DM-19] MPEG, Representative Renderer, <https://github.com/MPEGGroup/mpeg-3dg-renderer>

[DM-20] ISO/IEC 23090-29 Video-based dynamic mesh coding (V-DMC)

[N1] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 2021. NeRF: representing scenes as neural radiance fields for view synthesis. Commun. ACM 65, 1 (January 2022), 99–106. <https://doi.org/10.1145/3503250>

[N2] RABBY, AKM SHAHARIAR AZAD and Chengcui Zhang. “BeyondPixels: A Comprehensive Review of the Evolution of Neural Radiance Fields.” ArXiv abs/2306.03000 (2023): n. pag.

[N3] Daniel Duckworth, Peter Hedman, Christian Reiser, Peter Zhizhin, Jean-François Thibert, Mario Lučić, Richard Szeliski, and Jonathan T. Barron. 2024. SMERF: Streamable Memory Efficient Radiance Fields for Real-Time Large-Scene Exploration. ACM Trans. Graph. 43, 4, Article 63 (July 2024), 13 pages. https://doi.org/10.1145/3658193

[N4] Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding. ACM transactions on graphics (TOG), 41(4), 1-15.

[N5] Gao, Kyle et al. “NeRF: Neural Radiance Field in 3D Vision, A Comprehensive Review.” (2022).

[N6] G. Lafruit, Y. Liao, and G. Bang, “AhG on Implicit Neural Video Representations (INVR),” ISO/IEC JTC1/SC 29/WG04, M60641, Oct. 2022.G. Lafruit, Y. Liao, and G. Bang, “AhG on Implicit Neural Video Representations (INVR),” ISO/IEC JTC1/SC 29/WG04, M60641, Oct. 2022

[N7] Li, Sicheng et al. “NeRFCodec: Neural Feature Compression Meets Neural Radiance Fields for Memory-Efficient Scene Representation.” ArXiv abs/2404.02185 (2024): n. pag.

[N8] Dong-Ha Kim, Jun Young Jeong, Gwangsoon Lee, and Jae-Gon Kim "Compression method of NeRF model using NNC and VVC", Proc. SPIE 13164, International Workshop on Advanced Imaging Technology (IWAIT) 2024, 131642V (2 May 2024); https://doi.org/10.1117/12.3019533

[LF-1] Gershun A (1939) The light field. Moscow, 1936. Translated by Moon P, Timoshenko G in J Math Phys XVIII:51–151

[LF-2] Marc Levoy and Pat Hanrahan. 1996. Light field rendering. In Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (SIGGRAPH '96).

[LF-3] Google Light Stage X4, <https://www.fxguide.com/fxfeatured/exclusive-paul-debevec-and-the-light-stage-research-at-google/>

[LF-4] USC Lightstage X6, <https://vgl.ict.usc.edu/Data/LightStage/>

[LF-5] Zhou, Taotao, et al. "Relightable neural human assets from multi-view gradient illuminations." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.

[LF-6] Michael Broxton, John Flynn, Ryan Overbeck, Daniel Erickson, Peter Hedman, Matthew Duvall, Jason Dourgarian, Jay Busch, Matt Whalen, and Paul Debevec. 2020. Immersive light field video with a layered mesh representation. ACM Trans. Graph. 39, 4, Article 86 (August 2020), 15 pages. https://doi.org/10.1145/3386569.3392485

[LF-7] M. B. de Carvalho et al., "A 4D DCT-Based Lenslet Light Field Codec," 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 2018, pp. 435-439, doi: 10.1109/ICIP.2018.8451684.

[LF-8] ISO/IEC JTC 1/SC 29/WG 04, w24268, “Overview of lenslet video coding activities”, MPEG 147, Sapporo.

[LF-9] Daniel, Jamison R. et al. “Initial work on development of an open Streaming Media Standard for Field of Light Displays (SMFoLD).” SD&A (2018).

[LF-10] X. Min, J. Zhou, G. Zhai, P. Le Callet, X. Yang and X. Guan, "A Metric for Light Field Reconstruction, Compression, and Display Quality Evaluation," in IEEE Transactions on Image Processing, vol. 29, pp. 3790-3804, 2020, doi: 10.1109/TIP.2020.2966081.

[LF-11] ISO/IEC JTC 1/SC29/WG1 N100306, REQ "Use Cases and Requirements for Light Field Quality Assessment v5.0", 97th JPEG Meeting, Online, October 2022.

[LF-12] What you should know about light-field, <https://www.digitalmedia.fraunhofer.de/en/mediainformation/trendbrochures/trendbrochure-2021/what-you-should-know-about-light-field-.html>

[LF-13] "A Visual Introduction to the Past, Present, and Future of Light Field Technology", https://cubicleninjas.com/light-field-technology/

[LF-14] Ruben Verhack, "AI-Driven Breakthroughs in Image-Based Rendering: Light Fields, SMoE, Gaussian Splatting, NeRFs and beyond", <https://cubicleninjas.com/light-field-technology/>, Tech Posts Computer Graphics, <https://blog.datameister.ai/ai-driven-breakthroughs-image-based-rendering>, February 2024.

[GS-1] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkuehler, and George Drettakis. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Trans. Graph. 42, 4, Article 139 (August 2023), 14 pages. https://doi.org/10.1145/3592433

[GS-2] Wei, Meng et al. “Normal-GS: 3D Gaussian Splatting with Normal-Involved Rendering.” ArXiv abs/2410.20593 (2024): n. pag.

[GS-3] Özyeşil, Onur, et al. "A survey of structure from motion\*." Acta Numerica 26 (2017): 305-364.

[GS-5] T. Wu, Y.-J. Yuan, L.-X. Zhang, J. Yang, Y.-P. Cao, L.-Q. Yan, and L. Gao, “Recent advances in 3d gaussian splatting,” Computational Visual Media, pp. 1–30, 2024.

[GS-6] X. Lei, M. Wang, W. Zhou, and H. Li, “Gaussnav: Gaussian splatting for visual navigation,” arXiv preprint arXiv:2403.11625, 2024.

[GS-7] 3DGS.zip: A survey on 3D Gaussian Splatting Compression Methods, https://3dgs.zip/

[GS-8] Dalal, Anurag et al. “Gaussian Splatting: 3D Reconstruction and Novel View Synthesis: A Review.” IEEE Access 12 (2024): 96797-96820.

[GS-9] Dalal, Anurag & Hagen, Daniel & Robbersmyr, Kjell & Knausgård, Kristian. (2024). Gaussian Splatting: 3D Reconstruction and Novel View Synthesis, a Review. 10.48550/arXiv.2405.03417.

[GS-10] Nicolas Moenne-Loccoz, Ashkan Mirzaei, Or Perel, Riccardo de Lutio, Janick Martinez Esturo, Gavriel State, Sanja Fidler, Nicholas Sharp and Zan Gojcic; “3D Gaussian Ray Tracing: Fast Tracing of Particle Scenes”, ACM Transactions on Graphics and SIGGRAPH Asia, 2024.

[GS-11] Zhu, Huixin et al. “Scene reconstruction techniques for autonomous driving: a review of 3D Gaussian splatting.” Artif. Intell. Rev. 58 (2024): 30.

[GS-12] Jorge L. Charco, Angel D. Sappa, Boris X. Vintimilla, and Henry O. Velesaca. 2021. Camera pose estimation in multi-view environments: From virtual scenarios to the real world. Image Vision Comput. 110, C (Jun 2021). <https://doi.org/10.1016/j.imavis.2021.104182>

[GS-13] Huang, Zhentao and Minglun Gong. “Textured-GS: Gaussian Splatting with Spatially Defined Color and Opacity.” ArXiv abs/2407.09733 (2024): n. pag.

[GS-14] ISO/IEC 23090-14:2024/Amd.1:2025. "ISO/IEC 23090-14 2nd edition DAM 1 Support of MPEG-I immersive audio, scene understanding and other extensions"

[LS-1] Wang, Y., Lu, Z., Cao, P. et al. How Live Streaming Changes Shopping Decisions in E-commerce: A Study of Live Streaming Commerce. Comput Supported Coop Work 31, 701–729 (2022). <https://doi.org/10.1007/s10606-022-09439-2>

[LS-2] Xie, Junyuan et al. “Deep3D: Fully Automatic 2D-to-3D Video Conversion with Deep Convolutional Neural Networks.” European Conference on Computer Vision (2016).

[LS-3] Dumic, E. et al.. “Transmission of 3D Video Content. In: Assunção, P., Gotchev, A. (eds) 3D Visual Content Creation, Coding and Delivery.” Signals and Communication Technology (2019). Springer, Cham. https://doi.org/10.1007/978-3-319-77842-6\_8

[LS-4] Schierl, Thomas and Sam Narasimhan. “Transport and Storage Systems for 3-D Video Using MPEG-2 Systems, RTP, and ISO File Format.” Proceedings of the IEEE 99 (2011): 671-683.

[LS-5] 3GPP TR 26.905 V 18.0.0: "Mobile stereoscopic 3D video"

[LS-6] A. Banitalebi-Dehkordi, M. T. Pourazad and P. Nasiopoulos, "A human visual system-based 3D video quality metric," 2012 International Conference on 3D Imaging (IC3D), Liege, Belgium, 2012, pp. 1-5, doi: 10.1109/IC3D.2012.6615146.

[LS-7] Banitalebi-Dehkordi, Amin, Mahsa T. Pourazad, and Panos Nasiopoulos. "An efficient human visual system based quality metric for 3D video." Multimedia Tools and Applications 75, no. 8 (2016): 4187-4215.

[LS-8] Recommendation ITU-R BT.500-13, “Methodology for the subjective assessment of the quality of the television pictures”, 2012.

[LS-9] ISO/IEC JTC1/SC29/WG11 (MPEG), Document N12036, “Call for proposals on 3D video coding technology,” 96th MPEG meeting, Geneva, March 2011.

[LS-10] Q. Hyunh-Thu, P. L. Callet, and M. Barkowsky, “Video quality assessment: from 2D to 3D challenges and future trends,” IEEE 17th International Conference on Image Processing, (ICIP), pp.4025-4028, 2010.

[LS-11] Wei Bao, Wei Wang, Yuhua Xu, Yulan Guo, Siyu Hong, Xiaohu Zhang. InStereo2K: A large real dataset for stereo matching in indoor scenes. SCIENCE CHINA Information Sciences. 2020.

[LS-12] D. Scharstein, H. Hirschmüller, Y. Kitajima, G. Krathwohl, N. Nesic, X. Wang, and P. Westling. High-resolution stereo datasets with subpixel-accurate ground truth. In German Conference on Pattern Recognition (GCPR 2014), Münster, Germany, September 2014

[LS-13] HTM Codec Software, version 16.3, https://hevc.hhi.fraunhofer.de/svn/svn\_3DVCSoftware/branches/HTM-16.3-fixes/cfg/MV-HEVC/

[LS-14] V. Baroncini, K. Müller, and S. Shinya (editors), "MV-HEVC verification test report", JCT3V-N1001, May 2016.

[LS-15] VQEG 3DTV Group, Test Plan for Evaluation of Video Quality Models for Use with Stereoscopic Three-Dimensional Television Content, 2012

[LS-16] Recommendation ITU-R P.914, “Display requirements for 3D video quality assessment ”, 2016.

[LS-17] Recommendation ITU-R P.915, “Subjective assessment methods for 3D video quality”, 2016.

[LS-18] Recommendation ITU-R P.916, “Information and guidelines for assessing and minimizing visual discomfort and visual fatigue from 3D video”, 2016.

[LS-19] Guodong Chen, Sizhe Wang, Jacob Chakareski, Dimitrios Koutsonikolas, and Mallesham Dasari. 2025. Spatial Video Streaming on Apple Vision Pro XR Headset. In Proceedings of the 26th International Workshop on Mobile Computing Systems and Applications (HotMobile '25). Association for Computing Machinery, New York, NY, USA, 115–120. https://doi.org/10.1145/3708468.3711878

[LS-20] Sizhe Wang, Mingkun Liu, Mallesham Dasari, and Dimitrios Koutsonikolas. 2024. A First Look at Apple’s Stereoscopic Video and its Potential in Live Video Streaming for XR Headsets. In Proceedings of the 30th Annual International Conference on Mobile Computing and Networking (ACM MobiCom '24). Association for Computing Machinery, New York, NY, USA, 1617–1619. https://doi.org/10.1145/3636534.3697437

[Vol-1] KDDI, Transmission experiment using real-time codec compliant with the latest international standard of point cloud compression, https://www.kddi-research.jp/english/newsrelease/2023/012401.html

[Vol-2] SBTVD Forum, Brazilian Forum for digital terrestrial TV, https://forumsbtvd.org.br/tv3\_0/

[Vol-3] VFA, The Volumetric Format Association, https://www.volumetricformat.org/

[Vol-4] Arcturus, on-demand streaming of volumetric video, https://arcturus.studio/stream/

[Vol-5] 8i, stream volumetric video to any device, browser, or VR/AR headset, https://8i.com/stream/

[Vol-6] Broadpeak, V3C standardized content distribution at scale, https://broadpeak.tv/newsroom/mpeg-v3c-standardized-content-distribution-at-scale/

[Vol-7] Arcturus, playback solution with broad delivery capabilities, https://arcturus.studio/playback/

[Vol-8] 5G-MAG, V3C Immersive platform, https://5g-mag.github.io/Getting-Started/pages/v3c-immersive-platform/

[Vol-9] DVB project, DVB Study Mission Report S101, DVB Bluebook S101

[Vol-10] Ultra Video Group, Voxelized Point Cloud Dataset for Visual Volumetric Video-based Coding, https://ieeexplore.ieee.org/document/10178589

[Vol-11] ISO/IEC 23090-5 Visual Volumetric Video-based Coding (V3C) and Video-based Point Cloud Compression (V-PCC) – 2nd edition

[Vol-12] ISO/IEC 23090-10 Carriage of visual volumetric video-based coding data – 1st edition

[Vol-13] By C. Guede et al., IBC 2023 Tech Papers, https://www.ibc.org/technical-papers/ibc2023-tech-papers-efficient-delivery-and-rendering-on-client-devices-via-mpeg-i-standards-for-emerging-volumetric-video-experiences/10277.article

[Vol-14] Futuresource Consulting, Spotlight on HEVC, <https://www.interdigital.com/white_papers/spotlight-on-hevc-the-codec-of-choice-for-the-video-streaming-industry>

[Vol-15] Renderpeople, https://renderpeople.com/

[Vol-16] Volucap, https://volucap.com/

[Vol-16] MPEG WG7, Call for Proposals for Point Cloud Compression V2, https://mpeg.chiariglione.org/standards/mpeg-i/point-cloud-compression/call-proposals-point-cloud-compression-v2.html

[Vol-17] PCQM: A Full-Reference Quality Metric for Colored 3D Point Clouds, Gabriel Meynet, Yana Nehmé, Julie Digne, Guillaume Lavoué, https://hal.science/hal-02529668/document

[Vol-18] MPEG, mpeg-pcc-mmetric V1\_1\_7, https://github.com/MPEGGroup/mpeg-pcc-mmetric

[Vol-19] MPEG, Representative Renderer release 8.0, https://github.com/MPEGGroup/mpeg-3dg-renderer

[Vol-20] MPEG, Subjective verification test report for V-PCC, <https://www.mpeg.org/wp-content/uploads/mpeg_meetings/136_OnLine/w20992.zip>

[Vol-21] ISO/IEC 23090-10 Carriage of visual volumetric video-based coding data – 1st edition

[Vol-22] [XD Productions, https://xdprod.com/](https://xdprod.com/)

[Vol-23] Renderpeople free 4D People sample: <https://4dpeople.com/samples/>

[Vol-24] Renderpeople 4D People catalogue: <https://4dpeople.com/>

[Vol-25] [Ultra Video Group – UVG-VPC Dataset, https://ultravideo.fi/UVG-VPC/index.html](https://ultravideo.fi/UVG-VPC/index.html)

[Vol-26] MPEG, V-PCC test model tmc2 release R25.0, https://github.com/MPEGGroup/mpeg-pcc-tmc2

[Vol-27] MPEG, Random Access configuration, https://github.com/MPEGGroup/mpeg-pcc-tmc2/blob/master/cfg/

[Vol-28] MPEG, Subjective verification test report for V-PCC, https://www.mpeg.org/wp-content/uploads/mpeg\_meetings/136\_OnLine/w20992.zip

[Vol-29] MPEG, Rate configuration, https://github.com/MPEGGroup/mpeg-pcc-tmc2/tree/master/cfg/rate

[Vol-30] List of camera paths and splat blend parameter options for each tested sequence

[Vol-31] SBTVD, SBTVD TV 3.0 test report for video, https://forumsbtvd.org.br/wp-content/uploads/2021/12/SBTVD-TV\_3\_0-VC-Report.pdf

[Vol-32] https://www.kddi-research.jp/newsrelease/2025/020601.html

[Vol-33] CfP for Dynamic Mesh Coding, https://www.mpeg.org/wp-content/uploads/mpeg\_meetings/136\_OnLine/w21000.zip

[Vol-34] Ilola, L., Kondrad, L., Schwarz, S., & Hamza, A. (2022). An overview of the MPEG standard for storage and transport of visual volumetric video-based coding. Frontiers in Signal Processing, 2, 883943.

[Vol-35] Jing, Liqi. “Research on Video-based Dynamic Mesh Compression Technology and Proposed Improvements.” Advances in engineering research/Advances in Engineering Research, 2024, pp. 47–57, doi:10.2991/978-94-6463-518-8\_6.

[clause5-1] ITU-T Recommendation H.264 (08/2021): "Advanced video coding for generic audiovisual services".

[clause5-2] ITU-T Recommendation H.265 (08/2021): "High efficiency video coding".

[MD-4] ISO/IEC 23090-10:2022 (Amd1), “Information Technology — Coded Representation of Immersive media — Part 10: Carriage of Visual Volumetric Video-Based Coding Data”

[MD-5] Guede et al., IBC 2023, “Efficient Delivery and Rendering on Client Devices via MPEG-I Standards for Emerging Volumetric Video Experiences”. https://www.ibc.org/technical-papers/ibc2023-tech-papers-efficient-delivery-and-rendering-on-client-devices-via-mpeg-i-standards-for-emerging-volumetric-video-experiences/10277.article

[MD-6] Dziembowski, B. Kroon, J. Jung (Eds.), Common test conditions for MPEG immersive video, ISO/IEC JTC 1/SC 29/WG 04 N 0372, July 2023, Geneva.

[MD-7] D. Mieloch (Ed.), Verification test report of MPEG immersive video, ISO/IEC JTC 1/SC 29/WG 04 N 0341, April 2023, Antalya.

[MD-8] B. Brand, Michel Bätz, Joachim Keinert, Camorph: a toolbox for conversion between camera parameter conversions, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, url: https://github.com/Fraunhofer-IIS/camorph, 2022.

[MD-9] ISO/IEC 14496-15:2024: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format.

[MD-10] ISO/IEC 23090-12:2023: MPEG immersive video.

[MD-11] ITU-T H.265:2024 | ISO/IEC 23008-2:2024, Annex G: Multiview high efficiency video coding (MV-HEVC).

\* \* \* Second Change (new) \* \* \* \*

4.3.6.3.6 glTF as a Gaussian Splat format

Gaussian Splats (GS) provide an efficient representation for static and dynamic 3D scenes by compactly encoding local geometry and view-dependent appearance. To address interoperability, backward compatibility, and progressive rendering in various applications, integrating GS into the glTF (GL Transmission Format) standard [DM-7] is essential.

A structured method to store GS in glTF 2.0, supporting legacy devices, progressive downloads, and dynamic content via MPEG extensions is available in ISO/IEC 23090-14:2024/Amd.1:2025 [GS-14]. Gaussian Splats are proposed to be stored in glTF mainly using application-specific attributes, ensuring backward compatibility by enabling legacy receivers to interpret data as a traditional point cloud representation showing the base color for each splat.

Each Gaussian splat is defined by attributes derived from the INRIA Gaussian Splat PLY format, mapped directly to glTF attributes according to Table 4.3.6.3.6-1.

**Figure 4.3.6.3.6-1 Mapping of GS attributes to glTF primitive attributes via MPEG-SD-GS extension**

|  |  |
| --- | --- |
| GS Property | Corresponding glTF Attribute primitive via MPEG SD extensions |
| x, y, z | POSITION |
| f\_dc\_[0-2] | COLOR\_n |
| opacity | Alpha channel of COLOR\_n |
| rot\_[0-3] | \_\_MPEG\_GS\_ORIENTATION  (x,y,z,w) |
| scale\_[0-2] | \_MPEG\_GS\_SCALE |
| f\_rest\_[0-14] | \_MPEG\_GS\_SH\_COEFF\_R (R channel SH coeffs) |
| f\_rest\_[15-29] | \_MPEG\_GS\_SH\_COEFF\_G (G channel SH coeffs) |
| f\_rest\_[30-44] | \_MPEG\_GS\_SH\_COEFF\_B (B channel SH coeffs) |

The Gaussian Splats extension is added to glTF 2.0 primitive elements, explicitly supporting both static and dynamic Gaussian Splats leveraging MPEG timed media extensions. The primitive mode is set to 0 (POINTS), with the COLOR\_n attribute referencing Vec4 type, incorporating opacity.

To facilitate progressive download and rendering, Gaussian Splat attributes are structured hierarchically in buffer views according to significance and detail level:

- Initial buffer views store POSITION attributes followed immediately by COLOR\_n attributes, providing an initial coarse representation suitable for immediate visualization.

- Subsequent buffer views contain ORIENTATION and SCALE attributes.

- Spherical harmonics (SH) attributes for color are grouped by spherical harmonic order, enabling progressive refinement:

- Level-of-Detail 0 (LoD0): 3 base color components in COLOR\_n.

- Level-of-Detail 1 (LoD1): 1st-order SH, resulting in 9 coefficients.

- Level-of-Detail 2 (LoD2): 2nd-order SH, resulting in additional 15 coefficients.

- Level-of-Detail 3 (LoD3): 3rd-order SH, resulting in additional 21 coefficients.

Different attribute sets support progressive refinement:

- \_MPEG\_GS\_SH\_COEFF\_FIRST provides the 9 coefficients of 1st-order SH.

- \_MPEG\_GS\_SH\_COEFF\_SECOND provides the 15 coefficients of 2nd-order SH.

- \_MPEG\_GS\_SH\_COEFF\_THIRD provides the 21 coefficients of 3rd-order SH.

This hierarchical data organization enables efficient progressive streaming and immediate visual feedback on the receiver's device.

Reference tools will be developed for Gaussian Splats in glTF and scene description, providing comprehensive support for both static and dynamic Gaussian Splats. These reference tools will include utilities for converting INRIA PLY files into glTF or standardized MPEG scene descriptions extensions. Additionally, a renderer for the glTF GS format, potentially as an enhancement to the existing MPEG-3D renderer will also be made available in August 2025.

\* \* \* Third Change \* \* \* \*

4.3.6.3.7 Benefits and Limitations

4.3.6.3.7.1 Benefits

- Real-time Rendering with GPU acceleration.

- Accurate Reconstruction, it can capture the geometry accurately

- Explicit representation

- Ability to render complex scenes in real-time

- Interpretability of the representation, an explanation of the mathematical mechanism, i.e., the working principle, of 3DGS can help researchers analyze the complex relationships in 3D scene reconstruction technology and reveal the performance characteristics of 3DGS in depth. [GS-11]

- Gaussian Splatting can deliver high-quality, real-time visualizations.[GS-9].

- Gaussian Splatting has evolved to handle dynamic and deformable objects [GS-9].

- Gaussian Splatting can be applicable to various application space, such as digital avatars and SLAM [GS-9].

- Gaussian Splats can be stored in glTF 2.0, supporting legacy devices, progressive downloads, and dynamic content.

- Gaussian Splats facilitate progressive download and rendering, if GS attributes are structured hierarchically in buffer views according to significance and detail level as proposed for the glTF extension [GS-14].

4.3.6.3.7.2 Limitations

- There is a lack of industry agreement on the 3DGS format(s), due to no stable representation and compression format exists for static and dynamic 3DGS.

- Static and Dynamic 3DGS formats is evolving, multiple options are considered in current academic and industrial research. For dynamic, such research include modeling in 4 dimensions (i.e. temporal), time evolving 3DGS, and MLP predicted motion for 3DGS among others.

- High memory usage

- Not yet fully compatible with existing rendering pipelines

- Computation complexity [GS-9], the computational demands of handling large numbers of splats, especially for high-resolution rendering and complex scenes; it requires to process large datasets for training, which can also be time-consuming and resource-intensive.

- Edge artifacts [GS-9].

\* \* \* End of Changes \* \* \* \*