

BathyNeRF: Photogrammetrische Unterwasserrekonstruktion mit Neural Radiance Fields

Lukas Winiwarter¹, Boris Jutzi², Gottfried Mandlbauer³, Markus Brezovsky³, Anatol Günthner², & Frederik Schulte¹

Zusammenfassung

Das transnationale Forschungsprojekt BathyNeRF (DFG/FWF) an der TU Wien, dem KIT, und der Universität Innsbruck erforscht die Anwendung von Neural Radiance Fields für die Rekonstruktion von Unterwasser-Geometrien aus Photos, die von außerhalb des Gewässerkörpers (z.B. durch den Einsatz unbemannter Fluggeräte) aufgenommen wurden. Dafür muss die an der Luft-Wasser-Grenzschicht stattfindende Refraktion sowie die Absorption des Lichts in der Wassersäule explizit oder implizit modelliert werden. In diesem Beitrag werden die Ansätze des Forschungsprojektes vorgestellt, sowie Herausforderungen und erste Ergebnisse präsentiert.

Schlagwörter Bathymetrie · Refraktion · Snellius · Wasseroberfläche · Deep Learning

1 Einführung

In den vergangenen Jahren haben sich Radiance Fields als alternative Rekonstruktionsmethode zur Gewinnung von 3D-Geometrien aus konvergenten Aufnahmen insbesondere im Bereich der Computer Vision etabliert. Als wichtiger Vertreter der Radiance Fields nutzen Neural Radiance Fields (NeRFs, Mildenhall et al., 2020) ein neuronales Netz zur impliziten Repräsentation der 3D-Szene. Das Netz weist jedem Punkt (X, Y, Z) im Ortsraum eine Farbe (r, g, b) und eine Dichte (σ) zu. Farbe und Dichte werden entlang der Bildstrahlen akkumuliert, um für jedes Pixel einen Farbwert vorhersagen zu können. Durch überwachtetes Lernen werden orientierte Bilder als Trainingsdaten verwendet, um diese neuronalen Netze zu optimieren. Im Gegensatz zu klassischem Multi-View-Stereo (MVS) Dense Image Matching (DIM) ermöglichen NeRFs damit auch, solche Geometrien zu rekonstruieren, bei denen das Auffinden homologer Punkte prinzipiell herausfordernd ist, wie z.B. in Vegetationsbereichen.

Für die Berücksichtigung der Brechung wurden in der Literatur bereits verschiedene Ansätze entwickelt. NeRFrac (Zhan et al., 2023) modelliert eine refraktierende Oberfläche durch ein zusätzliches neuronales Netz, welches für jeden Bildstrahl angibt, an welcher Position die Refraktion auftritt.

Objekte, die die Szene repräsentieren, dürfen jedoch erst hinter dieser Oberfläche sein. REF²-NeRF (Kim et al., 2024) schätzen aus den 3D-Positionen einzelner Punkte entlang der Bildstrahlen einen *Offset*, also eine Verschiebung, die bei der Evaluierung des NeRFs angebracht wird, um die Farb- und Dichteinformationen zu erhalten.

In diesem Beitrag stellen wir darüber hinaus physikalisch motivierte Ansätze vor, die auch eine Repräsentation der Szene vor der refraktierenden Oberfläche erlauben. Diese unterscheiden sich in ihrer Komplexität und ihrem Grad der Abstraktion der Realität und werden in Abschnitt 2 vorgestellt.

Um die Qualität der Ergebnisse zu evaluieren, werden einerseits Bilddatensätze für bekannte Geometrien mittels Rendering simuliert, in späterer Folge sollen auch refraktionskorrigierte Airborne Laser Bathymetry (ALB)-Daten zur Validierung von mit Kameradrohnen erfassten Szenen dienen. Die Beurteilung der Ergebnisse erfolgt (a) über die Qualität rekonstruierter Bilddaten (Synthetisierung), wie bei NeRFs typisch, wobei sich hier die Maße des PSNR (Peak Signal-to-Noise Ratio), SSIM (Spatial Similarity Image Metric) und der LPIPS (Learned Perceptual Image Patch Similarity) etabliert haben; sowie (b) über die geometrische Qualität der aus dem NeRF extrahierten Geometrie, insbesondere den euklidischen

¹ Universität Innsbruck, Fakultät für technische Wissenschaften, Technikerstraße 13, 6020 Innsbruck, Österreich, E-Mail: [lukas.winiwarter|frederik.schulte]@uibk.ac.at

² Karlsruher Institut für Technologie, Institut für Photogrammetrie und Fernerkundung, 76128 Karlsruhe, Deutschland, E-Mail: [boris.jutzi|anatol.guenther]@kit.edu

³ TU Wien, Department für Geodäsie und Geoinformation, Wiedner Hauptstraße 7-10, 1040 Wien, Österreich, E-Mail: [gottfried.mandlbauer|markus.brezovsky]@tuwien.ac.at

Abstand zur Referenzgeometrie, und die Vollständigkeit der Rekonstruktion.

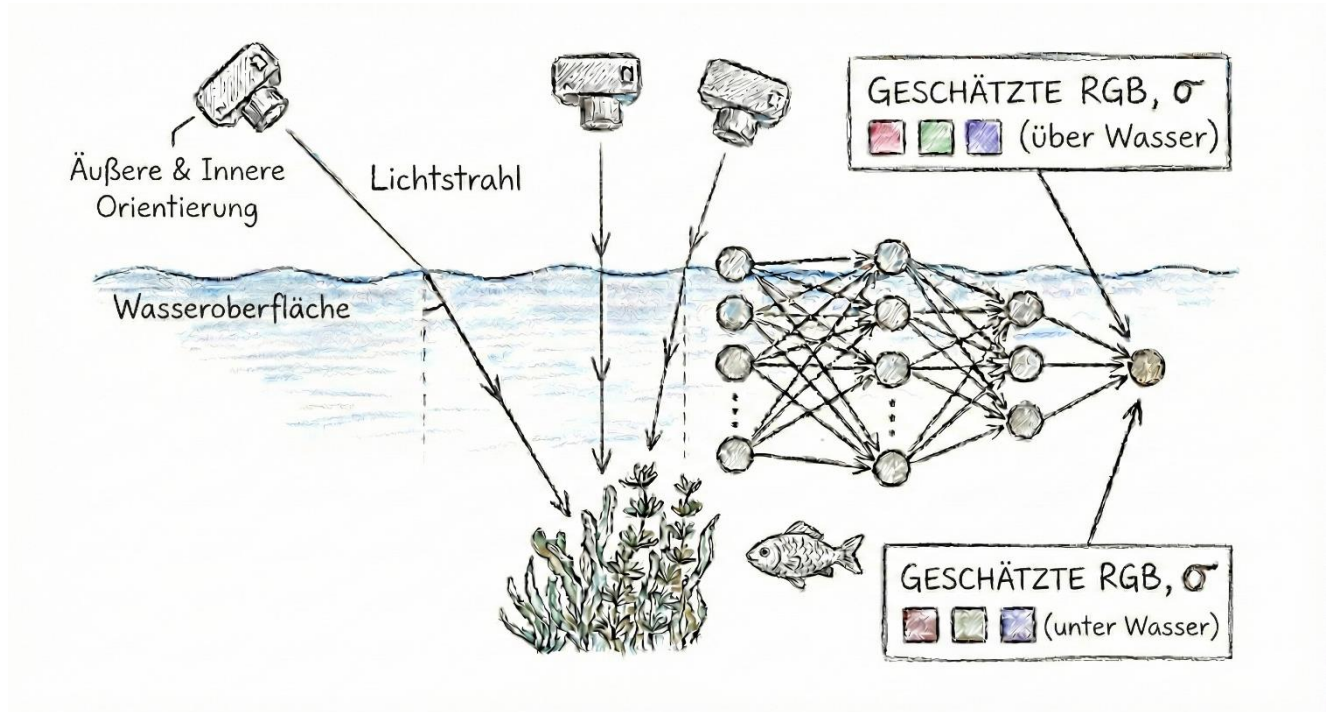


Abbildung 1 Konzeptabbildung BathyNeRF. Durch die korrekte Berücksichtigung der Strahlbrechung an der Wasseroberfläche wird der Tiefenfehler verringert und die geometrische Qualität der unter Wasser rekonstruierten Objekte verbessert. Grafik mit Gemini Pro 3/Nano Banana Pro erstellt.

2 Methoden

In diesem Abschnitt werden zunächst die notwendigen Vor- und Nachprozessierungsschritte besprochen, bevor auf die verschiedenen Experimente, die im Forschungsprojekt untersucht werden, eingegangen wird.

2.1.1 Orientierung

Die Orientierung der Aufnahmen ist eine Voraussetzung für die NeRF-Prozessierung. Im Falle der simulierten Daten kann die äußere und innere Orientierung direkt aus den Simulationseinstellungen abgegriffen werden und muss nicht geschätzt werden. Bei den realen Daten werden mittels Structure-from-Motion zunächst Pass- und Verknüpfungspunkte in den Aufnahmen detektiert, die in einem Bündelblockausgleich zur gleichzeitigen Schätzung der inneren und äußeren Orientierungen aller Aufnahmen genutzt werden.

2.1.2 Georeferenzierung

Für eine erfolgreiche Anwendung der entwickelten Methoden, sowie um den Vergleich der Ergebnisse mit externen Referenzdaten zu ermöglichen, ist eine Georeferenzierung der Daten unabdinglich. Da die verwendeten NeRF-Implementierungen typischerweise in lokalen, nicht-skalierten Koordinatensystemen arbeiten, ist dies auch für die simulierten Datensätze nötig. Dazu werden die datumsgebenden Informationen (Elemente der direkten Georeferenzierung bzw. Passpunktkoordinaten) zunächst aus einem erdfesten (ECEF) Koordinatensystem in ein lokales Tangentialkoordinatensystem transformiert. In diesem System bildet der Schwerpunkt der Kamerazentren den Ursprung und die lokale Zenitrichtung die positive Z-Achse. Die X-Achse zeigt nach Osten, und die Y-Achse ergänzt zu einem mathematisch positiven, rechtshändigen System. Die Transformation wird für jedes Untersuchungsgebiet ermittelt und mit den Daten gespeichert. Nach dem Export der Geometrien aus dem NeRF wird die Transformation wieder invertiert, und aus den erhaltenen ECEF-Koordinaten je nach Bedarf projizierte UTM- oder GK-Koordinaten berechnet.

2.2 Dual-NeRF-Ansatz mit ebener Trennfläche

Im ersten Ansatz wird der Objektraum in zwei Teile getrennt: einen Teil oberhalb der Wasseroberfläche, der Luft und alle Objekte, die nicht unter Wasser stehen, beinhaltet, und einen Teil unterhalb davon. Die Trennschicht wird dabei als horizontale, unbewegte Ebene bekannter Höhe angenommen.

Pro Teil wird nun ein eigenes NeRF trainiert. Trifft ein Bildstrahl bei der Rekonstruktion auf die Trennschicht, wird eine Brechung nach dem Gesetz von Snellius unter der Annahme von $n_{Luft}=1,0$ und $n_{Wasser}=1,333$ berechnet. Die Kombination aus den zwei NeRFs ermöglicht im Gegensatz zu REF²-NeRF (Kim et al., 2024) das geometrisch korrekte Rekonstruieren von Objekten oberhalb und unterhalb der Wasseroberfläche. Das gilt auch, wenn die Wasseroberfläche z.B. durch Äste oder Brücken teilweise verdeckt ist.

Abbildung 2 zeigt schematisch die Kombination der beiden Teil-NeRFs (NeRF_a und NeRF_b). Der Unterwasser-Anteil erzeugt ein virtuelles Pixel mit einer bestimmten Farbe an der Wasseroberfläche. Diese Information wird an das Überwasser-NeRF (NeRF_a) weitergegeben und bildet dort die Grundlage für die finale Pixelfarbe im Bild. Objekte, die zwischen Kamera und Wasseroberfläche vorhanden sind, können dabei die Information des Unterwasser-NeRFs ganz oder teilweise verdecken. Da unter Wasser eine wesentlich stärkere Dämpfung des Lichts als in Luft auftritt, die wellenlängenabhängig ist, wird für das Unterwasser-NeRF die Dämpfung für die Farbanteile Rot, Grün und Blau im Gegensatz zum klassischen NeRF-Ansatz separat geschätzt.

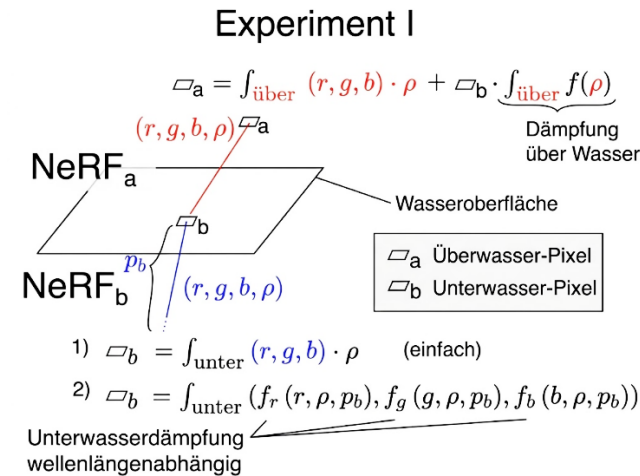


Abbildung 2 Schaubild zum Dual-NeRF-Ansatz mit ebener Trennfläche.

2.3 Dual-NeRF-Ansatz mit lokal geneigter Trennfläche

Die Darstellung der Wasseroberfläche als horizontale Ebene im vorigen Ansatz ist idealisiert und tritt in der Realität nur an windstillen Tagen und bei stehenden Gewässern auf. Um den Ansatz auf weitere reale Anwendungsgebiete zu erweitern, wird im zweiten Ansatz die Wasseroberfläche zwar mit einer konstanten Höhe, aber lokalen Neigungen angenommen. Die lokalen Neigungen sorgen dann entsprechend des Gesetzes von Snellius für eine lokale Richtungsänderung des gebrochenen Bildstrahls.

Als Modell für die lokal generierte Oberfläche dient ein Raster, bei dem pro Pixel die Komponenten des Oberflächennormalvektors gespeichert sind. Wenn nun bei der Auswertung des NeRFs ein Bildstrahl verfolgt wird, wird die Position, an der dieser auf die Wasseroberfläche trifft, bestimmt, das entsprechende Pixel ausgewertet, und die Oberflächenorientierung bei der Refraktion berücksichtigt.

Auch wenn davon auszugehen ist, dass die Oberfläche bei geneigten Elementen insgesamt nicht mehr planar ist, ist der Einfluss eines leichten Versatzes dieser Oberfläche im Gegensatz zur Verkippung vernachlässigbar. Für die Bestimmung des Schnittpunkts wird daher weiterhin ein Ebenenmodell genutzt. Der Ansatz lässt sich somit auch als Raster von kleinen, gegenüber der Horizontalen geneigten, Oberflächenelementen sehen, die die Refraktion lokal beeinflussen.

Abbildung 3 zeigt diesen Ansatz mit mehreren Bildstrahlen eines Photos, die oberhalb der Wasseroberfläche konfokal sind, dies aber aufgrund der unterschiedlichen Strahlbrechungen unter Wasser – wie auch bereits beim Fall mit ebener Trennfläche – nicht mehr sind. Die Kombination der beiden NeRFs findet analog zu Experiment I statt.

2.4 NeRF-Ansatz mit optischer Dichte

Im dritten Ansatz wird die Trennfläche zwischen Luft und Wasser nicht mehr explizit modelliert. Stattdessen wird in der NeRF-Regressionsaufgabe, bei der jedem Ort im Raum ein (r, g, b, σ) -Tupel zugewiesen wird, ein zusätzlicher Parameter für die optische Dichte bzw. die Brechungszahl n , sowie zwei Parameter für den lokalen Oberflächennormalvektor hinzugefügt. Die dritte Komponente des Normalvektors ergibt sich durch die Vorgabe $\left| \begin{pmatrix} n_x & n_y & n_z \end{pmatrix}^T \right|_2 = 1$. Damit ist die Regressionsaufgabe $f(X, Y, Z) = (r, g, b, \sigma, n, n_x, n_y)$

formuliert. Wird nun ein Bildstrahl verfolgt, so wird die Funktion stückweise ausgewertet. Das Verhältnis der Brechungsindizes von zwei aufeinanderfolgenden Auswertungen gibt dabei die Stärke der Refraktion, der geschätzte Oberflächennormalvektor die Richtung der Refraktion vor (siehe Abbildung 4). So kann der Bildstrahl iterativ verfolgt werden, bis ein Objekt getroffen wird und die NeRF-Komponenten (r, g, b, σ) eine entsprechende Farbe für das zu rekonstruierende Pixel erzeugen.

Experiment II

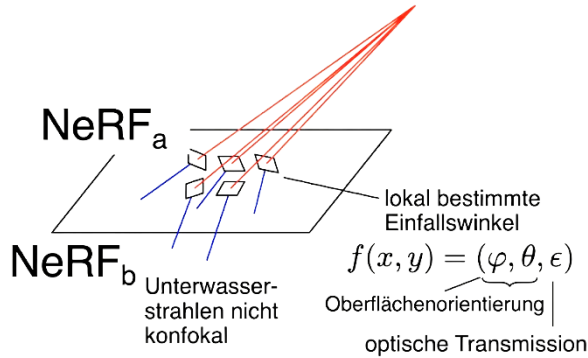


Abbildung 3 Schaubild zum Dual-NeRF-Ansatz mit lokal geeigneter Trennfläche.

3 Erste Ergebnisse

Um volle Kontrolle über die Einflüsse von Wasseroberfläche und Refraktion zu haben, wird zunächst ein rein simulierter Datensatz herangezogen. Dabei wird die Szene von Schulte et al. (2025) genutzt, die ein gesunkenes Schiff auf einem Schotterbett unter Wasser zeigt. Das Schiff weist dabei filigrane Strukturen (Täue und Takelagen) auf, welche besonders dazu geeignet sind, die geometrische Rekonstruktionsqualität und Auflösung zu beurteilen.

Schulte et al. (2025) haben gezeigt, dass bei einer fehlenden Berücksichtigung der Refraktion im NeRF die Szene zwar rekonstruiert werden kann, jedoch ein Tiefenfehler von – in diesem Fall – ca. 0,5 bis 1,5 m auftritt und die Rekonstruktion somit geometrisch unbefriedigend ist.

Zum aktuellen Zeitpunkt sind nur Ergebnisse des ersten Experiments vorhanden. Ein synthetisches Rendering der Szene ist in Abbildung 5 gezeigt. Es ist zu erkennen, dass der Zwei-Medien-Ansatz die Geometrie und Farbinformation der überspülten Objekte erlernen und repräsentieren kann.

Experiment III

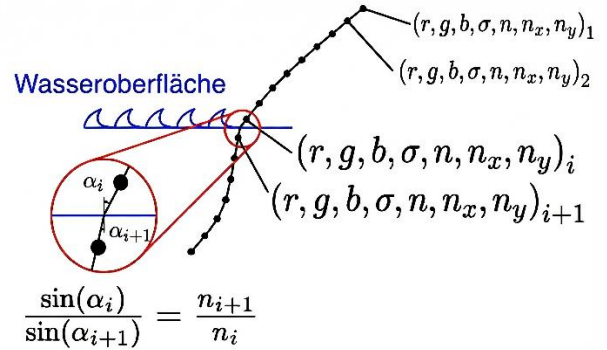


Abbildung 4 Schaubild zum NeRF-Ansatz mit modellierter optischer Dichte und Oberflächenorientierung.



Abbildung 5 Rekonstruierte Unterwasserszene mit zwei getrennten NeRFs und einer horizontalen Wasseroberfläche.

4 Diskussion

Sowohl die Ergebnisse der Literaturrecherche als auch eigene Arbeiten zeigen, dass die korrekte Berücksichtigung der Lichtrefraktion auch bei NeRF-Ansätzen von hoher Wichtigkeit ist. Insbesondere, wenn es darum geht, die Geometrie der Unterwasserszene oder filigrane Objektdetails zu rekonstruieren, muss diese explizit oder implizit modelliert werden.

Der erste Ansatz, der im Rahmen des BathyNeRF-Projekts vorgestellt wurde, modelliert die Wasseroberfläche als horizontale Ebene in bekannter Höhe. Mit diesem Ansatz konnten zufriedenstellende Ergebnisse erzielt werden. Im Gegensatz zu den in der Literatur präsentierten Ansätzen erlaubt unser Ansatz die gleichzeitige Rekonstruktion von Szenenteilen über und unter Wasser, was insbesondere in topographischen Anwendungsfällen (z.B. Fluss, Auwald;

siehe Brezovsky et al., 2025; Günthner et al., 2025) von Bedeutung ist.

5 Fazit & Ausblick

Die weiteren Experimente, die in diesem Beitrag kurz konzeptionell vorgestellt wurden, weisen eine zunehmend höhere Komplexität auf. Mit steigender Komplexität muss beim Trainieren der neuronalen Netze besonders darauf Acht gegeben werden, Korrelationen zwischen den zu bestimmenden Parametern aufzulösen (indem z.B. geeignete dekorrelierende Beobachtungen eingefügt werden) sowie das Modell physikalisch realistisch zu halten, indem ggf. Regularisierungsparameter eingeführt werden. Solche können beispielsweise genutzt werden, um lokal geschätzte Neigungen der Wasseroberfläche in einem plausiblen Bereich zu berücksichtigen.

Die weiteren Experimente sind jedoch näher an realistischen Anwendungen, da in topo-bathymetrischen Datensätzen selten von einer horizontalen, ebenen Wasseroberfläche ausgegangen werden kann. So kann der Weg von der Simulation zu realen Datensätzen Schritt für Schritt begangen werden.

6 Fördergeber

Das transnationale WEAVE-Forschungsprojekt BathyNeRF ist durch den Austrian Science Fund (FWF) [10.55776/PIN1353223] sowie durch die Deutsche Forschungsgemeinschaft (DFG) – 538522540 gefördert.

Literaturverzeichnis

- Brezovsky, M., Günthner, A., Schulte, F., Winiwarter, L., Jutzi, B., & Mandlbürger, G. (2025). Analysis of refraction aware Neural Radiance Fields for 3D reconstruction of through the water scenes. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-2-W10-2025*, 39–45. <https://doi.org/10.5194/isprs-archives-XLVIII-2-W10-2025-39-2025>
- Günthner, A., Brezovsky, M., Schulte, F., Winiwarter, L., Mandlbürger, G., & Jutzi, B. (2025). Exploring the Potential of Refractive NeRFs for Photogrammetric Bathymetry—First Application to UAV-based Data from the Pielach River. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-2-W10-2025*, 107–114. ISPRS WG II/7 - 3D Underwater Mapping from Above and Below – 3rd International Workshop - 8–11 July 2025, TU Wien, Vienna, Austria.

- <https://doi.org/10.5194/isprs-archives-XLVIII-2-W10-2025-107-2025>
- Kim, W., Fukiage, T., & Oishi, T. (2024). REF²-NeRF: Reflection and Refraction aware Neural Radiance Field. *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 7196–7203. <https://doi.org/10.1109/IROS58592.2024.10801830>
- Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020). *NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis* (arXiv:2003.08934). arXiv. <http://arxiv.org/abs/2003.08934>
- Schulte, F., Brezovsky, M., Günthner, A., Jutzi, B., Mandlbürger, G., & Winiwarter, L. (2025). Simulation and validation of underwater scenes for two-media optical 3D reconstruction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-2/W10-2025*, 271–278. <https://doi.org/10.5194/isprs-archives-XLVIII-2-W10-2025-271-2025>
- Zhan, Y., Nobuhara, S., Nishino, K., & Zheng, Y. (2023). *NeRFrac: Neural Radiance Fields through Refractive Surface*. 18402–18412. https://openaccess.thecvf.com/content/ICCV2023/html/Zhan_NeRFrac_Neural_Radiance_Fields_through_Refractive_Surface_ICCV_2023_paper.html