

# A Study on the Use of High Dynamic Range Imaging for Gaussian Splatting Methods: Are 8 bits Enough?

Valentina Piras<sup>1,2</sup>  Amedeo F. Bonatti<sup>3</sup>  Carmelo De Maria<sup>3</sup>  Paolo Cignoni<sup>2</sup>  Francesco Banterle<sup>2</sup> 

<sup>1</sup>Computer Science Department, University of Pisa, Italy

<sup>2</sup>ISTI-CNR, Italy

<sup>3</sup>Department of Information Engineering and Research Center E. Piaggio of the University of Pisa, Pisa, Italy

---

## Abstract

*The recent rise of Neural Radiance Fields (NeRFs)-like methods has revolutionized high-fidelity scene reconstruction, with 3D Gaussian Splatting (3DGS) standing out for its ability to generate photorealistic images while maintaining fast, efficient rendering. 3DGS delivers high-fidelity representations of complex scenes at any scale (from very small objects to entire cities), accurately capturing geometry, materials, and lighting, while meeting the need for fast and efficient rendering—crucial for applications requiring real-time performance. Although High Dynamic Range (HDR) technology, which enables the capture of comprehensive real-world lighting information, has been used in novel view synthesis, several questions remain unanswered. For example, does HDR improve the overall quality of reconstruction? Are 8 bits enough? Can tone mapped images be a balanced compromise regarding quality and details? To answer such questions, in this work, we study the application of HDR technology on the 3DGS method for acquiring real-world scenes.*

## CCS Concepts

• **Computing methodologies** → **Rendering; Computational photography; Reconstruction;**

---

## 1. Introduction

Neural Radiance Fields (NeRF)-like methods [MST\*20, BMT\*21, MESK22, FKYT\*22, KKLD23, HMR19, STH\*19] enable users to capture and represent reality with very high fidelity using simple equipment. Among these methods, 3D Gaussian Splatting (3DGS) [KKLD23] stands out for its ability to combine high-quality representations with fast and efficient rendering.

Before these methods, utilizing photogrammetry, capturing involved different time-consuming sessions for each component of the rendering pipeline: geometry, materials (i.e., textures, BRDFs, etc.), and lighting. Typically, each component needs its own separate campaign using expensive and bulky specialized equipment such as laser scanners, structured light scanners, LIDAR systems, high quality photographic setups, goniophotometers, etc.

Therefore, such revolutionizing technologies allow users to save time and money at the same time, as 3DGS and other NeRFs-like methods can be acquired using a DSLR camera or just a smartphone. This makes them highly valuable across various industries, including cinema for visual effects (e.g., bullet time, 360° camera rotation), product advertising, virtual tourism, and more.

Another technology that has revolutionized capturing and display is High Dynamic Range (HDR) imaging [MP95, DM97] that allow us to capture from very dark areas (e.g., details in shad-

ows at night) to very bright areas (e.g., a light bulb filament) of a scene. Even though we now have smartphones in our pockets that can capture and display some sort of HDR data, such technology has rarely been applied to 3DGS and other NeRF-like methods. Recently, novel display technologies [ZJY\*21] have shown an incredible match with reality when multi-view capture is paired with HDR imaging. Therefore, it is crucial to study the impact of HDR imaging in the field of neural rendering.

In this work, we apply HDR imaging in 3DGS capturing and processing (see Sec. 2.1 and Sec. 2.2 for details on these technologies, and Fig. 1 and Sec. 3.1 for an overview of our pipeline). Our main aim is to understand if classic Standard Dynamic Range (SDR) images are enough to achieve high-quality reconstructions or if the use of HDR imaging improves the overall quality. Specifically, we want to understand if tone mapping operators (TMOs), which compress and convert HDR content into SDR one, could be an acceptable compromise, given that working directly with HDR images is computationally expensive and thus requires significant hardware resources.

To summarize, the main contributions of this paper are:

- A new dataset for 3DGS rendering, acquired using HDR imaging, which is freely available online<sup>†</sup>;
- A novel study to determine the impact of tone mapping in 3DGS rendering in comparisons to single exposure and HDR data;
- A novel strategy for multi-view tone mapping for generating consistent renderings of HDR images.

## 2. Related Work

### 2.1. NeRF and 3D Gaussian Splatting

Mildenhall *et al.* [MST\*20] introduced an innovative approach that achieves state-of-the-art results for synthesizing novel views of complex static scenes from photographs and computer generated images. This work had a significant impact, paving the way for a surge of new research in the field of novel view synthesis and driving substantial progress in the area. This work has sparked many extensions with improved quality [BMT\*21] and reduced training time [MESK22, FKYT\*22]. However, all these extension cannot have at the same time real-time performance, high quality results, and fast training times.

3DGS [KKLD23] has emerged as the main solution in response to these challenges by introducing an innovative scene representation technique that manages real-time rendering at high-quality and competitive training times (i.e., less than a hour). Unlike coordinate-based implicit models [MST\*20, HMR19, STH\*19], 3DGS represent a scene employing a point cloud. Each  $k$ -th point is a 3D Gaussian,  $G_k$ , defined as:

$$G_k(\mathbf{x}) = \exp\left(-\frac{1}{2}(\mathbf{x} - \mu_k)^\top \Sigma_k^{-1}(\mathbf{x} - \mu_k)\right) \quad (1)$$

where  $\mu_k$  is the position (where  $G_k$  is centered), and  $\Sigma_k$  is an anisotropic  $3 \times 3$  covariance matrix determining the shape of  $G_k$ . To model appearance, we have a transparency value,  $\alpha_k$ , and spherical harmonics coefficients,  $S_k$ , to encode RGB view-dependent colors baking the appearance. For a given pixel  $(i, j)$  of the camera image plane, the final pixel color is computed as:

$$C(i, j) = \sum_{k=1}^N c_k \alpha_k G'_k(i, j) \prod_{l=1}^{k-1} (1 - \alpha_l G'_l(i, j)) \quad (2)$$

where  $N$  is the number of Gaussians,  $c_k$  is a RGB color computed using  $S_k$  for the given  $(i, j)$  pixel,  $G'$  is a Gaussian projected on the image plane. Note that Gaussians are sorted from back to front to maintain the correct transparency order.

3DGS represents a novel approach that integrates state-of-the-art visual quality with competitive training times: its tile-based splatting solution guarantees real-time rendering with state-of-the-art quality for 1080p resolution across various previously published datasets. This has pushed 3DGS to a huge popularity with many extensions introduced in a very short time [FXZ\*24a].

3D Gaussian Splatting (3DGS) has been applied in a range of fields, with research focusing on refining its efficiency, realism, and applicability across practical domains. It has seen use in robotics,

urban mapping, autonomous navigation, and virtual/augmented reality, among others, with studies exploring its capabilities for tasks such as reconstruction, manipulation, and perception [FXZ\*24b].

In light of these considerations, we chose 3DGS to explore the use of HDR imaging and neural radiance fields.

### 2.2. HDR Imaging

HDR imaging captures images with a broader range of brightness levels than SDR imaging. In a high-contrast scene with both bright areas and deep shadows, SDR images tend to lose details in the shadows, while bright areas appear faded or washed out. HDR imaging, with its expanded dynamic range, preserves details in both dark and bright regions, bringing the image closer to the full spectrum of colors and brightness levels seen in the real world.

Typically, HDR images are created by capturing multiple photographs of the same scene at different exposure times [DM97, LEPM22], as done in this work. Since HDR images require a higher bit depth than SDR images to represent the extended dynamic range, tone mapping operators (TMOs) are essential in HDR imaging during visualization. TMOs transform HDR images to SDR format so they can be viewed on standard monitors, preserving as much contrast and detail as possible despite the limited dynamic range of these displays.

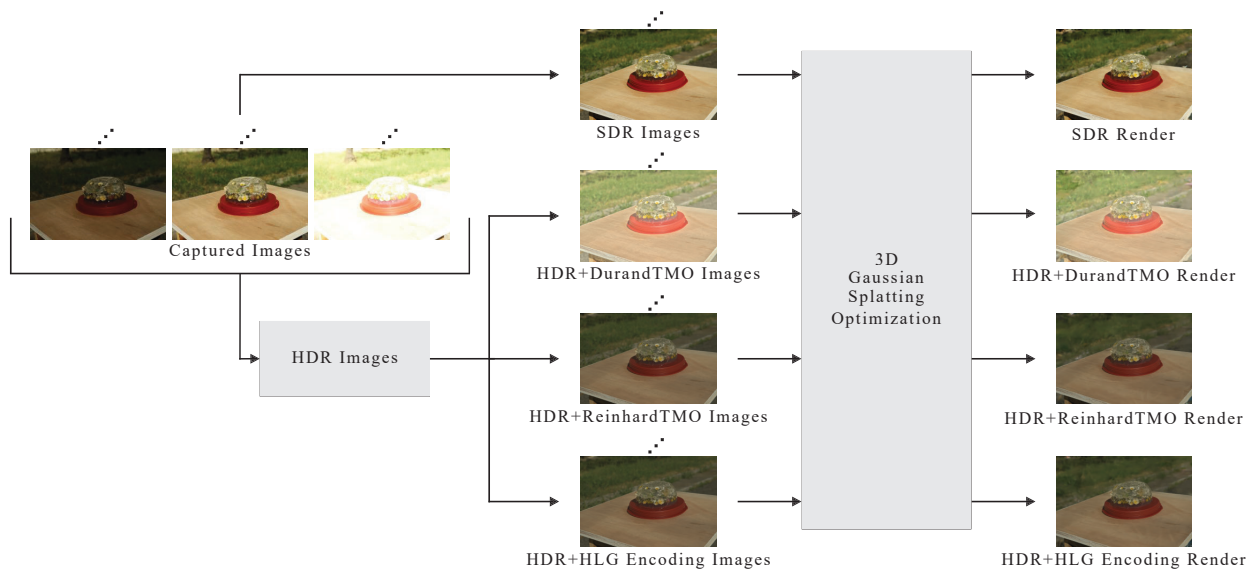
#### 2.2.1. Tone Mapping

Tone mapping has a vast literature [BADC17]. A tone mapping function converts an HDR image into an SDR one by applying an operator. We can broadly divide TMOs into two classes: *global operators* compress the luminance of each pixel applying the same function using global statistic from the image; and *local operators* compress the luminance applying a function that varies locally based on statistics computed on the pixel neighbors. Recently, deep learning-based solutions have started appearing [RSV\*20], where tone mapping is achieved by applying a series of convolutions, e.g., a UNet, using supervised and unsupervised training strategies [BA23]. Regarding multi-view tone mapping, research is very limited to stereo images only [SVRL19], which cannot be extended to be applied to different TMOs and works only on stereo images. Note that stereo images need to be properly captured with a custom setup or lens or to be rectified that is an expensive operation. In our work, we want to apply different TMOs (global and local) to medium-large number of images meant for 3DGS. Given the limited literature, we introduce a novel method for tone mapping HDR images for multi-view datasets that is designed to maintain consistency of values across different viewpoint (in which highlights move); see Section 3.3.

#### 2.2.2. HDR Encoding

Recently, novel standards have been proposed to encode HDR values for efficient content compression for HDRTV and smartphones. The main two encoding functions employed by vendors are the Perceptual Quantizer (PQ) [IR18] and Hybrid Gamma and Logarithm (HLG) [IR18]. PQ is a rational polynomial function for encoding HDR values in the range  $[10^{-4}, 10^4]$  cd/m<sup>2</sup>. HLG is a backward compatible encoding function working like a gamma function for

<sup>†</sup> <https://github.com/cnr-isti-vclab/3DGS-HDR>



**Figure 1:** This is the pipeline for our work on the “Jelly” scene captured in the wild. The first step is the photo campaign: for each viewing angle around the object of interest (the jelly in this scene), three photos are taken; one underexposed, one at the correct exposure, and one overexposed. The correctly exposed images represent the SDR set, while the three images together were used to create HDR images (one for each view). Two tone mapping operators (TMOs) were applied to the HDR images: Reinhard et al. TMO (ReinhardTMO) [RSSF02] and Durand and Dorsey TMO (DurandTMO) [DD02]. Additionally, HLG encoding was applied to the HDR images. This process generated four different sets of images, which were then used in the 3DGS optimization process [KKLD23] to create four distinct renders.

low luminance values and like a logarithm for high luminance values for a maximum of  $10^3$  cd/m<sup>2</sup>. Other encoding formats exist; for example, PU21 [MA21], but they are meant to support larger dynamic ranges. In this work, we selected the HLG format that is widely used in TV and mobile devices such as the Apple smartphones.

### 2.3. HDR in NeRF and 3DGS

Recently, researchers have started to exploit the potential of HDR technology, which refers to the set of techniques that produce HDR images from multiple SDR images taken at different exposures. Such application has shown the possibility to capture scenes with larger dynamic range than before improving the overall fidelity of the entire capturing process.

Mildenhall *et al.* [MHMB\*21] introduced RAW-NeRF. This method utilizes the RAW HDR image data and optimizes NeRF directly on linear RAW input in HDR color space. They demonstrated that reconstructing NeRF in RAW space makes it significantly more robust to noisy inputs and enables novel HDR view synthesis applications. However, this method is not without trade-offs. Most digital cameras can only save RAW images at full resolution with minimal compression. This leads to significant storage requirements. Sing *et al.* [SGM24] extended 3DGS applying it to 14-bit RAW inputs. These images are processed using a denoiser and bilinear demosaicing extending the dynamic range of captured scenes. In this way, challenging scenes can be reconstructed successfully without the typical failures when using 8-bit images. However, the method

cannot capture details in bright regions (e.g., light bulbs’ details) because more RAW may be required. Huang *et al.* [HZF\*23] proposed an end-to-end method called HDR-NeRF to recover HDR neural radiance fields from a set of SDR images with varying exposures at each viewpoint. This neural rendering system can generate novel HDR views and adjust the exposure of novel SDR views. It consists of two modules: an HDR radiance field that models the scene for radiance and densities and a tone mapper that models the camera response function (CRF) for colors. Although the HDR-NeRF method achieved high-quality results in terms of output quality, constructing an HDR radiance field is computationally intensive; i.e., it requires a day of computations on a NVIDIA V100 GPU machine. Concurrently to our work, Wang *et al.* [WWK\*24] extended 3DGS to handle HDR imaging and depth-of-field camera control. They achieve this by proposing a novel reconstruction framework that takes as input sparse images at varying viewpoints, exposure values, and depth of field.

Compared to our work, these methods focus on RAW files (12/14-stops), which do not cover the full dynamic range of a scene, or sparse sampling that may miss some view-dependent highlights. In our work, we explore how to reconstruct 3D scenes in HDR using 3DGS with exposure dense sampling [DM97] to faithfully recover the dynamic range of a scene.

### 3. Method

#### 3.1. Pipeline Overview

For this work, we acquired 9 scenes under both controlled conditions (in the lab) and uncontrolled conditions (in the wild) and we followed the pipeline of Fig 1. We captured SDR images at three different exposures for each viewing direction to construct HDR images (more details in Sec. 3.2).

Then, we generated 3DGS [KKLD23] representations employing the following sets of images: the SDR images, the HDR images with tone mapping applied using the Durand and Dorsey operator [DD02] (DurandTMO), the HDR images with tone mapping applied using the Reinhard et al. operator [RSSF02] (ReinhardTMO), and finally, the HDR images processed with the HLG encoding function (a very popular encoding format for mobile devices and TV); refer to Sec. 3.3 for further details on tone mapping operators.

Finally, we applied various metrics to all the renderings to assess their quality and similarity to the respective ground truth.

#### 3.2. Acquisition

For the acquisition phase, we used a Canon 5D DSLR camera mounted on a tripod to avoid alignment issues and ghost artifacts that can be caused when the camera is hand-held. We captured a total of 9 scenes: two under controlled lighting conditions, with spatial (checkerboard) and color (color checker) references, i.e., in the lab, and seven under uncontrolled lighting conditions, without spatial and color references, i.e., in the wild. Before the acquisitions in the lab and in the wild, we performed a white balance procedure based on the controlled lighting conditions in the lab, and on the uncontrolled lighting in the wild.

We used two liver phantoms as subjects in our lab scenes to explore the potential of applying 3DGS and HDR imaging to biomedical data, particularly in organ reconstruction. These liver-like structures were created to replicate the liver's color, visual appearance, and texture. The geometry of these phantoms approximates the dimensions and shape of a liver placed on the surgical table. Specifically, they are modeled as spherical caps with a base radius of 5.5 cm and a height of 3 cm. A mold was designed using CAD software (SolidWorks) to create these phantoms. The mold was then 3D-printed in ABS using the FDM approach, and Ecoflex 00-10 silicone was poured into it, with the addition of a softener to achieve the "squishy" consistency of a real liver.

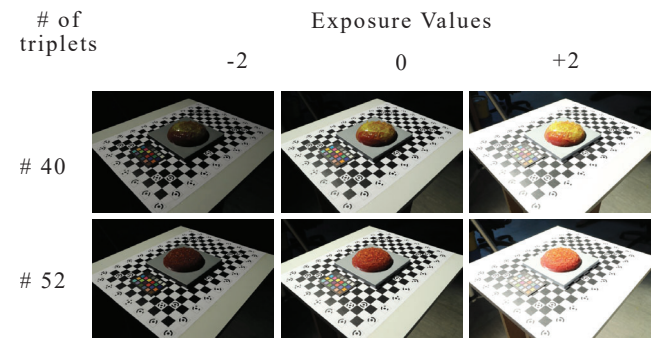
Two phantoms were created: one with the reddish-brown coloration of a healthy liver and one with the yellowish tones characteristic of a fatty liver (see Fig. 2). The phantoms were coated with a layer of PVA-PVP to replicate the visual appearance of the liver during the acquisition. Additionally, neutral-colored support was designed in CAD software (Fusion 360) and 3D-printed in PLA to hold the phantoms.



**Figure 2:** Liver phantoms: On the left is the fatty liver phantom, characterized by a yellowish color. On the right, there is the healthy liver phantom, characterized by a reddish-brown color.

The two phantoms were acquired under controlled conditions, using the following strategies: a checkerboard was placed underneath the phantoms to increase matching points in the photographs and to provide a scale reference within the scene; a color checker was positioned next to the subjects to ensure accurate color representation; and lighting similar to that of an operating room was used.

To create HDR images, we captured three images at each camera position with exposure values of -2-stop, 0-stop, and +2-stop: one underexposed, one overexposed, and one correctly exposed (i.e., the SDR image); see Fig. 3. We captured JPEG files to avoid the step of simulating a camera pipeline for the SDR image (i.e., 0-stop). Such image is typically employed in NeRF-like methods, so it was crucial to have real data in this step. To reduce distortions in the estimation of the camera response function (CRF) when merging different exposure for obtaining a HDR image, we pre-calibrated the CRF using an ad-hoc exposure stack (i.e., color-checker scene) using the HDR Toolbox [BAD17].



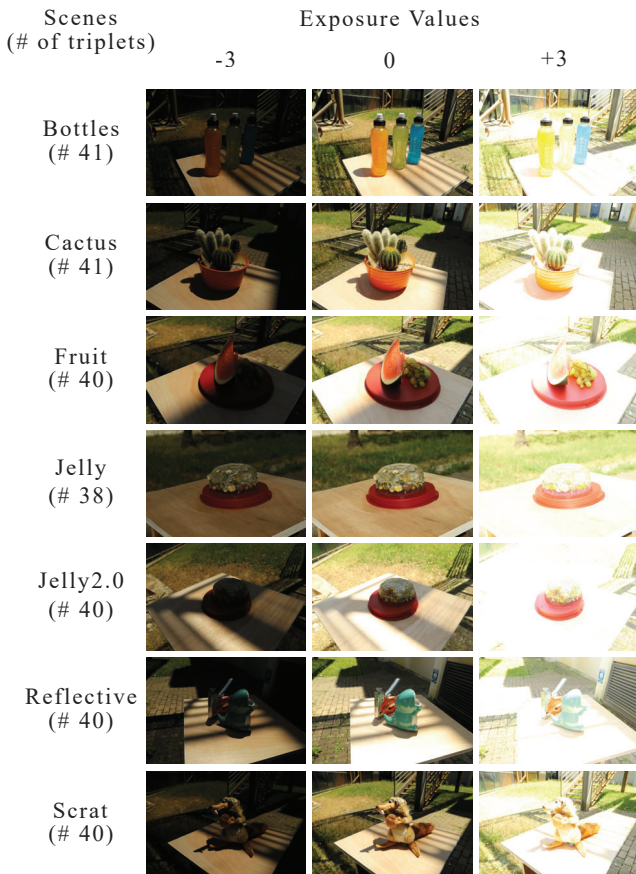
**Figure 3:** Triplet of images used to generate HDR images. For each scene captured in the lab, the underexposed image (exposure value -2-stop, left), the correctly exposed image (exposure value 0-stop, center), and the overexposed image (exposure value +2-stop, right) are shown.

For the scenes captured in the wild, we used various subjects with characteristics that make them very difficult, if not impossible, to capture using traditional photogrammetry techniques, and which also pose challenges for NeRF-like methods. Specifically,

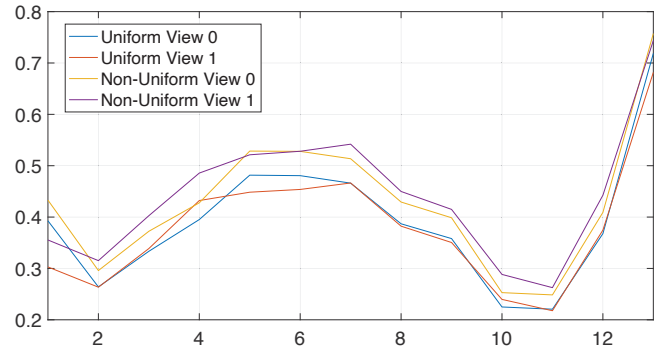
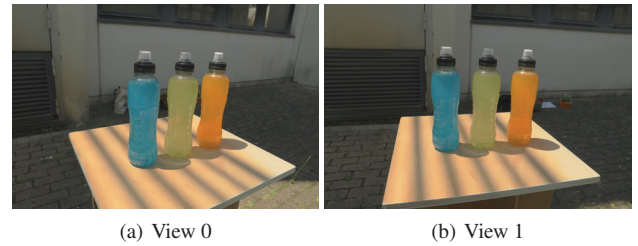
we captured bottles with colored liquids inside to evaluate the acquisition of semi-transparent objects; fruits to assess the acquisition of translucent objects; a cactus and a stuffed toy to test the acquisition of fuzz and hair; a gelatin dessert with fresh flowers inside to further assess transparent objects; and finally, a ceramic object and a stapler to evaluate reflective surfaces.

For the in the wild scenes, neither a checkerboard nor a color checker was used, and natural lighting was employed. These captures were performed in the summer, during the time of day when the sun was at its zenith, to minimize lighting variations during image capture, as NeRF-like methods are not robust to changes in illumination. The objects were placed partly in direct sunlight and partly in the shade to provide a wider range of lighting dynamics for the HDR images.

As in the lab, we captured three images at each camera position for HDR generation. However, this time, we used exposure values of -3-stop, 0-stop, and +3-stop, expecting a wider dynamic range; see Fig 4.



**Figure 4:** Triplet of images used to generate HDR images. For each scene captured in the wild, the underexposed image (exposure value -3-stop, left), the correctly exposed image (exposure value 0-stop, center), and the overexposed image (exposure value +3-stop, right) are shown.



**Figure 5:** An example showing consistency for the Bottles scenes from two different views (a) and (b). The bottom graph shows the plot of luminance values of 13 key points tracked in (a) and (b) for ReinhardTMO with our scheme (Uniform) and without it (Non-Uniform). On average, the Uniform scheme reduces luminance differences between the same key points in different views more than the Non-Uniform scheme.

### 3.3. Consistent Multi-View Tone Mapping

In our work, we want to understand if tone mapping is a valid alternative to HDR images or they perform better than SDR ones.

Similarly to video tone mapping [BAD17], if we tone map images from a multi-view dataset separately for each image, we will obtain inconsistent results. For example, Figure 5 (a) and (b) show the “Bottles” scene (from Figure 4) in two different views, these images represent the same object from a different view but view-independent diffuse reflections may have different values because they were independently tone-mapped. To minimize this effect, which may create views with varying appearances and introduce view-dependent effects when it is not the case, we propose a different strategy for gathering statistics.

The first step of our strategy is to compute global statistics on the luminance channel of each image in the multi-view dataset. These statistics are average, logarithmic average, minimum, maximum, key, etc. At this point, we need to pool a value for each global statistic to use it as input for the tone mapping phase. Different strategies can be applied. The use of minimum and maximum pooling should be avoided because they may create a few images that are extremely dark or bright. In our experiments, we found out that computing the median is a sound strategy. However, if the object has a “front-face” image, it is preferred, as visual impact, to select global statistics from that image.

We selected two popular solutions for tone mapping; i.e., Du-

rand and Dorsey operator [DD02] (DurandTMO), and Reinhard et al.’s operator [RSSF02] (ReinhardTMO). These two operators provide high-quality results, are straightforward to implement, and are typically employed as a baseline when doing tone mapping benchmarks. ReinhardTMO is a sigmoid function applied to the scaled input luminance based on the image statistics. DurandTMO works by reducing the luminance dynamic range of low frequencies in the image/video while keeping untouched the high frequencies. To avoid halos around edges, the bilateral filter is employed during frequency separation. Only low frequencies are compressed linearly in the logarithm domain. In our experiments, we applied our scheme for both ReinhardTMO and DurandTMO for computing global statistics such as the scaling factor, and the maximum white point. The graph in Figure 5 shows that our proposed schemes reduce inconsistency in luminance values. In this example, the mean absolute difference is 0.0273 when tone mapping each image separately (Non-Uniform), while this value lower down to 0.0205 when using our scheme (Uniform).

### 3.4. 3DGS Training

We decided to study four different types of generated 3DGS; i.e., using different input images. In particular, we created 3DGS representations for images using the 0-stop SDR image (this image is the best exposure determined by our camera), the ReinhardTMO (a sigmoid TMO), the DurandTMO (a local and logarithmic TMO), and the HLG HDR encoding (a very popular encoding format for mobile devices and TV); see Figure 6. After preparing all the image sets and extracting the undistorted images and SfM information generated by COLMAP [SF16], we trained the various models to obtain the Gaussian splats for each scene. In total, we had 36 trainings to be carried out; i.e., 9 scenes and 4 types of input images. We trained each model for 30,000 iterations on a single NVIDIA RTX 3090 GPU equipped with 24GB of memory. On average, the training lasted around 20 minutes and never exceeded 30 minutes.

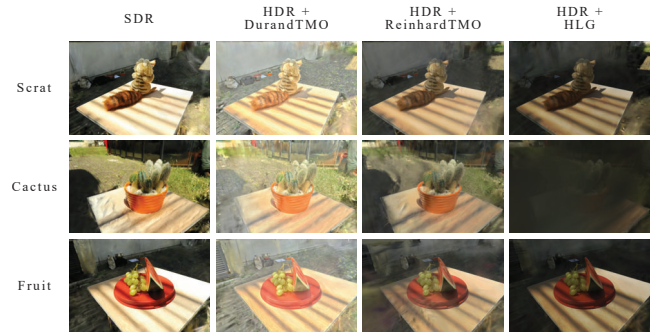


**Figure 6:** An example of the “Jelly2.0” scene showing the different types of input images used in our work: 0-stop SDR images, ReinhardTMO images, DurandTMO images, and HLG HDR images.

## 4. Results

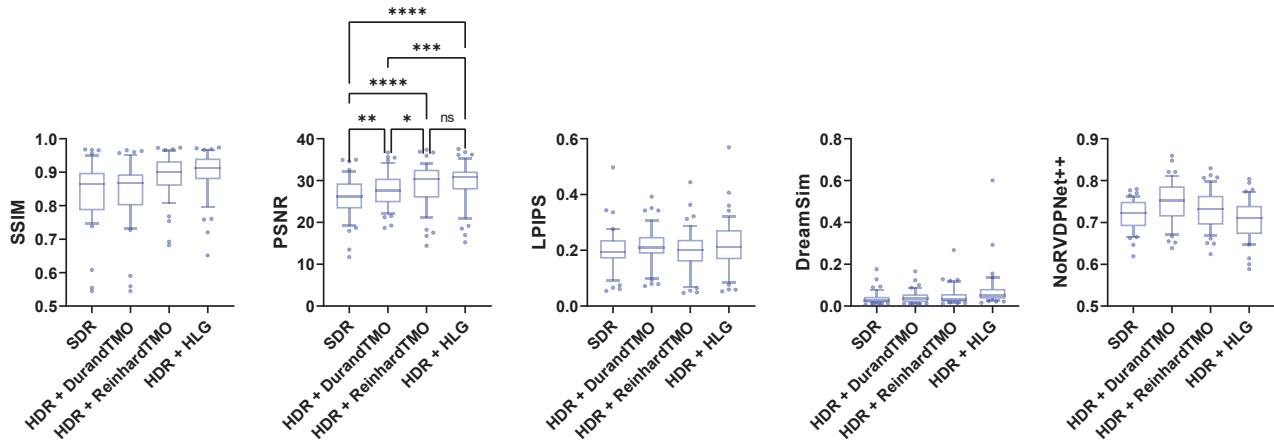
To evaluate the quality of the reconstructions for different scenes (see Fig. 11) based on their similarity to the ground truth (0-stop SDR images for SDR render, ReinhardTMO images for ReinhardTMO render, DurandTMO images for DurandTMO render and HLG HDR images for HLG HDR render), we used the following metrics with reference: the classic Peak signal-to-noise ratio (PSNR), SSIM [WBSS04], LPIPS [ZIE\*18], and DreamSim [FTS\*23]. Additionally, we used a no-reference metric: NoRVDPNet++ [BAM\*23].

3DGS renderings at the reference viewpoints in our dataset may have a large percentage of background pixels. Typically, the background may be modeled with a few Gaussians and this may result in low quality background. Note that a low-quality background depends on the scene and the method; see Fig. 9. To avoid low-quality backgrounds influencing negatively on the metrics, we have decided to apply a central crop (i.e., the central block in a  $3 \times 3$  grid) to each rendering. In this way, the background pixels are minimized reducing their influence in the final metrics computations.



**Figure 9:** An example showing that quality background depends on the scene and the method. The reconstruction of the “Scrat” scene using HDR + DurandTMO offers the best compromise between background blurring and shadow/reflection distortions, compared to other methods. The reconstruction of the “Fruit” scene using HDR + HLG has the darkest and smoothest background, compared to other methods. The reconstruction of “Cactus” scene using 0-stop SDR has the best-quality background, compared to other methods.

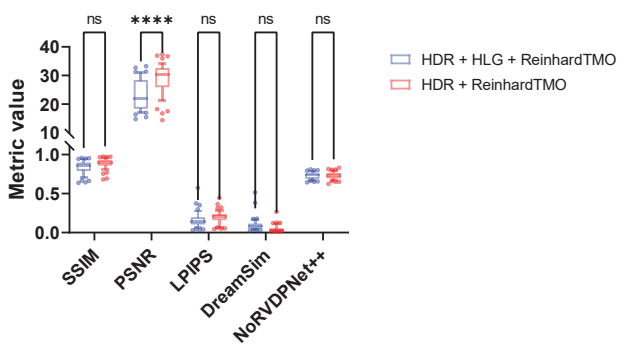
Table 1 reports all results for each scene, metric, and method. These results are summarized per method in Fig. 7. As a first step, we performed a two-way ANOVA test (Tukey’s multiple comparisons  $t$ -test using the Bonferroni correction, by comparing, within each column, the mean of the rows) using Graphpad Prism (version 9.1.0) on aggregated data. We ran this test to infer if statistically significant differences were present in the various metrics based on the tone mapping operation. Fig. 7 shows the results of this test, these underlines that most of the employed metrics (i.e., LPIPS, SSIM, DreamSim, and NoRVDPNet++) are not statistically significant. This result for SSIM, LPIPS, NoRVDPNet++ (which is a no-reference version of HDR-VDP2.2 [NMSC15]) is also confirmed in the recent literature [LWH\*24]. Unexpectedly, DreamSim, which should match perception in subjective experiments, had no statistically significant results. From PSNR results in Fig. 7, we can elicit that the most faithful reconstructions are the ones generated using HDR+HLG and HDR+ReinhardTMO input images. The least faithful reconstruction is the one generated by SDR, which is the classic method for generating 3DGS. This is an interesting finding because it shows HDR imaging improves the overall quality in reconstructions and it is favorable to be employed for capturing reality. Another interesting finding is the fact a global TMO such as ReinhardTMO has similar performance to HDR encoded data, but a local TMO (e.g., DurandTMO) has a lower quality.



**Figure 7:** Box-plot (10-90 percentile) showing the metrics (i.e. SSIM, PSNR, LPIPS, DreamSim, NoRVDPNet++) for the SDR images and the tone mapped ones (i.e., HDR + DurandTMO, HDR + ReinhardTMO, HDR + HLG). The results from the two-way ANOVA statistical test are shown only for the PSNR metric as the test highlighted no statistically significant difference for the other metrics. In the figure: ns: not significant, \*:  $p$ -value  $\leq 0.05$ , \*\*:  $p$ -value  $\leq 0.01$ , \*\*\*:  $p$ -value  $\leq 0.001$ , \*\*\*\*:  $p$ -value  $\leq 0.0001$ .

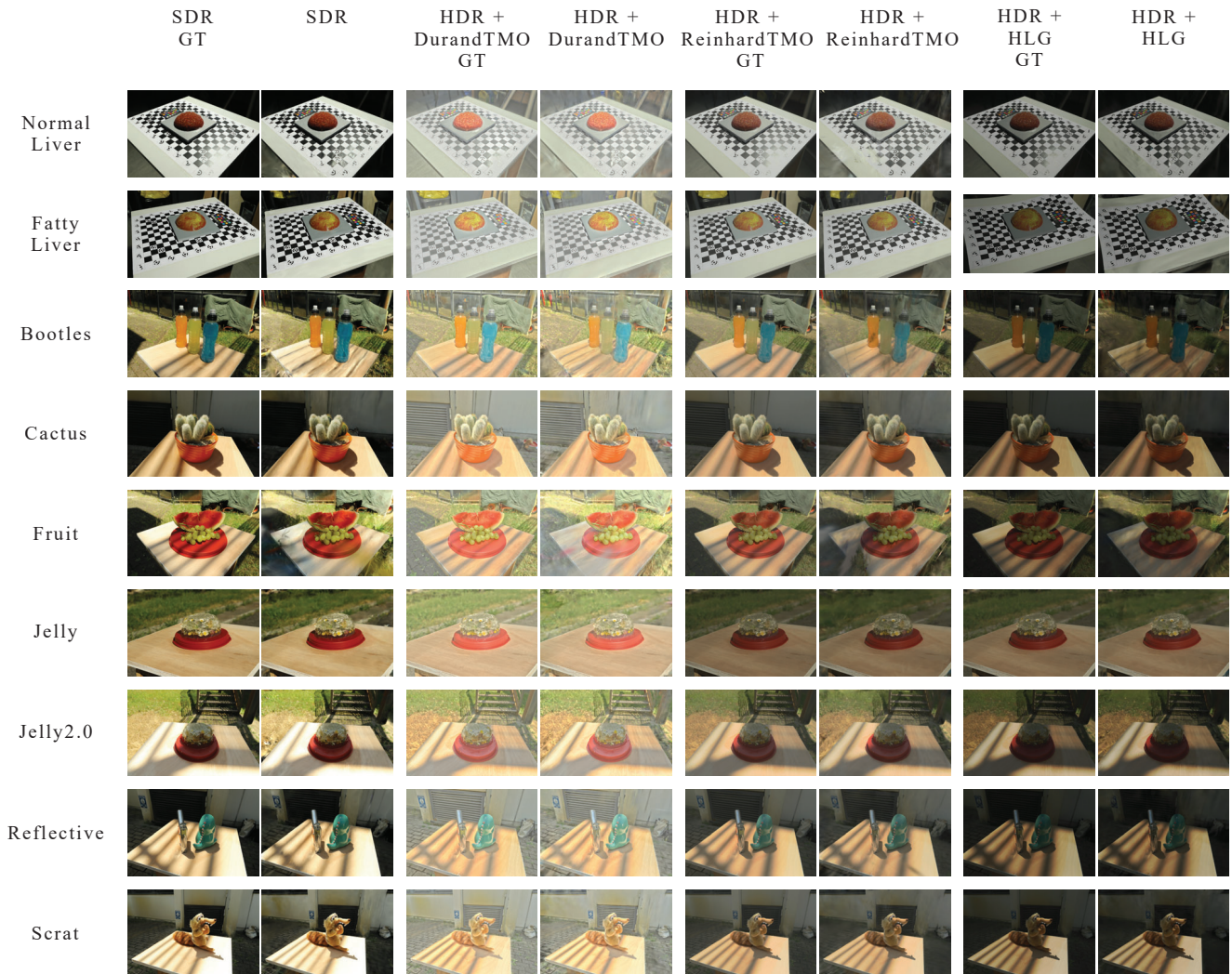


**Figure 8:** An example of the “Cactus” scene (zoomed) testing HDR+HLG+ReinhardTMO vs HDR+ReinhardTMO: On the left side, the 3DGS rendering generated using HDR+HLG+ReinhardTMO; in the middle, the reference ReinhardTMO image; on the right side, the 3DGS rendering generated using HDR+ReinhardTMO. Note that HDR+ReinhardTMO is closer to the reference in terms of matched tones and colors.



**Figure 10:** Box-plot (10-90 percentile) showing the metrics (SSIM, PSNR, LPIPS, DreamSim, NoRVDPNet++) for the two mappings HDR + ReinhardTMO and HDR + HLG + ReinhardTMO. In the figure: ns: not significant, \*\*\*\*:  $p$ -value  $\leq 0.0001$ .

At this point, we may conclude that HDR+HLG may be superior to HDR+ReinhardTMO. This is because we can decode HLG values, obtain HDR data, and tone map decoded renderings with a TMO of choice for a given production intent. To assert this, we ran a further evaluation test. As a first step, we computed renderings using HDR+HLG. Then, we decoded this rendering obtaining HDR values and tone mapped using ReinhardTMO. Finally, we compared these tone mapped images (HDR+HLG+ReinhardTMO) with ground truth tone mapped images using ReinhardTMO and against the 3DGS reconstructions HDR+ReinhardTMO. As before, we computed the previously selected metrics and tested their results using a two-way ANOVA test (Tukey’s multiple comparison  $t$ -test using the Bonferroni correction, by comparing each cell mean with the other cell mean in that row). From our statistical analysis, we obtained similar results; i.e., PSNR is the only statistically significant metric. Apart from this result, we found that HDR+ReinhardTMO has a better result than HDR+HLG+ReinhardTMO regarding image quality. Although this may be counterintuitive, this happens because we have a fitting process during a 3DGS reconstruction. For HDR+HLG, the fitting process may lose details that are present in HDR+ReinhardTMO images and vice versa. Therefore, we do not have an equivalent rendering quality when comparing HDR+HLG+ReinhardTMO and HDR+ReinhardTMO. This result is important because it shows that if our 3DGS needs to maintain details and the intent of a given TMO, we must first perform tone mapping and then proceed with the 3DGS reconstruction. HDR+HLG+ReinhardTMO still produces a reasonable quality (i.e., PSNR over 20), but we may lose some fidelity regarding tone and color reproduction; see Fig. 8.



**Figure 11:** The different types of input images used in our work for each scene: 0-stop SDR images, ReinhardTMO images, DurandTMO images, and HLG HDR images. For each scene, we show an example view with the ground truth (GT); i.e., the input of the 3DGS training, and their corresponding rendering.

## 5. Conclusions

We have presented a study that contributes to the ongoing exploration of the impact of HDR imaging using dense sampling for 3DGS. To achieve this, we have introduced a simple but effective novel scheme to enforce consistency when tone mapping multi-view HDR images representing the same scene. We have shown using our novel dataset that typically the use of HDR imaging is an advantage in terms of reconstruction. From visual comparisons, tone mapping emerges as a strategy to prefer for 3DGS when using HDR imaging especially when we want to maintain the intent and quality of a TMO. A further investigation needs to be carried out to fully understand which classes of TMOs are to be employed when dealing with 3DGS or other neural rendering methods. In future work, we would also like to extend the use of 3DGS and HDR to biomedical data such as laparoscopic reconstructions in which

specular and glossy highlights can be fully acquired using HDR imaging and a faithful rendering of dark areas is critical for the evaluation of organs and to achieve success in robotic or remote surgeries.

## Acknowledgements


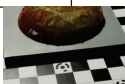



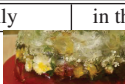
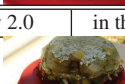
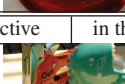
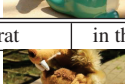
Amedeo F. Bonatti and Carmelo De Maria acknowledge the FoRe-Lab of the Department of Information Engineering of the University of Pisa.

## References

[BA23] BANTERLE F., ARTUSI A.: Modern High Dynamic Range Imaging at the Time of Deep Learning. In *Eurographics 2023 - Tutorials*



**Table 1:** This table presents SSIM, PSNR, LPIPS, DreamSim, and NoRVDPNet++ values for our dataset, calculated with a central crop applied to each scene. For each scene, we computed metrics values for single exposure (SDR), tone mapping using DurandTMO (HDR+DurandTMO), tone mapping using ReinhardTMO (HDR+ReinhardTMO), and HDR images encoded using HLG (HDR+HLG). The red color is for the best results, and the blue color is for the second best.

Cropped Scenes	Acquisition Conditions	Image Types	Metrics				
			SSIM $\uparrow$	PSNR $\uparrow$	LPIPS $\downarrow$	DreamSim $\downarrow$	NoRVDPNet++ $\uparrow$
Normal Liver 	in the lab	SDR	0.926 $\pm$ 0.040	29.896 $\pm$ 3.679	0.125 $\pm$ 0.066	0.020 $\pm$ 0.00751	72.163 $\pm$ 0.037
		HDR+DurandTMO	0.931 $\pm$ 0.033	32.680 $\pm$ 2.737	0.131 $\pm$ 0.057	0.016 $\pm$ 0.004	79.947 $\pm$ 0.032
		HDR+ReinhardTMO	0.942 $\pm$ 0.038	31.497 $\pm$ 5.287	0.106 $\pm$ 0.063	0.022 $\pm$ 0.013	77.264 $\pm$ 0.042
		HDR+HLG	0.951 $\pm$ 0.023	33.857 $\pm$ 2.541	0.113 $\pm$ 0.060	0.029 $\pm$ 0.011	74.773 $\pm$ 0.036
Fatty Liver 	in the lab	SDR	0.908 $\pm$ 0.054	27.758 $\pm$ 4.902	0.147 $\pm$ 0.069	0.029 $\pm$ 0.014	74.319 $\pm$ 0.015
		HDR+DurandTMO	0.923 $\pm$ 0.032	32.088 $\pm$ 2.912	0.151 $\pm$ 0.048	0.024 $\pm$ 0.016	79.046 $\pm$ 0.038
		HDR+ReinhardTMO	0.943 $\pm$ 0.025	32.670 $\pm$ 3.252	0.119 $\pm$ 0.052	0.022 $\pm$ 0.003	78.137 $\pm$ 0.019
		HDR+HLG	0.935 $\pm$ 0.029	30.523 $\pm$ 2.655	0.151 $\pm$ 0.062	0.059 $\pm$ 0.013	74.678 $\pm$ 0.020
Bottles 	in the wild	SDR	0.837 $\pm$ 0.057	25.745 $\pm$ 2.441	0.202 $\pm$ 0.030	0.031 $\pm$ 0.010	71.500 $\pm$ 0.029
		HDR+DurandTMO	0.848 $\pm$ 0.046	26.220 $\pm$ 1.160	0.225 $\pm$ 0.029	0.045 $\pm$ 0.012	72.420 $\pm$ 0.036
		HDR+ReinhardTMO	0.902 $\pm$ 0.035	28.585 $\pm$ 3.712	0.186 $\pm$ 0.023	0.036 $\pm$ 0.023	71.527 $\pm$ 0.038
		HDR+HLG	0.918 $\pm$ 0.036	28.351 $\pm$ 3.268	0.221 $\pm$ 0.062	0.076 $\pm$ 0.031	70.533 $\pm$ 0.049
Cactus 	in the wild	SDR	0.713 $\pm$ 0.096	22.640 $\pm$ 2.795	0.256 $\pm$ 0.063	0.050 $\pm$ 0.030	72.419 $\pm$ 0.050
		HDR+DurandTMO	0.677 $\pm$ 0.113	24.256 $\pm$ 2.395	0.278 $\pm$ 0.059	0.063 $\pm$ 0.035	74.466 $\pm$ 0.056
		HDR+ReinhardTMO	0.777 $\pm$ 0.070	25.154 $\pm$ 3.051	0.269 $\pm$ 0.066	0.076 $\pm$ 0.042	73.573 $\pm$ 0.050
		HDR+HLG	0.778 $\pm$ 0.077	24.205 $\pm$ 5.514	0.332 $\pm$ 0.113	0.173 $\pm$ 0.194	65.747 $\pm$ 0.048
Fruit 	in the wild	SDR	0.869 $\pm$ 0.037	27.916 $\pm$ 4.194	0.189 $\pm$ 0.028	0.029 $\pm$ 0.022	70.763 $\pm$ 0.036
		HDR+DurandTMO	0.863 $\pm$ 0.025	28.285 $\pm$ 3.787	0.239 $\pm$ 0.021	0.046 $\pm$ 0.032	73.070 $\pm$ 0.022
		HDR+ReinhardTMO	0.895 $\pm$ 0.034	28.229 $\pm$ 5.684	0.237 $\pm$ 0.027	0.061 $\pm$ 0.039	70.366 $\pm$ 0.032
		HDR+HLG	0.915 $\pm$ 0.018	32.097 $\pm$ 3.853	0.248 $\pm$ 0.045	0.078 $\pm$ 0.048	68.442 $\pm$ 0.019
Jelly 	in the wild	SDR	0.820 $\pm$ 0.033	26.042 $\pm$ 0.861	0.203 $\pm$ 0.027	0.033 $\pm$ 0.012	73.705 $\pm$ 0.027
		HDR+DurandTMO	0.830 $\pm$ 0.020	28.044 $\pm$ 1.308	0.205 $\pm$ 0.023	0.039 $\pm$ 0.012	77.120 $\pm$ 0.035
		HDR+ReinhardTMO	0.889 $\pm$ 0.010	30.737 $\pm$ 0.845	0.216 $\pm$ 0.031	0.047 $\pm$ 0.011	71.875 $\pm$ 0.023
		HDR+HLG	0.890 $\pm$ 0.011	30.725 $\pm$ 0.729	0.218 $\pm$ 0.033	0.052 $\pm$ 0.018	72.127 $\pm$ 0.031
Jelly 2.0 	in the wild	SDR	0.835 $\pm$ 0.061	25.203 $\pm$ 6.383	0.199 $\pm$ 0.041	0.041 $\pm$ 0.044	71.073 $\pm$ 0.031
		HDR+DurandTMO	0.831 $\pm$ 0.059	25.838 $\pm$ 4.383	0.225 $\pm$ 0.060	0.052 $\pm$ 0.057	73.666 $\pm$ 0.037
		HDR+ReinhardTMO	0.874 $\pm$ 0.039	26.945 $\pm$ 6.207	0.226 $\pm$ 0.033	0.047 $\pm$ 0.036	72.535 $\pm$ 0.034
		HDR+HLG	0.888 $\pm$ 0.052	29.537 $\pm$ 5.951	0.238 $\pm$ 0.044	0.054 $\pm$ 0.041	70.593 $\pm$ 0.028
Reflective 	in the wild	SDR	0.774 $\pm$ 0.120	21.422 $\pm$ 4.979	0.272 $\pm$ 0.114	0.072 $\pm$ 0.053	70.396 $\pm$ 0.023
		HDR+DurandTMO	0.810 $\pm$ 0.064	23.872 $\pm$ 2.846	0.261 $\pm$ 0.069	0.063 $\pm$ 0.020	70.774 $\pm$ 0.038
		HDR+ReinhardTMO	0.878 $\pm$ 0.064	26.455 $\pm$ 6.258	0.223 $\pm$ 0.114	0.080 $\pm$ 0.094	70.002 $\pm$ 0.032
		HDR+HLG	0.885 $\pm$ 0.065	26.947 $\pm$ 5.571	0.222 $\pm$ 0.094	0.111 $\pm$ 0.097	68.159 $\pm$ 0.038
Scrat 	in the wild	SDR	0.857 $\pm$ 0.058	26.797 $\pm$ 4.497	0.207 $\pm$ 0.042	0.031 $\pm$ 0.024	71.044 $\pm$ 0.047
		HDR+DurandTMO	0.863 $\pm$ 0.033	26.720 $\pm$ 3.809	0.216 $\pm$ 0.027	0.043 $\pm$ 0.019	72.679 $\pm$ 0.043
		HDR+ReinhardTMO	0.891 $\pm$ 0.043	29.374 $\pm$ 6.093	0.225 $\pm$ 0.052	0.043 $\pm$ 0.028	72.059 $\pm$ 0.048
		HDR+HLG	0.902 $\pm$ 0.050	29.211 $\pm$ 5.272	0.239 $\pm$ 0.050	0.059 $\pm$ 0.034	70.099 $\pm$ 0.057

(2023), Serrano A., Slusallek P., (Eds.), The Eurographics Association. doi:10.2312/egt.20231033. 2

[BADC17] BANTERLE F., ARTUSI A., DEBATTISTA K., CHALMERS A.: *Advanced High Dynamic Range Imaging: Theory and Practice (2nd Edition)*. AK Peters (CRC Press), Natick, MA, USA, July 2017. 2, 4, 5

[BAM\*23] BANTERLE F., ARTUSI A., MOREO A., CARRARA F., CIGNONI P.: NoR-VDPNet++: Real-Time No-Reference Image Quality Metrics. *IEEE Access* 11 (2023), 34544–34553. doi:10.1109/ACCESS.2023.3263496. 6

[BMT\*21] BARRON J. T., MILDENHALL B., TANCIK M., HEDMAN P., MARTIN-BRUALLA R., SRINIVASAN P. P.: Mip-nerf: A multi-scale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 5855–5864. doi:https://doi.org/10.48550/arXiv.2103.13415. 1, 2

[DD02] DURAND F., DORSEY J.: Fast bilateral filtering for the

display of high-dynamic-range images. *ACM Trans. Graph.* 21, 3 (2002), 257–266. URL: https://doi.org/10.1145/566654.566574, doi:10.1145/566654.566574. 3, 4, 6

[DM97] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (USA, 1997), ACM Press/Addison-Wesley Publishing Co., pp. 369–378. URL: https://doi.org/10.1145/258734.258884, doi:10.1145/258734.258884. 1, 2, 3

[FKYT\*22] FRIDOVICH-KEIL S., YU A., TANCIK M., CHEN Q., RECHT B., KANAZAWA A.: Plenoxels: Radiance fields without neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 5501–5510. doi:https://doi.org/10.48550/arXiv.2112.05131. 1, 2

[FTS\*23] FU\* S., TAMIR\* N., SUNDARAM\* S., CHAI L., ZHANG R., DEKEL T., ISOLA P.: Dreamsim: Learning new dimensions of human

- visual similarity using synthetic data. *arXiv:2306.09344* (2023). 6
- [FXZ\*24a] FEI B., XU J., ZHANG R., ZHOU Q., YANG W., HE Y.: 3d gaussian splatting as new era: A survey. *IEEE Transactions on Visualization and Computer Graphics* (2024), 1–20. doi:[10.1109/TVCG.2024.3397828](https://doi.org/10.1109/TVCG.2024.3397828). 2
- [FXZ\*24b] FEI B., XU J., ZHANG R., ZHOU Q., YANG W., HE Y.: 3d gaussian splatting as new era: A survey. *IEEE Transactions on Visualization and Computer Graphics* (2024), 1–20. doi:[10.1109/TVCG.2024.3397828](https://doi.org/10.1109/TVCG.2024.3397828). 2
- [HMR19] HENZLER P., MITRA N. J., RITSCHER T.: Escaping plato's cave: 3d shape from adversarial rendering. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2019), pp. 9984–9993. doi: <https://doi.org/10.48550/arXiv.1811.11606>. 1, 2
- [HZF\*23] HUANG X., ZHANG Q., FENG Y., LI H., WANG X., WANG Q.: Hdr-nerf: High dynamic range neural radiance fields, 2023. URL: <https://arxiv.org/abs/2111.14451>, arXiv:2111.14451. 3
- [IR18] ITU-R: ITU-R 2100-2: Image parameter values for high dynamic range television for use in production and international programme exchange, 2018. 2
- [KKL23] KERBL B., KOPANAS G., LEIMKÜHLER T., DRETTAKIS G.: 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics* 42, 4 (2023). doi: [10.1145/3592433](https://doi.org/10.1145/3592433). 1, 2, 3, 4
- [LEPM22] LECOQUAT B., EBOLI T., PONCE J., MAIRAL J.: High dynamic range and super-resolution from raw image bursts. *arXiv preprint arXiv:2207.14671* (2022). 2
- [LWH\*24] LIANG H., WU T., HANJI P., BANTERLE F., GAO H., MANTIUK R., ÖZTIRELI A. C.: Perceptual quality assessment of nerf and neural view synthesis methods for front-facing views. *Comput. Graph. Forum* 43, 2 (2024), i–iii. URL: <https://doi.org/10.1111/cgf.15036>, doi:[10.1111/cgf.15036](https://doi.org/10.1111/cgf.15036). 6
- [MA21] MANTIUK R. K., AZIMI M.: PU21: A novel perceptually uniform encoding for adapting existing quality metrics for HDR. In *Picture Coding Symposium, PCS 2021, Bristol, United Kingdom, June 29 - July 2, 2021* (2021), IEEE, pp. 1–5. URL: <https://doi.org/10.1109/PCS50896.2021.9477471>, doi:[10.1109/PCS50896.2021.9477471](https://doi.org/10.1109/PCS50896.2021.9477471). 3
- [MESK22] MÜLLER T., EVANS A., SCHIED C., KELLER A.: Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics (ToG)* 41, 4 (2022), 1–15. doi: <https://doi.org/10.48550/arXiv.2201.05989>. 1, 2
- [MHMB\*21] MILDENHALL B., HEDMAN P., MARTIN-BRUALLA R., SRINIVASAN P., BARRON J. T.: Nerf in the dark: High dynamic range view synthesis from noisy raw images, 2021. URL: <https://arxiv.org/abs/2111.13679>, arXiv:2111.13679. 3
- [MP95] MANN S., PICARD R. W.: Being "undigital" with digital cameras: Extending dynamic range by combining differently exposed pictures. In *In Proceedings of IS&T 46th annual conference* (May 1995), pp. 422–428. 1
- [MST\*20] MILDENHALL B., SRINIVASAN P. P., TANCIK M., BARRON J. T., RAMAMOORTHI R., NG R.: Nerf: Representing scenes as neural radiance fields for view synthesis. In *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part I* (2020), Vedaldi A., Bischof H., Brox T., Frahm J., (Eds.), vol. 12346 of *Lecture Notes in Computer Science*, Springer, pp. 405–421. URL: [https://doi.org/10.1007/978-3-030-58452-8\\_24](https://doi.org/10.1007/978-3-030-58452-8_24), doi:[10.1007/978-3-030-58452-8\\_24](https://doi.org/10.1007/978-3-030-58452-8_24). 1, 2
- [NMSC15] NARWARIA M., MANTIUK R. K., SILVA M. P. D., CALLET P. L.: HDR-VDP-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images. *J. Electronic Imaging* 24, 1 (2015), 010501. URL: <https://doi.org/10.1117/1.JEI.24.1.010501>, doi:[10.1117/1.JEI.24.1.010501](https://doi.org/10.1117/1.JEI.24.1.010501). 6
- [RSSF02] REINHARD E., STARK M. M., SHIRLEY P., FERWERDA J. A.: Photographic tone reproduction for digital images. *ACM Trans. Graph.* 21, 3 (2002), 267–276. URL: <https://doi.org/10.1145/566654.566575>, doi:[10.1145/566654.566575](https://doi.org/10.1145/566654.566575). 3, 4, 6
- [RSV\*20] RANA A., SINGH P., VALENZISE G., DUFAUX F., KOMODAKIS N., SMOLIC A.: Deep tone mapping operator for high dynamic range images. *IEEE Trans. Image Process.* 29 (2020), 1285–1298. URL: <https://doi.org/10.1109/TIP.2019.2936649>, doi:[10.1109/TIP.2019.2936649](https://doi.org/10.1109/TIP.2019.2936649). 2
- [SF16] SCHÖNBERGER J. L., FRAHM J.-M.: Structure-from-motion revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (2016). 6
- [SGM24] SINGH S., GARG A., MITRA K.: Hdrsplat: Gaussian splatting for high dynamic range 3d scene reconstruction from raw images. *BMVC* (2024). 3
- [STH\*19] SITZMANN V., THIES J., HEIDE F., NIESSNER M., WETZSTEIN G., ZOLLHOFER M.: Deepvoxels: Learning persistent 3d feature embeddings. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), pp. 2437–2446. doi: <https://doi.org/10.48550/arXiv.1812.01024>. 1, 2
- [SVRL19] SHARMA M., VENKATESH M. S., RAGAVAN G., LAL R.: A novel approach for multi-view 3d hdr content generation via depth adaptive cross trilateral tone mapping. In *2019 International Conference on 3D Immersion (IC3D)* (2019), pp. 1–8. doi:[10.1109/IC3D48390.2019.8975988](https://doi.org/10.1109/IC3D48390.2019.8975988). 2
- [WBSS04] WANG Z., BOVIK A., SHEIKH H., SIMONCELLI E.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 4 (2004), 600–612. doi:[10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861). 6
- [WWK\*24] WANG C., WOLSKI K., KERBL B., SERRANO A., BEMANA M., SEIDEL H., MYSZKOWSKI K., LEIMKÜHLER T.: Cinematic gaussians: Real-time HDR radiance fields with depth of field. *CoRR abs/2406.07329* (2024). URL: <https://doi.org/10.48550/arXiv.2406.07329>, arXiv:2406.07329, doi:[10.48550/ARXIV.2406.07329](https://doi.org/10.48550/ARXIV.2406.07329). 3
- [ZIE\*18] ZHANG R., ISOLA P., EFROS A. A., SHECHTMAN E., WANG O.: The unreasonable effectiveness of deep features as a perceptual metric. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018), pp. 586–595. doi:[10.1109/CVPR.2018.00068](https://doi.org/10.1109/CVPR.2018.00068). 6
- [ZJY\*21] ZHONG F., JINDAL A., YÖNTEM A. O., HANJI P., WATT S. J., MANTIUK R. K.: Reproducing reality with a high-dynamic-range multi-focal stereo display. *ACM Trans. Graph.* 40, 6 (dec 2021). URL: <https://doi.org/10.1145/3478513.3480513>, doi:[10.1145/3478513.3480513](https://doi.org/10.1145/3478513.3480513). 1