GeRaF: Neural Geometry Reconstruction from Radio Frequency Signals

Jiachen Lu[*], **Hailan Shanbhag**[*], **Haitham Al Hassanieh**
École Polytechnique Fédérale de Lausanne (EPFL)

Abstract

GeRaF is the first method to use neural implicit learning for near-range 3D geometry reconstruction from radio frequency (RF) signals. Unlike RGB or LiDAR-based methods, RF sensing can see through occlusion but suffers from low resolution and noise due to its *lensless* imaging nature. While lenses in RGB imaging constrain sampling to 1D rays, RF signals propagate through the entire space, introducing significant noise and leading to cubic complexity in volumetric rendering. Moreover, RF signals interact with surfaces via specular reflections, requiring fundamentally different modeling. To address these challenges, GeRaF (1) introduces filter-based rendering to suppress irrelevant signals, (2) implements a physics-based RF volumetric rendering pipeline, and (3) proposes a novel lensless sampling and lensless alpha blending strategy that makes full-space sampling feasible during training. By learning signed distance functions, reflectiveness, and signal power through MLPs and trainable parameters, GeRaF takes the first step towards reconstructing millimeter-level geometry from RF signals in real-world settings.

1 Introduction

Geometry reconstruction is a fundamental problem that enables a wide range of applications in fields such as virtual reality [14, 21, 32] and robotics [22, 29]. In recent years, neural reconstruction methods [55, 51, 40, 41, 35] have gained significant attention. A key advantage of these methods is their ability to represent the geometry of a scene continuously, which can help in many downstream tasks. However, vision-based sensors often struggle in environments with adverse weather conditions [67, 20, 37, 34] or even become completely unusable when objects are obscured by occlusions [2, 65, 66, 30, 57, 12, 13].

In contrast, radio frequency (RF) sensing, specifically wireless millimeter-wave (mmWave) sensing, has the ability to see *through* occlusions and remains robust under challenging visibility conditions and, unlike X-Ray, is not dangerous to humans [56], making it a compelling alternative for 3D reconstruction. This could open up a plethora of applications; for example, seeing whether items inside a box are the correct items, are damaged, or pose a threat, without having to ever open up the box. On the other hand, mmWave resolution is extremely low compared to vision. Fig. 1 shows an example of one radar image, which shows significantly less visual context than camera images provide, making it hard to directly extract 3D reconstruction from a radar image alone. In order to perform more complete 3D reconstruction, some works [2, 65, 30, 57, 13] have done 3D point cloud reconstruction or human body tracking by leveraging movement or emulating larger antenna apertures, however, the recovered point clouds from RF are still too sparse and noisy to reliably recover detailed geometry. Other works [6, 24] have proposed using neural implicit representations for RF sensing, aiming to concentrate scene information from heavy noise through neural optimization.

[*] Co-primary first authors, indicates equal contribution.

39th Conference on Neural Information Processing Systems (NeurIPS 2025).