

Neural Implicit Representations for 3D Synthetic Aperture Radar Imaging

Nithin Sugavanam¹ and Emre Ertin¹

Ohio State University, Columbus, OH 43210, USA {sugavanam.3,ertin}@osu.edu

Abstract. Synthetic aperture radar (SAR) is a tomographic sensor that measures 2D slices out of the 3D Spatial Fourier transform of the scene. In many operational scenarios, the measured set of 2D slices does not fill the 3D space in the Fourier domain, resulting in significant artifacts in the reconstructed imagery. Traditionally, simple priors, such as sparsity in the image domain, are used to regularize the inverse problem. In this paper, we review our recent work that achieves state-of-the-art results in 3D SAR imaging employing neural structures to model the surface scattering that dominates SAR returns. These neural structures encode the surface of the objects in the form of a signed distance function learned from the sparse scattering data. Since estimating a smooth surface from a sparse and noisy point cloud is an ill-posed problem, we regularize the surface estimation by sampling points from the implicit surface representation during the training step. We demonstrate the model’s ability to represent target scattering using measured and simulated data from single vehicles and a larger scene with a large number of vehicles. We conclude with future research directions calling for methods to learn complex-valued neural representations to enable synthesizing new collections from the volumetric neural implicit representation.

Keywords: Synthetic Aperture Radar, Tomographic Imaging, 3D Imaging, Signed Distance Functions, Neural Implicit Representations

1 Introduction

Synthetic aperture radar (SAR) is a tomographic sensor that measures 2D slices out of the 3D Spatial Fourier transform of the scene. Traditional 3D reconstruction techniques involve aggregating, and indexing phase history data in the spatial Fourier domain and applying an inverse 3D Fourier Transform to the data as shown in references [1,2]. Obtaining high-resolution imagery requires the data collected to be densely distributed in both the azimuth and elevation angle, which is often not satisfied in operational scenarios. For instance, the sampling in elevation dimension is sparse and non-uniform in the GOTCHA dataset [3,4]. To cope with the sparsely sampled data in elevation, regularized inversion methods [5] that combine non-uniform fast Fourier transform method to model the forward operator with regularization priors for the scene that promote structured solution such as sparsity [6], limited persistence in the viewing angle domain [7–9], vertical structures [10]. These regularization-based approaches promote dominant

scattering mechanisms [11, 12] and produce sparse point clouds in the spatial domain.

Furthermore, scattering from multiple internal reflections appears as scattering centers at locations away from the surface of the object. Additionally, the non-uniform sampling in the elevation domain leads to ambiguities in the height direction of the image domain. This sampling leads to aliased copies of the object. In this paper, we review our recent work that achieves state-of-the-art results in 3D SAR imaging employing neural structures to model the surface scattering that dominates SAR returns. These neural structures encode the surface of the objects in the form of a signed distance function learned from the sparse scattering data. Since estimating a smooth surface from a sparse and noisy point cloud is an ill-posed problem, we regularize the surface estimation by sampling points from the implicit surface representation during the training step. We demonstrate the model’s ability to represent target scattering using measured and simulated data from single vehicles and more significant scenes of hundreds of objects. We conclude with future research directions calling for methods to learn complex-valued neural representations to synthesize tomographic projections from previously unseen viewpoints/apertures.

2 Related Work

2.1 Classical 3D SAR Imaging

A common approach for 3D imaging is to formulate the inversion of the Fourier operator as an inverse problem imposing regularization to enforce sparsity in the spatial domain and promote correlation of scattering coefficients in the neighboring sub-apertures [13]. Reference [1] solves the recovery problem over individual sub-apertures and non-coherently integrates the result to obtain a wide-angle 3D representation of the object. The backscattered response of the object is also modeled as a superposition of the backscattered response of 3D canonical scattering mechanisms such as dihedral, trihedral, plate, cylinder, and top-hat [14]. The scattering behavior of these canonical reflectors has been derived as a function of the size of the scattering mechanism using the predictions from the Geometric theory of diffraction [15]. References [7, 10, 16] jointly model the sparsity in scattering center locations and the persistence of scattering coefficients in the azimuth domain. Alternatively, the imaging problem has been posed as an interferometric imaging problem using measurements obtained from multiple baselines [4], and [17]. The 3D non-uniform Fourier transform is approximated by a set of 2D non-uniform Fourier transform for each baseline for range and cross-range estimation and a 1D non-uniform Fourier transform for height estimation. The effect of the persistence of the target on the ambiguities in the point-spread function is presented in Reference [18]. The uncertainty in the localization of scattering centers is dictated by the persistence, which leads to ambiguities in the target localizations along the vertical direction projected along the viewing angle of elevation.

2.2 Neural Implicit Representations

Meanwhile, in computer vision, implicit neural representations, like Neural Radiance Fields (NeRF) and its derivatives [19,20], become popular volume rendering methods, boasting strong performance even when dealing with highly intricate objects. This volume rendering approach uses focal-plane camera geometry to sample multiple points along rays and perform composition of the colors of the sampled points to link the 3D model to 2D views. The traditional discrete representations of objects, scene geometry, and appearance using meshes and voxel grids scale poorly with the scene’s size. Recent developments utilize continuous functions parameterized by deep neural structures. These coordinate-based Deep-nets are trained to map the low-dimensional spatial coordinates to output a representation of shape or density for each spatial location. Coordinate-based Deep-Nets have been used to represent images [21], volume density [20], occupancy [22], and signed distance [23]. The signed distance function (SDF) is given by

$$SDF(\mathbf{p}) = \begin{cases} 0, & \mathbf{p} \in \Omega \\ +s, & \mathbf{p} \in \Omega^+ \\ -s, & \mathbf{p} \in \Omega^-, \end{cases} \quad (1)$$

where $s > 0$ and Ω represents the object boundary, Ω^+ represents the region outside the object and Ω^- represents the region inside the object.

Typically a Fourier feature layer is used as the input layer in the coordinate-based Deep-Nets that operate on the spatial coordinates. It has been shown that the Fourier feature mapping can be used to overcome the spectral bias of coordinate-based Deep-Nets towards low frequencies by allowing them to learn much higher frequency details in the geometry. Reference [21] showed that a random Fourier feature mapping with an appropriately chosen scale could dramatically improve the performance of coordinate-based Deep-Nets across many low-dimensional tasks in computer vision.

3 Denoising sparse point clouds using surface priors

We present a method to recover the point-cloud representation of the object from SAR phase history data. The measurements from the SAR are collected over azimuth angles $[\theta_{1,e}, \dots, \theta_{N_P,e}]$ and elevation angles $[\phi_{1,e}, \dots, \phi_{N_P,e}]$, with elevation passes $e = 1, \dots, N_{el}$ and frequency points $[f_1, \dots, f_{N_F}]$. The phase history data are denoted by $\mathbf{Y}_e = [\mathbf{Y}_{1,e} \mathbf{Y}_{2,e} \dots \mathbf{Y}_{N_P,e}]$ where $\mathbf{Y} = [\mathbf{Y}_1; \dots; \mathbf{Y}_{N_{el}}] \in \mathbb{C}^{N_F \times N_P \times N_{el}}$. We model the object as a collection of dominant scattering centers, and each scattering center is isotropic over a sub-aperture. We assume the pulses collected over different viewing angles are grouped in N_s sub-apertures over N_{el}

elevation passes. The backscattered signal in a sub-aperture m is defined as

$$\begin{aligned} \bar{Y}(f_i, \theta_{j,e}, \phi_{j,e}; \bar{\theta}_m, \bar{\phi}_m) &= \sum_{k=1}^K s_k^m \exp\left(-\frac{j4\pi f_i}{c} \right. \\ &\quad \left. (x_k \cos \phi_{j,e} \cos \theta_{j,e} + y_k \sin \theta_{j,e} \cos \phi_{j,e} + z_k \sin \phi_{j,e})\right), \\ &\quad + n(f_i, \theta_j, \phi_k) \end{aligned} \quad (2)$$

where $i = 1, \dots, N_F$, $j = 1, \dots, N_P$, $e = 1, \dots, N_{el}$, $m = 1, \dots, N_S$, $\bar{\theta}_m$ is the mean azimuth angle in sub-aperture m and $\bar{\phi}_m$ is the mean elevation angle in sub-aperture m . The K-sparse scattering centers are located at spatial coordinate $\mathbf{p}_k = [x_k, y_k, z_k]$ with sub-aperture dependent scattering coefficients $s_k^m \in \mathbb{C}$. The region of interest is discretized into $N_x \times N_y \times N_z$ voxels, and the resolution of these grid points is chosen according to the range resolution to get cube-like voxels. The measurement noise and modeling mismatches are combined in the complex-valued noise term $n(f_i, \theta_{j,e}, \phi_{j,e})$. The sub-aperture measurements can be expressed as

$$\mathbf{Y}(\bar{\theta}_m, \bar{\phi}_m) = \mathcal{F}_{3D}^m(\mathbf{S}_m) + \mathbf{N}, \quad (3)$$

where $\mathbf{Y}(\bar{\theta}_m, \bar{\phi}_m) \in \mathbb{C}^{N_F \times N_S \times N_{el}}$ is the sub-aperture phase history data, $\mathcal{F}_{3D}^m(\cdot)$ is the non-uniform Fourier transform [24] that maps the spatial coordinates to the K-space measurement space, and $\mathbf{S}_m \in \mathbb{C}^{N_x \times N_y \times N_z}$ are the complex-valued scattering coefficients for each spatial location in the region of interest.

Given the measurements, the inverse problem of recovering the scattering coefficients is solved using the regularized least squares method. The sparsity-promoting prior is used to enhance the dominant scattering centers and suppress the side lobes generated from the point spread function of the measurement operator. The regularized least squares problem is given by

$$\min_{\mathbf{S}_m} \|\mathbf{S}_m\|_1 \quad \text{Subject to } \|\mathbf{Y}(\bar{\theta}_m) - \mathcal{F}_{3D}^m(\mathbf{S}_m)\|^2 \leq \sigma^2, \quad (4)$$

where σ^2 is the modeling error energy. The recovered scattering centers have a sparse voxel representation for the sub-aperture m . The results from all the sub-apertures are combined non-coherently to get a joint-sparse representation given by

$$\mathbf{S} = \sum_{m=1}^{N_s} |\mathbf{S}_m|. \quad (5)$$

The resulting quantity S consists of the scattering coefficients of the point-cloud representation of the object in a regular voxel space. We denote each point in the voxel space with scattering coefficients above a minimum threshold belonging to the set \mathbf{P} . Specifically, we denote the point-cloud \mathbf{P} representing the object as

$$\mathbf{P} = \{\mathbf{p}_i = [x, y, z] \text{ if } |S([\mathbf{p}_i])| \geq \tau\}. \quad (6)$$

Next, we also find the viewing direction that corresponds to the maximum response of each point across different sub-apertures. We denote this set as \mathbf{V} and each \mathbf{v}_i is given by

$$\begin{aligned} m^* &= \operatorname{argmax}_m |\mathbf{S}_m(\mathbf{p}_i)| \\ \mathbf{v}_i &= [\cos(\bar{\theta}_{m^*}) \cos(\bar{\phi}_{m^*}); \sin(\bar{\theta}_{m^*}) \cos(\bar{\phi}_{m^*}); \sin(\bar{\phi}_{m^*})] \end{aligned} \quad (7)$$

The initial estimate of the normal $\mathbf{n}_i \in \mathcal{N}_P$ for each point $\mathbf{p}_i \in P$ is computed using a local principal component analysis. For each point, we find the nearest neighbors with a distance of at most $0.3m$. Next, we fit a plane using that subset of points, where the normal direction of the plane is assigned as the normal direction for the given point. Suppose the point \mathbf{p}_i has fewer than 3 neighbors in that neighborhood. We select the normal direction based on the sub-aperture with the maximum scattering coefficient at location \mathbf{p}_i . We utilize the set \mathbf{V} and assign $\mathbf{n}_i = \mathbf{v}_i$.

We utilize the signed distance function defined in Eq. (1) to represent the object and learn a coordinate-based MLP to predict the zero-level set of the SDF. Since the SAR point cloud is sparse, we use the concept of iso-points developed in Reference [25]. The iso-points set \mathcal{Q}_{iso} are the points sampled from the zero-level set of the neural network such that

$$\mathcal{Q}_{iso} = \{\mathbf{q} : f(\mathbf{q}, \Psi) = 0\} \quad (8)$$

The sampling process to obtain dense samples on the iso-surface involves the following steps: projection, uniform sampling, and up-sampling. The projection operation is performed on a randomly sampled point \mathbf{q} in the neighborhood of set P to project it back on the surface represented by a zero-level set of the neural network $f(\mathbf{q}; \Psi)$ using Newton’s iterations. Given a point \mathbf{q}_{k-1} , the k -th update step is given by

$$\mathbf{q}_k = \mathbf{q}_{k-1} - \pi \left(\frac{\mathbf{J}^T(\mathbf{q}_{k-1})}{\|\mathbf{J}(\mathbf{q}_{k-1})\|^2} f(\mathbf{q}_{k-1}) \right) \quad (9)$$

such that $\mathbf{J}(\mathbf{q}_i)$ is the Jacobian of the network parameters with respect to the spatial coordinates \mathbf{q}_i .

$$\pi(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|} \min(\|\mathbf{x}\|, \tau_0) \quad (10)$$

$\pi(\mathbf{x})$ is the clipping operation to avoid the possible noisy update step obtained from the non-smooth neural implicit surface representation, and τ_0 is the maximum preset threshold computed based on the bounding box size around the region of interest. Newton’s update steps are repeated until the $f(\mathbf{q}_k) \leq 1e-4$.

This projection operator still produces non-uniform samples from the surface and can lead to regions with no samples. We obtain uniform samples on the surface by moving the points away from the high-density region such that

$$\mathbf{q} = \mathbf{q} - \alpha \sum_{\mathbf{q}_k \in \mathcal{B}_\epsilon(\mathbf{q})} w(\mathbf{q}_i, \mathbf{q}) \frac{\mathbf{q}_i - \mathbf{q}}{\|\mathbf{q}_i - \mathbf{q}\|}, \quad (11)$$

where $\mathcal{B}_\epsilon(\mathbf{q})$ is the set of points in the ball of radius ϵ around the point \mathbf{q} . We additionally define the weighting factor $w(\mathbf{q}_i, \mathbf{q})$ as

$$w(\mathbf{q}_i, \mathbf{q}) = \exp\left(-\frac{\|\mathbf{q}_i - \mathbf{q}\|^2}{\sigma_p^2}\right), \quad (12)$$

such that distant points have reduced influence, where σ_p is the density bandwidth chosen per the scene size. Following the uniform sampling step to move samples away from the high-density region, the resulting set of points is projected on the zero-level set of the SDF using the iterations in Eq. (9).

Finally, a dense point cloud is created for a desired density using a simplified version of the edge-aware resampling (EAR) technique [26]. The points are resampled such that these points are pushed away from the edges to avoid discontinuity in the normal direction while penalizing the formation of clusters. The normal for each point in the iso-point set is computed using the normalized Jacobian given by $\mathbf{n}_i = \frac{J(\mathbf{q}_i)}{\|J(\mathbf{q}_i)\|}$. The estimated normal direction is used to update the iso-point cloud to ensure the points are away from edges by

$$\Delta\mathbf{q}_{edge} = \frac{\sum_{\mathbf{q}_k \in \mathcal{B}_\epsilon(\mathbf{q})} \phi(\mathbf{n}_i, \mathbf{q}_i - \mathbf{q})(\mathbf{q}_i - \mathbf{q})}{\sum_{\mathbf{q}_k \in \mathcal{B}_\epsilon(\mathbf{q})} \phi(\mathbf{n}_i, \mathbf{q}_i - \mathbf{q})}, \quad (13)$$

where $\phi(\mathbf{n}_i, \mathbf{q}_i - \mathbf{q}) = \exp\left(-\frac{\mathbf{n}_i^T(\mathbf{q}_i - \mathbf{q})}{\sigma_p^2}\right)$ is the anisotropic projection weight such that the normal computed for all the neighboring points \mathbf{q}_i is used to prefer points that lie on the same plane, enforce smoothness, and penalize edges. The following term penalizes the grouping of points and utilizes the weight function defined in Eq. (12).

$$\Delta\mathbf{q}_{repulsion} = 0.5 \frac{\sum_{\mathbf{q}_k \in \mathcal{B}_\epsilon(\mathbf{q})} w(\mathbf{q}_i, \mathbf{q})(\mathbf{q}_i - \mathbf{q})}{\sum_{\mathbf{q}_k \in \mathcal{B}_\epsilon(\mathbf{q})} w(\mathbf{q}_i, \mathbf{q})}. \quad (14)$$

Points that are not well-separated are given higher weights such that they distribute uniformly on the surface. The update step for each point is given by

$$\mathbf{q} = \mathbf{q} - \pi(\Delta\mathbf{q}_{repulsion}) - \pi(\Delta\mathbf{q}_{edge}). \quad (15)$$

Upsampling is performed by adding more samples to low-density point regions by adding new points to the set at these regions. A priority score to assess the regions is defined for each point \mathbf{q} as

$$P(\mathbf{q}) = \max_{\mathbf{q}_i \in \mathcal{B}_\epsilon(\mathbf{q})} \|\mathbf{q}_i - \mathbf{q}\| \quad (16)$$

$$\mathbf{q}^* = \operatorname{argmax}_{\mathbf{q}_i \in \mathcal{B}_\epsilon(\mathbf{q})} \|\mathbf{q}_i - \mathbf{q}\|. \quad (17)$$

The \mathbf{q}^* falls in the region of highest priority; therefore, we next find the highest priority neighbor of \mathbf{q}^* by $\mathbf{q}_i^* = \operatorname{argmax}_{\mathbf{q}_i \in \mathcal{B}_\epsilon(\mathbf{q}^*)} \|\mathbf{q}_i - \mathbf{q}^*\|$. A new point is inserted using the following asymmetric rule to avoid duplicate copies

$$\mathbf{q}_{new} = \frac{\mathbf{q}_i^* + 2\mathbf{q}^*}{3}. \quad (18)$$

After the upsampling step, the generated points are again projected back on the SDF using Newton’s iterations in Eq. (9).

3.1 Proposed solution

In this section, we present the network architecture to predict the signed distance function and discuss the loss function to learn the network parameters.

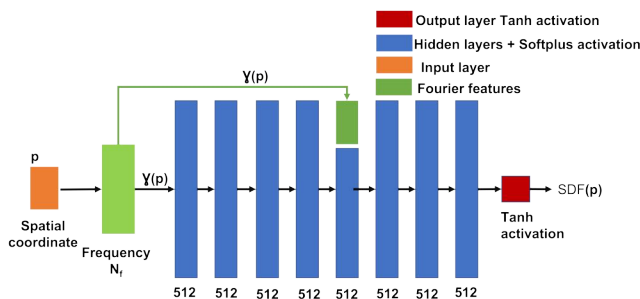


Fig. 1: The network architecture consists of the input layer Fourier layer and 8 linear layers with a softplus activation function and an output layer with a tanh activation function to predict the signed distance function.

Network architecture We utilize the network architecture shown in Figure 1 proposed in [23] and [25] to estimate the signed distance function. The input Fourier feature layers map the spatial coordinates to N_F frequencies. In the results section, we verify the effect of varying N_F with the scene complexity. The hidden layers have a latent-feature size of 512 with a soft-plus activation function. The transformed input layer is also fed in layer 4. The output layer consists of a linear map and the tanh activation function to predict the SDF. Finally, the weights and bias of the network is initialized with standard Gaussian variables.

Loss function The detected point clouds from SAR measurements are sparse and noisy, since these are aliased copies due to non-uniform elevation sampling. These point clouds are usually concentrated near edges and other dominant scattering mechanisms. We aim to reconstruct the 3D object by enforcing the surface prior and denoising the point cloud by estimating the network parameters to learn a 3D representation of the object of interest. The iso-points estimated from the implicit representation serve as a consistent and smooth approximation of the object that is refined during the training process. This regularization ensures the training procedure does not overfit to noise. Since the iso-points are

updated throughout the training process, we hypothesize the network learns the high-frequency signals governed by the underlying geometry of the object. We define three sets of point-clouds $\mathcal{Q}_{iso} = \{\mathbf{q} : f(\mathbf{q}; \Psi) = 0\}$ is the point-cloud consisting of iso-points, \mathcal{P} is the point-cloud consisting of the training set and \mathcal{Q}_b is the point-cloud consisting of the points sampled from the region of interest using a uniform probability distribution. Since the iso-points are uniformly distributed on the surface, these points can be used to augment the supervision in under-sampled areas by enforcing SDF to be 0 on the iso-points set. The 0 condition on the SDF is implemented through a sparsity promoting ℓ_1 regularization in Eqs. (19) and (20). The SDF for non-surface points is enforced to be non-zero by utilizing the exponential term from Eq. (21). The Eikonal regularizer in Eq. (22) enforces the output of the network to mirror the properties of a SDF.

$$L_{isoSDF} = \frac{1}{|\mathcal{Q}_{iso}|} \sum_{\mathbf{q} \in \mathcal{Q}_{iso}} |f(\mathbf{q})| \quad (19)$$

$$L_{on} = \frac{1}{|\mathcal{P}|} \sum_{\mathbf{p} \in \mathcal{P}} |f(\mathbf{p})| \quad (20)$$

$$L_{off} = \frac{1}{|\mathcal{Q}_b|} \sum_{\mathbf{q} \in \mathcal{Q}_b} \exp(-\alpha |f(\mathbf{p})|) \quad (21)$$

$$L_{Eik} = \frac{1}{|\mathcal{Q}_{background} \cup \mathcal{Q}_{iso}|} \sum_{\mathbf{q} \in \mathcal{Q}_b \cup \mathcal{Q}_{iso}} |1 - \|\mathbf{J}^T(\mathbf{q})\|| \quad (22)$$

We compute the iso-points' normals using a principal component analysis (PCA). The consistency between the Jacobian of the model and the normal estimated using a local neighborhood is evaluated using the cosine similarity as shown below

$$L_{isoNormal} = \frac{1}{|\mathcal{Q}_{iso}|} \sum_{\mathbf{q} \in \mathcal{Q}_{iso}} \left(1 - |SC(\mathbf{J}^T(\mathbf{q}), \mathbf{n}_{PCA}(\mathbf{q}))|\right) \quad (23)$$

$$L_{Normal} = \frac{1}{|\mathcal{P}|} \sum_{\mathbf{p} \in \mathcal{P}} \left(1 - |SC(\mathbf{J}^T(\mathbf{p}), \mathbf{n}_{PCA}(\mathbf{p}))|\right), \quad (24)$$

where the cosine similarity $SC(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|}$. The optimization objective is comprised of six parts:

$$L = \lambda_{isoSDF} L_{isoSDF} + \lambda_{isoNormal} L_{isoNormal} + \lambda_{eik} L_{eik} + \lambda_{onSDF} L_{onSDF} + \lambda_{normal} L_{normal} + \lambda_{offSDF} L_{offSDF}. \quad (25)$$

4 Results

We evaluate the proposed algorithm on the Civilian Vehicle Data Domes dataset presented in [27] and the GOTCHA parking lot measured datasets [3,4] using the entire circular aperture. We utilize 8 elevation passes of a SAR sensor with 640

MHz bandwidth corresponding to roughly 0.25 cm resolution. The point-cloud and surface normals are recovered using regularized SAR inversion methods described in [2]. The point-cloud and the estimated normal vectors are utilized to train the network to estimate the SDF. The surface reconstruction by denoising the point-clouds is solved using presented method. We qualitatively evaluate the performance of the surface reconstruction by varying the number of frequencies of the random Fourier feature layer at the input and the effect of utilizing iso-points in the training procedure.

Figure 2 illustrates the surface representation estimated by the network for a Jeep vehicle. We note that the points that are not close to the surface that arise from multi-path and non-uniform frequency domain sampling are treated as outliers and do not influence the surface estimation. We vary the number of frequencies N_f in the input layer. We consider $N_f = \{6, 9\}$ for evaluation. We notice that as N_f increases the edges are clearly defined and the flat surfaces are also modeled. We also note that fine details such as rear-view mirrors are missing for $N_f = 9$ compared with $N_f = 6$. Additionally, there are more artifacts and holes in the surface representation for the case of $N_f = 6$ compared to the case of $N_f = 9$. We also evaluate the benefit of utilizing iso-points in the training

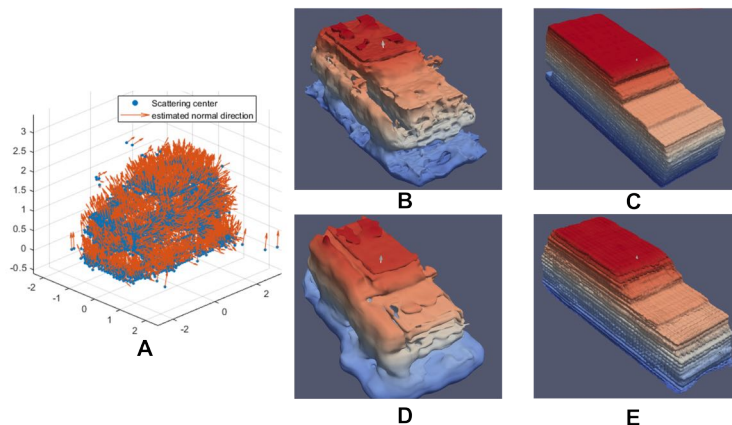


Fig. 2: Figure. A refers to the detected point-cloud and the estimated normal vectors from SAR phase-history measurements for Jeep 93 Vehicle using 8 elevation passes. Figure. B and Figure. C illustrate the recovered mesh using the SDF network after training when iso-points are not generated from the network to regularize the network optimization for the number of frequencies as 6 and 9 in the input layer, respectively. Figure. D and Figure. E illustrate the recovered mesh using the SDF network after training when iso-points are generated from the network to regularize the network optimization for the number of frequencies as 6 and 9 in the input layer, respectively.

procedure. The iso-points regularize the network training procedure and helps

in estimating the surface representations with more details and avoids artifacts as shown in Figures 2.

Figure 3 shows the surface recovered from the GOTCHA parking lot dataset. We utilize the entire 360 degree circular aperture to recover the point-cloud representation, creating sub-apertures with an azimuth span of 5 degrees. The number of random Fourier features in our reconstruction is set to $N_F = 9$. We observe that the proposed network architecture can successfully model the entire parking lot and capture the geometric details of individual vehicles. Further visualizations are provided in the supplementary materials.

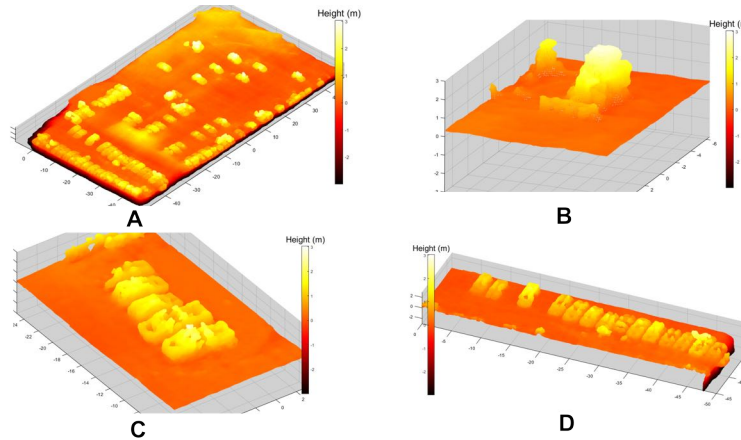


Fig. 3: Implicit surface recovery from the GOTCHA parking lot dataset. Figure A shows the entire parking lot, Figure B shows the backhoe from the dataset, and Figures C and D show the surface approximation of a set of cars in the parking lot.

5 Future Research Directions: Complex valued Neural Implicit Representations

The proposed framework in this paper allows denoising of the point clouds produced by the SAR sensor. However, complex amplitudes of the scatterers are lost along the way and therefore does not allow synthesis of novel views of the constructed objects. This objective can be achieved by extending the NeRF volume modeling strategy in two distinct ways to account for differences between Electro-optic (EO) and SAR phenomenology. First, the color map function has to be replaced with a view-dependent complex valued reflectance map with the view encoded by azimuth and elevation angles of the SAR platform. Second, the simple ray tracing based focal plane projections of the EO model has to be

replaced with orthographic projection of 3D scatterers of the object onto the 2D slant measurement plane. We note that the SAR projection operator is not a simple orthographic summation of the complex reflectivities, instead contributions are weighted with the phase terms proportional to their height from the slant plane, complicating the forward measurement model. This complex valued volumetric neural implicit representation can then be used to synthesize tomographic projections from previously unseen viewpoints/apertures.

References

1. Ferrara, M., Jackson, J.A., Austin, C.: Enhancement of multi-pass 3D circular SAR images using sparse reconstruction techniques. In: Algorithms for Synthetic Aperture Radar Imagery XVI. Volume 7337., International Society for Optics and Photonics (2009) 733702
2. Austin, C.D., Ertin, E., Moses, R.L.: Sparse signal methods for 3-D radar imaging. *IEEE Journal of Selected Topics in Signal Processing* **5**(3) (2010) 408–423
3. Ertin, E., Austin, C.D., Sharma, S., Moses, R.L., Potter, L.C.: Gotcha experience report: Three-dimensional SAR imaging with complete circular apertures. In: Algorithms for synthetic aperture radar imagery XIV. Volume 6568., International Society for Optics and Photonics (2007) 656802
4. Austin, C.D., Ertin, E., Moses, R.L.: Sparse multipass 3d sar imaging: Applications to the gotcha data set. In: Algorithms for Synthetic Aperture Radar Imagery XVI. Volume 7337., SPIE (2009) 19–30
5. Sugavanam, N., Ertin, E.: Recovery guarantees for MIMO radar using multi-frequency LFM waveform. In: 2016 IEEE Radar Conference (RadarConf), IEEE (2016) 1–6
6. Sugavanam, N., Ertin, E.: Interrupted SAR imaging with limited persistence scattering models. In: 2017 IEEE Radar Conference (RadarConf), IEEE (2017) 1770–1775
7. Sugavanam, N., Ertin, E.: Limited persistence models for SAR automatic target recognition. In: Algorithms for Synthetic Aperture Radar Imagery XXIV. Volume 10201., International Society for Optics and Photonics (2017) 102010M
8. Sugavanam, N., Ertin, E., Burkholder, R.: Approximating bistatic SAR target signatures with sparse limited persistence scattering models. In: Int. Conf. on Radar, Brisbane. (2018)
9. Sugavanam, N., Ertin, E., Burkholder, R.: Compressing bistatic SAR target signatures with sparse-limited persistence scattering models. *IET Radar, Sonar & Navigation* **13**(9) (2019) 1411–1420
10. Sugavanam, N., Ertin, E.: Models of anisotropic scattering for 3d sar reconstruction. In: 2022 IEEE Radar Conference (RadarConf22), IEEE (2022) 1–6
11. Sugavanam, N., Baskar, S., Ertin, E.: High resolution mimo radar sensing with compressive illuminations. *IEEE Transactions on Signal Processing* **70** (2022) 1448–1463
12. Agarwal, T., Sugavanam, N., Ertin, E.: Sparse signal models for data augmentation in deep learning atr. In: 2020 IEEE Radar Conference (RadarConf20), IEEE (2020) 1–6
13. Scarnati, T., Jamora, J.R.: Three-dimensional object reconstruction from sparse multi-pass SAR data. In: Algorithms for Synthetic Aperture Radar Imagery XXVIII. Volume 11728., International Society for Optics and Photonics (2021) 117280I

14. Jackson, J.A., Moses, R.L.: Synthetic aperture radar 3d feature extraction for arbitrary flight paths. *IEEE Transactions on Aerospace and Electronic Systems* **48**(3) (2012) 2065–2084
15. Jackson, J.A., Rigling, B.D., Moses, R.L.: Canonical scattering feature models for 3d and bistatic sar. *IEEE Transactions on Aerospace and Electronic Systems* **46**(2) (2010) 525–541
16. Varshney, K.R., Cetin, M., Fisher III, J.W., Willsky, A.S.: Joint image formation and anisotropy characterization in wide-angle SAR. *Proc. SPIE Algorithms for Synthetic Aperture Radar Imagery XIII* **6237** (2006) 62370D–62370D–12
17. Ertin, E., Moses, R.L., Potter, L.C.: Multibaseline ifsar for 3d target reconstruction. In: *Algorithms for Synthetic Aperture Radar Imagery XV*. Volume 6970., SPIE (2008) 60–70
18. Moore, L.J., Potter, L.C.: Three-dimensional resolution for circular synthetic aperture radar. *Proc. SPIE* **6568** (2007) 656804–656804–12
19. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. In: *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I* 16, Springer (2020) 405–421
20. Wang, Z., Wu, S., Xie, W., Chen, M., Prisacariu, V.A.: Nerf-: Neural radiance fields without known camera parameters. *arXiv preprint arXiv:2102.07064* (2021)
21. Tancik, M., Srinivasan, P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J., Ng, R.: Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems* **33** (2020) 7537–7547
22. Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., Geiger, A.: Occupancy networks: Learning 3d reconstruction in function space. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. (2019) 4460–4470
23. Park, J.J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S.: DeepSDF: Learning continuous signed distance functions for shape representation. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. (2019) 165–174
24. Greengard, L., Lee, J.Y.: Accelerating the nonuniform fast fourier transform. *SIAM review* **46**(3) (2004) 443–454
25. Yifan, W., Wu, S., Oztireli, C., Sorkine-Hornung, O.: Iso-points: Optimizing neural implicit surfaces with hybrid representations. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2021) 374–383
26. Huang, H., Wu, S., Gong, M., Cohen-Or, D., Ascher, U., Zhang, H.: Edge-aware point set resampling. *ACM transactions on graphics (TOG)* **32**(1) (2013) 1–12
27. Dungan, K.E., Austin, C., Nehrbass, J., Potter, L.C.: Civilian vehicle radar data domes. In: *Algorithms for synthetic aperture radar Imagery XVII*. Volume 7699., International Society for Optics and Photonics (2010) 76990P