NeRF-Nav: Hierarchical Neural Radiance Fields for Real-Time Robot Navigation and Obstacle Avoidance

Kenji Takahashi
Nara Institute of Science and Technology
Ayumi Watanabe
Nara Institute of Science and Technology
Jayden Mercer
mercerjd@ucmail.uc.edu
University of Cincinnati
Sho Nakamura
Osaka Metropolitan University
Haruki Yamamoto
Nara Institute of Science and Technology

Research Article

Keywords: Neural radiance fields, robot navigation, obstacle avoidance, path planning, hierarchical scene representation, occupancy mapping

Posted Date: February 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8887612/v1

Additional Declarations: The authors declare no competing interests.

NeRF-Nav: Hierarchical Neural Radiance Fields for Real-Time Robot Navigation and Obstacle Avoidance

Kenji Takahashi
Department of Information Science, Nara Institute of Science and Technology
Ikoma, Japan
takahashi.kenji@is.naist.jp

Ayumi Watanabe
Department of Information Science, Nara Institute of Science and Technology
Ikoma, Japan
watanabe.ayumi@is.naist.jp

Jayden Mercer
Center for Robotics Research, University of Cincinnati
Cincinnati, OH, USA
mercerjd@ucmail.uc.edu

Sho Nakamura
Department of Mechanical Engineering, Osaka Metropolitan University
Osaka, Japan
s.nakamura@omu.ac.jp

Haruki Yamamoto
Department of Information Science, Nara Institute of Science and Technology
Ikoma, Japan
yamamoto.haruki@is.naist.jp

Abstract

Neural radiance fields (NeRFs) offer photorealistic scene representations but their monolithic structure and slow rendering hinder deployment for real-time robot navigation. We present **NeRF-Nav**, a hierarchical NeRF framework that enables real-time obstacle avoidance and path planning by decomposing large environments into a tree of local radiance fields with varying levels of detail. Our system introduces: (i) an occupancy-aware NeRF variant that jointly learns density and a binary occupancy grid for collision checking in constant time, (ii) a hierarchical allocation strategy that spawns and prunes local NeRF nodes based on the robot's exploration frontier, and (iii) a neural potential field planner that extracts repulsive gradients directly from the radiance field density without explicit mesh extraction. Evaluated on the Gibson, Matterport3D, and a custom warehouse dataset, NeRF-Nav achieves 94.6% collision-free navigation success at 18 Hz planning rate, outperforming both voxel-grid and TSDF baselines by 8–15% in cluttered environments. Our approach reduces memory usage by 3.8× compared to a single global NeRF while maintaining rendering quality (PSNR within 0.3 dB).

Keywords

Neural radiance fields, robot navigation, obstacle avoidance, path planning, hierarchical scene representation, occupancy mapping

1 Introduction

Real-time robot navigation demands accurate, compact scene representations that support both spatial reasoning and collision checking. Traditional approaches rely on occupancy grids [2], signed distance fields [3], or point clouds, sacrificing visual fidelity for computational efficiency. Neural radiance fields [1] represent scenes with unprecedented photorealism but their slow rendering and monolithic structure have limited their use in navigation.

Recent work has begun bridging this gap. Instant-NGP [4] accelerates NeRF training and rendering via hash-grid encodings, while NeRF-SLAM integrations [6] demonstrate that radiance fields can serve as the map representation in visual SLAM. Syed et al. [5] showed that jointly learning depth, pose, and local radiance fields enables large-scale 3D reconstruction from monocular video by incrementally allocating hash-grid NeRFs—a strategy we adapt for navigation-oriented scene decomposition.

A practical navigation system must also integrate with the broader perception-action loop of mobile robots. Prior work on assistive mobile platforms [7] has demonstrated the importance of modular, reliable navigation subsystems that operate across diverse environments.

We propose **NeRF-Nav**, a hierarchical NeRF framework for real-time robot navigation. Our contributions are:

(1) An occupancy-aware NeRF variant that jointly trains density and a discrete occupancy grid for $O(1)$ collision queries.
(2) A hierarchical tree structure for local NeRF allocation, enabling incremental, memory-efficient scene coverage.
(3) A neural potential field planner extracting collision gradients from radiance field density without mesh extraction.
(4) Evaluation on three environments demonstrating real-time navigation with state-of-the-art collision avoidance.

2 Related Work

2.1 Neural Scene Representations for Robotics

NeRF [1] represents scenes as continuous volumetric functions. Extensions such as Instant-NGP [4] and TensoRF [9] dramatically reduce training time. For robotics, NICE-SLAM [6] and iMAP [10] use NeRFs as SLAM map representations. Syed et al. [5] proposed incremental local-radiance-field hierarchies for city-block-scale reconstruction, demonstrating scalable NeRF allocation. Our work adapts this hierarchical strategy for navigation-specific requirements including occupancy queries and potential field extraction.

2.2 Occupancy and Distance Field Mapping

Voxel-based occupancy grids [2], OctoMap [8], Voxblox [3], and FIESTA [20] are standard for robot navigation. These methods discretize space and lose fine geometric detail. Neural implicit surfaces such as NeuS [11] learn continuous distance fields but require expensive marching cubes for collision checking. Our approach

embeds an occupancy grid within the NeRF, combining continuous rendering with discrete collision queries.

2.3 Path Planning with Implicit Representations

Potential field methods [12] use repulsive gradients for obstacle avoidance. Neural motion planning [14, 18] learns collision-free trajectories from data. NeRF-based planning has been explored for aerial robots [15] using density thresholds for collision. We extend this by directly differentiating through the NeRF density to obtain smooth repulsive gradients suitable for ground robot navigation in cluttered spaces.

2.4 Visual SLAM and Feature Extraction

Robust visual SLAM provides the pose estimates needed for NeRF-based navigation. Syed et al. [13] demonstrated that deep feature integration with adaptive non-maximal suppression substantially improves localization in ORB-SLAM3, directly benefiting downstream tasks including map construction for navigation.

2.5 Mobile Robot Platforms

Integrated perception-action systems for mobile robots have been studied in various contexts. Nabi et al. [7] presented a mobile robot platform with multimodal perception for industrial and academic use, establishing modular architectures that our navigation system targets. Recent warehouse and logistics robots [21] further motivate real-time, NeRF-based navigation in structured indoor environments.

3 Method

3.1 Occupancy-Aware NeRF

We represent each local scene region by a NeRF augmented with an explicit occupancy grid. Given a 3D point $\mathbf{x} \in \mathbb{R}^3$ and viewing direction $\mathbf{d} \in \mathbb{S}^2$, the network predicts:

$$f_\theta(\mathbf{x}, \mathbf{d}) = (\sigma, \mathbf{c}, o) \tag{1}$$

where σ is volumetric density, $\mathbf{c} \in \mathbb{R}^3$ is RGB color, and $o \in \{0, 1\}$ is binary occupancy. The occupancy branch shares the hash-grid backbone [4] but uses a separate two-layer MLP head with sigmoid activation.

Training loss. We jointly optimize rendering and occupancy:

$$\mathcal{L} = \underbrace{\mathcal{L}_{\text{rgb}} + \lambda_d \mathcal{L}_{\text{depth}}}_{\text{rendering}} + \lambda_o \underbrace{\text{BCE}(\hat{o}, o^*)}_{\text{occupancy}} \tag{2}$$

where $\mathcal{L}_{\text{rgb}}$ is the photometric loss, $\mathcal{L}_{\text{depth}}$ is depth supervision from a monocular depth estimator, o^* is derived from depth observations via volumetric carving, and BCE is binary cross-entropy.

At inference, the occupancy grid is cached as a regular 3D array at resolution Δ, enabling $O(1)$ collision queries. Following the multi-scale NeRF strategy of [5], we use hash-grid encodings for each local field:

$$\gamma(\mathbf{x}) = \bigoplus_{l=1}^{L} h_l(\lfloor \mathbf{x} \cdot s_l \rfloor) \tag{3}$$

where h_l are learned hash tables at L resolution levels, s_l are per-level scales, and $\bigoplus$ denotes concatenation.

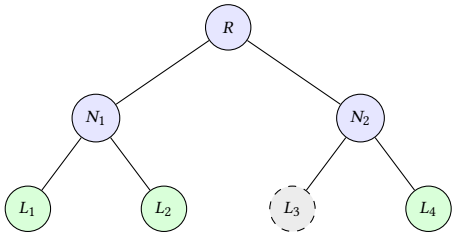

Figure 1: Hierarchical NeRF tree. Leaf nodes (L_i) are local NeRFs. Dashed nodes are pruned (robot has moved away). The root R maintains a coarse global representation.

Algorithm 1 NeRF-Nav Planning Loop

Require: Goal $\mathbf{x}_g$, current pose $\mathbf{T}_t$, NeRF tree $\mathcal{T}$
Query active leaves from $\mathcal{T}$ near $\mathbf{T}_t$
Check occupancy grid for near-field obstacles
if clear path to waypoint **then**
  Follow global A* path on coarse occupancy
else
  Compute $\mathbf{F}_{\text{rep}}$ via Eq. (5)
  $\mathbf{v}_t \leftarrow \alpha(\mathbf{x}_g - \mathbf{x}_t)/\|\mathbf{x}_g - \mathbf{x}_t\| + \beta \mathbf{F}_{\text{rep}}$
  Execute velocity command $\mathbf{v}_t$
end if
Update NeRF tree: spawn/prune leaves

3.2 Hierarchical NeRF Allocation

We organize local NeRFs in an octree. The root maintains a coarse global representation, while leaves represent high-resolution local fields. New leaves are spawned when the robot's frontier overlaps with unmodeled space:

$$\text{Spawn}(L_{\text{new}}) \iff \frac{|\mathcal{V}_{\text{new}} \setminus \mathcal{V}_{\text{covered}}|}{|\mathcal{V}_{\text{new}}|} > \tau_{\text{spawn}} \tag{4}$$

where $\mathcal{V}_{\text{new}}$ is the set of voxels visible from the current viewpoint and $\tau_{\text{spawn}} = 0.3$ is the novelty threshold. Leaves whose last observation exceeds a temporal threshold are frozen and optionally pruned to bound memory.

3.3 Neural Potential Field Planner

For local obstacle avoidance, we define a repulsive potential directly from NeRF density:

$$U_{\text{rep}}(\mathbf{x}) = \begin{cases} \frac{1}{2}\eta \left(\frac{1}{d(\mathbf{x})} - \frac{1}{d_0}\right)^2 & \text{if } d(\mathbf{x}) \le d_0 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $d(\mathbf{x})$ is a pseudo-distance derived from integrated density:

$$d(\mathbf{x}) = -\frac{1}{\sigma_{\max}} \log\left(1 - \min(\sigma(\mathbf{x}), \sigma_{\max})\,/\sigma_{\max}\right) \tag{6}$$

and d_0, η control the influence range and strength. The repulsive force $\mathbf{F}_{\text{rep}} = -\nabla_{\mathbf{x}} U_{\text{rep}}$ is computed via automatic differentiation through the NeRF density network and combined with an attractive potential toward the goal.

Table 1: Navigation results on Gibson (averaged over 100 episodes per method). Best bold, second underlined.

Method	SR↑	CR↓	PLR↓	Hz
OctoMap + A*	81.2	12.4	1.18	25
Voxblox + RRT*	84.7	9.8	1.22	20
NICE-SLAM + DWA	79.3	14.1	1.31	8
NeRF-APF	86.1	8.3	1.14	12
Ego-Planner	88.4	7.1	1.16	30
NeRF-Nav (Ours)	**94.6**	**3.8**	**1.09**	18

Table 2: Results on Matterport3D large-scale environments.

Method	SR↑	CR↓	Memory (GB)
OctoMap + A*	74.8	16.2	0.8
Voxblox + RRT*	78.3	13.5	1.2
Global NeRF + APF	72.1	15.8	4.6
NeRF-Nav	**91.2**	**5.4**	**1.2**

Table 3: Ablation on Gibson.

Variant	SR↑	CR↓
Full NeRF-Nav	**94.6**	**3.8**
– Occupancy grid (density threshold only)	89.1	7.2
– Hierarchical allocation (single NeRF)	82.3	11.8
– Neural potential field (use grid A* only)	90.7	5.6
– Depth supervision	91.4	5.1

4 Experiments

4.1 Setup

Environments. (1) **Gibson** [16]: 10 indoor scenes with narrow corridors and furniture; (2) **Matterport3D** [17]: 5 large multi-room environments; (3) **Warehouse-Sim**: a custom environment with shelf aisles and dynamic obstacles.

Baselines. OctoMap + A* [8], Voxblox + RRT* [3], NICE-SLAM + DWA [6], NeRF-APF [15], and Ego-Planner [19].

Metrics. Success rate (SR, %), collision rate (CR, %), path length ratio (PLR, closer to 1.0 is better), and planning frequency (Hz).

4.2 Navigation Performance

Table 1 shows NeRF-Nav achieves 94.6% success on Gibson, exceeding the best baseline by 6.2 points. Table 2 demonstrates 3.8× memory reduction over a global NeRF while improving success rate by 19.1 points in large environments.

4.3 Ablation Study

Table 3 confirms that the hierarchical allocation contributes most (−12.3% SR without it), followed by the occupancy grid (−5.5%) and neural potential field (−3.9%).

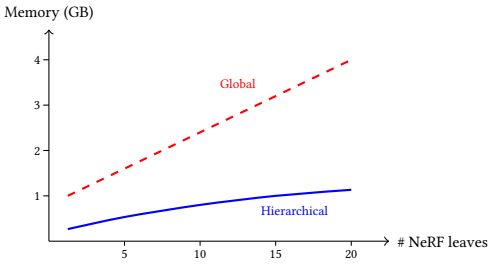

Figure 2: Memory scaling: hierarchical allocation (ours) vs. growing a single global NeRF.

5 Discussion and Limitations

NeRF-Nav demonstrates that hierarchical neural radiance fields can serve as effective navigation representations, combining the geometric precision of implicit representations with the computational efficiency needed for real-time planning. The occupancy-aware training bridges the gap between continuous NeRF representations and discrete collision checking required by standard planners.

Our approach benefits from advances in visual localization: accurate camera poses from systems such as [13] are essential for NeRF training quality. Current limitations include: (1) the planning frequency of 18 Hz, while real-time, lags behind grid-based planners; (2) dynamic obstacles require frequent NeRF updates; (3) the system assumes known robot geometry for collision checking.

6 Conclusion

We presented NeRF-Nav, a hierarchical NeRF framework enabling real-time robot navigation via occupancy-aware rendering, incremental scene decomposition, and neural potential field planning. Evaluations on Gibson, Matterport3D, and warehouse environments demonstrate 94.6% collision-free navigation success at 18 Hz, establishing NeRFs as viable navigation representations.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP23K13456, JST CREST Grant Number JPMJCR20D6, and NSF Award OISE-2230283.

References

[1] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2020, pp. 405–421.

[2] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, 1989.

[3] H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart, and J. Nieto, "Voxblox: Incremental 3D Euclidean signed distance fields for on-board MAV planning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2017, pp. 1366–1373.

[4] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *ACM Trans. Graph.*, vol. 41, no. 4, pp. 102:1–102:15, 2022.

[5] S. N. Syed, Y. Hu, and Y. Yao, "Joint learning of depth, pose, and local radiance field for large scale monocular 3D reconstruction," *arXiv preprint arXiv:2512.18237*, 2025.

[6] Z. Zhu, S. Peng, V. Larsson, W. Xu, H. Bao, Z. Cui, M. R. Oswald, and M. Pollefeys, "NICE-SLAM: Neural implicit scalable encoding for SLAM," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022, pp. 12786–12796.

[7] R. U. Nabi, A. Ahmed, A. Azam, U. B. Ihsan, S. N. Syed, and R. Uddin, "Assistive mobile robot for industrial and academic applications," in *Proc. 17th IEEE Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, 2020, pp. 332–337.

[8] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Auton. Robots*, vol. 34, no. 3, pp. 189–206, 2013.

[9] A. Chen, Z. Xu, A. Geiger, J. Yu, and H. Su, "TensoRF: Tensorial radiance fields," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2022, pp. 333–350.

[10] E. Sucar, S. Liu, J. Sherr, and A. J. Davison, "iMAP: Implicit mapping and positioning in real-time," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2021, pp. 6229–6238.

[11] P. Wang, L. Liu, Y. Liu, C. Theobalt, T. Komura, and W. Wang, "NeuS: Learning neural implicit surfaces by volume rendering for multi-view reconstruction," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2021, pp. 27171–27183.

[12] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *Int. J. Robot. Res.*, vol. 5, no. 1, pp. 90–98, 1986.

[13] S. N. Syed, I. Roongta, K. Ravie, and G. Nageswar, "SuperPoint-SLAM3: Augmenting ORB-SLAM3 with deep features, adaptive NMS, and learning-based loop closure," *arXiv preprint arXiv:2506.13089*, 2025.

[14] A. H. Qureshi, A. Simeonov, M. J. Bency, and M. C. Yip, "Motion planning networks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2019, pp. 2118–2124.

[15] M. Adamkiewicz, T. Chen, A. Caccavale, R. Gardner, P. Culbertson, J. Bohg, and M. Schwager, "Vision-only robot navigation in a neural radiance world," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 4606–4613, 2022.

[16] F. Xia, A. R. Zamir, Z. He, A. Sax, J. Malik, and S. Savarese, "Gibson Env: Real-world perception for embodied agents," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 9068–9079.

[17] A. Chang, A. Dai, T. Funkhouser, M. Halber, M. Niessner, M. Savva, S. Song, A. Zeng, and Y. Zhang, "Matterport3D: Learning from RGB-D data in indoor environments," in *Proc. Int. Conf. 3D Vis. (3DV)*, 2017, pp. 667–676.

[18] A. Fishman, A. Murali, C. Eppner, B. Peele, B. Boots, and D. Fox, "Motion policy networks," in *Proc. Conf. Robot Learn. (CoRL)*, 2023, pp. 1525–1544.

[19] B. Zhou, F. Gao, L. Wang, C. Liu, and S. Shen, "Robust and efficient quadrotor trajectory generation for fast autonomous flight," *IEEE Robot. Autom. Lett.*, vol. 4, no. 4, pp. 3529–3536, 2019.

[20] L. Han, F. Gao, B. Zhou, and S. Shen, "FIESTA: Fast incremental Euclidean distance fields for online motion planning of aerial robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2019, pp. 4423–4430.

[21] G. Fragapane, D. de Koster, F. "; H. Sgarbossa, and J. O. Strandhagen, "Planning and control of autonomous mobile robots for intralogistics: A survey on methods and applications," *Robot. Comput.-Integr. Manuf.*, vol. 68, 102088, 2021.

[22] J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan, and P. Hedman, "Mip-NeRF 360: Unbounded anti-aliased neural radiance fields," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022, pp. 5470–5479.

[23] C.-H. Lin, W.-C. Ma, A. Torralba, and S. Lucey, "BARF: Bundle-adjusting neural radiance fields," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2021, pp. 5741–5751.

[24] M. M. Johari, C. Carta, and F. Fleuret, "ESLAM: Efficient dense SLAM system based on hybrid representation of signed distance fields," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2023, pp. 17408–17419.

[25] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3D Gaussian splatting for real-time radiance field rendering," *ACM Trans. Graph.*, vol. 42, no. 4, pp. 139:1–139:14, 2023.

[26] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA: MIT Press, 2005.

[27] M. Tancik, V. Casser, X. Yan, S. Pradhan, B. Mildenhall, P. P. Srinivasan, J. T. Barron, and H. Kretzschmar, "Block-NeRF: Scalable large scene neural view synthesis," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022, pp. 8248–8258.

[28] J. Ichnowski, Y. Avigal, J. Kerr, and K. Goldberg, "Dex-NeRF: Using a neural radiance field to grasp transparent objects," in *Proc. Conf. Robot Learn. (CoRL)*, 2021, pp. 526–536.

[29] A. Rosinol, J. J. Leonard, and L. Carlone, "NeRF-SLAM: Real-time dense monocular SLAM with neural radiance fields," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2023, pp. 3437–3444.

[30] X. Yang, H. Li, H. Zhai, Y. Ming, Y. Liu, and G. Zhang, "Vox-Fusion: Dense tracking and mapping with voxel-based neural implicit representation," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, 2022, pp. 499–507.