DART: Implicit Doppler Tomography for Radar Novel View Synthesis

Tianshu Huang[*1] John Miller[*12] Akarsh Prabhakara[1] Tao Jin[1] Tarana Laroia[1]
Zico Kolter[12] Anthony Rowe[12]
[1]Carnegie Mellon University [2]Bosch Research
{tianshu2, jmiller4, aprabhak, taojin, tlaroia, zkolter, agr}@andrew.cmu.edu

Abstract

Simulation is an invaluable tool for radio-frequency system designers that enables rapid prototyping of various algorithms for imaging, target detection, classification, and tracking. However, simulating realistic radar scans is a challenging task that requires an accurate model of the scene, radio frequency material properties, and a corresponding radar synthesis function. Rather than specifying these models explicitly, we propose DART — Doppler Aided Radar Tomography, a Neural Radiance Field-inspired method which uses radar-specific physics to create a reflectance and transmittance-based rendering pipeline for range-Doppler images. We then evaluate DART by constructing a custom data collection platform and collecting a novel radar dataset together with accurate position and instantaneous velocity measurements from lidar-based localization. In comparison to state-of-the-art baselines, DART synthesizes superior radar range-Doppler images from novel views across all datasets and additionally can be used to generate high quality tomographic images.[1]

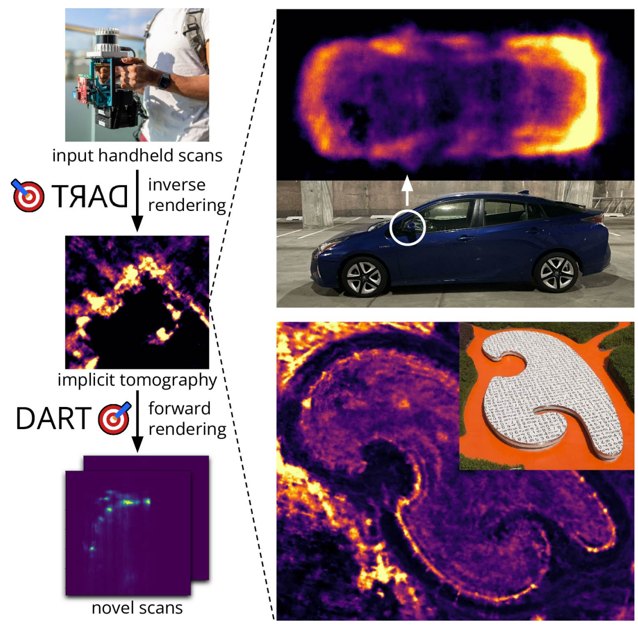

Figure 1. DART uses scans from a handheld radar to learn an implicit tomography of a scene in order to accurately render scans from novel viewpoints (left). DART's implicit tomography can also be sampled to map the radar properties of a scene (right).

1. Introduction

Driven by advances in the automotive industry, miniaturized millimeter wave (mmWave) radar chips are becoming cheaper and more ubiquitous. Boasting a high range resolution and the ability to penetrate light materials, mmWave radars have proven effective in many application domains including collision avoidance and driver assistance in automobiles [14, 36, 57, 64, 65], through-occlusion imaging in airport scanners [30, 68], and vision-denied tracking and mapping [2, 9, 22, 37, 43].

Because designing, testing, and deploying new radar systems in the real world can be costly, many rapid prototyping pipelines heavily rely on simulation. Modern radar simulation tools normally require the user to manually specify the geometry and characteristics of the scene, including all material properties [3]. While other sensors (e.g. lidar) can be used to scan an environment and produce a mesh or voxel map, they cannot capture radar-specific material properties that are crucial for generating realistic radar scans. Thus, in practice, this results in greatly simplified environment models due to the difficulty of meticulously surveying a scene and generating (or annotating) a model by hand.

We envision a more intelligent, data-driven approach to scene modeling for radar simulation where a user can carry a handheld radar sensor through a static environment and automatically generate a model suitable for accurate simulation of that environment. To this end, we frame the radar simulation problem as one of *novel view synthesis*: using several radar measurements of a scene to simulate what a

[*]Equal Contribution.
[1]Our implementation, data collection platform, and collected datasets can be found via our project site: https://wiselabcmu.github.io/dart/.

radar would see from a new pose. Such a system would not only accelerate the development and testing of new algorithms across a variety of environmental conditions, but also open the door to a myriad of new inference techniques in radar sensing such as localization, mapping, imaging, and recognition which rely on accurate forward rendering models and could greatly benefit from realistic radar models.

Novel View Synthesis Neural Radiance Fields (NeRFs) [48] have revolutionized novel view synthesis, leading to an explosion in interest in graphics and beyond. By leveraging a (neural) *implicit* scene representation instead of explicitly modeling scene geometry, textures, and materials, NeRFs are able to capture and reproduce visual intricacies such as specularity, translucency, reflections, and complex occlusions. This results in a 3D scene capture and rendering system that boasts an unprecedented level of photorealism.

Drawing inspiration from the success of NeRFs, we formulate an analogous problem for mmWave radar imaging. Our method, *Doppler-Aided Radar Tomography* (DART), takes a similar approach by implicitly capturing material properties from input scans which are reproduced when the model is sampled from a novel viewpoint. Though our model is implicit, we can also generate an explicit tomographic image by sampling along a voxel grid, which we use to show that DART is not simply memorizing the input data, but is in fact *learning* the geometry and material properties of the scene (Fig. 1).

Key Challenges Applying NeRF's implicit scene modeling paradigm to the radar domain presents substantial challenges. We derive a rendering model from the ground up that appropriately reflects the unique nature of radar wave propagation. In NeRF, rendering each pixel involves integrating samples along a 1D ray, following a pinhole camera model [48]. However, radar waves propagate *radially* from the antenna. Even after range-azimuth-elevation processing, each radar pixel corresponds to a coarse 2D region of space, as the elevation-azimuth resolution of compact mmWave radars tends to be relatively poor[2]. One key insight is to choose a radar representation space — range-Doppler — which greatly reduces angular ambiguity in one dimension under the assumption that the scene is static and the radar is moving with a known velocity [26]. This presents additional systems challenges, as the sensor platform needs to be moving and its velocity must be measured accurately alongside its position and orientation.

Even with the dimensionality reduction afforded by range-Doppler processing, rendering a single radar pixel involves integrating samples along a *circle*, rather than a ray

[2] For context, these radars have angular resolutions on the order of 15°, orders of magnitude worse than cameras (≈0.01°) [58, 76]

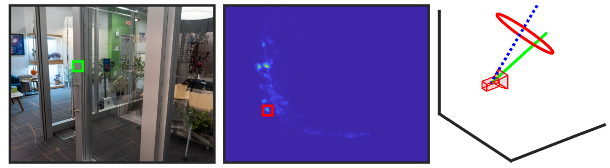

Figure 2. NeRF's pinhole camera model renders a pixel (left) by integrating along a ray (right, green), while DART's range-Doppler model renders a pixel (middle) by integrating along a velocity-dependent (right, blue) circle (right, red).

(Fig. 2). However, appropriately capturing occlusion effects requires that the nearest ranges are processed first due to occlusion caused by objects closer to the radar. Additionally, the size of the integration arc grows as the distance from the radar increases, resulting in an effective decrease in sampling density for points further from the radar that needs to be accounted for. Through careful modeling of these effects and a clever sampling scheme prioritizing sample re-use, we derive a computationally efficient forward rendering function that produces realistic novel radar scans.

Contributions We propose DART: Doppler-Aided Radar Tomography, which implicitly learns a tomographic representation of the world in order to accurately synthesize radar range-Doppler images. To summarize our contributions:

1. We formulate the problem of radar novel view synthesis from implicit reflectance and transmittance maps using range-Doppler images.
2. Using a NeRF-inspired technique, we explicitly formulate the forward rendering of range-Doppler radar images and implicitly invert it using gradient descent to learn a neural-implicit representation.
3. We construct a data collection rig and collect novel radar imaging datasets with accurate position and instantaneous velocity along with reference lidar point clouds.
4. We evaluate DART across a range of scenarios and show that it out-performs the state-of-the-art, quantitatively and qualitatively, in both its synthetic radar renderings and its implicit imaging of scenes.

Limitations Since we rely on Doppler, our method is limited to static scenes, and requires accurate velocity estimates and a constantly moving radar. While motion is intrinsic to our method, we believe that it is reasonable to require movement during scanning. Poor velocity estimates or non-static scenes can cause DART to perform poorly; we hope to relax these limitations in the future.

2. Related Work

2.1. Radar Simulation

Model-Based Approaches *Model-based* methods use a physics and environment models to simulate the propaga-

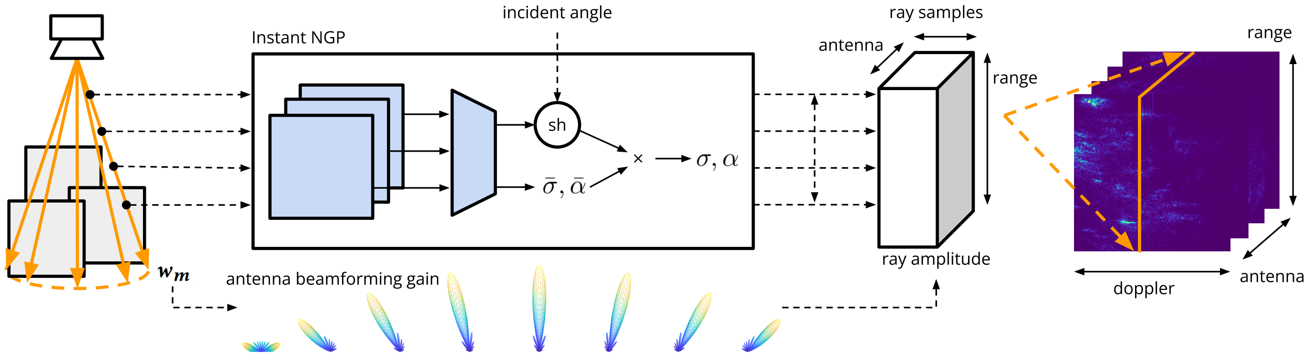

Figure 3. DART tackles radar novel view synthesis by learning a neural implicit map of the world from a trajectory of radar measurements. We make key radar-specific decisions in choosing (1) a high quality radar representation space — **Range-Doppler**, (2) a world model that captures radar interactions — σ **and** α **with spherical harmonics coefficients**, (3) a network architecture to model our desired representation — **Instant NGP**, and (4) an optimized radar rendering and training methodology — **Range-Doppler specific rendering**.

tion of radar signals through the environment using some combination of ray tracing [3, 11, 25, 66, 67], finite element modeling (FEM) [13, 45], or finite-difference time domain (FDTD) simulation [16, 19, 75]. While simulators can replicate complex scene dynamics (e.g. occlusion, path loss, multipath, non-Lambertian effects), they make no attempt to infer the environment structure from sensor data, and their accuracy is limited by the user's ability to create a radar-realistic model of the environment.

Data-Driven Approaches *Data-driven* methods use real sensor scans to build an environment model. *Sparse* methods use constant false alarm rate detection (CFAR) to detect discrete reflectors in the environment [15, 49, 63]. On the other hand, *dense* methods divide the environment into an explicit voxel grid and infer the radar properties of each cell.

Dense methods can be further divided into coherent and incoherent aggregation. If a fixed (e.g. linear or circular) trajectory or sub-wavelength-accurate pose estimates are available, Synthetic Aperture Radar (SAR) can be used [46, 50, 52, 56, 81, 82]; however, this is impractical for a mobile platform over large areas. Instead, sensor readings (with high angular resolution via many antennas or SAR on smaller pieces of a trajectory) can also be aggregated in an *incoherent* manner, which has been referred to as multiview 3D reconstruction [33–35] and radargrammetry [12].

2.2. Machine Learning Methods in Radar

Many classical radar problems such as radar super-resolution [10, 17, 20, 21, 23, 53, 54, 72], odometry [2, 43], mapping [42], activity recognition [39, 70, 77, 80], and object classification [32, 69, 85] have been applied to cheaper, lighter, and more compact radar systems using machine learning. We now seek to solve the novel view synthesis problem from compact, low resolution radars while implicitly creating a higher resolution map.

2.3. Neural Radiance Fields

Instead of defining an *explicit* inverse imaging algorithm that recovers a representation of the scene from sensor readings, Neural Radiance Fields [48] *implicitly* invert a forward rendering function through stochastic gradient descent. This requires the following components:

1. **World model**: NeRF defines the world as an RGB color and transparency for each position and viewing angle; subsequent works have generalized this to handle anti-aliasing [5], different cameras, and lighting [47, 73].
2. **World representation**: Beyond neural networks [48] or voxel grids [40], more recent works have explored spatial hash tables [51] as well as function decompositions for view angle dependence [18, 83].
3. **Rendering function and Model Inversion**: NeRFs model each pixel as a ray and ray-trace the radiance field. The invertibility of this rendering function is crucial: by assuming that each pixel is a ray, the NeRF is "supervised" by one RGB image pixel per ray, allowing NeRF to "solve" for the few opaque points along the ray.

We innovate on these key enablers of NeRFs in order to apply this approach to mmWave radars. By applying NeRF techniques to radar, we hope to leverage the extensive body of neural radiance field literature, while also unlocking the potential of neural-implicit representations.

Beyond Visual Fields The success of NeRFs has inspired numerous other efforts to apply the same general principle to other sensors, including spatial audio [44], imaging sonar [55, 59], LIDAR simulations [27], and RSSI (Received Signal Strength Indicator) mapping [84]. NeRFs have also been applied to radar [29, 71] for camera-like high-resolution Synthetic Aperture Radar instead of the compact and inexpensive radars we explore in this paper.

3. DART: Doppler-Aided Radar Tomography

While our overall approach is inspired by Neural Radiance Fields, the physics of radar presents several new challenges. We make the following key design decisions (Fig. 3):

1. We first choose a radar measurement representation space — range-Doppler — that overcomes the poor spatial resolution of compact radars (Sec. 3.1, 3.2).
2. We then choose a model to account for radar-specific effects of electromagnetic wave interaction which are important for realistic view synthesis such as specularity, ghost reflections and partial occlusions (Sec. 3.3).
3. Finally, to effectively train and learn neural implicit maps for radars, we choose a network architecture for an *adaptive grid* world representation, design a range-Doppler *rendering* method, and propose key rendering optimizations (Sec. 3.3 — 3.4).

3.1. Range-Doppler Representation

Unlike cameras, radars are active sensors which illuminate a scene by transmitting a radio frequency waveform. Upon processing reflections received from objects in the scene, radars can perceive the world in 3 dimensions — range, azimuth, and elevation — as a heatmap indicating the reflectivity of objects at that 3D coordinate [60, 61].

However, while bulky mechanical radars or large solid-state radar arrays can provide azimuth and elevation resolution close to typical cameras, modern inexpensive and compact solid-state radars feature small antenna arrays which make them far inferior in the azimuth and elevation axes [28]. As a result, these compact radars can only generate coarse heatmaps ($>15°$ resolution) in the azimuth and elevation axes, causing each range-azimuth-elevation bin to point to a coarse region of 3D space which is far less sharp than a ray from a camera pixel [38, 41, 76].

To achieve better angular resolution, radars can instead leverage the Doppler effect: objects moving at different relative velocities to the radar have different Doppler velocities, which can be measured by examining the residual phase of the range-azimuth-elevation heatmap [79]. Crucially, in a static scene, these relative velocities depend on not just the relative speed between the radar and the world, but also the relative azimuth and elevation angle between objects and the radar, with each Doppler corresponding to a cone in space [60]. Because of the fine range and Doppler resolutions, Doppler greatly reduces the ambiguity of each bin in 3D space down to a thin ring (Fig. 4), which we further reduce by making a *thinness* argument across the range and Doppler axis in order to simplify integration down to a circle for radar rendering (Sec. 3.4).

3.2. Radar Pre-Processing

mmWave radars use a waveform called Frequency Modulated Continuous Wave (FMCW), and measure a continuous time signal; we then convert these signals into range-Doppler-antenna heatmaps. To summarize key points of our radar processing pipeline (Appendix A.1):

- **Undesirable Range-Doppler Side Lobes**: A single reflective object can create sidelobes that bleed into several range-Doppler bins and mask off weaker objects [61, 86]. Rather than forcing DART to model this, we use a Hann weighting window along both range and Doppler axis to mitigate this effect (Appendix A.1).
- **Multiple Antennas**: We perform range-Doppler processing on each of the eight transmit-receive (TX/RX) pairs in our radar. During our rendering process (Sec. 3.4), we apply the antenna gain and array factor for each TX/RX pair (Fig. 3), emphasizing 8 sections of the field of view. While our sense of high-quality azimuth-elevation information still stems from leveraging Doppler, this provides some coarse directional information.

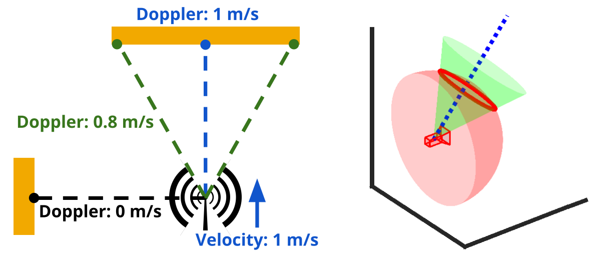

Figure 4. Doppler arises due to differences in relative velocities between points with different relative angles to the radar (left). Each range value (red) corresponds to a sphere, while each Doppler value corresponds to a cone (green). The intersection forms the range-Doppler pixel (see Fig. 2).

3.3. DART's World Model

If we had an accurate model of the world and the electromagnetic wave interaction for all objects in the world, we could just apply this model to the region defined by each range-Doppler pixel to calculate its value. However, due to the complex nature of real-world scenes and interactions, both tasks are highly difficult and typically impractical. Instead, we model these properties in a data-driven way, representing the reflectance and transmittance using a view-dependent neural network-based approach.

Modeling RF Reflectivity Modeling mmWave material interactions is one of the most challenging factors of radar view synthesis. From the perspective of radar, points in space have two key properties: reflectance (the proportion of energy that reflects back), and transmittance (the proportion of energy that continues past) [60]. However, millimeter waves also interact with objects differently depending on incidence angles [4]; for example, metal surfaces can be specular and may be invisible from certain view points. As such, we model each physical point with a reflectance and transmittance value, each of which depend on the incident angle. We formalize this as

$$\sigma : \mathbb{R}^6 \longrightarrow \mathbb{R}, \qquad \alpha : \mathbb{R}^6 \longrightarrow [0,1], \tag{1}$$

which model the reflectance σ and transmittance α as a function of the position ($\mathbb{R}^3$) and incident angle ($\mathbb{R}^3$) of an incoming radar wave, and allows DART to model a wide range of radar phenomena such as partial occlusions, specularity, and ghost reflections (Appendix A.2).

World Representation While voxel-based approaches are highly effective for learning visual radiance fields [18, 83], radar images have a much poorer elevation and azimuth resolution compared to cameras even after exploiting the Doppler axis. This magnifies the difference in spatial resolution that σ and α can be resolved for between close and far ranges. Moreover, unlike cameras, our angular resolution is variable at all scales — be it at a trajectory level, frame-to-frame level or even within a frame (Sec. 3.1). Similar to NeRFs [48], we turn to neural implicit representations as a means of creating an "adaptive" grid, and base our model on the Instant Neural Graphics Primitive[3] [51].

Unlike most visual NeRFs, we do not provide the incident angle as an input to the neural network [74]. Instead, our architecture (visualized in the center block of Fig. 3) outputs a "base" reflectance $\bar{\sigma}$ and transmittance $\bar{\alpha}$, as well as shared spherical harmonics coefficients [83] which are applied to the incident angle as an inner product. In addition to computational advantages, this allows us to directly interpret $(\bar{\sigma}, \bar{\alpha})$ as spherical integrals of our learned reflectance and transmittance functions (Appendix A.3).

We also find that the output activation function on σ and α is critical for numerical stability and performance. Since σ is unbounded[4], we apply a linear activation to σ. Then, to constrain $\alpha \in [0,1]$, we apply the activation function

$$f(\alpha) = \exp(\max(0, \alpha)), \tag{2}$$

which we pair with a custom gradient estimator to handle initialization instability (Appendix A.4).

3.4. Radar Rendering and Model Training

We train σ and α using a differentiable mapping which generates a multi-antenna range-Doppler heatmap from a given (σ, α) network; we refer to this as *radar rendering*. Unlike visual NeRFs, DART must account for a range of physical effects in addition to occlusion including path attenuation, antenna gain patterns, and the radar-specific Doppler axis.

[3][51] implicitly creates an adaptive world grid by using many spatial hash tables with geometrically increasing resolutions, and resolves the output with a small neural network; we use the same general architecture.

[4]σ can be negative; however, since the observed range-Doppler-antenna heatmaps cannot be negative, $\sigma < 0$ will always increase both train and validation loss, so allowing this does not cause overfitting.

Ray Tracing Consider a single "ray" emitted from a radar at position $\boldsymbol{x}$ and orientation (rotation matrix) $\boldsymbol{A}$ at an incidence angle $\boldsymbol{w}$. As the ray travels through space up to the maximum range of the processed (range, Doppler, antenna) image, each point $\boldsymbol{x} + r_i\boldsymbol{w}$ at range r_i receives a signal of amplitude u_i, which is attenuated by a factor of r_i due to free space path loss. Each point then reflects a signal of amplitude $u_i\sigma(\boldsymbol{t}_i)$ back towards the radar, and propagates an amplitude of $u_i\alpha(\boldsymbol{t}_i)$ onwards. As reflected signals return to the radar, the signal loses an additional attenuation factor of r_i, while also suffering from occlusion from $\forall j < i : \alpha(\boldsymbol{t}_j)$.

Sampling $r_1 \ldots r_{N_r}$ discretely along the range bins of the processed heatmap across antennas, the radar return amplitude $C(i, k, \boldsymbol{w})$ for ray $\boldsymbol{w}$ at range bin i and antenna k is

$$C(i,k,\boldsymbol{w}) = g_k(\boldsymbol{A}^{-1}\boldsymbol{w})\frac{\sigma(\boldsymbol{x}+r_i\boldsymbol{w})}{r_i^2}\prod_{i'=1}^{i-1}\alpha(\boldsymbol{t}_{i'})^2, \tag{3}$$

where $g_k(\boldsymbol{A}^{-1}\boldsymbol{w})$ is the antenna beamforming gain antenna k at angle $\boldsymbol{w}$ (specified relative to the radar orientation $\boldsymbol{A}$).

Doppler Integration For a given pose with radar position $\boldsymbol{x}$, velocity $\boldsymbol{v}$, and orientation $\boldsymbol{A}$, we evaluate the return $\boldsymbol{Y}(r_i, d_j, k) \in \mathbb{R}$ at each range-Doppler-antenna bin (r_i, d_j, k), synthesizing a view-specific, multi-antenna range-Doppler heatmap. Since the doppler velocity is measured as $d_j = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$, we integrate the return C along the thin ring corresponding to each bin as:

$$\boldsymbol{Y}(r_i, d_j, k) \propto \frac{r_i}{||\boldsymbol{v}||_2}\int_{\langle \boldsymbol{w},\boldsymbol{v}\rangle = d_j,\ ||\boldsymbol{w}||_2=1} C(i,k,\boldsymbol{w})\, d\boldsymbol{w} \tag{4}$$

Note that we need to correct for the varying width of the discrete bins as a function of range and radar speed by dividing by the speed $||\boldsymbol{v}||_2$ and multiply by r_i (Appendix A.5).

Approximating this integral as a sum over M random directions $\boldsymbol{w}_1 \ldots \boldsymbol{w}_M$ such that $\langle \boldsymbol{w}, \boldsymbol{v} \rangle = d_j$, we multiply by an additional factor of r_i to correct for the circumference of the range-Doppler intersection as r_i increases. This yields

$$\boldsymbol{Y}(r_i, d_j, k) \propto \frac{r_i^2}{M||\boldsymbol{v}||_2}\sum_{m=1}^{M} C(i,k,\boldsymbol{w_m}). \tag{5}$$

Optimized Rendering As the (σ, α) field function must be evaluated for every sample, efficient sampling is critical to computational efficiency. Thus, a naive approach of treating each (range, Doppler, antenna) "pixel" as an independent sample as is standard practice in NeRFs would be computationally prohibitive, requiring the field to be sampled (range, Doppler, antenna, range integration, Doppler integration) times to render a single image. As such, we aggressively reuse samples of σ and α by rendering all bins with the same Doppler simultaneously (Appendix A.6).

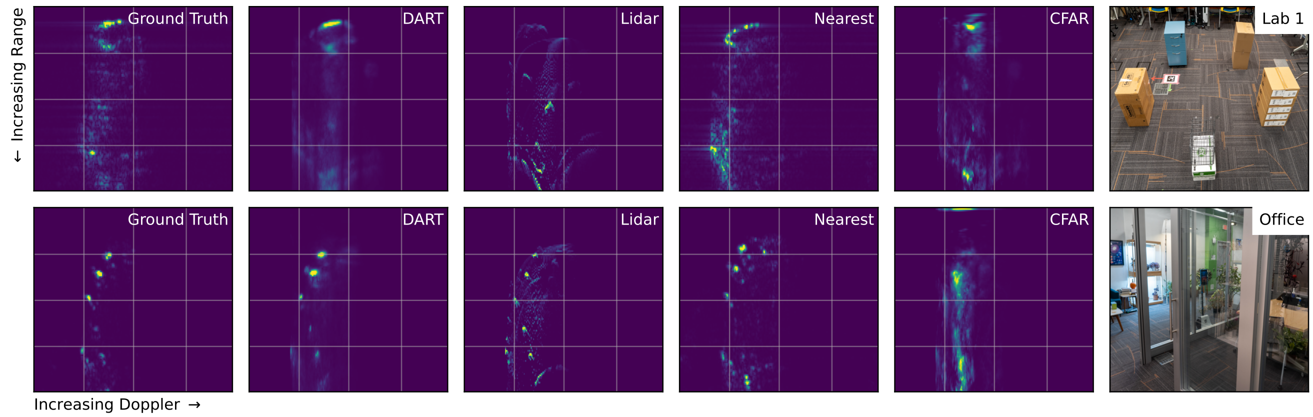

Figure 5. Example (validation) range-Doppler frames and descriptive photos of our method and baselines. DART accurately reproduces the overall radar image, though it lacks the resolution to resolve smaller weak reflectors. *Lidar* can model weak reflectors, but cannot accurately scale them due to a lack of radar-specific information, while *Nearest* produces radar-realistic but inaccurate images since exhaustively measuring all possible poses is impractical. Finally, *CFAR* cannot model transmittance or measure the "volume" of a point.

Training We train our (σ, α) field function using stochastic gradient descent with the Adam [31] optimizer and a $l1$ (i.e. mean-absolute-error) loss. For details about our training process and other hyperparameters, see Appendix A.7.

4. Experiments

We constructed a handheld data collection rig with a mmWave radar and a lidar used for localization[5] (Fig. 6; Appendix B.1). We used this to collect 12 traces ranging from 5 to 15 minutes long in a diverse set of environments including a lab space, townhouse, high-rise apartment, and an early 20th century house (Appendix B.2).

4.1. Baselines

We implement three baselines for radar novel view synthesis and mapping, a *model-based* approach and two *data-driven* approaches (see Sec. 2.1).

- **Lidar Scan-Based Simulator:** We use lidar scans to create an occupancy grid, which we then use in a raytracing radar simulator (assuming occupied grids have a fixed constant reflectance and no transmittance, similar to [3] without material annotations). This baseline represents the standard practice in radar simulation [3, 13, 66, 75].
- **Nearest Neighbor:** We implement a naive nearest-neighbor baseline which finds the training point with the closest (position, velocity) to the novel viewpoint. While simple, this has the advantage of using radar data to "simulate" images compared to our lidar-based simulator [7].
- **CFAR Point Cloud Aggregation:** CFAR is a commonly used adaptive algorithm in radar systems to detect target returns against a background of noise, clutter and interference [49]. We use the Matlab Phased Array System Toolbox [1] to detect radar-reflective targets, de-project those targets into 3D points using Bartlett direction-of-arrival estimation [6], then reproject the points according to the novel pose. This approach is similar to our lidar baseline in that it uses point cloud aggregation, but is better able to capture radar-specific scene properties.

[5]Note that while we use lidar to obtain pose estimates using Cartographer [24], any accurate 3D SLAM system can be used.

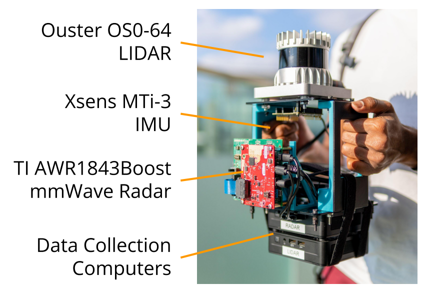

Figure 6. Handheld data collection system; see Appendix B.1.

For additional details on our baselines, see Appendix B.3.

4.2. Metrics

We apply our model to a holdout test set consisting of the last 20% of each trace. We then compute the SSIM [78] of the test images and the effective sample size-corrected standard error (SE) for the mean SSIM; see Appendix B.4. We also compute the SSIM values of 25/30/35dB-equivalent Gaussian noise to help quantify our SSIM values.

5. Results

DART synthesizes significantly more accurate radar images than each baseline on all traces collected, while using minimal training. We also demonstrate DART's ability to sample tomographic images from its implicit map which are more dense than CFAR point clouds, and more faithfully reproduce radar characteristics than lidar scans.

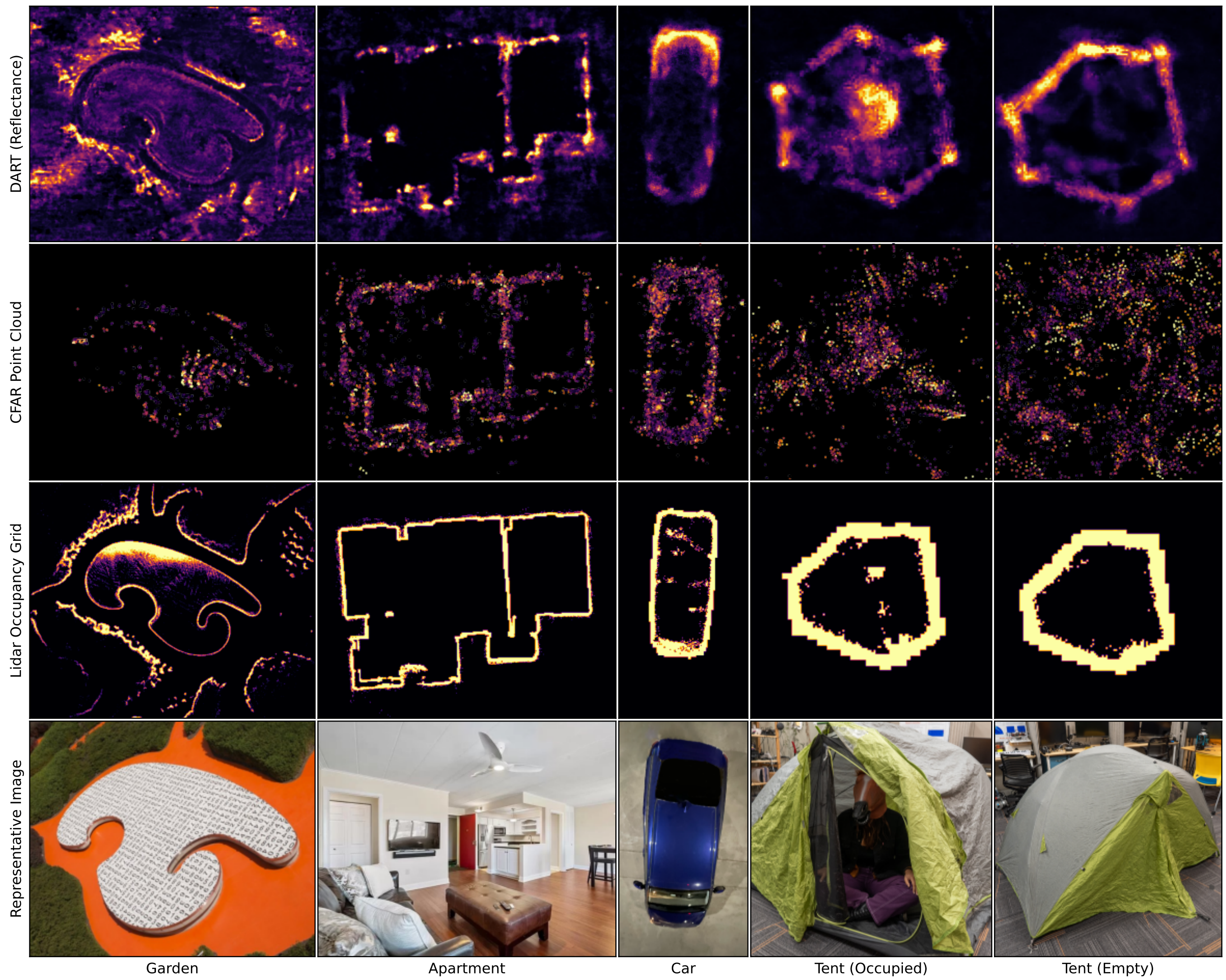

Figure 7. Comparison of DART (top) with CFAR (middle) and a Lidar occupancy grid for reference. While CFAR struggles with cluttered scenes and creates point clouds which are both noisy and sparse, DART creates relatively clear maps which capture radar-specific properties on both ourdoors (Garden) and indoors (Apartment) environments. DART can also image objects with relative clarity (Car), including resolving objects partially occluded by radar-transparent surfaces (Tent — Occupied/Empty).

Training Time We train DART for 3 epochs on each dataset using a RTX 4090 GPU, taking between 1-2$\times$ the data collection time ($\approx$ 10 minutes) of each dataset[6]; this indicates the potential of real-time training with future algorithmic and computing hardware improvements.

Ablations Each part of DART's design significantly improves its accuracy, including view dependence using spherical harmonics and our dynamic grid representation (Tab. 1). For additional ablations, see Appendix C.1.

[6]Training time is not directly proportional to the dataset length: since Doppler bins are not observed when the radar speed is less than the Doppler velocity of that bin, we omit these bins, decreasing the training time. See Appendix B.2 for the length and training time of each dataset.

5.1. Comparison with Baselines

DART synthesizes far more accurate radar images than each baseline on all traces in our dataset (Appendix C.2), with the Lidar-based simulator and Nearest Neighbor baselines performing the worst, and CFAR-based simulation in between. DART is also significantly better than each baseline when evaluated as a whole (Tab. 1).

To understand the performance differences between DART and each baseline, we selected two example range-Doppler images from our dataset (Fig. 5):

- *Lidar-based simulation* (Lidar) can accurately identify reflector positions, but cannot correctly scale their radar return due to the lack of radar-specific material properties.
- *Nearest-Neighbor* (Nearest) approaches can, by defini-

Method	Mean SSIM	SSIM Improvement
DART	0.636 ± 0.012	—
Lidar	0.463 ± 0.005	0.174 ± 0.013
Nearest	0.468 ± 0.006	0.168 ± 0.012
CFAR	0.545 ± 0.007	0.091 ± 0.006
No View Dep.	0.614 ± 0.015	0.022 ± 0.005
20cm Grid	0.591 ± 0.015	0.046 ± 0.004

Table 1. Mean SSIM and SSIM improvement of DART over each baseline (and select ablations) across our dataset along with 95% confidence intervals; see Appendix C.2 for a breakdown by dataset.

tion, generate radar-realistic images. However, measuring all possible (position, orientation, velocity) poses is impractical, leading to "misplaced" images which do not vary continuously over different poses.

- *Constant False Alarm Rate* (CFAR) is commonly used to generate point clouds from radar images. Compared to lidar point clouds, CFAR point clouds are sparse and low-resolution, but capture radar specific properties not measured in lidar. However, CFAR cannot provide any notion of the *size* of each point or its transmittance, which requires the point or grid size to be manually tuned, leading to either excessively sparse or blurry images.

DART therefore achieves its efficacy by using a domain-appropriate sensor and carefully selecting a representation which allows it to use all available sensor information.

5.2. Tomography and Mapping

While DART is not designed primarily as an *explicit* tomography or mapping tool, we can sample the implicit representation[7] to create a (σ, α) reflectance and transmittance grid. This also allows us to verify that DART truly learns the mmWave properties of a scene (and does not simply memorize and interpolate the training data).

Material Properties Example We created an evaluation scene with 5 different boxes. DART is able to learn the unique reflectance and transmittance properties of each materials, which we visualize through tomographic reflectance and transmittance maps (Fig. 8). For additional examples from our datasets, see Appendix C.3.

Comparison with Baselines In addition to creating more accurate radar simulations, DART can also produce more accurate and dense maps than CFAR. Fig. 7 shows several examples comparing tomographic maps of reflectance learned by DART with corresponding slices of the point cloud generated by CFAR. While not as sharp as lidar scans, DART produces reasonably clear maps which capture the radar-specific properties of each scene.

[7] To address view dependence, we analytically take the spherical integral of σ and α at each point; see Appendix A.3 for details.

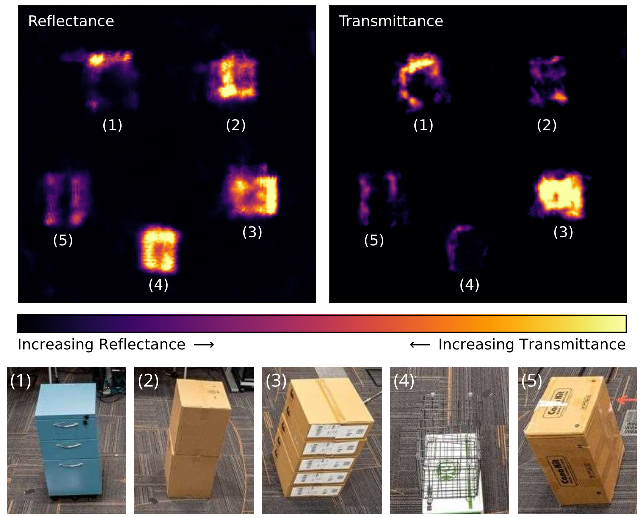

Figure 8. Tomographic images of 5 boxes made from different materials: (1) a metal filing cabinet which appears less reflective (due to specularity), but blocks radar waves; (2) an empty box which reflects radar waves but does not block them; (3) a stack of boxes containing electronics equipment which both reflect and block radar waves; (4) a highly reflective metal mesh with large holes that allow radar to penetrate it; and (5) a different empty box which neither reflects nor blocks radar waves.

6. Conclusion

We present DART: Doppler Aided Radar Tomography, a NeRF-inspired radar novel view synthesis algorithm which learns an implicit tomographic map from range-Doppler images, and demonstrate its effectiveness against state-of-the-art baselines. We derive a physics-based rendering model for radar from first principles, and construct an end-to-end system for learning an implicit scene representation and generate realistic novel radar views. While DART provides a strong baseline for future work, many opportunities remain to apply lessons learned from visual NeRFs; given the rapid pace of innovation in NeRF, these opportunities will likely multiply in the coming years. We also currently make a number of assumptions – such as a static scene and the availability of accurate ground-truth pose – which could be relaxed as has been done with visual NeRFs, enabling a single-chip radar solution for localization, mapping, and imaging. Finally, as we add mmWave radar to the repertoire of NeRF-enabled sensing technologies, this furthers the potential for multimodal implicit mapping in the future.

References

[1] Constant false alarm rate (cfar) detection. Url: https://www.mathworks.com/help/phased/ug/constant-

false-alarm-rate-cfar-detection.html; Accessed: 2023-11-15. 6, 8

[2] Yasin Almalioglu, Mehmet Turan, Chris Xiaoxuan Lu, Niki Trigoni, and Andrew Markham. Milli-rio: Ego-motion estimation with low-cost millimetre-wave radar. *IEEE Sensors Journal*, 21(3):3314–3323, 2020. 1, 3

[3] Stefan Auer, Richard Bamler, and Peter Reinartz. Raysar - 3d sar simulator: Now open source. 2016. 1, 3, 6

[4] Kshitiz Bansal, Keshav Rungta, Siyuan Zhu, and Dinesh Bharadia. Pointillism: Accurate 3d bounding box estimation with multi-radars. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, pages 340–353, 2020. 4

[5] Jonathan T Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P Srinivasan. 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, 2021. 3

[6] M. S. Bartlett. Periodogram analysis and continuous spectra. *Biometrika*, 37(1/2):1–16, 1950. 6

[7] Nitin Bhatia et al. Survey of nearest neighbor techniques. *arXiv preprint arXiv:1007.0085*, 2010. 6

[8] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018. 5

[9] Sarah H Cen and Paul Newman. Precise ego-motion estimation with millimeter-wave radar under diverse and challenging conditions. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6045–6052. IEEE, 2018. 1

[10] Yuwei Cheng, Jingran Su, Mengxin Jiang, and Yimin Liu. A novel radar point cloud generation method for robot environment perception. *IEEE Transactions on Robotics*, 2022. 3

[11] C. J. Coleman. A ray tracing formulation and its application to some problems in over-the-horizon radar. *Radio Science*, 33(4):1187–1197, 1998. 3

[12] Michele Crosetto and F Pérez Aragues. Radargrammetry and sar interferometry for dem generation: validation and data fusion. In *SAR workshop: CEOS committee on earth observation satellites*, page 367, 2000. 3

[13] Qingyun di and Miaoyue Wang. Migration of ground-penetrating radar data method with a finite-element and dispersion. *Geophysics*, 69, 2004. 3, 6

[14] Fangqiang Ding, Andras Palffy, Dariu M. Gavrila, and Chris Xiaoxuan Lu. Hidden gems: 4d radar scene flow learning using cross-modal supervision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9340–9349, 2023. 1

[15] Christopher Doer and Gert F. Trommer. Yaw aided radar inertial odometry uisng manhattan world assumptions. In *2021 28th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS)*, pages 1–10, 2021. 3

[16] Jacques R. Ernst, Hansruedi Maurer, Alan G. Green, and Klaus Holliger. Full-waveform inversion of crosshole radar data based on 2-d finite-difference time-domain solutions of maxwell's equations. *IEEE Transactions on Geoscience and Remote Sensing*, 45(9):2807–2828, 2007. 3

[17] Shiwei Fang and Shahriar Nirjon. Superrf: Enhanced 3d rf representation using stationary low-cost mmwave radar. In *International Conference on Embedded Wireless Systems and Networks (EWSN)...*, page 120. NIH Public Access, 2020. 3

[18] Sara Fridovich-Keil, Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. Plenoxels: Radiance fields without neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5501–5510, 2022. 3, 5, 2

[19] C.M. Furse, S.P. Mathur, and O.P. Gandhi. Improvements to the finite-difference time-domain method for calculating the radar cross section of a perfectly conducting target. *IEEE Transactions on Microwave Theory and Techniques*, 38(7): 919–927, 1990. 3

[20] Xiangyu Gao, Sumit Roy, and Guanbin Xing. Mimo-sar: A hierarchical high-resolution imaging algorithm for mmwave fmcw radar in autonomous driving. *IEEE Transactions on Vehicular Technology*, 70(8):7322–7334, 2021. 3

[21] Andrew Geiss and Joseph C Hardin. Radar super resolution using a deep convolutional neural network. *Journal of Atmospheric and Oceanic Technology*, 37(12):2197–2207, 2020. 3

[22] Junfeng Guan, Sohrab Madani, Suraj Jog, Saurabh Gupta, and Haitham Hassanieh. Through fog high-resolution imaging using millimeter wave radar. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. 1

[23] Junfeng Guan, Sohrab Madani, Suraj Jog, Saurabh Gupta, and Haitham Hassanieh. Through fog high-resolution imaging using millimeter wave radar. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11464–11473, 2020. 3

[24] Wolfgang Hess, Damon Kohler, Holger Rapp, and Daniel Andor. Real-time loop closure in 2d lidar slam. In *2016 IEEE international conference on robotics and automation (ICRA)*, pages 1271–1278. IEEE, 2016. 6, 8

[25] Nils Hirsenkorn, Paul Subkowski, Timo Hanke, Alexander Schaermann, Andreas Rauch, Ralph Rasshofer, and Erwin Biebl. A ray launching approach for modeling an fmcw radar system. In *2017 18th International Radar Symposium (IRS)*, pages 1–10, 2017. 3

[26] Franz Hlawatsch. *Time-frequency analysis and synthesis of linear signal spaces: time-frequency filters, signal detection and estimation, and Range-Doppler estimation*. Springer Science & Business Media, 1998. 2

[27] Shengyu Huang, Zan Gojcic, Zian Wang, Francis Williams, Yoni Kasten, Sanja Fidler, Konrad Schindler, and Or Litany. Neural lidar fields for novel view synthesis. *arXiv preprint arXiv:2305.01643*, 2023. 3

[28] Cesar Iovescu and Sandeep Rao. The fundamentals of millimeter wave sensors. *Texas Instruments*, pages 1–8, 2017. 4, 1

[29] JR Jamora, Dylan Green, Ander Talley, and Thomas Curry. Utilizing sar imagery in three-dimensional neural radiance fields-based applications. In *Algorithms for Synthetic Aperture Radar Imagery XXX*, page 1252002. SPIE, 2023. 3

[30] Mahfuza Khatun, Hani Mehrpouyan, David Matolak, and Ismail Guvenc. Millimeter wave systems for airports and short-range aviation communications: A survey of the current channel models at mmwave frequencies. In *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, pages 1–8, 2017. 1

[31] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 6

[32] Atsutake Kosuge, Satoshi Suehiro, Mototsugu Hamada, and Tadahiro Kuroda. mmwave-yolo: A mmwave imaging radar-based real-time multiclass object recognition system for adas applications. *IEEE Transactions on Instrumentation and Measurement*, 71:1–10, 2022. 3

[33] Jaime Laviada, Ana Arboleya-Arboleya, Yuri Álvarez, Borja González-Valdés, and Fernando Las-Heras. Multiview three-dimensional reconstruction by millimetre-wave portable camera. *Scientific reports*, 7(1):6479, 2017. 3

[34] Jaime Laviada, Ana Arboleya-Arboleya, and Fernando Las-Heras. Multistatic millimeter-wave imaging by multiview portable camera. *IEEE Access*, 5:19259–19268, 2017.

[35] Jaime Laviada, Miguel Lopez-Portugues, Ana Arboleya-Arboleya, and Fernando Las-Heras. Multiview mm-wave imaging with augmented depth camera information. *IEEE Access*, 6:16869–16877, 2018. 3

[36] Peizhao Li, Pu Wang, Karl Berntorp, and Hongfu Liu. Exploiting temporal relations on radar perception for autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 17071–17080, 2022. 1

[37] Yu-Jhe Li, Jinhyung Park, Matthew O'Toole, and Kris Kitani. Modality-agnostic learning for radar-lidar fusion in vehicle detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 918–927, 2022. 1

[38] Yu-Jhe Li, Shawn Hunt, Jinhyung Park, Matthew O'Toole, and Kris Kitani. Azimuth super-resolution for fmcw radar in autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 17504–17513, 2023. 4

[39] Jaime Lien, Nicholas Gillian, M Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. Soli: Ubiquitous gesture sensing with millimeter wave radar. *ACM Transactions on Graphics (TOG)*, 35(4):1–19, 2016. 3

[40] Lingjie Liu, Jiatao Gu, Kyaw Zaw Lin, Tat-Seng Chua, and Christian Theobalt. Neural sparse voxel fields. *Advances in Neural Information Processing Systems*, 33:15651–15663, 2020. 3

[41] Yunfei Long, Daniel Morris, Xiaoming Liu, Marcos Castro, Punarjay Chakravarty, and Praveen Narayanan. Radar-camera pixel depth association for depth completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12507–12516, 2021. 4

[42] Chris Xiaoxuan Lu, Stefano Rosa, Peijun Zhao, Bing Wang, Changhao Chen, John A Stankovic, Niki Trigoni, and Andrew Markham. See through smoke: robust indoor mapping with low-cost mmwave radar. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*, pages 14–27, 2020. 3

[43] Chris Xiaoxuan Lu, Muhamad Risqi U Saputra, Peijun Zhao, Yasin Almalioglu, Pedro PB De Gusmao, Changhao Chen, Ke Sun, Niki Trigoni, and Andrew Markham. milliego: single-chip mmwave radar aided egomotion estimation via deep sensor fusion. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, pages 109–122, 2020. 1, 3

[44] Andrew Luo, Yilun Du, Michael Tarr, Josh Tenenbaum, Antonio Torralba, and Chuang Gan. Learning neural acoustic fields. In *Advances in Neural Information Processing Systems*, pages 3165–3177. Curran Associates, Inc., 2022. 3

[45] M. Malinen and P. Råback. *Elmer finite element solver for multiphysics and multiscale problems*. Forschungszentrum Juelich, 2013. 3

[46] Babak Mamandipoor, Greg Malysa, Amin Arbabian, Upamanyu Madhow, and Karam Noujeim. 60 ghz synthetic aperture radar for short-range imaging: Theory and experiments. In *2014 48th Asilomar Conference on Signals, Systems and Computers*, pages 553–558. IEEE, 2014. 3

[47] Ricardo Martin-Brualla, Noha Radwan, Mehdi S. M. Sajjadi, Jonathan T. Barron, Alexey Dosovitskiy, and Daniel Duckworth. NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. In *CVPR*, 2021. 3

[48] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021. 2, 3, 5

[49] G. Minkler and J. Minkler. CFAR: The principles of automatic radar detection in clutter. *NASA STI/Recon Technical Report A*, 90:23371, 1990. 3, 6

[50] Mohammadreza Mostajabi, Ching Ming Wang, Darsh Ranjan, and Gilbert Hsyu. High-resolution radar dataset for semi-supervised learning of dynamic objects. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2020. 3

[51] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics (ToG)*, 41(4):1–15, 2022. 3, 5, 2

[52] Akarsh Prabhakara, Vaibhav Singh, Swarun Kumar, and Anthony Rowe. Osprey: a mmwave approach to tire wear sensing. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*, pages 28–41, 2020. 3

[53] Akarsh Prabhakara, Diana Zhang, Chao Li, Sirajum Munir, Aswin C Sankaranarayanan, Anthony Rowe, and Swarun Kumar. Exploring mmwave radar and camera fusion for high-resolution and long-range depth imaging. In *2022*

IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 3995–4002. IEEE, 2022. 3

[54] Akarsh Prabhakara, Tao Jin, Arnav Das, Gantavya Bhatt, Lilly Kumari, Elahe Soltanaghaei, Jeff Bilmes, Swarun Kumar, and Anthony Rowe. High resolution point clouds from mmwave radar. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023. 3

[55] Mohamad Qadri, Michael Kaess, and Ioannis Gkioulekas. Neural implicit surface reconstruction using imaging sonar. *arXiv preprint arXiv:2209.08221*, 2022. 3

[56] Kun Qian, Zhaoyuan He, and Xinyu Zhang. 3d point cloud generation with millimeter-wave radar. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(4), 2020. 3

[57] Kun Qian, Shilin Zhu, Xinyu Zhang, and Li Erran Li. Robust multimodal vehicle detection in foggy weather using complementary lidar and radar signals. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 444–453, 2021. 1

[58] Julien Rebut, Arthur Ouaknine, Waqas Malik, and Patrick Pérez. Raw high-definition radar for multi-task learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 17021–17030, 2022. 2

[59] Albert W Reed, Juhyeon Kim, Thomas Blanford, Adithya Pediredla, Daniel C Brown, and Suren Jayasuriya. Neural volumetric reconstruction for coherent synthetic aperture sonar. *arXiv preprint arXiv:2306.09909*, 2023. 3

[60] Mark A Richards, Jim Scheer, William A Holm, and William L Melvin. *Principles of Modern Radar*. Citeseer, 2010. 4, 2, 5

[61] Mark A Richards et al. *Fundamentals of radar signal processing*. Mcgraw-hill New York, 2005. 4, 2

[62] Christian P Robert, George Casella, and George Casella. *Monte Carlo statistical methods*. Springer, 1999. 8

[63] Hermann Rohling. Radar cfar thresholding in clutter and multiple target situations. *IEEE Transactions on Aerospace and Electronic Systems*, AES-19(4):608–621, 1983. 3

[64] Mark E Russell, Arthur Crain, Anthony Curran, Richard A Campbell, Clifford A Drubin, and Willian F Miccioli. Millimeter-wave radar sensor for automotive intelligent cruise control (icc). *IEEE Transactions on microwave theory and techniques*, 45(12):2444–2453, 1997. 1

[65] Nicolas Scheiner, Florian Kraus, Fangyin Wei, Buu Phan, Fahim Mannan, Nils Appenrodt, Werner Ritter, Jurgen Dickmann, Klaus Dietmayer, Bernhard Sick, and Felix Heide. Seeing around street corners: Non-line-of-sight detection and tracking in-the-wild using doppler radar. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. 1

[66] Christian Schöffmann, Barnaba Ubezio, Christoph Böhm, Stephan Mühlbacher-Karrer, and Hubert Zangl. Virtual radar: Real-time millimeter-wave radar sensor simulation for perception-driven robotics. *IEEE Robotics and Automation Letters*, 6(3):4704–4711, 2021. 3, 6

[67] Christian Schüßler, Marcel Hoffmann, Johanna Bräunig, Ingrid Ullmann, Randolf Ebelt, and Martin Vossiek. A realistic radar ray tracing simulator for large mimo-arrays in automotive environments. *IEEE Journal of Microwaves*, 1(4):962–974, 2021. 3

[68] David M Sheen, Douglas L McMakin, and Thomas E Hall. Cylindrical millimeter-wave imaging technique for concealed weapon detection. In *26th AIPR Workshop: Exploiting New Image Sources and Sensors*, pages 242–250. SPIE, 1998. 1

[69] Xian Shuai, Yulin Shen, Yi Tang, Shuyao Shi, Luping Ji, and Guoliang Xing. millieye: A lightweight mmwave radar and camera fusion system for robust object detection. In *Proceedings of the International Conference on Internet-of-Things Design and Implementation*, pages 145–157, 2021. 3

[70] Akash Deep Singh, Sandeep Singh Sandha, Luis Garcia, and Mani Srivastava. Radhar: Human activity recognition from point clouds generated through a millimeter-wave radar. In *Proceedings of the 3rd ACM Workshop on Millimeter-wave Networks and Sensing Systems*, pages 51–56, 2019. 3

[71] William Snyder, Stephen DelMarco, Dylan Snover, Amit Bhatia, and Scott Kuzdeba. Extending neural radiance fields (nerf) for synthetic aperture radar (sar) novel image generation. In *Synthetic Data for Artificial Intelligence and Machine Learning: Tools, Techniques, and Applications*, pages 268–276. SPIE, 2023. 3

[72] Yue Sun, Honggang Zhang, Zhuoming Huang, and Benyuan Liu. Deeppoint: A deep learning model for 3d reconstruction in point clouds via mmwave radar. *arXiv preprint arXiv:2109.09188*, 2021. 3

[73] Matthew Tancik, Vincent Casser, Xinchen Yan, Sabeek Pradhan, Ben Mildenhall, Pratul P Srinivasan, Jonathan T Barron, and Henrik Kretzschmar. Block-nerf: Scalable large scene neural view synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8248–8258, 2022. 3

[74] Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Justin Kerr, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, et al. Nerfstudio: A modular framework for neural radiance field development. *arXiv preprint arXiv:2302.04264*, 2023. 5, 2

[75] F.L. Teixeira, Weng Cho Chew, M. Straka, M.L. Oristaglio, and T. Wang. Finite-difference time-domain simulation of ground penetrating radar on dispersive, inhomogeneous, and conductive soils. *IEEE Transactions on Geoscience and Remote Sensing*, 36(6):1928–1937, 1998. 3, 6

[76] *AWR1843AOP Single-chip 77- and 79-GHz FMCW mmWave Sensor Antennas-OnPackage (AOP)*. Texas Instruments, 2021. 2, 4, 5

[77] Yuheng Wang, Haipeng Liu, Kening Cui, Anfu Zhou, Wensheng Li, and Huadong Ma. m-activity: Accurate and real-time human activity recognition via millimeter wave radar. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8298–8302. IEEE, 2021. 3

[78] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 6

[79] Christoph Wasserzier, Josef G. Worms, and Daniel W. O'Hagan. How noise radar technology brings together active sensing and modern electronic warfare techniques in a combined sensor concept. In *2019 Sensor Signal Processing for Defence Conference (SSPD)*, pages 1–5, 2019. 4

[80] Hongfei Xue, Qiming Cao, Yan Ju, Haochen Hu, Haoyu Wang, Aidong Zhang, and Lu Su. M4esh: mmwave-based 3d human mesh construction for multiple subjects. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, pages 391–406, 2022. 3

[81] Hiroyoshi Yamada, Takumi Kobayashi, Yoshio Yamaguchi, and Yuuichi Sugiyama. High-resolution 2d sar imaging by the millimeter-wave automobile radar. In *2017 IEEE Conference on Antenna Measurements & Applications (CAMA)*, pages 149–150. IEEE, 2017. 3

[82] Muhammet Emin Yanik and Murat Torlak. Near-field mimo-sar millimeter-wave imaging with sparsely sampled aperture data. *Ieee Access*, 7:31801–31819, 2019. 3

[83] Alex Yu, Ruilong Li, Matthew Tancik, Hao Li, Ren Ng, and Angjoo Kanazawa. Plenoctrees for real-time rendering of neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5752–5761, 2021. 3, 5, 2

[84] Xiaopeng Zhao, Zhenlin An, Qingrui Pan, and Lei Yang. Nerf2: Neural radio-frequency radiance fields. *arXiv preprint arXiv:2305.06118*, 2023. 3

[85] Zihao Zhao, Yuying Song, Fucheng Cui, Jiang Zhu, Chunyi Song, Zhiwei Xu, and Kai Ding. Point cloud features-based kernel svm for human-vehicle classification in millimeter wave radar. *IEEE Access*, 8:26012–26021, 2020. 3

[86] Yi Zhou, Lulu Liu, Haocheng Zhao, Miguel López-Benítez, Limin Yu, and Yutao Yue. Towards deep radar perception for autonomous driving: Datasets, methods, and challenges. *Sensors*, 22(11), 2022. 4