ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XI-2-2026
XXV ISPRS Congress 2026 “From Imagery to Understanding”, Commission II, 4–11 July 2026, Toronto, Canada

Refraction-Aware Gaussian Splatting for Shallow Water Bathymetry from UAV Imagery

Taiki Uno[1], Sohei Kobayashi[2]

[1] Kyoto University, Graduate School of Engineering, Kyoto, Japan - unotaiki55@gmail.com
[2] Kyoto University, Disaster Prevention Research Institute, Uji, Japan - kobayashi.sohei.7a@kyoto-u.ac.jp

Keywords: Bathymetry, Gaussian Splatting, Photogrammetry, Refraction Correction, UAV Imagery

Abstract

Unmanned Aerial Vehicles (UAV)-based photogrammetry provides an efficient solution for shallow water bathymetry, yet its accuracy is fundamentally constrained by light refraction at the air-water interface, which violates the central geometric assumptions of traditional photogrammetry. Existing approaches, ranging from empirical corrections and iterative post-processing to black-box deep learning, often compromise geometric fidelity, physical interpretability, or generalization. We address this challenge through Refraction-Aware Gaussian Splatting (RA-GS), which embeds a physically rigorous two-media refraction model directly into the Gaussian Splatting (GS) reconstruction pipeline. Rather than relying on computationally expensive per-pixel ray tracing, we formulate an analytical parameter transformation that maps the true underwater position, scale, and opacity of each Gaussian to their apparent states observed through a planar refractive interface. Through this fully differentiable transformation, true underwater 3D geometry and photorealistic appearance are jointly optimized by directly minimizing the photometric error within the standard GS framework. This approach relies solely on RGB imagery, eliminating the need for external depth priors or deep learning networks. Using a physically based, ray-traced synthetic riverbed dataset, we isolate and explicitly correct refractive distortions. Our method achieves a geometric F1-score of 94% (10 cm threshold at 10 m depth) and produces high-quality novel view synthesis with a PSNR of 25.9 dB and SSIM of 0.93. Field experiments on real UAV data corroborate the practical utility for high-precision bathymetric mapping under calm-surface conditions. By resolving the fundamental refractive difficulty, the proposed framework provides a physically grounded, computationally efficient, and practically useful solution for next-generation photogrammetric bathymetry.

1. Introduction

1.1 Background

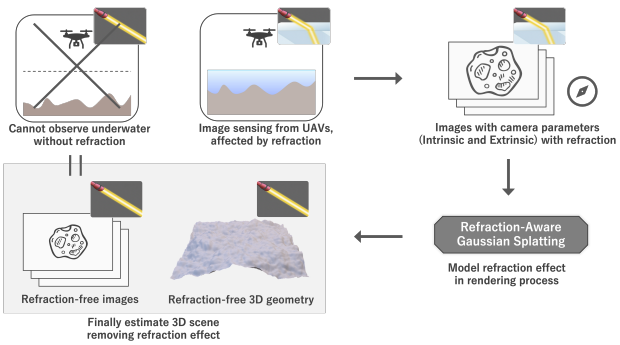

Figure 1. Overview of our proposed method for photogrammetric bathymetry. Our Refraction-Aware Gaussian Splatting takes multi-view images distorted by refraction (top right) as input. By explicitly modeling refraction within the rendering pipeline, it reconstructs a geometrically accurate and refraction-free 3D scene (bottom left).

Accurate bathymetric information is crucial for applications such as flood forecasting (Grimaldi et al., 2018) or river morphology monitoring (Hemmelder et al., 2018). Traditional ship-borne sonar (Giordano et al., 2015), while effective in deeper waters, is limited in shallow environments due to grounding risk, narrow swath coverage, and high operational costs. Airborne LiDAR Bathymetry (ALB) offers improved coverage (Saylam et al., 2018), but the required hardware and flight operations remain expensive and difficult to deploy at high temporal frequency.

In recent years, Unmanned Aerial Vehicle (UAV) photogrammetry has become a widely used, low-cost, and efficient technique for terrestrial mapping (Bemis et al., 2014, Gomez and Purdie, 2016, Iglhaut et al., 2019). Extending it to derive underwater topography—known as photogrammetric bathymetry—is therefore an attractive alternative (He et al., 2024). However, when images are acquired from above the water surface, they are affected by several complex optical phenomena, including specular highlights, wave-induced distortions, scattering, and most critically, refraction at the air-water interface. Among these, refraction is the principal challenge because conventional photogrammetry and SfM assume that light travels in a straight line (Furukawa and Ponce, 2010, Schönberger and Frahm, 2016). As a result, naively applying standard pipelines leads to systematic depth errors and inconsistent multi-view geometry.

1.2 Related Work

The fundamental challenge in photogrammetric bathymetry is the ”two-media problem,” wherein refraction at the air-water interface inherently violates the assumption of straight-line light propagation underlying the standard pinhole camera model (Rinner, 1948). Early methods attempted to mitigate these multi-view inconsistencies via post-hoc empirical corrections, such as uniform depth scaling (Woodget et al., 2014) or iterative, view-dependent ray tracing (Dietrich, 2016). However, the former lacks theoretical rigor in a multi-view context and explicitly ignores horizontal refractive displacements. Conversely, while view-dependent ray tracing models the physics more accurately, it requires iterative computational overhead.

More recently, data-driven approaches have leveraged deep learning for image-domain undistortion (Agrafiotis et al., 2020)

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-605-2026

or direct depth regression (Mandlburger et al., 2021). While capable of approximating complex non-linear refractive mappings, these networks are often limited by site-specific inductive biases, limited cross-domain generalization, and a strong dependency on expensive, large-scale ground-truth annotations such as ALB.

Several recent methodologies strive to explicitly integrate physical refraction models directly into the 3D reconstruction pipeline. For instance, Refraction-Aware SfM (R-SfM) (Makris et al., 2024) successfully incorporates Snell's law into bundle adjustment but yields predominantly sparse point clouds, necessitating a reliance on auxiliary learning-based techniques for densification. Concurrently, implicit neural representations like NeRFrac (Zhan et al., 2023) address dynamic surface refraction through latent embeddings. Yet, these volumetric rendering techniques introduce prohibitive computational bottlenecks and frequently lack strict metric interpretability.

1.3 Refraction-Aware Gaussian Splatting

To address these limitations, this work leverages recent advancements in differentiable rendering to propose Refraction-Aware Gaussian Splatting (RA-GS) (Figure 1), a novel 3D reconstruction framework that explicitly integrates a physical refraction model directly into the Gaussian Splatting (GS) pipeline (Kerbl et al., 2023). While GS represents scenes explicitly as a set of anisotropic Gaussians which offer rapid optimization, real-time rendering, and highly interpretable spatial geometry, it inherently assumes a single uniform medium.

The core contribution of our work lies in formulating the optical observation of submerged Gaussians through the air-water interface as an analytical parameter transformation (Figure 2). Rather than relying on computationally prohibitive pixel-wise ray tracing, our method projects each Gaussian's physical properties, specifically its spatial position, scale, and opacity, into an apparent state. This formulation enables highly efficient, refraction-aware rasterization.

Furthermore, because this analytical transformation is fully differentiable, it seamlessly integrates into the standard GS optimization framework, such as the original 3D Gaussian Splatting or recent 2D variants (Huang et al., 2024). Consequently, the true underwater 3D geometry is directly learned and estimated by minimizing the photometric error between the input multi-view imagery and the rendered views. As a result, our framework achieves both rigorous geometric accuracy and photorealistic visual fidelity using solely RGB images, completely bypassing the reliance on expensive external ground-truth data typically required by data-driven deep learning approaches.

Experiments conducted through 3DGS on a synthetic riverbed dataset demonstrate that our method delivers photorealistic rendering with a PSNR exceeding 25 dB and precise geometric reconstruction with an F1-score above 90%. Furthermore, a real-world field demonstration conducted based on 2DGS validates the method's capability for high-precision, in-situ bathymetric surveying. These results highlight our approach as a physically grounded and computationally efficient solution for next-generation photogrammetric bathymetry.

The code and datasets are available at https://github.com/unotaiki/Refractive-Aware-Gaussian-Splatting.

2. Methods

Our goal is to reconstruct a refraction-free 3D representation from aerial images that are distorted by the air–water interface. To achieve this, we incorporate a physically grounded refractive transformation into the GS rendering pipeline. For each view, every Gaussian is mapped from its true underwater parameters ($\boldsymbol{p}$, $\boldsymbol{s}$, α) to an apparent set ($\boldsymbol{p}'$, $\boldsymbol{s}'$, α') via a differentiable three-component transformation: position correction (Section 2.1), scale correction (Section 2.2), opacity correction (Section 2.3).

The proposed RA-GS framework operates under the assumption of a calm, planar water surface with a known interface elevation and a constant refractive index. Taking multi-view imagery with pre-calibrated camera intrinsics and extrinsics as input, the pipeline optimizes a set of Gaussian primitives to yield a dense, refraction-free geometry of the in-situ underwater scene. To explicitly isolate and resolve the fundamental geometric distortions governed by Snell's law, complex radiometric phenomena such as light attenuation and scattering, and specular surface reflections are deliberately excluded from the current optical model.

2.1 Position Correction

We assume a camera located at $(0, 0, H)$, a planar and calm water surface at $z = 0$, and an underwater Gaussian center $\boldsymbol{p} = (x, y, z)$, where $z < 0$. Because refraction is rotationally symmetric around the viewing direction, we reduce the geometry to the rz-plane using the radial distance $r = \sqrt{x^2 + y^2}$.

Let s be the radial coordinate of the ray–interface intersection ($0 \le s < r$). Let θ_w and θ_a denote, respectively, the ray angles with respect to the interface normal in water and air. With refractive index n for water (air assumed to be 1), Snell's law reads:

$$\sin\theta_a = n \sin\theta_w \tag{1}$$

By relating these angles to the geometry of the triangles formed by the camera, the interface, and the underwater point, we obtain a quartic equation in s (triangle similarity between $\triangle OAI$ and $\triangle IPC$):

$$\begin{aligned}(1-n^2)s^4 + 2(n^2-1)rs^3 \\ + ((1-n^2)r^2 - n^2H^2 + z^2)s^2 \\ + 2n^2H^2rs - n^2H^2r^2 = 0.\end{aligned} \tag{2}$$

The physically valid root s can be computed numerically using Newton iterations or Ferrari's closed-form solution.

Once s is known, the geometric relations ($\triangle IPC$, $\triangle IP'C'$, and $\triangle I'P'C'$):

$$\begin{aligned} r - s &= -z\tan\theta_w \\ r' - s &= z'\tan\theta_a \\ -z'\tan(\theta_a + \Delta\theta_a) &= (r'-s) + (-\Delta s)\end{aligned} \tag{3}$$

combined with Eq. (1) and the derivative of Snell's law:

$$\frac{d\theta_w}{d\theta_a} = \frac{1}{n}\frac{\cos\theta_a}{\cos\theta_w} \tag{4}$$

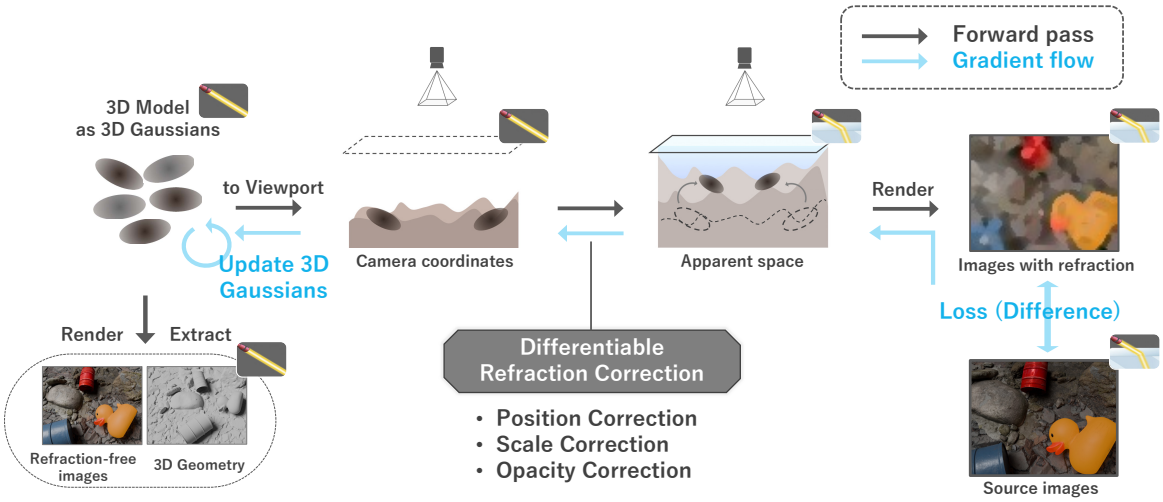

Figure 2. Overview of Refraction-Aware Gaussian Splatting. A 3D scene represented by Gaussian primitives is transformed into an apparent space for each camera view by applying a differentiable refractive transformation consisting of position, scale, and opacity corrections. These transformations explicitly model the light bending at the air-water interface. The transformed Gaussians are rendered and compared with the input images, which inherently contain refractive distortions. The reconstruction loss is backpropagated to optimize the original Gaussians, enabling the recovery of a refraction-free 3D model and photorealistic novel views.

lead to closed-form expressions for the apparent position (Nassar, 1994, Missailidis and Guilhon, 2025):

$$\begin{cases} r' = r + (n^2 - 1)\, z \tan^3 \theta_w, & \text{(5a)} \\ z' = \dfrac{1}{n}\left(\dfrac{\cos\theta_a}{\cos\theta_w}\right)^3 z. & \text{(5b)} \end{cases}$$

Because $z < 0$ and $\theta_a > \theta_w$, these relations guarantee $r' < r$ and $|z'| < |z|$; i.e., underwater points appear both shallower and closer to the optical axis (Figure 3). For $r = 0$ (nadir view), the formula simplifies to the well-known approximation $z' = z/n$.

To integrate the refractive transformation into the differentiable rasterization pipeline, we must compute the Jacobian of the apparent position mapping for accurate backpropagation. The apparent Jacobian J_{app}, which maps the physical coordinates (r, z) to the apparent coordinates (r', z'), is analytically formulated as follows:

$$J_{\text{app}} = \frac{\partial(r', z')}{\partial(r, z)} = \begin{pmatrix} 1 + \dfrac{3(n^2-1) z \tan^2\theta_w \sec^2\theta_w}{\Delta} & (n^2-1)\tan^3\theta_w \left(1 + \dfrac{3z\sec^2\theta_w}{\Delta}\right) \\ \dfrac{3z'\tan\theta_w}{\Delta}\left(1 - n^2\dfrac{\cos^2\theta_w}{\cos^2\theta_a}\right) & \dfrac{z'}{z} + \dfrac{3z'\tan^2\theta_w}{\Delta}\left(1 - n^2\dfrac{\cos^2\theta_w}{\cos^2\theta_a}\right) \end{pmatrix} \quad (6)$$

where Δ is defined as:

$$\Delta = nH\frac{\cos\theta_w}{\cos^3\theta_a} - z\sec^2\theta_w. \quad (7)$$

2.2 Scale Correction

The nonlinear refractive mapping $\mathbf{p} \rightarrow \mathbf{p}'$ induces view-dependent spatial compression (Figure 4). Without compensatory scaling, this distortion causes uncorrected Gaussians to blur and unnaturally expand at the surface (Figure 5).

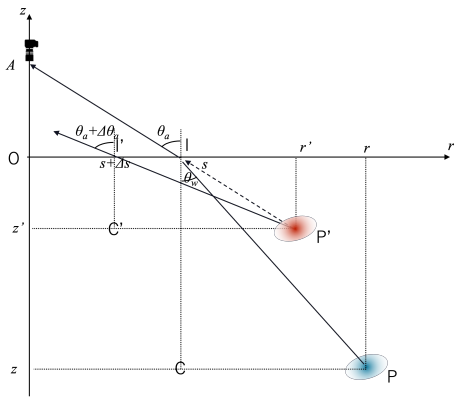

Figure 3. Illustration of the apparent position P' versus the true position P in the rz-plane. Even though the actual object is located at the blue Gaussian position P, it appears shifted toward the red Gaussian position P' when observed through the refractive interface.

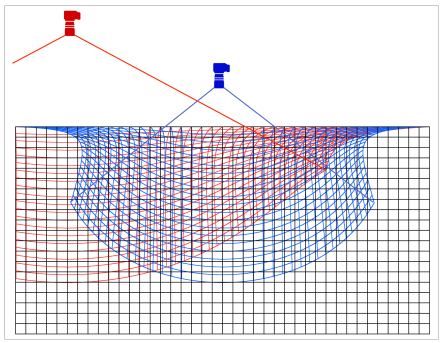

Figure 4. Space compression from the camera view. The space is distorted by the camera position. Each colored grid represents the corresponding compressed space of the entire grid in camera coordinates (shown in black.)

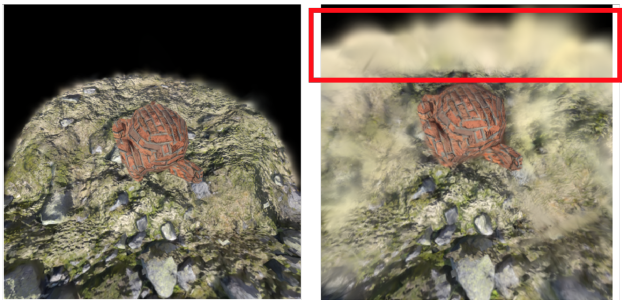

Figure 5. Refractive scaling artifacts. Left: Ground truth rendering without a refractive interface. Right: Rendering with only position correction, exhibiting severe expansion artifacts at large incidence angles (red box).

To mitigate these artifacts, we quantify the local anisotropic distortion by extracting directional scale factors from the apparent Jacobian J_{app}. Specifically, the scale factor for the x-axis is given by the norm of the transformed basis vector, $s_x = \|J_{\text{app}} \boldsymbol{e}_x\|$, with s_y and s_z computed analogously.

While the determinant $|\det J_{\text{app}}|$ serves as a mathematically strict volumetric measure, empirical testing revealed that it overcompensates at large incidence angles, forcing Gaussians to shrink excessively. Instead, we formulate an isotropic scale correction factor S based on the geometric mean of the directional scales:

$$S = (s_x s_y s_z)^{1/3} \tag{8}$$

The corrected Gaussian scale is simply updated as:

$$\boldsymbol{s}' = S\boldsymbol{s} \tag{9}$$

This isotropic regularization effectively eliminates refractive expansion artifacts while maintaining training stability and preventing extreme primitive deformation.

2.3 Opacity Correction

Observing scenes through a refractive interface inherently reduces measured radiance due to etendue conservation (Winston et al., 2018), scaling the outgoing radiance by a factor of $1/n^2$ (Figure 6). In standard GS, optimizing against these physically attenuated target images frequently causes the model to converge to artificially low opacities. While these semi-transparent Gaussians minimize photometric loss by explaining the low observed brightness, they produce geometrically ambiguous, blurry, and under-constrained reconstructions.

To counteract this degradation, we introduce an opacity regularization. By explicitly scaling the apparent opacity of submerged Gaussians proportionally to the radiance loss:

$$\alpha' = \frac{1}{n^2}\alpha \tag{10}$$

we effectively reduce the opacity budget available to each primitive. While this serves as a principled heuristic rather than a strict volumetric scattering model, it explicitly penalizes trivial semi-transparent solutions. Consequently, the optimizer is forced to reconstruct the scene using a denser, spatially coherent set of primitives, significantly reducing floaters and improving overall geometric fidelity.

3. Experiments

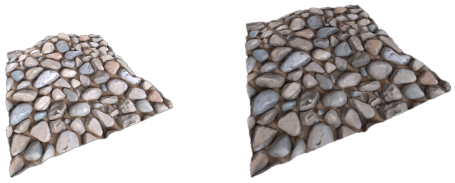

Figure 6. Comparison of path-traced renderings in a synthetic riverbed dataset (Section 3). The right image, viewed through a water surface, exhibits substantial radiance reduction compared to the left image rendered without refraction.

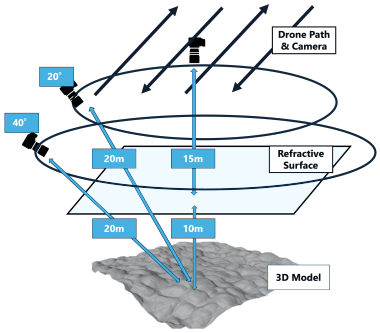

Figure 7. Overview of the simulated data capture setup.

Dataset To isolate geometric refractive effects from confounding radiometric phenomena such as water turbidity and specular surface reflections, we generated a controlled synthetic dataset using the Blender Cycles path tracer (Community, 2025) (Figure 7). The scene features a displaced gravel riverbed at an average depth of 10 m beneath a flat water surface. The dataset comprises 90 training images from a simulated pinhole camera (70° field of view, 800 × 800 pixels). To obtain robust multi-view coverage and evaluate the limits of refractive distortion, we rendered nadir views with 85% forward and side overlap, together with oblique views at 20° and 40° inclination; the viewpoint azimuth advanced in 10° steps around the target for full coverage.

Implementation We implemented RA-GS by extending the 3DGS implementation in the open-source gsplat framework (Ye et al., 2025). The refractive transformation is integrated into the differentiable rasterizer as a custom PyTorch autograd function (Paszke et al., 2019). Camera intrinsics and extrinsics are treated as known priors; the refractive index is fixed at 1.33, and the water surface is modeled as a horizontal plane at elevation 0 m with known position and its normal. Gaussians are initialized on a regular grid on a horizontal plane at 10 m depth. We optimize for 30,000 iterations on a single NVIDIA RTX 3090 GPU with the standard 3DGS adaptive density control hyperparameters.

Geometry Extraction To quantitatively evaluate bathymetric accuracy, we extract a height map from the optimized Gaussians. Raw points are sampled in proportion to Gaussian volume (Stuart et al., 2025) and filtered to remove floating artifacts via local-plane outlier filtering (Girardeau-Montaut et al., 2016). The cleaned point cloud is discretized onto a 2D horizontal grid; the median height in each cell yields the final surface-aligned point cloud. For this synthetic dataset, we use a 5 cm grid spacing and tune filtering to balance surface smoothness

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-605-2026 | © Author(s) 2026. CC BY 4.0 License.

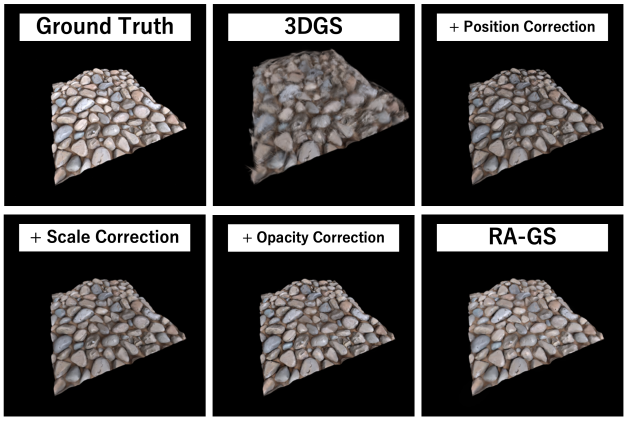

Figure 8. Qualitative comparison of appearance reconstruction. We show ground truth, the baseline 3DGS, and ablations of individual correction components.

with detail preservation; in general, grid resolution and filter settings are adjusted per dataset.

4. Results

The evaluation is performed on the synthetic dataset described in Section 3, where all training images contain distortions induced by a planar air–water interface. Our goal is to assess how effectively the proposed refractive corrections recover both appearance and geometry relative to ground truth.

4.1 Appearance Evaluation

Following established protocol in Novel View Synthesis (NVS), we evaluate similarity between rendered images from unseen viewpoints and ground truth. We employ three widely used metrics: the Peak Signal-to-Noise Ratio (PSNR) for pixel-wise accuracy; the Structural Similarity Index Measure (SSIM) (Wang et al., 2004) for local structure; and the Learned Perceptual Image Patch Similarity (LPIPS) (Zhang et al., 2018), which uses deep features from a convolutional neural network (CNN) to approximate human perception. To fully characterize the impact of refraction, metrics are reported under two conditions (Table 1): (1) rendering with the refractive interface and comparing to refracted ground truth, and (2) rendering without the interface and comparing to refraction-free ground truth.

Across all metrics, Gaussian center position correction provides the largest performance gain, consistent with its role in resolving refractive ray-path deviations. SSIM, which is particularly relevant for downstream tasks such as riverbed classification, rises from 0.68 (baseline) to above 0.95 with position correction alone. Opacity correction further improves perceptual quality, achieving the lowest LPIPS among all ablations. However, when evaluated in the non-refractive rendering condition, the PSNR remains considerably lower—approximately 13 dB below the refractive case. This is expected: as discussed in Section 2.3, the uniform opacity reduction preserves background visibility but does not physically model the reduction in radiance occurring at the refractive interface. Consequently, small luminance deviations are heavily penalized by PSNR despite being visually negligible. Scale correction exhibits a mixed effect. Although geometrically beneficial (as shown in Table 2), isotropic covariance correction introduces slight blurring in fine textures,

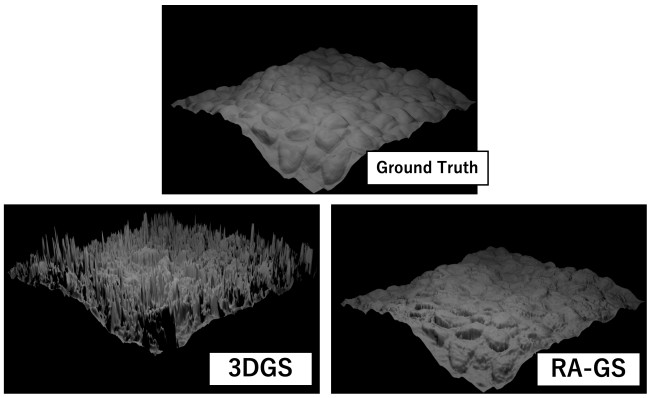

Figure 9. Qualitative geometry comparison. Top: ground truth geometry rendered in the simulator. Bottom-left: extracted geometry without correction. Bottom-right: extracted geometry with all corrections applied. Each image is rendered from an identical viewpoint under the same lighting conditions. The ground truth and the corrected geometry agree closely in position and shape; the uncorrected geometry is noisy and biased toward shallower depths.

marginally reducing SSIM in non-refractive rendering. Nonetheless, all three corrections together yield a well-balanced representation with competitive perceptual fidelity (Figure 8).

Despite a training time roughly three times that of standard 3DGS, the absolute runtime remains practical and can be further reduced via a dedicated CUDA implementation.

4.2 Geometry Evaluation

Accurate reconstruction of underwater terrain from aerial imagery is fundamentally challenging due to refractive ray bending. We evaluate geometric accuracy using Chamfer distance and the F1 score, two standard metrics for point-cloud comparison (Seitz et al., 2006). Chamfer distance captures the average nearest-neighbor deviation between reconstructed and ground-truth surfaces, while the F1 score balances precision and recall under a specified tolerance. Following practical requirements for high-resolution bathymetric mapping (Agrafiotis et al., 2020), we adopt 10 cm and 30 cm thresholds as realistically stringent criteria for operational use.

As shown in Table 2, the baseline 3DGS fails in the refractive scenario, yielding a CD of 8.477 m and near-zero F1 scores. Introducing position correction alone resolves most of the refractive displacement, reducing the CD by two orders of magnitude to 0.110 m. Scale correction further refines spatial consistency, bringing the CD down to 0.033 m and improving the 10 cm F1 score from 76.51% to 85.79%. Opacity correction has a smaller geometric influence but contributes modest improvements in both precision and recall.

The full model (Ours), combining all three corrections, achieves the best performance, with a CD of only 0.011 m and F1 scores of 94.00% (10 cm) and 99.12% (30 cm). These results are obtained at water depths up to 10 m, demonstrating robustness to severe refractive distortion. Qualitative comparisons in Figure 9 confirm these findings: the corrected geometry closely matches the ground truth in position and shape, whereas the uncorrected reconstruction collapses into a shallow, noisy surface.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-605-2026 | © Author(s) 2026. CC BY 4.0 License.

Table 1. Quantitative appearance evaluation under two rendering conditions: (1) rendering with the refractive interface and comparing to refractive ground truth, and (2) rendering without the interface and comparing to refraction-free ground truth. We show the ground truth, the baseline 3DGS, and ablation results where each component (scale or opacity) is added on top of position correction. The full method (Ours) applies position, scale, and opacity correction simultaneously. All models are trained from refracted images only.

Method	Correction Components			Render w/ Refraction			Render w/o Refraction			Stats	
	Position	Scale	Opacity	$SSIM^{\uparrow}$	$PSNR^{\uparrow}$	$LPIPS^{\downarrow}$	$SSIM^{\uparrow}$	$PSNR^{\uparrow}$	$LPIPS^{\downarrow}$	# Gaussians (K)	Time (min)
3DGS	✗	✗	✗	-	-	-	0.682	14.12	0.292	40.0	6.8
Position	✓	✗	✗	0.980	37.56	0.023	0.953	21.49	0.039	23.1	10.9
+ Scale	✓	✓	✗	0.964	37.21	0.023	0.936	22.97	0.043	28.0	21.0
+ Opacity	✓	✗	✓	0.984	38.44	0.018	0.954	25.27	0.033	38.8	11.5
RA-GS (Ours)	✓	✓	✓	0.981	38.42	0.017	0.933	25.97	0.046	41.3	21.9

Table 2. Quantitative geometry evaluation using Chamfer distance (CD) and F1 score at 10 cm and 30 cm thresholds. All models are trained on refracted images.

Method	Correction Components				10cm			30cm		
	Position	Scale	Opacity	$CD^{\downarrow}$ (m)	$Precision^{\uparrow}$ (%)	$Recall^{\uparrow}$ (%)	$F1^{\uparrow}$ (%)	$Precision^{\uparrow}$ (%)	$Recall^{\uparrow}$ (%)	$F1^{\uparrow}$ (%)
3DGS	✗	✗	✗	8.477	0.03	0.20	0.06	0.12	2.52	0.23
Position	✓	✗	✗	0.110	69.80	84.65	76.51	82.80	99.95	90.57
+ Scale	✓	✓	✗	0.033	81.10	91.06	85.79	93.03	99.97	96.37
+ Opacity	✓	✗	✓	0.054	79.29	90.80	84.66	90.64	99.99	95.09
RA-GS (Ours)	✓	✓	✓	0.011	91.56	96.58	94.00	98.26	100.00	99.12

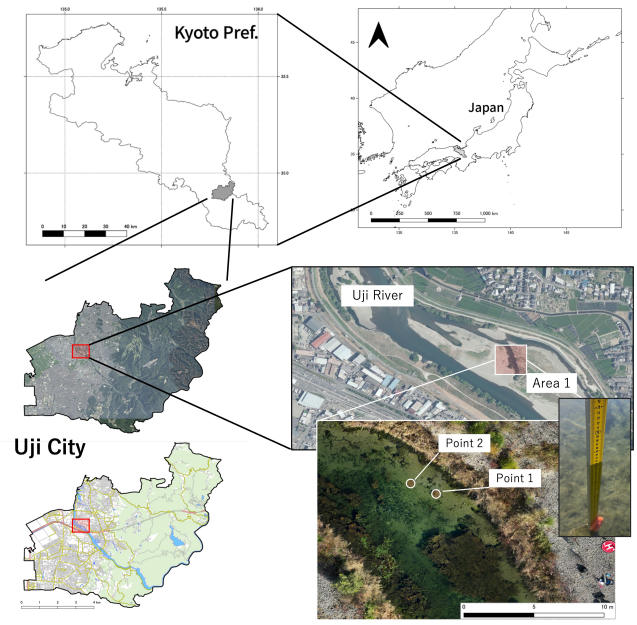

Figure 10. Geographical overview of the field demonstration site. The study area lies along a shallow side channel of the Uji River in Uji City, Kyoto Prefecture, Japan. The detailed UAV orthophoto (bottom right) shows the test site (Area 1) and the locations of two submerged verification targets (Point 1 and Point 2) used for quantitative depth evaluation.

5. Field Demonstration

5.1 Data Acquisition

To assess practical applicability under uncontrolled outdoor conditions, we conducted a field demonstration. While large-scale statistical benchmarks remain future work, this experiment provides evidence that the proposed pipeline is feasible for in-situ UAV bathymetry when the water surface is calm and relatively clear.

Imagery was acquired with a DJI Mavic 3 Enterprise UAV (DJI, 2023) equipped with an RTK-GNSS module. A polarizing filter was used to attenuate specular reflections from the air-water interface. The study site was a side channel of the Uji River with maximum depth of approximately 1 m. We selected this location for its tranquil surface and high water clarity, which help isolate refractive geometry errors from confounding effects such as turbidity and wave-induced distortion (Figure 11).

Camera intrinsics and extrinsics, treated as fixed priors during RA-GS optimization, were estimated with SfM using RealityScan (RealityScan: A Photogrammetric System for 3D Reconstruction, n.d.). To reduce refraction-induced artifacts in pose estimation, the water surface was segmented and masked with DINO.txt (Jose et al., 2025), a vision–language model built on the DINOv3 architecture (Siméoni et al., 2025). This prompt-driven segmentation affords zero-shot adaptation across diverse scenes and reduces manual preprocessing. The global coordinate frame was aligned by fitting a rigid transform that maps shoreline points to the plane $z = 0$, so the still-water surface serves as a known reference datum.

Flights were performed at 12 m altitude above the channel, yielding an approximate ground sampling distance (GSD) of 13 mm/pixel. The dataset comprises 188 images: 108 nadir views with 85 % forward and side overlap, and 80 oblique views captured manually on a circular trajectory around the target, with pitch angles between approximately 30° and 45°. For RA-GS training, images were downsampled by a factor of four, giving an effective resolution of about 1300×1000 pixels.

5.2 Implementation

We integrated the refractive transformation into 2D Gaussian Splatting (2DGS) (Huang et al., 2024) using the open-source gsplat implementation (Ye et al., 2025). Following the original 2DGS formulation, we apply the depth-distortion and normal-consistency regularization terms to encourage smooth surface geometry.

The initial point cloud was obtained from COLMAP Multi-View Stereo (Schönberger et al., 2016). As in our synthetic experiments and standard photogrammetric bathymetry practice, initial depths were scaled by the nominal refractive index of water, $n = 1.33$ (Westaway et al., 2001, Woodget et al., 2014).

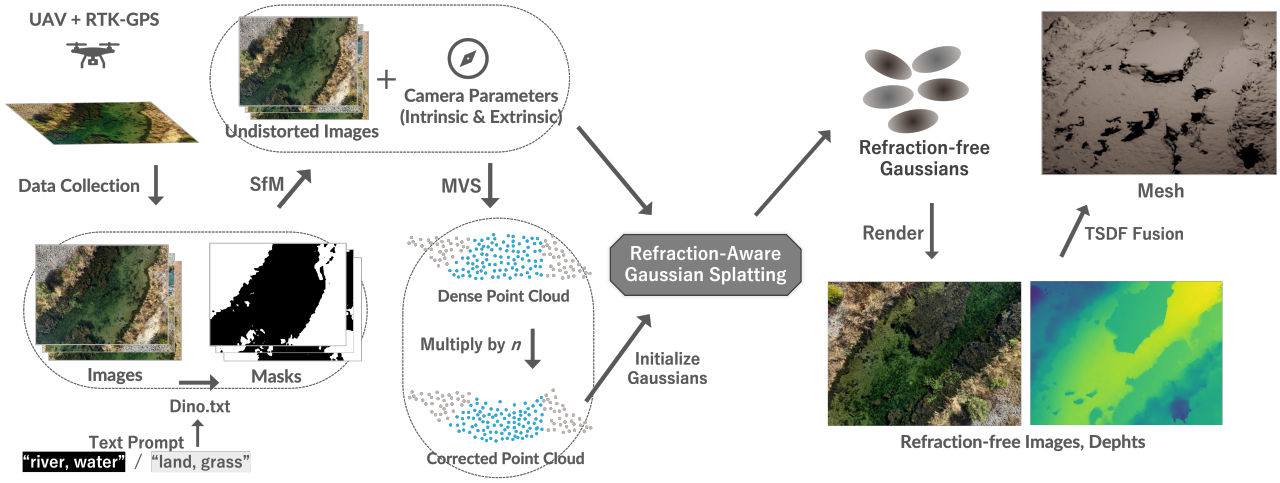

Figure 11. Overview of the Refraction-Aware Gaussian Splatting (RA-GS) pipeline for real-world UAV imagery. For reliable Structure-from-Motion (SfM) camera alignment, water surfaces are segmented and masked automatically from text prompts using DINO.txt. The Multi-View Stereo (MVS) point cloud is scaled by the refractive index of water (n) to initialize RA-GS optimization, following common shallow-water practice. The optimized refraction-free Gaussian model provides high-fidelity novel view synthesis (RGB and depth) and dense riverbed mesh extraction via truncated signed distance function (TSDF) fusion.

Table 3. Depth estimates and errors at two submerged verification targets (Point 1 and Point 2). “+ Refractive scaling” denotes post hoc depth scaling by $n = 1.33$. Mean and relative errors aggregate the two points.

	Point 1		Point 2		Mean Error	Relative Error
Method	Depth [cm]	Error [cm]	Depth [cm]	Error [cm]	[cm]	[%]
Measured (Ground Truth)	61.0	–	71.0	–	–	–
COLMAP MVS	43.7	17.3	51.2	19.8	18.6	1.54
+ Refractive Scaling	58.3	2.7	68.3	3.7	3.2	0.267
2DGS	39.0	22.0	44.9	26.1	24.1	2.00
+ Refractive Scaling	51.2	9.8	60.0	11.0	10.4	0.867
RA-GS (Ours)	60.7	0.3	70.7	0.3	0.3	**0.025**

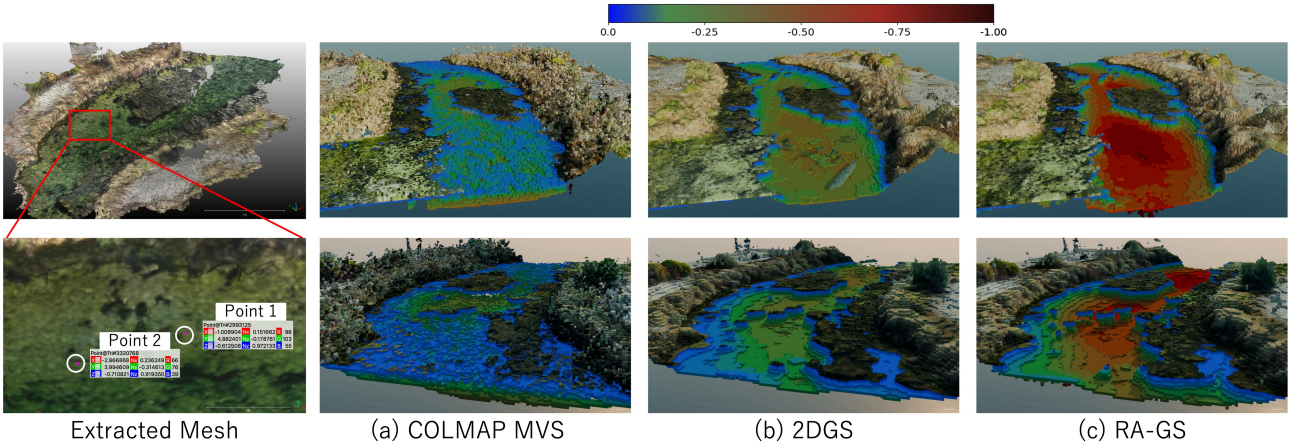

Figure 12. Left: The high-resolution RA-GS mesh with centimeter-scale markers clearly visible. Right: Qualitative comparison of dense riverbed reconstructions from the field dataset. Depth (m) follows the color bar; the still-water surface is at $z = 0$ m. For clarity, meshes are voxelized at 8 cm. (a) **COLMAP MVS** produces a fragmented, noisy surface because multi-view correspondence through the interface is inconsistent. (b) **2D Gaussian Splatting (2DGS)** yields a smoother mesh but systematically underestimates depth, visible as overly shallow colors when refraction is ignored in the reconstruction model. (c) **RA-GS (Ours)** explicitly models refraction and recovers a dense bathymetric surface consistent with the in-situ measurements.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-605-2026 | © Author(s) 2026. CC BY 4.0 License.

Optimization ran for 30,000 iterations with default 2DGS hyperparameters. We applied refractive correction only to submerged Gaussians, i.e., those satisfying $p_z < z_{\text{water level}}$, thereby jointly optimizing above-water and underwater regions. Dense meshes were extracted by truncated signed distance function (TSDF) fusion (Curless and Levoy, 1996) via the Open3D API (Open3D: A Modern Library for 3D Data Processing, n.d.). Consistent with the evaluation in the 2DGS paper, median depth maps were used as TSDF input. Given a GSD of 13 mm, we set the TSDF voxel size to 3 cm (approximately 3× GSD) and the truncation distance to 10 cm (approximately 3× voxel size) (Figure 11).

5.3 Evaluation and Results

Two verification targets were placed on the riverbed and their water depths were measured in situ with a ruler. Because underwater RTK-GNSS fixes are impractical, target horizontal positions were identified on the high-resolution RA-GS mesh, where the centimeter-scale markers were clearly visible (Figure 12 left). To reduce sensitivity to local noise, reported depths were computed as the mean of four samples around the base of each target.

Table 3 summarizes depth accuracy relative to COLMAP MVS and 2DGS. Relative error is defined as the mean absolute depth error divided by the flight altitude (12 m). For COLMAP MVS and 2DGS, we report both raw MVS/GS depths and depths after uniform scaling by $n = 1.33$ (Westaway et al., 2001, Woodget et al., 2014).

RA-GS estimates depths of 60.7 cm and 70.7 cm at Point 1 and Point 2, respectively, compared with ground truth of 61.0 cm and 71.0 cm; the mean absolute error is 0.3 cm, or 0.025 % of the flight height. This relative error is on the same order as the 0.02 % figure reported for multi-view refractive correction with explicit ray tracing (Dietrich, 2016), while RA-GS additionally delivers a dense surface representation rather than sparse bathymetric samples.

Even after refractive scaling, the baselines remain biased shallow: corrected COLMAP MVS attains a mean absolute error of 3.2 cm, and corrected 2DGS 10.4 cm. These gaps indicate that uniform depth scaling does not remove the multi-view inconsistency induced by refraction at a planar interface. Embedding a refraction-aware observation model in the reconstruction stage is therefore important for metrically reliable UAV bathymetry from above-water imagery.

6. Conclusion

We introduced Refraction-Aware Gaussian Splatting (RA-GS), a novel framework that integrates a physically rigorous two-media refraction model directly into the Gaussian Splatting optimization process. By deriving a fully differentiable analytical transformation that maps underwater Gaussians to their refracted appearances in aerial imagery, our method enables refraction-free 3D reconstruction without relying on post-hoc empirical depth scaling or data-driven priors. Extensive experiments on a physically based synthetic riverbed dataset demonstrate that RA-GS substantially mitigates refractive distortions, achieving high-fidelity novel view synthesis and geometric recovery. Furthermore, our in-situ field demonstration validates the method’s capability to deliver precise bathymetric geometries on real-world datasets.

6.1 Limitations and Future Work

We assume a calm water surface and geometric refraction through a planar interface; specular reflections, volumetric attenuation and scattering, and Fresnel mixing are not modeled and appear as residual photometric error in practice. Camera extrinsics, water-surface height, and refractive index (n=1.33) were treated as known or fixed. At sites with weak terrestrial control, structure-from-motion can yield inaccurate poses; such settings motivate tighter integration with refraction-aware bundle adjustment (Makris et al., 2024). Although implicit differentiation supplies gradients with respect to extrinsics and n, joint optimization of pose and index will require careful schedules and constraints to remain stable.

We see three concrete extensions. First, validation on larger and more diverse synthetic and real-world datasets should be paired with radiometric underwater priors (e.g., WaterSplatting (Li et al., 2025), SeaSplat (Yang et al., 2025)) so that appearance and geometry are disentangled under turbid or strongly attenuating water. Second, the isotropic scale correction in Section 2.2 can be replaced by a full anisotropic covariance correction, $\Sigma^{3D}_{\text{app}} = J_{\text{app}} \Sigma^{3D} J^{\top}_{\text{app}}$, to better match the view-dependent stretching induced by refraction. Third, the radiance reduction at the air-water interface follows from etendue conservation ($L_a = L_w / n^2$); combining this with gamma-corrected sRGB yields a consistent per-Gaussian photometric scaling $\boldsymbol{c}'_k = n^{-2/\gamma} \boldsymbol{c}_k$, which enables the color parameters to represent the true in-situ albedo, rather than baking the refraction-induced darkening into them.

Overall, this work takes a significant step toward resolving the long-standing challenge of refractive distortion in photogrammetric bathymetry. By tightly coupling optical physics with explicit 3D representations, RA-GS lays the mathematical foundation for accurate depth mapping in shallow aquatic environments and opens new directions for physically grounded reconstruction in refractive media.

References

Agrafiotis, P., Karantzalos, K., Georgopoulos, A., Skarlatos, D., 2020. Correcting Image Refraction: Towards Accurate Aerial Image-Based Bathymetry Mapping in Shallow Waters. *Remote Sensing*, 12(2). https://www.mdpi.com/2072-4292/12/2/322.

Bemis, S. P., Micklethwaite, S., Turner, D., James, M. R., Akciz, S., Thiele, S. T., Bangash, H. A., 2014. Ground-based and UAV-Based photogrammetry: A multi-scale, high-resolution mapping tool for structural geology and paleoseismology. *Journal of Structural Geology*, 69, 163–178. http://dx.doi.org/10.1016/j.jsg.2014.10.007.

Community, B. O., 2025. Blender - a 3D modelling and rendering package. Blender Foundation, Stichting Blender Foundation, Amsterdam.

Curless, B., Levoy, M., 1996. A volumetric method for building complex models from range images. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, SIGGRAPH96, ACM, 303–312.

Dietrich, J. T., 2016. Bathymetric Structure-from-Motion: extracting shallow stream bathymetry from multi-view stereo photogrammetry. *Earth Surface Processes and Landforms*, 42(2), 355–364. http://dx.doi.org/10.1002/esp.4060.

DJI, 2023. DJI Mavic 3 Enterprise Series User Manual. Accessed: 2026-01-25.

Furukawa, Y., Ponce, J., 2010. Accurate, Dense, and Robust Multiview Stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8), 1362–1376. http://dx.doi.org/10.1109/tpami.2009.161.

Giordano, F., Mattei, G., Parente, C., Peluso, F., Santamaria, R., 2015. Integrating Sensors into a Marine Drone for Bathymetric 3D Surveys in Shallow Waters. *Sensors*, 16(1), 41. http://dx.doi.org/10.3390/s16010041.

Girardeau-Montaut, D. et al., 2016. CloudCompare. *France: EDF R&D Telecom ParisTech*, 11(5), 2016.

Gomez, C., Purdie, H., 2016. UAV- based Photogrammetry and Geocomputing for Hazards and Disaster Risk Monitoring – A Review. *Geoenvironmental Disasters*, 3(1). http://dx.doi.org/10.1186/s40677-016-0060-y.

Grimaldi, S., Li, Y., Walker, J. P., Pauwels, V. R. N., 2018. Effective Representation of River Geometry in Hydraulic Flood Forecast Models. *Water Resources Research*, 54(2), 1031–1057. http://dx.doi.org/10.1002/2017wr021765.

He, J., Zhang, S., Cui, X., Feng, W., 2024. Remote sensing for shallow bathymetry: A systematic review. *Earth-Science Reviews*, 258, 104957. http://dx.doi.org/10.1016/j.earscirev.2024.104957.

Hemmelder, S., Marra, W., Markies, H., De Jong, S. M., 2018. Monitoring river morphology & bank erosion using UAV imagery–A case study of the river Buëch, Hautes-Alpes, France. *International Journal of Applied Earth Observation and Geoinformation*, 73, 428–437. http://dx.doi.org/10.1016/j.jag.2018.07.016.

Huang, B., Yu, Z., Chen, A., Geiger, A., Gao, S., 2024. 2d gaussian splatting for geometrically accurate radiance fields. *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*, SIGGRAPH ’24, ACM, 1–11.

Iglhaut, J., Cabo, C., Puliti, S., Piermattei, L., O’Connor, J., Rosette, J., 2019. Structure from Motion Photogrammetry in Forestry: a Review. *Current Forestry Reports*, 5(3), 155–168. http://dx.doi.org/10.1007/s40725-019-00094-3.

Jose, C., Moutakanni, T., Kang, D., Baldassarre, F., Darcet, T., Xu, H., Li, D., Szafraniec, M., Ramamonjisoa, M., Oquab, M., Siméoni, O., Vo, H. V., Labatut, P., Bojanowski, P., 2025. Dinov2 meets text: A unified framework for image- and pixel-level vision-language alignment. *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 24905–24916.

Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G., 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics*, 42(4), 1–14. http://dx.doi.org/10.1145/3592433.

Li, H., Song, W., Xu, T., Elsig, A., Kulhanek, J., 2025. WaterSplatting: Fast Underwater 3D Scene Reconstruction using Gaussian Splatting. *3DV*.

Makris, A., Nicodemou, V. C., Alevizos, E., Oikonomidis, I., Alexakis, D. D., Roussos, A., 2024. Refraction-Aware Structure from Motion for Airborne Bathymetry. *Remote Sensing*, 16(22), 4253. http://dx.doi.org/10.3390/rs16224253.

Mandlburger, G., Kölle, M., Nübel, H., Soergel, U., 2021. BathyNet: A Deep Neural Network for Water Depth Mapping from Multispectral Aerial Images. *PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 89(2), 71–89. http://dx.doi.org/10.1007/s41064-021-00142-3.

Missailidis, G. T., Guilhon, I., 2025. Determination of the position of an image formed by flat refracting surfaces at large angles of incidence: three different didactic approaches. *Revista Brasileira de Ensino de Física*, 47. http://dx.doi.org/10.1590/1806-9126-rbef-2024-0417.

Nassar, A. B., 1994. Apparent Depth. *Physics Teacher*, 32(9), 526–29.

Open3D: A Modern Library for 3D Data Processing, n.d.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. Pytorch: An imperative style, high-performance deep learning library.

RealityScan: A Photogrammetric System for 3D Reconstruction, n.d.

Rinner, H., 1948. Problems of two medium photogrammetry. *Photogrammetric Engineering*, 35(2), 275–282.

Saylam, K., Hupp, J. R., Averett, A. R., Gutelius, W. F., Gelhar, B. W., 2018. Airborne lidar bathymetry: assessing quality assurance and quality control methods with Leica Chiroptera examples. *International Journal of Remote Sensing*, 39(8), 2518–2542. http://dx.doi.org/10.1080/01431161.2018.1430916.

Schönberger, J. L., Frahm, J.-M., 2016. Structure-from-motion revisited. *Conference on Computer Vision and Pattern Recognition (CVPR)*.

Schönberger, J. L., Zheng, E., Frahm, J.-M., Pollefeys, M., 2016. *Pixelwise View Selection for Unstructured Multi-View Stereo*. Springer International Publishing, 501–518.

Seitz, S., Curless, B., Diebel, J., Scharstein, D., Szeliski, R., 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*, 1, 519–528.

Siméoni, O., Vo, H. V., Seitzer, M., Baldassarre, F., Oquab, M., Jose, C., Khalidov, V., Szafraniec, M., Yi, S., Ramamonjisoa, M., Massa, F., Haziza, D., Wehrstedt, L., Wang, J., Darcet, T., Moutakanni, T., Sentana, L., Roberts, C., Vedaldi, A., Tolan, J., Brandt, J., Couprie, C., Mairal, J., Jégou, H., Labatut, P., Bojanowski, P., 2025. DINOv3.

Stuart, L. A. G., Morton, A., Stavness, I., Pound, M. P., 2025. 3dgs-to-pc: 3d gaussian splatting to dense point clouds. *2025 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, IEEE, 3789–3798.

Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600–612. http://dx.doi.org/10.1109/tip.2003.819861.

Westaway, R. M., Lane, S. N., Hicks, D. M., 2001. Remote sensing of clear-water, shallow, gravel-bed rivers using digital photogrammetry. *Photogrammetric Engineering and Remote Sensing*, 67(11), 1271–1282.

Winston, R., Jiang, L., Ricketts, M., 2018. Nonimaging optics: a tutorial. *Advances in Optics and Photonics*, 10(2), 484. http://dx.doi.org/10.1364/aop.10.000484.

Woodget, A. S., Carbonneau, P. E., Visser, F., Maddock, I. P., 2014. Quantifying submerged fluvial topography using hyperspatial resolution UAS imagery and structure from motion photogrammetry. *Earth Surface Processes and Landforms*, 40(1), 47–64. http://dx.doi.org/10.1002/esp.3613.

Yang, D., Leonard, J. J., Girdhar, Y., 2025. Seasplat: Representing underwater scenes with 3d gaussian splatting and a physically grounded image formation model. *2025 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 7632–7638.

Ye, V., Li, R., Kerr, J., Turkulainen, M., Yi, B., Pan, Z., Seiskari, O., Ye, J., Hu, J., Tancik, M. et al., 2025. gsplat: An open-source library for Gaussian splatting. *Journal of Machine Learning Research*, 26(34), 1–17.

Zhan, Y., Nobuhara, S., Nishino, K., Zheng, Y., 2023. Nerfrac: Neural radiance fields through refractive surface. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, 18356–18366.

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., Wang, O., 2018. The unreasonable effectiveness of deep features as a perceptual metric. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE.