Benchmarking Neural Rendering Approaches for 3D Reconstruction of Underwater Environments

Salvatore Mario Carota[1,*,a], Alessandro Privitera[1,*,b], Daniele Di Mauro[2,c], Antonino Furnari[1,2,d], Giovanni Maria Farinella[1,2,e] and Francesco Ragusa[1,2,f]

[1] *Department of Mathematics and Computer Science, University of Catania, Viale Andrea Doria, 6, Catania, Italy*

[2] *Next Vision s.r.l., Viale Andrea Doria, 6, Catania, Italy*

{*antonino.furnari, giovanni.farinella, francesco.ragusa*}*@unict.it*, {*salvatore.carota, prvlsn01s02c351v*}*@studium.unict.it,*

Keywords: Underwater 3D Reconstruction, Neural Rendering, 3D Gaussian Splatting.

Abstract: We tackle the problem of 3D reconstruction of underwater scenarios using neural rendering techniques. We propose a benchmark adopting the SeaThru-NeRF dataset, performing a systematic analysis by comparing several established methods based on NERF and 3D Gaussian Splatting through a series of experiments. The results were evaluated both quantitatively, using various 2D and 3D metrics, and qualitatively, through a user survey assessing the fidelity of the reconstructed images. This serves to provide critical insight into how to select the optimal techniques for 3D reconstruction of underwater scenarios. The results indicate that, in the context of this application, among the algorithms tested, NeRF-based methods performed better in both mesh generation and novel view synthesis than the 3D Gaussian Splatting based methods.

1 INTRODUCTION

3D reconstruction is a classic computer vision task that has become ubiquitous across various scientific fields, including archaeological inspections (De Reu et al., 2014), biological studies (Correia and Brito, 2023; Irschick et al., 2022), and architectural projects (Münster et al., 2024; Cui et al., 2024). An area of particular interest, due to its diverse applications ranging from biological assessment to archaeological discovery, is underwater 3D reconstruction, which poses unique challenges due to several critical differences compared to reconstructing non-underwater scenes. Image captured in underwater environments differ significantly because of the presence of water, which alters the behavior of light (Li et al., 2019; Islam et al., 2020; Zhang and Johnson-Roberson, 2023; Hou et al., 2020). These differences include variations in lighting, optical distortions, and limited visibility. Together, these factors create a complex set of challenges for accurate 3D reconstruction (Akkaynak and Treibitz, 2019). Overcoming these challenges is highly beneficial for many fields. In underwater heritage conservation, 3D reconstruction enables the inspection of artifacts and structures without risking damage or compromising their integrity (Memet, 2008; Perez-Alvaro, 2023). This technology not only aids in protecting cultural assets but also allows for their presentation to a wider audience, such as in virtual museums. Additionally, underwater environmental sciences can benefit from advancements in 3D reconstruction technologies to monitor coral reef health by detecting changes over time. Detailed 3D models enable marine biologists to study complex habitats, providing deeper insights into ecological interactions (Zhang et al., 2023; Adamczak et al., 2019; Kaandorp, 1993). 3D reconstructions of underwater environments can also be utilized in video games, movies, and virtual and augmented reality applications to enhance user experiences. A great example of its usage in the field of culture is a project called "First Life" [1], which enables the visitor to become a virtual voyager, traveling through sub-

[a] https://orcid.org/0009-0008-6431-9156
[b] https://orcid.org/0009-0001-3507-8233
[c] https://orcid.org/0000-0002-4286-2050
[d] https://orcid.org/0000-0001-6911-0302
[e] https://orcid.org/0000-0002-6034-0432
[f] https://orcid.org/0000-0002-6368-1910
*These authors share first authorship.

[1] https://www.nhm.ac.uk/discover/news/2015/june/dive-back-in-time-with-david-attenborough-s-first-life.html

Carota, S. M., Privitera, A., Di Mauro, D., Furnari, A., Farinella, G. M. and Ragusa, F.
Benchmarking Neural Rendering Approaches for 3D Reconstruction of Underwater Environments.
DOI: 10.5220/0013381200003912

In *Proceedings of the 20th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2025) - Volume 2: VISAPP*, pages 766-773
ISBN: 978-989-758-728-3; ISSN: 2184-4321

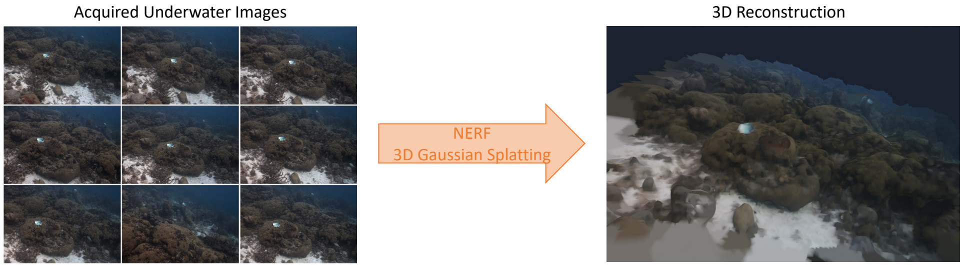

Figure 1: From a set of acquired images in underwater environments, the task is to reconstructs the 3D model of the environment.

merged landscapes and creating for themselves a new perspective on the history of life on Earth. Recently, several solutions have dealt with the problem of 3D environment reconstruction, but a big gap remains in underwater environments reconstruction. Photogrammetry (Schönberger and Frahm, 2016), Neural Radiance Fields (NeRF) (Levy et al., 2023), and 3D Gaussian splatting (3DGS) (Kerbl et al., 2023) techniques have played an important role in enhancing and introducing new methods for reconstruction of 3D models.

In this work, we present a benchmark for 3D Reconstruction of underwater environments (see Figure 1) using the state-of-the-art SeaThru-NeRF dataset introduced in (Levy et al., 2023). The benchmark compares models based on NeRF and 3DGS, at the same time we test how underwater enhancing techniques perform as a preprocessing step in the context of neural rendering. The results were evaluated both quantitatively, using a variety of evaluation metrics at both the rendering and mesh generation levels, and qualitatively, through a user survey. We found out that NeRF-based models are slightly better suited for the task compared to 3DGS-based methods, which nonetheless remain highly promising.

The contributions of this work are: 1) We conducted a systematic analysis of NeRF-based and 3DGS-based methods for underwater environment reconstruction, providing insights into the performance of the tested methods; 2) We analyzed how enhancement techniques can improve the reconstruction of 3D models; 3) We quantitatively evaluated both new view synthesis task and 3D mesh reconstruction; 4) We conducted a qualitative study on the accuracy of reconstructed 3D models through questionnaires administered to a total of 40 subjects.

2 RELATED WORK

Our work builds on prior research in underwater datasets, underwater image enhancement, 3D reconstruction, and neural rendering, which will be briefly described in the following sections.

Underwater Datasets. In literature several dataset were proposed (Li et al., 2019; Islam et al., 2020; Zhang and Johnson-Roberson, 2023; Hou et al., 2020; Akkaynak and Treibitz, 2019), some of them are real, i.e. capturing real scenes, others are synthetic, i.e. images are crafted in some way to solve a particular task. They are used for various purposes, i.e. enhancement, 3d reconstruction, robotics. Among them: UIEB (Li et al., 2019) consisting of 950 real-world underwater images with different natural and artificial lighting conditions. UFO-120 (Islam et al., 2020) contains 1,500 paired samples splitted in training and validation sets and 120 paired samples for benchmark evaluation. Each shot is provided with a high-resolution ground truth version, its distorted low-resolution version, and a saliency map mask. BNU (Zhang and Johnson-Roberson, 2023) includes images captured in a 1.3m-deep tank and in Lake Erie. The JPEG images were post-processed and camera poses were calculated using COLMAP. SUID (Hou et al., 2020), is a synthetic dataset produced by applying special effects that simulate underwater conditions in terrestrial images. SeaThru and the following work from the same research group SeaThru-NERF (Akkaynak and Treibitz, 2019; Levy et al., 2023) contains underwater scenes captured in three different sea with a total of 29, 20 and 18 images respectively. We chose this dataset for training the models due to its diverse range of scenarios and number of images.

Underwater Image Enhancement and Restoration. Underwater image enhancement is the task to reduce or remove water effects on images recorded underwater. WaterGAN (Li et al., 2017) is a color correction model based on Generative Adversarial Networks (GAN). The generator estimates the attenuation, backscatter, and camera characteristics of underwater images. The model is trained with both underwater and non-underwater images in order to create synthetic underwater images. In (Cho et al., 2020) they used GANs for image correction and enhancement through the image-translation technique.

Figure 2: Images belonging to the Seathru dataset adopted for the proposed benchmark. The dataset contains three different underwater scenes: Red Sea (left), Caribbean Sea (center) and Pacific Ocean (right).

The model is trained with underwater images in order to capture textures and details of underwater images, the losses used are reconstruction loss, laplacian loss and perception loss. Furthermore, Semi-UIR (Huang et al., 2023) is a semi-supervised underwater image restoration framework based on the mean-teacher model, designed to incorporate unlabeled data into network training. The student model learns from labeled data, while the teacher model guides the training process on unlabeled images by generating reliable "pseudo-labels." Experimental results on both full-reference and no-reference underwater benchmarks show significant improvements in both quantitative and qualitative performance over SOTA methods. We used this method for the enhancement preprocessing step.

Multi-view surface reconstruction is the process that creates a 3D surface starting from a set of images taken from different angles, exploiting points correspondences between images to estimate the shape of an object. Several works adopted volumetric grid methods for reconstructing multi-view surfaces (Boent and Pula, 1999; Kutulakos and Seitz, 2000; Laurentini, 1994; Szeliski, 1993; Seitz and Dyer, 1999). Other works focused on the techniques of cloud point-based techniques (Furukawa and Ponce, 2009; Galliani et al., 2015; Schoenberger et al., 2016; Tola et al., 2012). Its influence has, however been even more extensive among dense reconstruction methods: the Poisson surface reconstruction algorithm (Kazhdan et al., 2006a) along with its screened version (Kazhdan and Hoppe, 2013). Machine learning methods adopting deep learning for enhancing multi-view surface reconstruction techniques have also recently been explored (Chen et al., 2019; Huang et al., 2018; Yao et al., 2018).

Neural Radiance Fields (NeRF) is a pioneering approach that was initiated by (Mildenhall et al., 2020) The idea is to use a neural network to implicitly model a scene from a set of images annotated with a pose. The model thus learns the behavior of light and the geometry of a scene enabling the use of such model to generate novel views. Several variants have been developed to extend the possibilities of NeRF (Barron et al., 2021; Alex et al., 2021; Kai et al., 2020; Sun et al., 2022).

Neural surface reconstruction is a technique that involves using neural networks to learn complex surfaces of 3D objects or scenes in a continuous and highly detailed manner. Various methods have been proposed for this task, utilizing volumetric grid-based methods for scene reconstruction (Niemeyer et al., 2020; Oechsle et al., 2021). The authors of (Lior et al., 2020) uses Signed Distance Functions (SDF) to implicitly model surfaces by defining them as the zero level set of SDF. Similar NeRF-based methods have further been extended to surface reconstruction, works such as (Wang et al., 2021; Lior et al., 2021; Darmon et al., 2022; Fu et al., 2022; Yue et al., 2022) have moved the frontiers in extending the original NeRF framework towards high-fidelity surface modeling. There are also point cloud-based techniques (Fu et al., 2022; Zhang et al., 2022), which accomplished good reconstructions taking in input sparser data points.

3D Gaussian Splatting (3DGS) was proposed in (Kerbl et al., 2023). The method is a change in perspective compared to NeRF, 3D scenes are represented in an explicit way: the scene is modelled as a collection of 3D Gaussian functions distributed in the space then they are "splatted" in 2D in order to match the set of images, together with the corresponding cameras calibrated by Structure from Motion, taken in input. The core of the approach is the optimization step, where a dense set of 3D Gaussians accurately representing the scene is created. In addition to positions and covariance, it also optimizes Spherical Harmonics coefficients representing color of each Gaussian to correctly capture the view-dependent appearance of the scene. The optimization of these parameters is interleaved with steps that adaptevely control the density of the Gaussians to better represent the scene. The optimization takes full advantage of standard GPU-accelerated frameworks and adds custom CUDA kernels, following recent best practices (Alex et al., 2021; Sun et al., 2022). The projection method implements a tile-based rasterizer for Gaussian splats inspired by recent software rasterization approaches (Lassner and Zollhofer, 2021). The rasterization pipeline is fully differentiable, and given the projection to 2D can rasterize anisotropic splats similar to previous 2D splatting methods (Kopanas et al., 2021). As for NeRF, various method are improving over the initial, such as Splatfacto-W (Xu

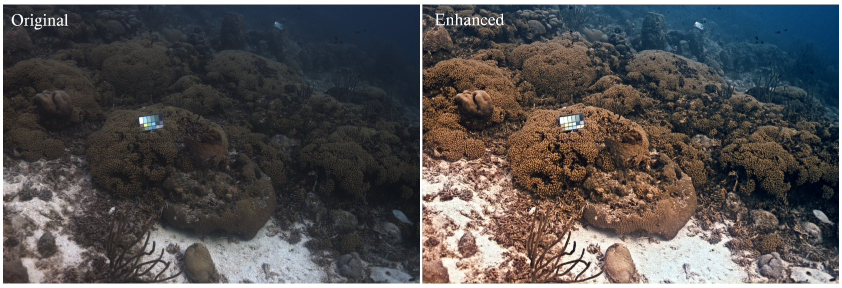

Figure 3: Original image sample (left) vs. its enhanced version (right).

et al., 2024), which improves results in presence of unconstrained photo collections. Its key contributions include latent appearance modeling, efficient transient object handling, and precise background modeling.

Underwater 3D Reconstruction. Two major techniques are used for 3D reconstruction: image-based (Levy et al., 2023; Weidner et al., 2017; Jordt et al., 2016) and laser-scanner-based (Bartolini et al., 2005). Image-based methods are highly cost-effective, while laser-scanner-based methods require expensive equipment and typically take a long time to acquire data. In particular SeaThru-NeRF (Levy et al., 2023) is a NeRF-based method specifically designed for underwater scenes, with the unique capability of modelling the solid objects present in the scene and the medium.

3 BENCHMARK

Dataset. We perform our benchmark using the dataset presented in (Levy et al., 2023). It contains underwater scenes captured in three different seas (see Figure 2): the Red Sea (Eilat, Israel), the Caribbean Sea (Curacao), and the Pacific Ocean (Panama), with a total of 29, 20, and 18 images respectively. The images were acquired as RAW images using a Nikon D850 SLR camera in a Nauticam underwater housing with a dome port to avoid refractions. The images were resized to an average size of 900×1400 and white-balanced with a 0.5% clipping per channel to remove extremely noisy pixels. Finally, COLMAP (Schönberger and Frahm, 2016) was used to extract the camera poses.

Task. We consider the problem of generate a 3D model from a set of underwater images, where the same scene is captured from different viewpoints.

Models. We trained the following models: Neuralangelo (Li et al., 2023), Seathru-NeRF (Levy et al., 2023), Splatfacto (Kerbl et al., 2023), and Splatfacto-W (Xu et al., 2024). We also performed a COLMAP dense reconstruction which is our baseline. We transformed the outputs of each model into 3D meshes using different methods. For Neuralangelo, we used the Marching Cubes algorithm (Lorensen and Cline, 1987). SeaThru-NeRF, on the other hand, was processed using Poisson Surface Reconstruction (Kazhdan et al., 2006b). For Splatfacto and Splatfacto-W, we employed TSDF (Truncated Signed Distance Function) implemented in dn-splatter (Turkulainen et al., 2024). Lastly, COLMAP dense reconstruction was also processed using Poisson Surface Reconstruction (Kazhdan et al., 2006b). We also evaluated the opportunity to do image enhancement before using Neural Rendering Models (see Figure 3). The enhanced dataset was created by applying the SEMI-UIR algorithm, trained on SUID dataset, to the images of SEATHRU-NeRF. Results that consider enhanced images are indicated with *+enh*.

We finally imported these meshes into Blender[2] to generate renders (see Figure 4) using camera paths exported directly from Nerfstudio[3].

Quantitative Evaluation. To evaluate the quality of rendered images, we adopted various metrics: MUSIQ (Ke et al., 2021) gives a score of the perceived quality of an image-very similar to human judgments as well as UCIQE (Yang and Sowmya, 2015) and UIQM (Panetta et al., 2016) are specifically designed for quality assessment in underwater images. We also evaluated the mesh generation quality using the Hausdorff distance (Cignoni et al., 1998). In particular, we used the implemented version in MeshLab[4]. This allowed us to make a quantitative comparison of the meshes resulting from the various algorithms with respect to a dense reconstruction from COLMAP. We computed the metric bidirectionally, obtaining a symmetric version by taking the maximum value. Differently from the other metrics, which evaluate 2D results, this distance evaluates the accuracy of the reconstruction of the 3D geometry of the environment.

Qualitative Evaluation. To evaluate the quality of

[2]https://www.blender.org/
[3]https://docs.nerf.studio/
[4]https://www.meshlab.net/

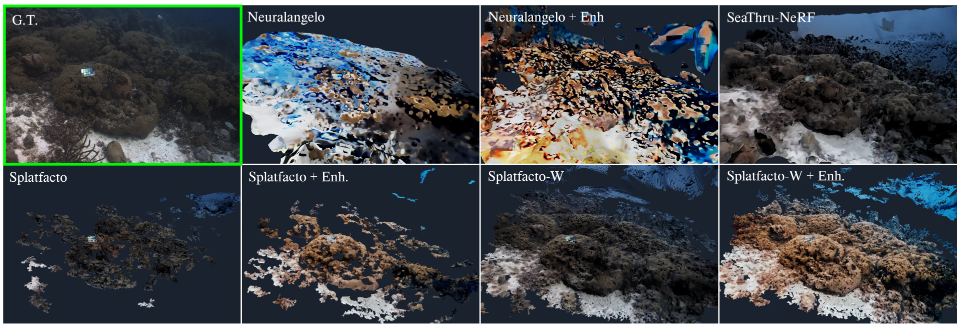

Figure 4: The image shows 3D model renders (Curaçao scene) created by four different algorithms and rendered in Blender (the green border indicates the ground truth).

renderings obtained by the different reconstruction models, we designed a survey that has been administered to 40 people. Participants were asked to rate each 3D reconstruction on a scale from 1 to 5, where a score of 1 indicates the lowest quality and 5 represents the best quality.

4 RESULTS

Table 1 shows the average results derived from calculating the different 2D metrics across all images rendered from a set of sampled viewpoints along a camera path, evaluated for each scene and model. For the metric MUSIQ: NeRF-Based methods lead the scoreboard with 2 best results (Neuralangelo-Panama 61.513, SeaThru-NeRF-Readsea 65.612) and one second best. For the metric UCIQE. NeRF-Based methods lead the scoreboard with 3 best results (Neuralangelo-Panama 3.461, Neuralangelo-Readsea 5.769, Neuralangelo-Enh-Curaçao 4.62) and 3 second best. Finally for UIQM: gaussian splatting method Splatfacto-W+enh is the best performer. We can state that 2 metrics over 3 show an advantage of NeRF-based metrics. Table 2 compares the rendered outputs to a reference image calculating the PSNR among them. The results show that Seathru-NeRF and COLMAP are the leading models in terms of rendering accuracy, Splatfacto-W also performs well (see Mean column: 15.8329 vs. 15.8203 vs 15.4112). Here we have a second proof that a NeRF based method is favorable to gaussian splatting ones. Finally 3D Mesh distances using Hausdorff distance are reported in Table 3. Seathru-NeRF achieves excellent performance because it can produce a good reconstruction, thanks to the ability to distinguish between medium and objects. SeaThru-NeRF is closely followed by Splatfacto-W+enh (0.0482 vs 0.0715), here, we can observe how enhancement plays a crucial role. All the quantitative measures show a clear advantage by NeRF on gaussian splatting, in particular with

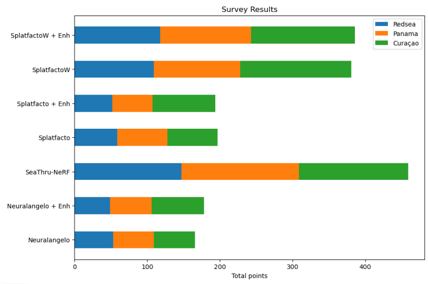

Figure 5: Qualitative evaluation results.

SeaThru-NeRF leading the scoreboard for 3D mesh generation and PSNR results. For the qualitative assessment, from the survey results we noticed that SeaThru-NeRF consistently outperformed other models, emerging as the best reconstruction model in two out of the three scenes. Splatfacto-W + enh. was identified as the second-best model overall, also achieving high ratings across most scenes. This ranking highlights the relative strengths of these two models in producing high-quality renderings for diverse scenes. Figure 4 shows some examples of the reconstructions provided to users for evaluation.

These tables collectively provide insight into the relative strengths of each model. In our view, Seathru-NeRF and Splatfacto-W achieve superior results primarily due to their specialized features. Seathru-NeRF is specifically designed for underwater environments, making it particularly effective in handling underwater visual challenges. Splatfacto-W, benefits from advanced background modeling and effective handling of transient objects, both of which enhance its rendering accuracy and adaptability.

Qualitative results in Figure 5 are coherent with quantitative ones: SeaThru-NeRF consistently outperformed other models, emerging as the best reconstruction model in two out of the three scenes. Splatfacto-W + Enh was identified as the second-best model overall, also achieving high ratings across most scenes. Some caveats: the survey was not randomized, thus some bias could be present.

Table 1: MUSIQ (Ke et al., 2021), UCIQE (Yang and Sowmya, 2015) and UIQM (Panetta et al., 2016) evaluation of new view syntesis, best results in bold, second best underlined. Each value represents the average of multiple rendered image values.

	Curaçao			Panama			Redsea		
Model	**MUSIQ**	**UCIQE**	**UIQM**	**MUSIQ**	**UCIQE**	**UIQM**	**MUSIQ**	**UCIQE**	**UIQM**
Neuralangelo	46.336	2.411	2.138	**61.513**	**3.461**	2.475	49.186	**5.769**	1.351
Neuralangelo+enh.	50.824	**4.62**	1.95	58.524	2.232	**2.868**	56.885	4.367	1.911
SeaThru-NeRF	57.412	4.427	2.145	55.432	0.852	1.927	**65.612**	0.654	2.021
Splatfacto	59.274	1.896	1.703	49.674	1.391	1.617	61.045	1.38	1.560
Splatfacto-w	61.470	2.836	2.478	55.694	1.123	2.073	64.632	1.069	1.983
Splatfacto +enh.	59.742	1.733	1.843	56.055	1.443	1.623	59.860	1.404	1.567
Splatfacto-w+enh.	**64.042**	1.940	**2.564**	58.066	1.185	2.260	63.105	1.099	**2.052**
Colmap-Poisson	45.350	1.132	1.712	51.647	0.983	1.838	65.289	0.802	1.410

Table 2: Average PSNR, rendered image is compared to original image or to the original preprocessed when enhancement is in place, best results in bold, second best underlined.

Method	Curaçao	Panama	Redsea	Mean
Neuralangelo	10.6242	11.3654	9.4266	10.4721
Neuralangelo+enh.	10.0941	10.1605	9.7282	9.9943
SeaThru	16.1436	**16.7314**	**14.6238**	**15.8329**
Splactfacto	17.3554	15.1125	10.7050	14.3910
Splactfacto-W	16.4640	16.1713	13.5984	15.4112
Splactfacto+enh.	10.5768	10.2170	9.8972	10.2303
Splactfacto-W+enh.	11.1514	11.3191	11.6711	11.3805
Colmap-Poisson	**18.4726**	16.4224	12.566	15.8203

Table 3: Mean Normalized Haussdorf distance (Cignoni et al., 1998) between COLMAP reconstruction and NERF/3DGS based reconstruction, best results in bold, second best underlined.

Method	Curaçao	Panama	Redsea	Mean
Neuralangelo	0.1538	0.0969	0.1387	0.1298
Neuralangelo+enh.	0.1706	0.0819	0.1379	0.1301
SeaThru-NeRF	**0.0287**	0.0652	**0.0508**	**0.0482**
Splatfacto	0.0906	**0.0466**	0.1394	0.0922
Splatfacto-W	0.0753	0.0539	0.0919	0.0737
Splatfacto+enh.	0.1347	0.0818	0.1203	0.1123
Splatfacto-W+enh.	0.0809	0.0517	0.0818	0.0715
Colmap-Poisson	0	0	0	0

5 CONCLUSION

In this work, we presented a benchmark to perform 3D reconstruction of underwater scenes using Neural Rendering techniques. Quantitative analysis shows promising results, Neural Rendering model are on par with SfM only when they take care of modeling the medium (SeaThru-NeRF) or when they take care to model different camera settings (Splatfacto-W). In future work we will focus on improving 3DGS with medium modelling similar to Seathru-NeRF.

ACKNOWLEDGEMENTS

This research has been supported by Next Vision s.r.l. [5] and by the project Neural Rendering & Edge AI Platform for 4D synthetic Twins generation during Underwater Navigation & Exploration (NEPTUNE) PNRR MUR Project CUP J53D23020140005 COR 18115262 - Spoke 3 Robotics and AI for Socio-economic Empowerment (RAISE).

REFERENCES

Adamczak, S. K., Pabst, A., McLellan, W. A., and Thorne, L. H. (2019). Using 3d models to improve estimates of marine mammal size and external morphology. *Frontiers in Marine Science*, 6.

Akkaynak, D. and Treibitz, T. (2019). Sea-thru: A method for removing water from underwater images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1682–1691.

Alex, Y., Fridovish-Keil, S., Matthew, T., Qinhong, C., Benjamin, R., and Angjoo, K. (2021). Plenoxels: Radiance fields without neural networks. *arXiv preprint arXiv:2112.05131*.

Barron, J. T., Mildenhall, B., Tancik, M., Hedman, P., Martin-Brualla, R., and Srinivasan, P. P. (2021). Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5855–5864.

Bartolini, L., De Dominicis, L., Ferri de Collibus, M., Fornetti, G., Guarneri, M., Paglia, E., Poggi, C., and Ricci, R. (2005). Underwater three-dimensional imaging with an amplitude-modulated laser radar at a 405 nm wavelength. *Applied optics*, 44(33):7130–7135.

Boent, J. S. and Pula, P. (1999). Probabilistic voxelized volume reconstruction. In *Proceedings of International Conference on Computer Vision (ICCV)*, volume 2.

Chen, R., Han, S., Xu, J., and Su, H. (2019). Point-based multi-view stereo network. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1538–1547.

[5] https://www.nextvisionlab.it/

Cho, Y., Jang, H., Malav, R., Pandey, G., and Kim, A. (2020). Underwater image dehazing via unpaired image-to-image translation. *International Journal of Control, Automation and Systems*, 18:605–614.

Cignoni, P., Rocchini, C., and Scopigno, R. (1998). Metro: measuring error on simplified surfaces. In *Computer Graphics Forum*, volume 17, pages 167–174. Blackwell Publishers.

Correia, H. A. and Brito, J. H. (2023). 3d reconstruction of human bodies from single-view and multi-view images: A systematic review. *Computer Methods and Programs in Biomedicine*, 239:107620.

Cui, D., Wang, W., Hu, W., Peng, J., Zhao, Y., Zhang, Y., and Wang, J. (2024). 3d reconstruction of building structures incorporating neural radiation fields and geometric constraints. *Automation in Construction*, 165:105517.

Darmon, F., Bascóle, B., Devaux, J.-C., Souhila, P., and Aubry, M. (2022). Improving neural implicit surfaces geometry with patch warping. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6260–6269.

De Reu, J., De Smedt, P., Herremans, D., Van Meirvenne, M., Laloo, P., and De Clercq, W. (2014). On introducing an image-based 3d reconstruction method in archaeological excavation practice. *Journal of Archaeological Science*, 41:251–262.

Fu, Q., Sun, Q., Yew, T.-W., and Tiao, W. (2022). Geoneus: Geometry-consistent neural implicit surfaces learning for multi-view reconstruction. *arXiv preprint arXiv:2205.15848*.

Furukawa, Y. and Ponce, J. (2009). Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8):1362–1376.

Galliani, S., Lasinger, K., and Schindler, K. (2015). Massively parallel multiview stereopsis by surface normal diffusion. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 873–881.

Hou, G., Zhao, X., Pan, Z., Yang, H., Tan, L., and Li, J. (2020). Benchmarking underwater image enhancement and restoration, and beyond. *IEEE Access*, 8:122078–122091.

Huang, P.-H., Kopf, J., Ahuja, N., Bleyer, M., Lenz, J., and Xu, J.-B. (2018). Deepmvs: Learning multi-view stereopsis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2821–2830.

Huang, S., Wang, K., Liu, H., Chen, J., and Li, Y. (2023). Contrastive semi-supervised learning for underwater image restoration via reliable bank. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18145–18155.

Irschick, D. J., Christiansen, F., Hammerschlag, N., Martin, J., Madsen, P. T., Wyneken, J., Brooks, A., Gleiss, A., Fossette, S., Siler, C., Gamble, T., Fish, F., Siebert, U., Patel, J., Xu, Z., Kalogerakis, E., Medina, J., Mukherji, A., Mandica, M., Zotos, S., Detwiler, J., Perot, B., and Lauder, G. (2022). 3d visualization processes for recreating and studying organismal form. *iScience*, 25(9):104867.

Islam, M. J., Luo, P., and Sattar, J. (2020). Simultaneous Enhancement and Super-Resolution of Underwater Imagery for Improved Visual Perception. In *Robotics: Science and Systems (RSS)*, Corvalis, Oregon, USA.

Jordt, A., Köser, K., and Koch, R. (2016). Refractive 3d reconstruction on underwater images. *Methods in Oceanography*, 15-16:90–113. Computer Vision in Oceanography.

Kaandorp, J. A. (1993). 2d and 3d modelling of marine sessile organisms. In Crilly, A. J., Earnshaw, R. A., and Jones, H., editors, *Applications of Fractals and Chaos*, pages 41–61, Berlin, Heidelberg. Springer Berlin Heidelberg.

Kai, Z., Gernot, R., Noah, S., and Vladlen, K. (2020). Nerf++: Analyzing and improving neural radiance fields. *arXiv preprint arXiv:2010.07492*.

Kazhdan, M., Bolitho, M., and Hoppe, H. (2006a). Poisson surface reconstruction. In *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, pages 61–70.

Kazhdan, M., Bolitho, M., and Hoppe, H. (2006b). Poisson surface reconstruction. In *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, SGP '06, page 61–70, Goslar, DEU. Eurographics Association.

Kazhdan, M. and Hoppe, H. (2013). Screened poisson surface reconstruction. *ACM Transactions on Graphics (ToG)*, 32(3):1–13.

Ke, J., Wang, Q., Wang, Y., Milanfar, P., and Yang, F. (2021). Musiq: Multi-scale image quality transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5148–5157.

Kerbl, B., Kopanas, G., Leimkühler, T., and Drettakis, G. (2023). 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):139–1.

Kopanas, G., Philip, J., Leimkühler, T., and Drettakis, G. (2021). Point-based neural rendering with per-view optimization. In *Computer Graphics Forum*, volume 40, pages 29–43. Wiley Online Library.

Kutulakos, K. N. and Seitz, S. M. (2000). A theory of shape by space carving. *International Journal of Computer Vision*, 38(3):199–218.

Lassner, C. and Zollhofer, M. (2021). Pulsar: Efficient sphere-based neural rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1440–1449.

Laurentini, A. (1994). The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2):150–162.

Levy, D., Peleg, A., Pearl, N., Rosenbaum, D., Akkaynak, D., Korman, S., and Treibitz, T. (2023). Seathru-nerf: Neural radiance fields in scattering media. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 56–65.

Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., and Tao, D. (2019). An underwater image enhancement benchmark dataset and beyond. *IEEE Transactions on Image Processing*, 29:4376–4389.

Li, J., Skinner, K. A., Eustice, R. M., and Johnson-Roberson, M. (2017). Watergan: Unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robotics and Automation letters*, 3(1):387–394.

Li, Z., Müller, T., Evans, A., Taylor, R. H., Unberath, M., Liu, M.-Y., and Lin, C.-H. (2023). Neuralangelo: High-fidelity neural surface reconstruction. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Lior, Y., Yoni, K., Dror, M., Matan, A., Meirav, R., and Yaron, L. (2020). Multiview neural surface reconstruction by disentangling geometry and appearance. *Advances in Neural Information Processing Systems*, 33:2492–2503.

Lior, Y., Yoni, K., Dror, M., Matan, A., Meirav, R., and Yaron, L. (2021). Volume rendering of neural implicit surfaces. *Advances in Neural Information Processing Systems*, 34:4805–4815.

Lorensen, W. E. and Cline, H. E. (1987). Marching cubes: A high resolution 3d surface construction algorithm. *SIGGRAPH Comput. Graph.*, 21(4):163–169.

Memet, J.-B. (2008). Conservation of underwater cultural heritage: characteristics and new technologies. *Museum International*, 60(4):42–49.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. (2020). Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*.

Münster, S., Apollonio, F. I., Blümel, I., Fallavollita, F., Foschi, R., Grellert, M., Ioannides, M., Jahn, P. H., Kurdiovsky, R., Kuroczyński, P., Lutteroth, J.-E., Messemer, H., and Schelbert, G. (2024). Handbook of digital 3d reconstruction of historical architecture. page 204.

Niemeyer, M., Mescheder, L., Oechsle, M., and Geiger, A. (2020). Differentiable volumetric rendering: Learning implicit 3d representations without 3d supervision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3504–3515.

Oechsle, M., Peng, S., and Geiger, A. (2021). Unisurf: Unifying neural implicit surfaces and radiance fields for multi-view reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5589–5598.

Panetta, K., Gao, C., and Agaian, S. (2016). Human-visual-system-inspired underwater image quality measures. *IEEE Journal of Oceanic Engineering*, 41(3):541–551.

Perez-Alvaro, E. (2023). Underwater cultural heritage and the sustainable development goals. *Blue Papers*, 2(2).

Schoenberger, J. L., Zheng, E., Frahm, J.-M., and Pollefeys, M. (2016). Pixelwise view selection for unstructured multi-view stereo. In *Proceedings of the European Conference on Computer Vision*, pages 501–518.

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

Seitz, S. M. and Dyer, C. R. (1999). Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision*, 35(2):151–173.

Sun, C., Sun, M., and Chen, H.-T. (2022). Direct voxel grid optimization: Super-fast convergence for radiance fields reconstruction. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5459–5469.

Szeliski, R. (1993). Rapid octree construction from image sequences. *CVGIP: Image Understanding*, 58(1):23–32.

Tola, E., Strecha, C., and Fua, P. (2012). Efficient large-scale multi-view stereo for ultra high-resolution image sets. *Machine Vision and Applications*, 23(5):903–920.

Turkulainen, M., Ren, X., Melekhov, I., Seiskari, O., Rahtu, E., and Kannala, J. (2024). Dn-splatter: Depth and normal priors for gaussian splatting and meshing.

Wang, Y., Skorokhodov, I., Theobalt, P., and Wonka, P. (2021). Hf-neus: Improved surface reconstruction using high-frequency details. *Advances in Neural Information Processing Systems*, 34:19220–19230.

Weidner, N., Rahman, S., Li, A. Q., and Rekleitis, I. (2017). Underwater cave mapping using stereo vision. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5709–5715. IEEE.

Xu, C., Kerr, J., and Kanazawa, A. (2024). Splatfacto-w: A nerfstudio implementation of gaussian splatting for unconstrained photo collections. *arXiv preprint arXiv:2407.12306*.

Yang, M. and Sowmya, A. (2015). An underwater color image quality evaluation metric. *IEEE Transactions on Image Processing*, 24(12):6062–6071.

Yao, Y., Luo, Z., Li, S., Fang, T., and Quan, L. (2018). Mvsnet: Depth inference for unstructured multi-view stereo. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 767–783.

Yue, Z., Peng, S., Niemeyer, M., Sattler, T., and Geiger, A. (2022). Exploring monocular geometric cues for neural implicit surface reconstruction. *arXiv preprint arXiv:2206.00665*.

Zhang, C., Zhou, H., Christiansen, F., Hao, Y., Wang, K., Kou, Z., Chen, R., Min, J., Davis, R., and Wang, D. (2023). Marine mammal morphometrics: 3d modeling and estimation validation. *Frontiers in Marine Science*, 10.

Zhang, J., Shao, Y., Li, T., Fang, D. M., Tsian, Y., and Quan, L. (2022). Critical regularizations for neural surface reconstruction in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6270–6279.

Zhang, T. and Johnson-Roberson, M. (2023). Beyond nerf underwater: Learning neural reflectance fields for true color correction of marine imagery. *IEEE Robotics and Automation Letters*, 8(10):6467–6474.