PPISP-Local: Local Tone Mapping for Physically Plausible 3D Gaussian Splatting of Heritage Ceilings

Anthony Wong[1] **and Cheng Wang**[1]

[1]School of Computing Sciences, University of East Anglia, Norwich, United Kingdom

E-mail: Cheng.C.Wang@uea.ac.uk

Abstract. Mobile-phone-based 3D Gaussian Splatting (3DGS) is increasingly attractive for heritage digitization because it supports rapid, non-invasive acquisition in difficult environments. However, large indoor heritage scenes such as decorated ceilings often exhibit strongly unbalanced illumination, with side-window lighting, long capture trajectories, and automatic mobile camera processing causing view-dependent appearance inconsistencies. These effects are only partially handled by existing global photometric correction methods. We present PPISP-Local, a compact extension of a physically plausible image signal processing (ISP) correction model for radiance-field reconstruction. The method retains global exposure, vignetting, colour correction, and camera response modelling, and adds a low-capacity, luminance-only local tone-mapping module to better approximate mobile-phone ISP behaviour in unevenly lit indoor scenes. We evaluate the method on a high-ceiling heritage dataset and compare vanilla 3DGS, PPISP, and PPISP-Local on held-out ceiling views. PPISP-Local achieves the best reconstruction quality on this challenging smartphone dataset, suggesting that weak local tone adaptation is beneficial for large indoor heritage reconstruction when global photometric models are insufficient.

1 Introduction

Digital imaging has become central to cultural heritage documentation, especially where direct physical access is limited. In the case of elevated historic ceilings, digitization must balance safety, coverage, fidelity, and practicality. Recent digital heritage study on the Long Gallery ceiling at Blickling Hall [4] showed that a mobile-phone-driven 3DGS workflow can provide a useful base reconstruction for a large heritage interior, while preserving real-time navigation and contextual visualization in virtual reality. That case study involved a 37.5 m long and 4.8 m high historical ceiling, recorded non-invasively using 4K iPhone 12 video, subsampled to 994 frames for 3DGS reconstruction. The same study also reported that indoor lighting in the gallery is highly uneven, dominated by windows on one side, and that smartphone video with automatic tone mapping was more balanced than DSLR video for this scene.

These observations expose a broader imaging problem for heritage reconstruction. Even when geometry is static, the image appearance across views can vary substantially because of automatic exposure, white balance, camera response, and other image signal processing (ISP) operations. Such variation is particularly pronounced in mobile-phone video, where the camera often adapts aggressively to changing scene brightness during a long capture path. In large indoor heritage scenes, this can lead to inconsistent radiance estimation, reduced texture fidelity, and unstable novel-view synthesis.

Recent work on physically plausible photometric correction for radiance fields has addressed this issue by modelling image formation explicitly. PPISP introduces a differentiable ISP pipeline with exposure offset, chromatic vignetting, colour correction, and camera response function (CRF), together with a controller for predicting per-view photometric parameters at novel viewpoints [3]. This is a strong step toward interpretable appearance compensation. However, PPISP itself notes an important limitation: it does not model localized tone mapping, which is common in modern phone cameras [3]. In practice, mobile ISP often performs local luminance adaptation in addition to global exposure, especially in scenes with bright windows and darker architectural interiors.

This limitation is directly relevant to heritage imaging. High ceilings, long corridors, and side-lit interiors are common in historic houses, galleries, churches, and halls. In such spaces, uneven illumination is not a nuisance but part of the imaging reality. If the camera locally lifts shadows or compresses highlights while traversing the scene, a purely global correction model may leave residual appearance inconsistency. This motivates the present paper.

We propose PPISP-Local, a compact extension of PPISP designed for heritage scenes captured by mobile phone. The key idea is to keep the physically plausible global ISP model, but add a *small, luminance-only local tone-mapping residual*. This module is intentionally weak: it is low resolution, bounded, smooth, and residual around identity. The goal is not to replace the global model with a flexible image harmonizer, but to capture the low-frequency local adaptation that often appears in mobile-phone ISP. We evaluate it on the Blickling Hall high-ceiling dataset, comparing 3DGS, PPISP, and PPISP-Local on held-out ceiling views. Our contribution is therefore twofold: first, an imaging-method contribution in the form of a physically grounded local tone-mapping extension; second, an application contribution showing why such an extension matters for imaging for heritage in large indoor environments.

2 Related Work

2.1 Imaging and reconstruction for heritage ceilings

Ceiling digitization in heritage settings is challenging because of scale, distance, access constraints, and the need for non-invasive workflows. MacDonald *et al* used photometric stereo for damage visualization on a heritage ceiling, demonstrating fine deformation analysis without physical intervention [8, 9]. Guarneri *et al* employed multi-wavelength laser scanning for a frescoed ceiling in Palazzo Barberini, addressing large-scale capture under difficult illumination [5]. Reinoso-Gordo *et al* combined laser scanning and photogrammetry to document deformation and ornament in a sixteenth-century wooden ceiling [13]. More generally, optical 3D sensing has become a major tool in heritage recording, with strong overviews provided by Remondino and by Scopigno *et al* [12, 14].

Recent digital heritage study proposed a level-of-detail workflow for the Blickling Hall Long Gallery ceiling, using 3DGS for base reconstruction and structured-light scanning for selected regions of interest [4]. That paper established two facts relevant here. First, 3DGS is practical for a large, elevated indoor heritage scene captured from mobile video. Second, the capture environment exhibits severe lighting imbalance, and mobilephone video produced more balanced results than DSLR video under the same constraints [4]. The present work builds directly on that heritage imaging scenario.

2.2 3D Gaussian Splatting in heritage imaging

3D Gaussian Splatting (3DGS) enables efficient real-time rendering using explicit anisotropic Gaussian primitives [7]. It has quickly become attractive for immersive visualization and heritage documentation because it combines photorealistic appearance with practical performance. Recent heritage applications include object-scale digitization with integrated segmentation [2], large architectural documentation with UAV imagery [16], and broader discussions of Gaussian splatting as an immersive heritage medium [6]. Edwards *et al* showed that 3DGS outperformed an MVS+Poisson baseline on the Blickling ceiling under long-range video-based capture, while preserving real-time navigation suitable for VR [4].

Despite these advantages, 3DGS remains sensitive to photometric inconsistency in the input imagery. This is especially relevant for mobile-phone capture in indoor heritage scenes, where automatic imaging pipelines may change over the course of a scan.

2.3 *Photometric correction for radiance fields*

Appearance inconsistency in radiance field reconstruction has been addressed with a range of correction models. NeRF-W introduced per-image appearance embeddings for unconstrained photo collections [10]. Urban Radiance Fields used explicit affine transformations to improve robustness under real-world variation [11]. More expressive methods such as BilaRF model spatially varying appearance using bilateral processing [15]. ADOP also incorporated explicit image-formation parameters such as exposure, white balance, and camera response [1].

PPISP is especially relevant to our work because it formulates photometric correction as a physically plausible ISP pipeline with distinct modules for exposure, vignetting, colour correction, and CRF, and it emphasises the need to predict per-view parameters for novel viewpoints without access to the target image [3]. Importantly for our motivation, PPISP also identifies localized tone mapping as a limitation of the global model [3]. This is the point we address here.

3 Method

3.1 *Baseline PPISP formulation*

PPISP models the final image as a post-processing pipeline applied to the rendered radiance image. In the global version, the radiance image L is passed through exposure correction, vignetting, colour correction, and CRF to produce the final RGB output [3]. Exposure and colour correction are frame-specific, while vignetting and CRF are camera-shared. The pipeline can be trained jointly with the scene representation and then paired with a controller that predicts per-frame parameters for novel views [3].

3.2 *PPISP-Local*

We retain the global PPISP structure and add one extra module between colour correction and CRF:

$$L \to \mathcal{E} \to \mathcal{V} \to \mathcal{C} \to \mathcal{T}_{loc} \to \mathcal{G}.$$

Here $\mathcal{E}$ is exposure, $\mathcal{V}$ is vignetting, $\mathcal{C}$ is colour correction, $\mathcal{T}_{loc}$ is the new local tone-mapping module, and $\mathcal{G}$ is the CRF.

The local module is designed to model weak, low-frequency local luminance adaptation without introducing arbitrary spatial recolouring. Let I_{cc} denote the image after colour correction. We compute luminance as

$$Y(u) = w_R I_{cc}^R(u) + w_G I_{cc}^G(u) + w_B I_{cc}^B(u),$$

where u is a pixel location and (w_R, w_G, w_B) are fixed luminance weights.

For each frame f, we introduce a low-resolution local latent map $M_f \in \mathbb{R}^{h \times w}$, with default resolution 8×8. The map is bilinearly upsampled to full image resolution and converted into a bounded residual modulation field:

$$\Delta_f(u) = \alpha \tanh(\widetilde{M}_f(u)),$$

where $\widetilde{M}_f$ is the upsampled map and α is a small scalar controlling the strength of the local tone adjustment. The luminance is then modified as $Y'_f(u) = Y(u)(1 + \Delta_f(u))$.

To preserve chromaticity, the colour-corrected RGB image is rescaled by the luminance ratio:

$$I_{loc}(u) = I_{cc}(u) \frac{Y'_f(u)}{Y(u) + \epsilon}.$$

The local tone-mapping operator is therefore $\mathcal{T}_{loc}(I_{cc}; M_f) = I_{ltm}$.

This design has four useful properties. First, it reduces to the original PPISP formulation when $M_f = 0$. Second, it remains luminance-only, avoiding local colour hallucination. Third, it is low capacity because the map is coarse and smooth. Fourth, its effect is bounded by α, preventing the module from acting as a strong local image editor.

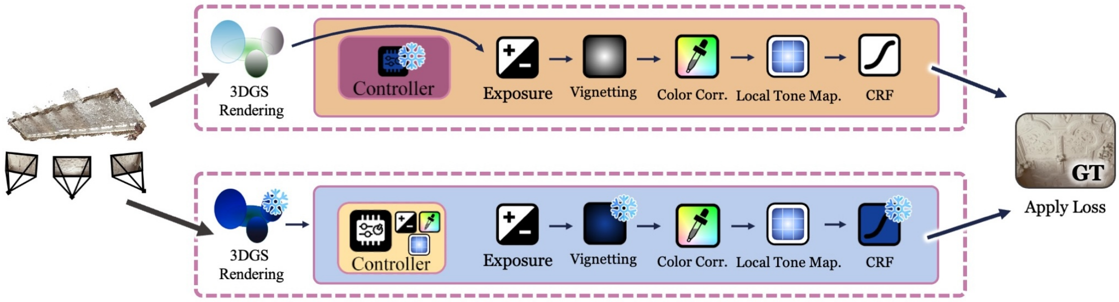

Figure 1: Overview of PPISP-Local. The method keeps the physically plausible global PPISP pipeline and inserts a bounded luminance-only local tone-mapping module before the CRF.

3.3 Training

We follow the same two-stage logic as PPISP. In Stage 1, the scene representation, the global PPISP parameters, and the per-frame local tone maps are jointly optimised on the training views. The reconstruction loss is the same photometric loss used by the base renderer, applied after the full PPISP-Local pipeline. We add regularisation on the local maps to keep them weak and smooth:

$$\mathcal{L}_{loc} = \lambda_{id}\|M_f\|_2^2 + \lambda_{tv}\mathrm{TV}(M_f) + \lambda_{amp}\|\Delta_f\|_2^2.$$

In Stage 2, the scene and shared ISP terms are frozen, and a controller is trained to predict the per-frame exposure, colour correction, and local map coefficients from the rendered radiance image, enabling application to novel views.

4 Experiments and Results

4.1 Dataset and evaluation setup

We evaluate our method on the Blickling Hall Long Gallery ceiling dataset, a large-scale indoor heritage scene introduced in prior work [4]. The ceiling spans 37.5 m in length and 4.8 m in height, and was captured using 4K iPhone 12 video under strongly unbalanced illumination conditions. The video was subsampled to 994 frames for Structure-from-Motion and 3DGS reconstruction, as shown in Figure 2.

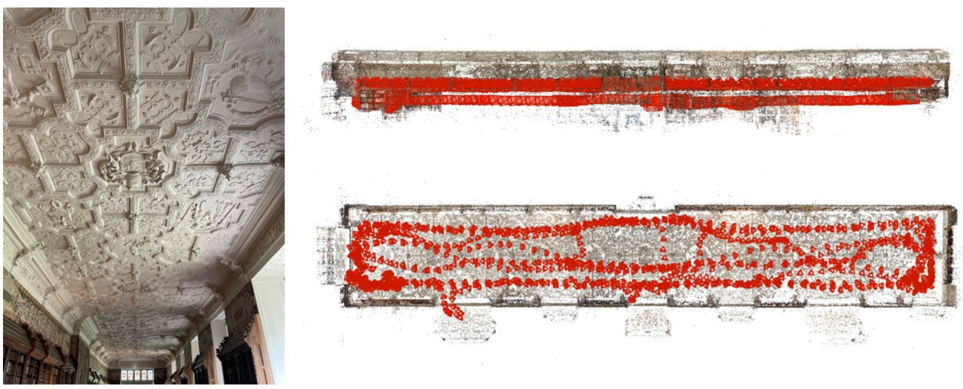

Figure 2: Left is a photo of the long ceiling. Right is the camera pose estimation and sparse point cloud initialization (side view and overhead trajectory).

This dataset is particularly challenging due to (i) strong side-window lighting, (ii) long capture trajectories with changing viewpoints, and (iii) automatic mobile ISP effects such as exposure adaptation and tone mapping. These characteristics make it well suited for evaluating photometric consistency in radiance-field reconstruction. In particular, we are interested in how well each method handles view-dependent appearance changes caused by uneven illumination and mobile ISP behaviour.

4.2 Quantitative results

Table 1 reports quantitative results for both training and validation splits. Across all metrics, PPISP improves substantially over vanilla 3DGS, confirming the importance of explicit photometric modelling. More importantly, PPISP-Local consistently outperforms PPISP on both splits. Although the numerical gains are moderate, they are consistent across all metrics and reflect improved appearance stability rather than simple global brightness shifts. These results indicate that a weak local tone-mapping component provides measurable benefits beyond global ISP correction in this challenging indoor setting.

Table 1: Quantitative comparison on the Blickling ceiling dataset. Results are reported on both the training and held-out validation splits. Higher is better for PSNR and SSIM, while lower is better for LPIPS.

	Train			Val		
Method	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
3DGS(gsplat)	19.1289	0.6745	0.4930	17.8560	0.6454	0.5198
PPISP	20.2568	0.6978	0.4401	19.1009	0.6701	0.4625
PPISP-Local	**20.5510**	**0.7007**	**0.4352**	**19.5770**	**0.6739**	**0.4565**

4.3 Qualitative results

Figure 3 shows qualitative comparisons on the rendered whole views from 3DGS reconstruction. Vanilla 3DGS (gsplat) exhibits noticeable brightness inconsistencies and loss of detail in regions affected by uneven illumination. PPISP improves overall balance, but residual inconsistencies remain, particularly in shadowed or window-facing regions. PPISP-Local further improves visual consistency by better preserving detail across illumination gradients. In particular, it reduces overexposed highlights and recovers structure in darker regions, leading to more coherent appearance across viewpoints.

Figure 4 visualizes the predicted local tone maps. These maps are smooth and low-frequency, confirming that the model captures weak spatial adaptation rather than performing arbitrary local editing. As shown in the examples, the tone maps correlate with illumination variation in the scene, modulating luminance in a controlled and physically plausible manner.

4.4 Discussion of results

The improvements observed with PPISP-Local are consistent with the imaging characteristics of the dataset. As smartphone capture already performs local tone adaptation during acquisition, a purely global correction model cannot fully account for this behaviour, leading to residual inconsistencies.

By introducing a constrained local luminance adjustment, PPISP-Local better matches the effective image formation process of mobile cameras. Importantly, the gains are achieved without introducing high-capacity spatial appearance models, preserving the interpretability and stability of the original PPISP formulation.

Figure 3: Qualitative comparison on the rendered whole ceiling views. All methods trained only for 10000 steps, with PPISP and PPISP-Local trained at stage one for 8000 steps, then the controller activated for distillation.

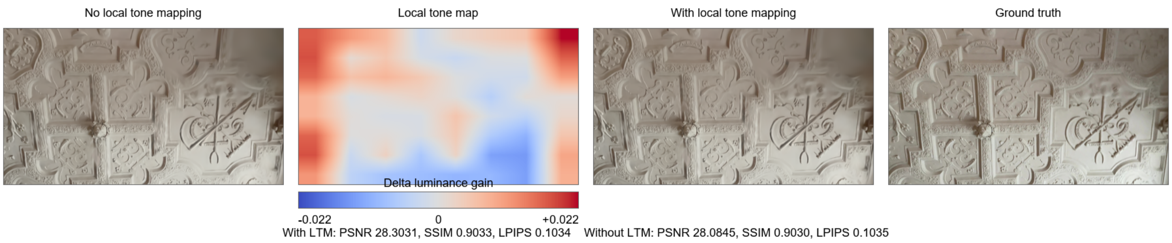

Figure 4: Visualisation of local tone maps predicted by PPISP-Local. These maps remain smooth and low-frequency, reflecting weak local luminance adaptation.

5 Conclusion

We presented PPISP-Local, a local tone-mapping extension of PPISP for 3DGS reconstruction in heritage interiors captured with mobile phones. The method was motivated by the Blickling Hall high-ceiling dataset, where smartphone capture proved practical under highly uneven indoor lighting. By adding a weak local luminance adaptation module to a physically plausible global ISP pipeline, PPISP-Local better matches the effective imaging behaviour of mobile cameras in this setting.

We evaluate the method on held-out ceiling views, comparing vanilla 3DGS, PPISP, and PPISP-Local. PPISP-Local achieves the best reconstruction performance on this heritage scene, supporting the idea that local tone adaptation is beneficial for large indoor heritage reconstruction when illumination is unbalanced and mobile phone ISP plays a substantial role.

More broadly, this work suggests that heritage imaging pipelines can benefit from physically grounded photometric correction methods that are aware not only of optics and exposure, but also of weak local ISP behaviour in practical mobile capture.

Acknowledgements

The work is funded by UEA's AHRC Impact Accelerator Award (AH/X003442/1) and Royal Society ISPF International Collaboration Award (ICA/R2/252130).

References

[1] Darius Rückert, Linus Franke, and Marc Stamminger. 2022. ADOP: Approximate Differentiable One-Pixel Point Rendering. *ACM Transactions on Graphics* 41, 4, Article 99 (July 2022), 14 pages. https://doi.org/10.1145/3528223.3530122

[2] Mahtab Dahaghin, Myrna Castillo, Kourosh Riahidehkordi, Matteo Toso, and Alessio Del Bue. 2024. Gaussian Heritage: 3D Digitization of Cultural Heritage with Integrated Object Segmentation. *arXiv preprint arXiv:2409.19039*. https://doi.org/10.48550/arXiv.2409.19039

[3] Isaac Deutsch, Nicolas Moënne-Loccoz, Gavriel State, and Zan Gojcic. 2026. PPISP: Physically-Plausible Compensation and Control of Photometric Variations in Radiance Field Reconstruction. *arXiv preprint arXiv:2601.18336*. https://doi.org/10.48550/arXiv.2601.18336

[4] Jacob Edwards, Stephen D. Laycock, Thomas Roebuck, YingLiang Ma, and Cheng Wang. 2025. Level-of-Detail Digitization of High Ceilings in Virtual Reality. In *Digital Heritage 2025 – 4th International Congress*, Siena, Italy.

[5] Massimiliano Guarneri, Simone Ceccarelli, Mario Ferri De Collibus, Massimo Francucci, and Massimiliano Ciaffi. 2019. Multi-Wavelengths 3D Laser Scanning for Pigment and Structural Studies on the Frescoed Ceiling "The Triumph of Divine Providence". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-2/W15 (2019), 549–554. https://doi.org/10.5194/isprs-archives-XLII-2-W15-549-2019

[6] Osama Jamil and AnnMarie Brennan. 2025. Immersive Heritage through Gaussian Splatting: A New Visual Aesthetic for Reality Capture. *Frontiers in Computer Science* 7 (2025), Article 1515609. https://doi.org/10.3389/fcomp.2025.1515609

[7] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics* 42, 4, Article 139 (July 2023), 14 pages. https://doi.org/10.1145/3592433

[8] Lindsay MacDonald, Ian Gibb, Stuart Robson, and Constantina Vlachou-Mogire. 2012. High Art: Visualising Damage on a Heritage Ceiling. In *Electronic Visualisation and the Arts (EVA 2012)*, London, UK, 309–316. https://doi.org/10.14236/ewic/EVA2012.53

[9] Lindsay MacDonald, Ali Hosseininaveh Ahmadabadian, and Stuart Robson. 2014. Photogrammetric Analysis of a Heritage Ceiling. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5 (2014), 379–383. https://doi.org/10.5194/isprsarchives-XL-5-379-2014

[10] Ricardo Martin-Brualla, Noha Radwan, Mehdi S. M. Sajjadi, Jonathan T. Barron, Alexey Dosovitskiy, and Daniel Duckworth. 2021. NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 7210–7219.

[11] Konstantinos Rematas, Andrew Liu, Pratul P. Srinivasan, Jonathan T. Barron, Andrea Tagliasacchi, Thomas Funkhouser, and Vittorio Ferrari. 2022. Urban Radiance Fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 12932–12942.

[12] Fabio Remondino. 2011. Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. *Remote Sensing* 3, 6 (2011), 1104–1138. https://doi.org/10.3390/rs3061104

[13] Juan Francisco Reinoso-Gordo, Antonio Gámiz-Gordo, and Pedro Barrero-Ortega. 2021. Digital Graphic Documentation and Architectural Heritage: Deformations in a 16th-Century Ceiling of the Pinelo Palace in Seville (Spain). *ISPRS International Journal of Geo-Information* 10, 2 (2021), Article 85. https://doi.org/10.3390/ijgi10020085

[14] Roberto Scopigno, Paolo Cignoni, Nico Pietroni, Marco Callieri, and Matteo Dellepiane. 2017. Digital Fabrication Techniques for Cultural Heritage: A Survey. *Computer Graphics Forum* 36, 1 (2017), 6–21. https://doi.org/10.1111/cgf.12781

[15] Yuehao Wang, Chaoyi Wang, Bingchen Gong, and Tianfan Xue. 2024. Bilateral Guided Radiance Field Processing. *ACM Transactions on Graphics* 43, 4, Article 148 (July 2024). https://doi.org/10.1145/3658148

[16] Yingwen Yu, Edward Verbree, Peter van Oosterom, and Uta Pottgiesser. 2025. 3D Gaussian Splatting for Modern Architectural Heritage: Integrating UAV-Based Data Acquisition and Advanced Photorealistic 3D Techniques. *AGILE: GIScience Series* 6 (2025), Article 51. https://doi.org/10.5194/agile-giss-6-51-2025