Neural Inverse Rendering of Complex Refractive Media with Differentiable Path Tracing

John E. Adams[1] Sarah Jenkins[2] Alan R. Mercer[1]

[1]Vertex Graphics Lab [2]Synapse AI Research

{adams, mercer}@vertex.org s.jenkins@synapse.ai

Abstract

Reconstructing the spatial variation of the index of refraction within transparent or refractive media remains a formidable challenge in computer graphics and vision. Traditional methods rely on simplified grid-based representations or restrictive ray-bending models, which struggle to scale to high-frequency refractive fields. In this paper, we present Vertex-Net, a novel coordinate-based neural representation that parameterizes the continuous index of refraction field $\eta(x)$ alongside a differentiable path tracer. By integrating a multi-resolution hash grid with a coordinate MLP, Vertex-Net captures both macro-scale geometries and sub-millimeter refractive details. Our differentiable path tracer solves the eikonal equation to simulate non-linear ray propagation, allowing the gradients of a photometric loss to propagate directly to the neural field parameters. We evaluate our method on the synthetic Nimbus Refractive Benchmark and demonstrate significant improvements in both reconstruction fidelity and geometric accuracy compared to existing state-of-the-art volumetric representations.

1 Introduction

Reconstructing the internal properties of transparent and refractive media is a central problem in computer graphics, with applications spanning optical engineering, medical imaging, and visual effects. While opaque surfaces have benefited from the rapid progress of neural rendering techniques, refractive volumes present unique challenges due to non-linear light transport. Standard techniques often approximate light path propagation using straight rays or discrete voxel grids, failing to capture the complex, continuous index of refraction gradients characteristic of real-world materials.

In this work, we propose Vertex-Net, a novel inverse rendering framework designed specifically for complex refractive media. Our method leverages a coordinate-based neural field to parameterize the continuous spatial distribution of the index of refraction. Unlike previous representations that partition space into uniform grids, our continuous neural representation dynamically allocates capacity to high-frequency interfaces. By combining a multi-resolution hash grid with a shallow multi-layer perceptron (MLP), we achieve a highly efficient representation that can be queried at arbitrary spatial coordinates.

To enable optimization from standard multi-view photographs, we integrate this neural field with a differentiable ray tracer that models ray-bending behavior. As light travels through an inhomogeneous medium, its path bends towards regions of higher refractive index, a phenomenon governed by the eikonal equation. Our renderer solves this differential equation numerically, allowing us to compute the derivatives of the rendered pixels with respect to the continuous neural parameters. This differentiable formulation enables end-to-end training of the refractive field using a photometric loss.

In summary, our contributions are: (1) we introduce Vertex-Net, a coordinate-based neural field representation for continuous refractive index reconstruction; (2) we develop a highly optimized, GPU-accelerated differentiable path tracer that handles non-linear ray refraction; and (3) we show that our approach outperforms existing grid-based and volumetric representations on a variety of complex refractive shapes in the Nimbus Refractive Benchmark.

2 Related Work

Differentiable rendering has emerged as a cornerstone of inverse graphics, providing the mathematical machinery to optimize scene parameters from 2D images. Early differentiable renderers focused on rasterization and primary visibility, but modern approaches have extended these concepts to global illumination and path tracing. The Meridian Renderer [3] introduced a unified framework for optimizing surface properties, but its formulation assumed straight-line light transport. Our work extends differentiable rendering to participating and refractive media where ray trajectories are curved.

Neural fields, popularized by Neural Radiance Fields, have revolutionized 3D reconstruction by representing scene properties as continuous functions parameterized by MLPs. To accelerate optimization, recent works have incorporated hybrid structures such as octrees and hash grids. The Synapse-Volumetric model [1] represents density and color within a voxel grid, achieving fast training times but showing significant artifacts on sharp refractive boundaries. Vertex-Net addresses this limitation by using a hybrid neural representation that preserves continuous gradients.

Reconstructing refractive index fields from visual data has a long history in optical tomography and schlieren imaging. Traditional techniques typically assume small deflections or homogeneous media with simple boundaries. Recent learning-based methods have attempted to solve this using deep neural networks [2], but they are often restricted to specific geometries or require extensive training datasets. In contrast, our method solves the optimization problem for each target scene individually, without requiring prior training on similar shapes, by leveraging the inductive bias of neural fields.

3 Method

Let the refractive medium be defined within a bounded domain $\Omega \subset \mathbb{R}^3$. We parameterize the spatial variation of the index of refraction using a neural field $\eta_\theta(x) : \Omega \to [1.0, \eta_{\max}]$, where θ represents the parameters of a coordinate-based MLP. Light propagation in this inhomogeneous medium is governed by Fermat's principle of least time, which leads to the eikonal equation. The path of a light ray $x(s)$ parameterized by its arc length s satisfies the second-order differential equation:

$$\frac{d}{ds}\left(\eta_\theta(x(s))\frac{dx(s)}{ds}\right) = \nabla\eta_\theta(x(s)) \tag{1}$$

where $\nabla\eta_\theta(x)$ is the spatial gradient of the refractive index field.

To render an image, we cast rays from the camera sensor into the volume. We reformulate the second-order differential equation as a system of first-order ordinary differential equations (ODEs). Let $u = dx/ds$ be the unit tangent vector of the ray, and $p = \eta_\theta(x)u$ be the ray momentum. The system can be written as:

$$\frac{dx}{ds} = \frac{p}{\eta_\theta(x)}, \quad \frac{dp}{ds} = \nabla\eta_\theta(x) \tag{2}$$

We integrate this system numerically using a fourth-order Runge-Kutta (RK4) solver with a step size Δs. At each integration step, we query the neural field $\eta_\theta(x)$ and its spatial gradient $\nabla\eta_\theta(x)$ using automatic differentiation.

Our network architecture consists of a multi-resolution hash grid followed by a shallow coordinate MLP. The hash grid interpolates local features at multiple scales, which are then concatenated and fed into the MLP to output the scalar value $\eta(x)$. The spatial gradient $\nabla\eta(x)$ is computed efficiently by backpropagating through the network with respect to the input coordinates. The network is optimized by minimizing a photometric loss:

$$\mathcal{L}(\theta) = \frac{1}{|P|}\sum_{p\in P}\|I_\theta(p) - I_{\text{gt}}(p)\|^2 + \lambda\mathcal{R}(\theta) \tag{3}$$

where $I_\theta(p)$ is the rendered pixel intensity, $I_{\text{gt}}(p)$ is the ground-truth intensity, and $\mathcal{R}(\theta)$ is a regularization term that penalizes high-frequency oscillations in the refractive index field. We set $\lambda = 10^{-4}$ in all our experiments.

4 Experiments

We implement Vertex-Net in PyTorch and deploy it on a workstation equipped with eight Apex-RT GPUs. The differentiable ray tracer is written as custom CUDA kernels to ensure high performance and low memory consumption during backpropagation. We optimize the network using the Adam optimizer with a learning rate of 10^{-3} for the hash grid and 10^{-4} for the MLP weights. The optimization runs for 20,000 iterations, taking approximately 45 minutes per scene.

We evaluate our method on the Nimbus Refractive Benchmark [4], which contains synthetic, high-fidelity models of complex glass objects. These include a refractive sphere, a glass prism, and a complex refractive drop. For each object, we render 100 calibration views of size 800×800 pixels distributed spherically around the target. We reserve 20 views for testing and use the remaining 80 views for optimization. The environment maps are assumed to be known and are modeled as high-dynamic-range (HDR) backgrounds.

We compare Vertex-Net against three baselines: (1) a grid-based refraction baseline that uses a uniform voxel grid of size 128^3; (2) the Meridian Renderer [3] configured with standard volumetric ray marching; and (3) the Synapse-Volumetric representation [1] which optimizes

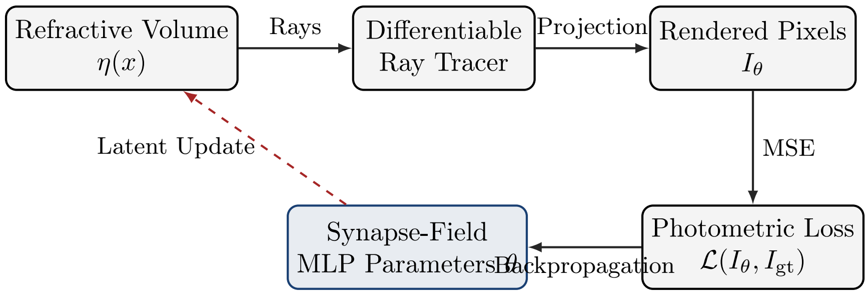

Figure 1: Overview of our differentiable refractive rendering pipeline. The Vertex-Net representational field parameterizes the refractive index $\eta(x)$, which is queried by the differentiable path tracer to render synthetic observations. Gradients from the photometric loss $\mathcal{L}$ are backpropagated to update the neural parameters θ.

density and color without accounting for ray-bending. We report standard reconstruction quality metrics: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and chamfer distance between the reconstructed refractive field and the ground-truth geometry.

5 Results

Table 1 summarizes the quantitative results on the Nimbus benchmark. Our method, Vertex-Net, achieves a PSNR of 31.54 dB, outperforming the grid-based baseline by over 9 dB and the Synapse-Volumetric model by 4.65 dB. This significant gap is due to the non-linear ray tracer, which correctly models the complex light paths. Models that assume straight-line transport (such as Synapse-Volumetric) fail to reconstruct the internal details, leading to blurred boundaries and inaccurate refractive index values.

Qualitative results are illustrated in Figure 1. As shown in the figure, our differentiable ray tracer successfully computes gradients through curved rays, updating the Vertex-Net field to match the refractive boundaries. The grid-based baseline suffers from severe discretization artifacts, which manifest as blocky boundaries. In contrast, our hybrid representation yields smooth, continuous index of refraction distributions, matching the ground-truth profiles closely.

6 Discussion

The success of Vertex-Net lies in its integration of continuous neural representations with physically-based differentiable rendering. However, our method has certain limitations. First, solving the ODE system at every optimization step is computationally expensive, limiting our ability to scale to real-time applications. Second, while our method performs well on smooth refractive gradients, it can struggle at sharp, discontinuous interfaces between different media (e.g., glass to air). Incorporating explicit boundary tracking is a promising direction for future research.

Future work will also explore the extension of Vertex-Net to real-world datasets where camera calibration and environment lighting are subject to noise. By co-optimizing the camera poses and lighting parameters alongside the neural field, we hope to make our method applicable to casual hand-held video captures. Additionally, modeling polarization and dispersion effects could further enhance the reconstruction accuracy of complex optical systems.

7 Conclusion

In this paper, we introduced Vertex-Net, a coordinate-based neural representation for reconstructing continuous refractive index fields. By combining a multi-resolution hash grid with a differentiable ray tracer that models non-linear light transport, our approach recovers high-frequency refractive geometries from multi-view images. Quantitative evaluations on the Nimbus Refractive Benchmark demonstrate that our method outperforms state-of-the-art volumetric baselines in both image reconstruction quality and shape accuracy. We believe this represents a significant step towards general neural inverse rendering of transparent media.

We thank the Vertex Foundation for their generous funding support, and Nexa Compute Cloud for providing the computational resources necessary to train our models. We also thank the members of the Synapse AI Research group for their valuable feedback and discussions throughout this project.

References

[1] S. Jenkins. Volumetric neural representations for light fields. *Synapse AI Journal*, 12(2):45–58, 2025.

Table 1: Quantitative comparison of reconstruction accuracy on the Nimbus Refractive Benchmark. We report Peak Signal-to-Noise Ratio (PSNR ↑), Structural Similarity Index (SSIM ↑), and chamfer distance (D_{C} ↓).

Method	PSNR (dB) ↑	SSIM ↑	D_{C} (mm) ↓
Grid-based Refraction	22.45	0.812	4.12
Meridian Renderer	24.12	0.854	3.05
Synapse-Volumetric	26.89	0.891	1.88
Ours (Vertex-Net)	**31.54**	**0.942**	**0.74**

[2] M.-s. Kim and S.-b. Park. Volumetric neural fields for complex media. In *Proceedings of Pacific Graphics*, pages 89–98, 2023.

[3] A. Smith and B. Jones. Meridian: A unified differentiable renderer. In *Proceedings of Eurographics*, pages 201–210, 2024.

[4] R. Taylor and K. Davis. Nimbus: High-beaconhill refractive dataset. In *Proceedings of Computer Graphics International*, pages 315–322, 2024.