INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 7
MPEG CODING FOR 3D GRAPHICS AND HAPTICS

ISO/IEC JTC 1/SC 29/WG 7 m78302
October 2026, Hangzhou

Title: [GSC][JEE6.1] Proposal of Campus Sequence for Large-Scale Scene

Author: GyeongHo Choi, Seung-Sik Choi, Geon Hee Han, Seung-Mi Lee, Jong-Beom Jeong (Kunsan National University, KSNU)

Abstract

This contribution introduces the Campus dataset as a large-scale Gaussian splatting content designed for the purpose of evaluating and advancing GSC activities. Captured using drone imagery, the dataset focuses on the campus of Kunsan National University (KSNU) in South Korea, particularly centered around the Digital Information Building and its surrounding area. The scene was reconstructed and trained using gsplat. Subjective and objective quality analyses are presented based on rendered views and objective measurements.

1 Introduction

MPEG GSC JEE 6.1 was established as a collaborative activity between WG 4, WG 5/JVET, and WG 7 with the objective of preparing datasets for the GSC exploration [1]. According to Mandate 2 of the JEE 6.1, there is an ongoing solicitation for new test material to address underrepresented use cases, with a particular emphasis on large-scale content. Furthermore, Mandate 4 requires the analysis of newly proposed content to recommend candidate materials for potential integration into the common test conditions (CTC).

In response to these mandates and the associated call for content [2], this contribution introduces a new large-scale dataset named the ***Campus*** sequence. Designed to reflect real-world complexity, this dataset addresses the need for extensive outdoor test material in campus level. By providing this sequence, we aim to support the generation of I-3DGS models and facilitate reliable evaluations of GSC technologies on large-scale environments, directly contributing to the objectives of the JEE 6.1 data preparation activities.

2 Capturing Setup

To construct a robust large-scale 3D representation, visual data collection was conducted centering on the Digital Information Building at KSNU. We utilized a DJI Mavic 4 Pro unmanned aerial vehicle (UAV) for the photographic survey. Specifically, the onboard telephoto lens—featuring a 1/1.5-inch 50-megapixel CMOS sensor, an f/2.8 aperture, and a 168 mm equivalent focal length—was operated to record all imagery in 4K ultra-high definition.

Figure 1 describes the acquisition setup. The flight trajectory was deliberately designed as expansive circular orbits surrounding the main architectural target. These orbital paths were executed at two distinct elevation levels of 50 m and 70 m to ensure comprehensive multi-angle coverage of the structure. Furthermore, dense foliage and surrounding trees inevitably cast shadows and blocked certain viewing angles. To compensate for these occlusions, supplementary shots were specifically gathered in those shaded zones to guarantee complete geometric and textural integrity.

A notable characteristic of this sequence is its extensive exterior glass windows and rooftop solar panels. These architectural materials produce intense specular highlights and severe reflections during daylight. Consequently, this dataset naturally encompasses highly complex non-Lambertian surfaces, making it an exceptionally challenging and valuable asset for evaluating view-dependent rendering capabilities.

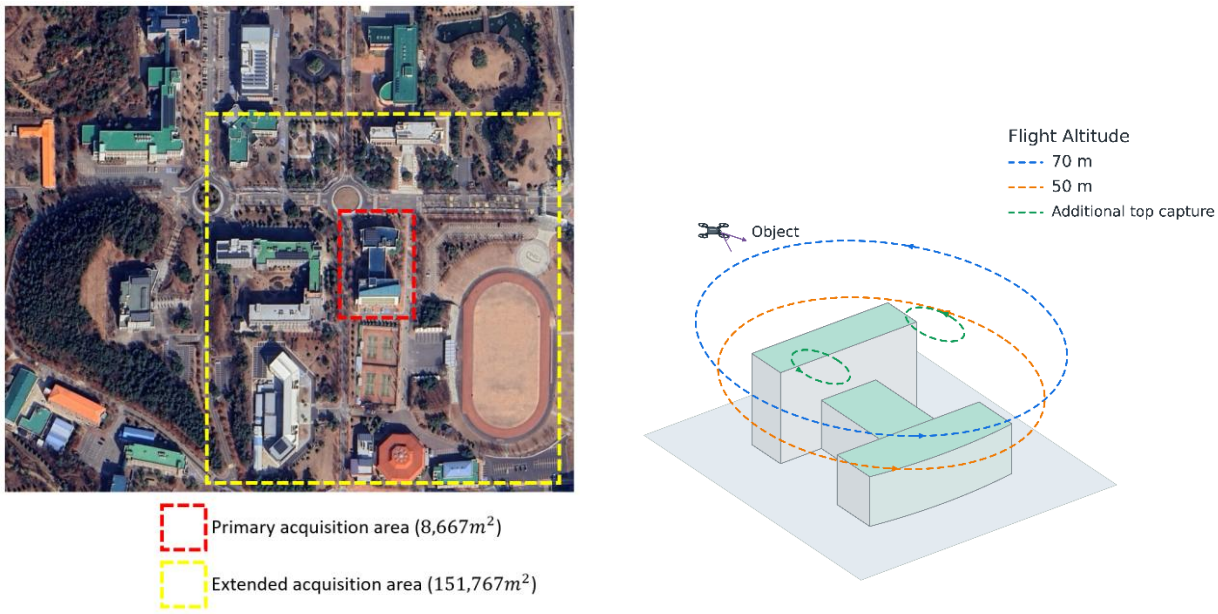
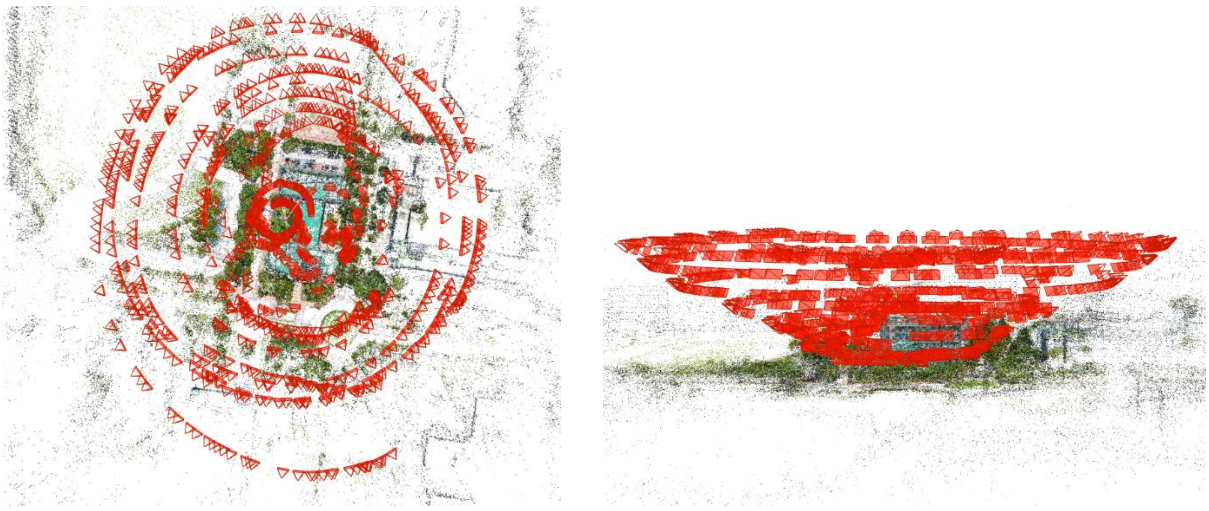


Figure 1. Acquisition setup. (a) primary and extend acquisition areas, (b) drone flight path

A total of 900 images of 3840x2160 resolution were extracted from the drone-captured video and utilized for Colmap processing. Figure 2 illustrates the Colmap reconstruction results. The global mapper configuration was applied [3], generating a total of 477,050 points. Figure 3 presents snapshots of the *Campus* sequence. Thirty images were selected from the total of 900 images for visualization, and this combination was also utilized in the subset setting described in the following section.



(a) (b)

Figure 2. Colmap of *Campus* sequence. (a) top view, (b) side view



Figure 3. Snapshot images of *Campus* sequence (30 out of 900).

3 Gaussian Splats

Following the experimental conditions of the *Library* sequence, the Gaussian splats were generated using gsplat v1.4.0 and trained for 60,000 iterations [4]. Two RTX 5090 GPUs were

utilized for the training, resulting in a total of 4,318,404 Gaussians. Figure 4 to 7 describe noticeable enlarged sections of ground-truth image, rendered image, and error map with PSNR values. Figure 4 (View 8) and Figure 5 (View 16) show the rooftop solar panels and tree areas, respectively, achieving relatively high PSNRs of 25.40 dB and 25.52 dB despite their complex patterns. In contrast, Figure 6 (View 232) displays the building's exterior wall with glass windows and trees, showing a drop in PSNR to 21.57 dB. As stated in Section 2, the error map confirms that rendering errors significantly increased due to the strong non-Lambertian reflections from the glass windows and the geometric complexity of the foliage. It is worth noting that the View 232 contains extended acquisition area, and the PSNR values were still over 20.00dB. Figure 7 (View 648) also records a PSNR of 22.20 dB, with high rendering errors occurring around the lawn and landscaped areas cast with tree shadows. On average, a PSNR of 23.74 dB, an SSIM of 0.7511, and an LPIPS of 0.2819 were observed across all views.

Figure 8 shows the distribution of RGB-PSNR values across all view indices. It can be observed that the rendering quality varies significantly depending on the viewpoint. The PSNR on the graph exhibits a distinct deviation ranging from approximately 20 dB to 27 dB. This variation is primarily caused by illumination inconsistencies, shadow effects, and strong view-dependent characteristics induced by materials such as glass windows and solar panels. Therefore, this viewpoint-dependent quality variation should be considered as a crucial factor for evaluation.



Figure 4. Ground-truth image, rendered image, and error map for view index 8 (PSNR: 25.40dB).



Figure 5. Ground-truth image, rendered image, and error map for view index 16 (PSNR: 25.52dB).



Figure 6. Ground-truth image, rendered image, and error map for view index 232 (PSNR: 21.57dB).



Figure 7. Ground-truth image, rendered image, and error map for view index 648 (PSNR: 22.20dB).

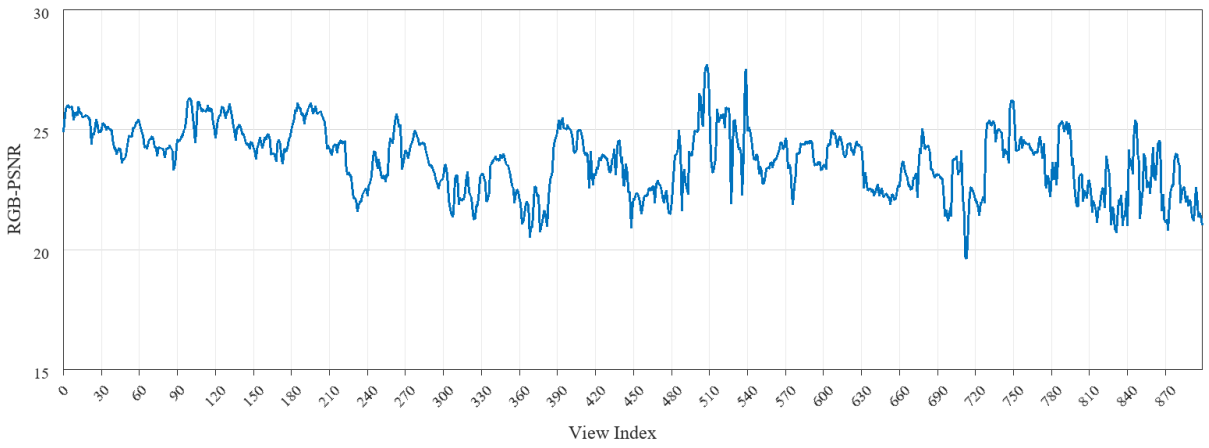


Figure 8. RGB-PSNR per view index.

4 Pose Trace

To support performance evaluation and subjective evaluation at arbitrary views, pose traces are provided for the *Campus* sequence. A camera path encompassing the entire scene was first generated using SuperSplat, and the camera poses were subsequently recovered by localizing this path within the Colmap results. These poses were then transformed to match the gsplat coordinate system, and smoothing as well as inter-frame pose interpolation were applied to generate the final pose traces. Figure 9 demonstrates the pose trace, including camera trajectory and a rendered view. This contribution provides mpeg-3d-renderer-compliant pose trace file, with the rendered pose trace using the renderer.



(a) **(b)**

Figure 9. (a) Visualization of pose trace paths, (b) rendered pose trace

5 Coded Results with GSC VSEI, G-PCC, and V-PCC

To evaluate the coding performance of the proposed sequence, experiments were conducted based on JEE 6.2 [5]. The used softwares are as follows: (i) GSC VSEI reference framework version GSC-2.0 [6], (ii) G-PCC test model (TM) version release-v3.0-rc2 [7], and (iii) V-PCC TM version release-27.0.1 [8].

For all softwares, the C1 test conditions were used. For evaluation, the mpeg-gsc-metrics version release_mpeg155 was applied [9]. For GSC VSEI, the evaluation was performed using the HEVC test model (HM). To reduce the computational complexity of the evaluation, the *Campus* sequence provides two camera parameters: full view and subset. The full view condition contains all the 900 views, and the subset condition has 30 views. The content is available in MPEG content server [10].

Table 1, 2, 3, and 4 shows the coded results of GSC VSEI, G-PCC, and V-PCC, respectively. Experimental results proved that the Subset condition can represent the Full condition, while reducing the computational complexity. As shown in Table 3 and 4, the bytes between Full and Subset condition are different. This is because the current V-PCC utilizes importance-based Gaussian pruning which uses the camera parameters. Therefore, for Full and Subset conditions, 2.5K and 2.4K Gaussians remain after pruning, respectively.

Table 1. GSC VSEI coded results. Full View: 900 views, Subset: 30 views.

		Full View			Subset			Size Reduction
Rate	**Bytes**	**RGB-PSNR**	**YUV-PSNR**	**YUV-SSIM**	**RGB-PSNR**	**YUV-PSNR**	**YUV-SSIM**	
R1	63,951,494	36.66	40.69	0.9933	36.98	40.91	0.9938	×15.94
R2	53,258,740	34.60	38.72	0.9904	34.81	38.88	0.9911	×19.14
R3	40,926,593	31.42	35.73	0.9836	31.52	35.81	0.9846	×24.90
R4	28,231,660	27.23	31.70	0.9658	27.35	31.80	0.9677	×36.10
R5	17,010,366	21.50	26.13	0.9157	21.67	26.27	0.9186	×59.91
Original	**1,019,144,822**	-	-	-	-	-	-	×1.00

Table 2. G-PCC coded results. Full View: 900 views, Subset: 30 views.

		Full View			Subset			Size Reduction
Rate	**Bytes**	**RGB-PSNR**	**YUV-PSNR**	**YUV-SSIM**	**RGB-PSNR**	**YUV-PSNR**	**YUV-SSIM**	
R1	39,492,556	31.97	36.14	0.9867	32.12	36.20	0.9877	×25.81
R2	62,116,904	37.95	41.90	0.9950	37.87	41.76	0.9954	×16.41
R3	94,616,699	43.18	47.13	0.9977	42.95	46.86	0.9978	×10.77
R4	134,073,294	46.37	50.51	0.9984	45.99	50.16	0.9985	×7.60
Original	**1,019,144,822**	-	-	-	-	-	-	×1.00

Table 3. V-PCC coded results (Full View).

		Full View			Size Reduction
Rate	**Bytes**	**RGB-PSNR**	**YUV-PSNR**	**YUV-SSIM**	
R1	16,614,169	27.06	31.20	0.9641	×61.34
R2	21,032,662	31.27	35.34	0.9807	×48.46
R3	27,613,857	34.88	38.98	0.9887	×36.91
R4	35,920,874	36.19	40.58	0.9917	×28.37
Original	**1,019,144,822**	-	-	-	×1.00

Table 4. V-PCC coded results (Subset).

		Full View			Size Reduction
Rate	**Bytes**	**RGB-PSNR**	**YUV-PSNR**	**YUV-SSIM**	
R1	15,783,113	27.19	31.34	0.9665	×64.57
R2	19,978,824	31.41	35.50	0.9822	×51.01
R3	26,191,913	35.18	39.22	0.9895	×38.91
R4	34,002,094	36.51	40.90	0.9923	×29.97
Original	**1,019,144,822**	**-**	**-**	**-**	×1.00

6 Conclusion

This contribution introduced the *Campus* sequence as a novel large-scale GSC dataset to support the ongoing GSC JEE 6.1 data preparation activities. Centered on the Digital Information Building at KSNU, the dataset was captured via multi-altitude drone flights to encompass complex real-world characteristics, most notably severe non-Lambertian reflections from extensive glass windows and solar panels.

The sequence was successfully reconstructed and trained using gsplat, producing approximately 4.3 million Gaussians. Given its challenging view-dependent properties and large-scale outdoor nature, the *Campus* sequence serves as a highly valuable asset for assessing the scalability, compression efficiency, and rendering robustness of GSC technologies. Therefore, it is recommended to include the *Campus* sequence into the CTC to facilitate further exploration and advancement in future GSC standardization activities.

7 License

Individuals and organizations using this dataset agree that the video sequences, derived images, and all related intellectual property rights remain the property of the respective owners listed below. These materials may only be used for the purpose of developing, testing, and promoting technology standards and for academic research. Additionally, official confirmation has been obtained from Kunsan National University that such usage for education and research purposes is permitted. Acknowledgement and reference to the following document are required in all documents that report any usage of the materials:

G. H. Choi, S. -S. Choi, G. H. Han, S. -M. Lee, J. -B. Jeong. [GSC][JEE6.1] Proposal of Campus Sequence for Large-Scale Scene

ISO/IEC JTC1/SC29/WG7 MPEG 3DGH/m78302, Hangzhou, October 2026.

The respective owners make no warranties with respect to the materials and expressly disclaim any warranties regarding their fitness for any purpose.

Owner of the dataset: Kunsan National University (KSNU), PRISM Lab. (54150) 558, Daehak-ro, Gunsan-si, Jeonbuk-do, South Korea

Contact person:

- GyeongHo Choi (cgh3578@kunsan.ac.kr)
- Jong-Beom Jeong (jongbeomjeong@kunsan.ac.kr)

8 Acknowledgement

This research was supported by the MSIT(Ministry of Science, ICT), Korea, under the National Program for AI university (Kunsan National University) (No. RS-2026-0-00330) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation) in 2026.

9 References

[1] "JEE 6.1 on data preparation", Standard ISO/IEC JTC1/SC29 WG5, document n424, July 2026.

[2] "Call for content for Gaussian splats coding", Standard ISO/IEC JTC1/SC29 WG5, document n422, July 2026.

[3] L. Pan, D. Barath, M. Pollefeys, J. L. Schönberger, "Global Structure-from-Motion Revisited", European Conference on Computer Vision (ECCV), 2024.

[4] R. Koo, Y. G. Kim, I. Yang, S. H. Ahn, E. S. Ryu, H. Lee, G. Bang, "[GSC][JEE6.1 related] Proposals of Library sequence", Standard ISO/IEC JTC1/SC29 WG4, document m74011, October 2025.

[5] "JEE 6.2 on anchor generation", Standard ISO/IEC JTC1/SC29 WG5, document n425, September 2026.

[6] https://vcgit.hhi.fraunhofer.de/jvet-tuc/gsc-vsei-framework.git

[7] https://git.mpeg.expert/MPEG/3dgh/g-pcc/software/tm/mpeg-pcc-tmc13-gs

[8] https://git.mpeg.expert/MPEG/3dgh/v-pcc/software/mpeg-pcc-tmc2

[9] https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/mpeg-gsc-metrics

[10] https://content.mpeg.expert/data/Explorations/GSC/m78302_Campus/