Radar Fields: An Extension of Radiance Fields to SAR

Thibaud Ehret[1]   Roger Marí[1]   Dawa Derksen[2]   Nicolas Gasnier[2]   Gabriele Facciolo[1]

[1]Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, France
[2]CNES, Toulouse, France

Abstract

Radiance fields have been a major breakthrough in the field of inverse rendering, novel view synthesis and 3D modeling of complex scenes from multi-view image collections. Since their introduction, it was shown that they could be extended to other modalities such as LiDAR, radio frequencies, X-ray or ultrasound. In this paper, we show that, despite the important difference between optical and synthetic aperture radar (SAR) image formation models, it is possible to extend radiance fields to radar images thus presenting the first "radar fields". This allows us to learn surface models using only collections of radar images, similar to how regular radiance fields are learned and with the same computational complexity on average. Thanks to similarities in how both fields are defined, this work also shows a potential for hybrid methods combining both optical and SAR images.

1. Introduction

Many of the latest breakthroughs in the domain of 3D reconstruction have been inspired by Neural Radiance Fields (NeRF) [37]. By definition, NeRFs are a variant of Neural Fields that learn a 3D representation of the scene based on the measurement of radiance, given a set of optical images. In the past years, NeRFs have been reshaped, reworked and adapted to various use cases. In particular, their application domain has been extended far beyond computer vision [20, 55].

This diversity of application domains is largely due to the flexibility of the Neural Fields paradigm. Neural Fields represent a path forward for including physics into the Neural Network learning process, a longstanding aim in AI research. The common goal of all of the Neural Field research works cited in the following section is to learn a high-dimensional representation (e.g. a 3D shape) from a set of low-dimensional data (e.g. 2D images). From a physical point of view, the low-dimensional measurements can be seen as boundary conditions for a set of partial differential equations (PDE) that describe the formation process of the observations. In this setting learning amounts to an

(a) Regular radiance fields
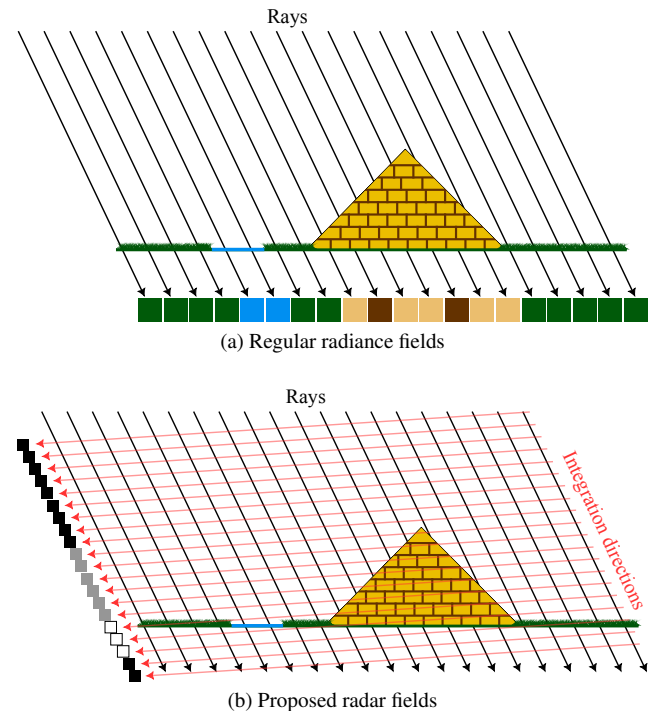

(b) Proposed radar fields

Figure 1. The difference between the integration process of a classic optical radiance field and the proposed radar field. While the radiance field integrates along each rays (in black), radar fields integrates across rays (in red) in a given azimuth plane (see Figure 2). Similarly to radiance fields, radar fields can be optimized from multiple SAR images to recover the underlying surface of the scene.

iterative optimization process to find a high-dimensional representation that simultaneously satisfies both the PDE and the boundary conditions. The learning process generally involves starting with a randomly (or arbitrarily) initialized scene. A set of low-dimensional observations are then computed using a forward "rendering" process, based on a discretized form of the PDE. These observations are confronted to the real measurements, yielding a loss. The loss is back-propagated onto the representation to improve its coherence with the observations. This process is repeated

until convergence is reached.

In this work, we focus on using Synthetic Aperture Radar (SAR) data instead of optical images to achieve 3D reconstruction. We introduce Radar Fields, a new framework for learning a volumetric model based on multiple SAR measurements of a scene. Since the wavelengths used in SAR experience minimal interactions with atmospheric gases or aerosols, the propagation medium is virtually transparent and non-absorbing. Thus, unlike optical sensors, SAR has day-night cloud-penetrating imaging capacities. Unlike interferometric SAR (InSAR), the proposed methodology does not prescribe the geometric configuration of the acquisitions. It could therefore be used with opportunistic acquisitions. We believe that the main impact of this research will be for global 3D reconstruction, due to the wide-scale availability and high quality of the latest SAR Earth Observation data from missions such as TerraSAR-X [38] or Capella [17, 49]. The methodology that we present is general and could equally be applied to aerial SAR or to other radar-based measurement systems.

To summarize, our contributions are the following:
1. The definition of radar fields, a differentiable rendering framework adapted to radar image synthesis.
2. A new surface model model definition, inspired by digital surface models, that can easily be used in frameworks like radar fields for remote sensing applications.
3. Adapting key training techniques, like NeRF random depth sampling or being able to train despite noisy data, to radar fields.

2. Related works

Our work is not the first to exploit SAR images to produce 3D models. In fact, space-borne SAR sensors are currently used to produce large scale Digital Surface Models (DSM) such as Copernicus DEM [1]. The TanDEM-X mission [61] uses the principle of radar interferometry to estimate the Earth's surface topography. The derived Copernicus DEM product describes the elevation of the terrain and built-up land elements with an average error of 2-4 meters. The spatial resolution of the product ranges from 10-30 meters.

An alternative technique to interferometry, known as SAR tomography [44, 48], consists in using stacks of 2D Synthetic Aperture Radar (SAR) images to estimate the surface height with a finer spatial resolution (1-2m). This technique involves first estimating the covariance matrix in each radar cell. In a second step, spectral analysis is used to invert this covariance matrix, resulting in an approximation of the surface height distribution. The resulting DSM achieves a mean altitude error ranging from 1-2 meters in urban areas. However, SAR Tomography commonly suffers from missing parts in urban surfaces, particularly in flat areas such as ground or rooftops. Moreover, the spatial heterogeneity of the height distribution near abrupt altitude tran-

sitions can result in a blurry reconstruction [40].

Other works propose deep learning-based single-image height estimation methods using a Convolutional Neural Network encoder-decoder architecture [41, 42], taking a SAR image as input and directly predicting depth maps. These methods provide fast inference times, but the training requires large amounts of ground-truth 3D data synchronized with multi-view SAR acquisitions to generalize to different viewing angles or landscapes.

Neural fields have seen a number of applications in optical-based satellite imagery, starting with the seminal work S-NeRF [15] followed by Sat-NeRF [34], and recently, EO-NeRF [35]. S-NeRF initially tackled the problem of varying lighting conditions; Sat-NeRF introduced a latent transient vector to disentangle temporary objects such as cars from the static scene, and EO-NeRF achieved state-of-the-art 3D reconstruction quality by explicitly modeling the shadow effects and learning per-image color corrections. Another recent work, SparseSat-NeRF [58], reduced the need for a large number of input images using depth priors based on stereo-photogrammetry. Season-NeRF [19] models seasonal variations to enable temporal interpolation between different acquisition dates. Sat-Mesh [39] replaced the NeRF model with an implicit surface representation, yielding impressive visual DSM quality, but not disentangling the transient objects or generalizing to unseen solar angles. RS-NeRF [54] and SatensoRF [59] have successfully reduced the optimization time, but like Sat-Mesh do not generalize to unseen solar angles. In the domain of aerial images, issues such as large-scale data [51, 56] and image co-registration [8] have also been addressed with NeRF approaches.

Since the arrival of NeRFs, researchers in various fields have also successfully applied the Neural Fields paradigm to entirely different data than optical images. For example, in the domain of radiative transfer and light modeling, these include applications to LiDAR [23, 50], Time-of-flight cameras [2] and non-line of sight cameras [47]. Impressive results have also been achieved using event based cameras, which detect a difference in pixel intensity rather than the absolute pixel intensity [24, 27, 33, 46]. In other regions of the electromagnetic spectrum, applications have risen for 3D reconstruction based on X-rays [9, 10] and Radio Frequency [60]. The equations of light propagation being similar in nature to the equations of sound propagation, the same paradigm can be adapted to the acoustic setting [7], with applications in ultra-sound [31, 53] and sonar [43]. Finally, the principle of implicit learning has also yielded exploratory works in geodesy [25] and gravitational lensing [30].

3. Introduction to radiance fields

NeRF [37] represents a static scene as a continuous volumetric function $\mathcal{F}$. This function can, for example, be encoded by a fully-connected neural network (MLP) [37], a voxel decomposition [26] or a tensor interpolation [6]. $\mathcal{F}$ predicts the emitted RGB color $\mathbf{c} = (r, g, b)$ and a non-negative scalar volume density σ at a 3D point $\mathbf{x} = (x, y, z)$ of the scene seen from a viewing direction $\mathbf{d} = (d_x, d_y, d_z)$, i.e.

$$\mathcal{F} : (\mathbf{x}, \mathbf{d}) \mapsto (\mathbf{c}, \sigma). \qquad (1)$$

Multi-view consistency is encouraged by restricting the network to predict the volume density σ based only on the spatial coordinates $\mathbf{x}$, while allowing the color $\mathbf{c}$ to be predicted as a function of both $\mathbf{x}$ and and other view-dependent features, such as the viewing direction $\mathbf{d}$.

Given a set of input views and their camera poses, $\mathcal{F}$ is optimized by rendering the color of individual rays traced across the scene and intersecting the known pixels. Each ray $\mathbf{r}$ is defined by a point of origin $\mathbf{o}$ and a direction vector $\mathbf{d}$. The color $\mathbf{c}(\mathbf{r})$ of a ray $\mathbf{r}(t) = \mathbf{o} + t\mathbf{d}$ is computed as

$$\mathbf{c}(\mathbf{r}) = \sum_{i=1}^{N} T_i \alpha_i \mathbf{c}_i. \qquad (2)$$

The rendered color $\mathbf{c}(\mathbf{r})$ results from the weighted integration of the colors $\mathbf{c}_i$ predicted at N different points sampled along the ray $\mathbf{r}$. Each point $\mathbf{x}_i$ in $\mathbf{r}$ is obtained as $\mathbf{x}_i = \mathbf{o} + t_i \mathbf{d}$, where t_i is the depth step.

The contribution of each point to the rendering equation (2) follows a light transmittance model based on the opacity α_i and transmittance T_i values, which follow geometry defined by the density σ

$$\alpha_i = 1 - \exp(-\sigma_i(t_{i+1} - t_i)); \quad T_i = \prod_{j=1}^{i-1}(1 - \alpha_j). \quad (3)$$

The opacity α_i increases with σ_i and is the probability that a point belongs to a non-transparent surface. The transmittance T_i is the probability that light arrives without hitting previous opaque points in the ray $\mathbf{r}$.

Given (3), the depth $d(\mathbf{r})$ observed in the direction of a ray $\mathbf{r}$ can be rendered in a similar manner to (2) [14, 45] as

$$d(\mathbf{r}) = \sum_{i=1}^{N} T_i \alpha_i t_i. \qquad (4)$$

The radiance field is optimized by minimizing the mean squared error (MSE) between the rendered color and the actual color of the input views:

$$\sum_{\mathbf{r} \in \mathcal{R}} \|\mathbf{c}(\mathbf{r}) - \mathbf{c}_{\text{GT}}(\mathbf{r})\|_2^2, \qquad (5)$$

where $\mathbf{c}_{\text{GT}}(\mathbf{r})$ is the observed color of the pixel intersected by the ray $\mathbf{r}$, and $\mathbf{c}(\mathbf{r})$ is the color predicted by the radiance field using (2). $\mathcal{R}$ is the set of rays in each input batch.

Surface models. The concept of radiance fields has been extended to surface models [13, 18, 32, 52, 57]. Instead of considering the scene as continuous volumetric function $\mathcal{F}$, the scene is represented as a surface $\mathcal{S}$. This surface is characterized by an implicit function $\mathcal{F}$ such that

$$\mathcal{S} = \left\{ \boldsymbol{x} \in \mathbb{R}^3 \mid \mathcal{F}(\boldsymbol{x}) = 0 \right\}. \qquad (6)$$

Denoting d the minimum distance between $\boldsymbol{x}$ and $\mathcal{S}$, $\mathcal{F}$ is defined as a signed distance function (SDF) with

$$\mathcal{F}(\boldsymbol{x}) = \begin{cases} d(\boldsymbol{x}, \mathcal{S}) & \text{if } \boldsymbol{x} \text{ is outside } \mathcal{S}, \\ -d(\boldsymbol{x}, \mathcal{S}) & \text{if } \boldsymbol{x} \text{ is inside } \mathcal{S}. \end{cases} \qquad (7)$$

This property is enforced using an additional loss term to enforce the Eikonal property of the SDF

$$\mathcal{L}_{Eikonal} = \sum_{x \in \Omega} \left(\|\nabla \mathcal{F}(\boldsymbol{x})\|_2^2 - 1 \right)^2 \qquad (8)$$

with Ω containing points sampled near the surface as well as uniformly in the entire volume.

The optimization of a surface based model is similar to that of a regular radiance field. The surface model is transformed into a volumetric model using a function that produces a pseudo-density from the SDF value at a given point. For example, in VolSDF [57], the transformation function Ψ is defined with the cumulative distribution function of the Laplace function

$$\Psi(d) = \begin{cases} \frac{1}{2} \exp\left(-\frac{d}{\beta}\right) & \text{for } d \geqslant 0 \\ 1 - \frac{1}{2} \exp\left(\frac{d}{\beta}\right) & \text{for } d < 0, \end{cases} \qquad (9)$$

where α and β are optimized with the rest of the model. The rendering operation remains unchanged, as (2) and (3) use the pseudo-density

$$\sigma_i = \Psi\left(\mathcal{F}\left(\boldsymbol{x}_i\right)\right). \qquad (10)$$

Representing the geometry with an implicit function $\mathcal{F}$ opens the door to exploit additional information, such as the normal of the surface, i.e. $n(\boldsymbol{x}) = \nabla \mathcal{F}(\boldsymbol{x})$ when $\nabla \mathcal{F}(\boldsymbol{x}) \neq \mathbf{0}$, which is prevented by the Eikonal constraint.

4. Radar fields

The models presented in Section 3 have been proposed for optical images. We propose here to extend the concept of radiance fields to SAR images.

Introduction to the SAR image formation model. A SAR imaging device is an active system. It is usually composed of an antenna mounted on a moving spaceborne or airborne platform that emits electromagnetic pulses in a side-looking geometry and then receives echoes that are backscattered by the ground surface. The satellite flight direction is called along-track or azimuth, while the direction of the beam is called across-track or range. For each azimuth pulse emitted, the satellite receives the echoes, demodulates and samples them by range, ordered by acquisition time, and stores them as complex numbers separating the amplitude and phase components of the signal. The amplitude is related to the ground reflectivity and angle, while the phase contains information about the satellite-target distance or *path delay* (modulo the wavelength). The resulting image is said to be the raw format. Because of the large footprint of the electromagnetic beam and the length of the pulse, in this format, the response from a single scatterer (*i.e.* a ground reflective element) on the ground is spread over many pixels and the spatial resolution is very low. The resolution is largely improved through a process known as focusing [11], which yields a complex-valued image equivalent to employing short pulses (for range) and a narrow beam perpendicular to the motion of the satellite known as zero-doppler plane (corresponding to azimuth). Unlike optical images, SAR images represent the distance to reflectors in the range direction. The SAR acquisition process is illustrated in Figure 2. If two reflectors are at the same distance from the satellite (in the same azimuth plane, *e.g.* the red plane in Fig. 2), the response will be the sum of both, resulting in foreshortening, layover and shadow effects. Luckily, this range image can be analogously interpreted as a projected view of the scene from a direction perpendicular to the incidence angle and illuminated by the radar pulse (as illustrated in Figure 1). Note however that this perpendicular projection should not account for occlusions.

Definition of radar fields. Using the notations from Section 3, we want to learn the function $\mathcal{F}$ corresponding to the surface model of the scene using a collection of SAR images. In this section, we derive the model for a given azimuth plane. This is not a limitation since the image is equivalent to a zero-doppler acquisition for each azimuth plane as mentioned in the previous paragraph, except in the case of double or multiple bounces, which we discuss in Section 6.

Consider a set of rays $(\boldsymbol{r}_j)$, with respective origin and direction $(\boldsymbol{o}_j, \boldsymbol{v}_j)$, in the given azimuth plane and $(d_i)_{i \in [\![1,N]\!]}$ the distance sampling associated to these rays. Let $s_{\boldsymbol{r}_j}(d_i)$ be the sensed signal associated to $\boldsymbol{r}_j$ at distance d_i. Following Section 3, this yields

$$s_{\boldsymbol{r}_j}(d_i) = -T_{i,j}\alpha_{i,j}\left\langle \boldsymbol{v}_j | \boldsymbol{n}_{i,j} \right\rangle^{\theta}, \qquad (11)$$
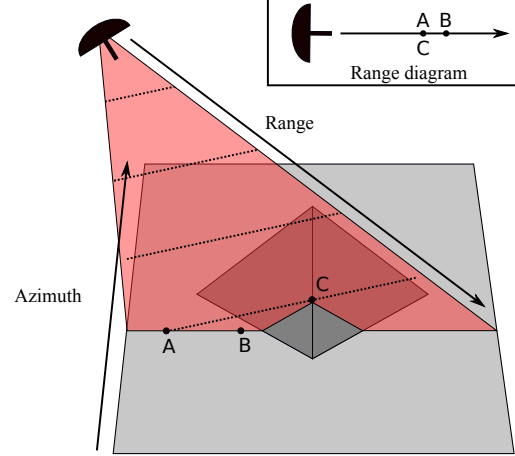

Figure 2. Acquisition process of a SAR image: At each azimuth position an electromagnetic pulse is sent from the antenna. The red plane illustrates the azimuth plane associated to a pulse. Points A and C, which are at the same distance from the sensor, are projected to the same point while B is projected on a different point (as seen in the range diagram).

where the definition of $T_{i,j}$ and $\alpha_{i,j}$, for a fixed j, is the same as in (3), $\boldsymbol{n}_{ij} = \nabla\mathcal{F}(\boldsymbol{o}_j + d_i\boldsymbol{v}_j)$, and where the coefficient θ is used to model the specularity of the reflection [3, 4]. Indeed, in the case of SAR images, there is no color: the intensity of the reflected signal is characterized by the surface. Traditionally in SAR simulation [3, 4], the reflected signal s is given by

$$s = -\left\langle \boldsymbol{v} | \boldsymbol{n} \right\rangle^{\theta} \qquad (12)$$

with $\boldsymbol{v}$ the incidence direction and $\boldsymbol{n}$ the normal to the surface. Often a purely Lambertian model, corresponding to $\theta = 1$, is assumed. Larger values of θ are used to represent specular surfaces. The parameter θ can be learned at the same time as the location of the surface.

As described in the previous paragraph and represented in Figure 1, the integration is performed *across* rays for SAR images and not *along* rays as in optical images: for a given azimuth, reflectors that are at the same distance are projected into the same pixel. To satisfy this condition, the rendering operation (2) must be replaced by

$$s(d_i) = \sum_{\boldsymbol{r}_j} s_{\boldsymbol{r}_j}(d_i) = -\sum_{\boldsymbol{r}_j} T_{i,j}\sigma_{i,j}\left\langle \boldsymbol{v}_j | \boldsymbol{n}_{i,j} \right\rangle^{\theta}. \quad (13)$$

Note how the role of indices i and j are inverted between Eq. (2) and Eq. (13) and how **c** depends on the ray while s depends only the distance. While the position of rays is well defined by the sensor configuration in the optical case (see Sec. 3), this is not the case for SAR. Conversely, the sampling of the $\boldsymbol{x}_{i,j}$ is perfectly defined in the SAR case – following the range sampling – but is flexible in the opti-

cal case. Finally, computing the results for all the azimuth planes gives the complete SAR image.

Differences between radiance and radar fields. While radiance and radar fields are conceptually similar, we list some important differences here. They are also summarized in Table 1. The first difference is that to generate a single SAR pixel, it is necessary to compute all rays corresponding to its azimuth plane. Thankfully, computing a single pixel provides the result for the entire range line at the same time. This means that while rendering a single pixel requires more computations, the overall amount of computations necessary to learn a radar field remains the same, on average, as the one needed to learn a radiance field. In practice the only difference is how rays are batched. While there is no requirement on how to batch rays in the radiance field case, it is necessary to batch all rays of a given azimuth plane together during computation for radar fields. Note that with additional assumptions on the scene, it is possible to slightly relax this requirement. Indeed, if an approximate position of the surface is known *a priori*, then it would be possible to split the distance sampling into multiple smaller batches that can be computed separately. Note that no such assumption is made for the experiments presented in Section 5.

A key technique to improve the performance in NeRF [37] is to introduce some noise when sampling along rays instead of using a fixed uniform sampling. This is not possible with a radar field because the sampling step along rays is defined by the range sampling. It is however possible to transpose this idea to ray sampling since, as shown in Eq. (13), the final signal does not require a specific ray position. as long as $\boldsymbol{o}_j$ stays in its corresponding azimuth plane and that it does not impact the range sampling. If we denote by $\boldsymbol{w}_j$ the vector orthogonal to $\boldsymbol{v}_j$ inside the azimuth plane, then the perturbed ray origin is defined by

$$\tilde{\boldsymbol{o}}_j = \boldsymbol{o}_j + n\boldsymbol{w}_j \quad \text{with } n \sim \mathcal{N}(0,1). \quad (14)$$

While optical images have a good SNR in normal illumination conditions, SAR images are fundamentally noisy. They exhibit speckle noise, which has been studied and shown to follow a complex Wishart distribution [21, 22]. The distribution p of the noisy sample covariance matrix C of dimension d corresponding to the noise-free covariance Σ in a L looks configuration is

$$p(\boldsymbol{C}) = \frac{L^{Ld}|\boldsymbol{C}|^{L-d}}{\Gamma_d(L)|\boldsymbol{\Sigma}|^L} \exp\left(-L\operatorname{Tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{C}\right)\right), \quad (15)$$

with $\Gamma_d(L) = \pi^{d(d-1)/2} \prod_{k=1}^{d} \Gamma(L-k+1)$ and Γ the gamma function. When working with single-channel single-look intensity images, the sample covariance matrix reduces to an intensity I and this model can be simplified into a multiplicative speckle noise with regard to the noise-free reflectance R, such that

$$I = n \times R \quad \text{with} \quad n \sim \Gamma(1,1). \quad (16)$$

	Radiance field	Radar Field
Complexity	height × width × samples	azimuth × range × samples (on average)
Integration	Along the ray	Across the ray
Distance sampling	Variable (along a ray)	Fixed (defined by data)
Ray sampling	Fixed (defined by data)	Variable (in a given azimuth plane)
Noise	Noise-free	Follows Eq. 16

Table 1. Main differences between regular radiance fields and the proposed radar fields.

Learning the surface despite the noise is not a problem in practice. Indeed, comparing the noiseless generated view to the noisy samples is sufficient. This is similar to Noise2Noise [29] that shows that it is possible to learn a denoising network despite having only noisy samples. SAR2SAR [12] showed that this idea can also be applied to SAR data and [16] that it is possible to learn the noiseless images of a scene using a single burst of images, even when the noise has a bias.

Surface model implementation. In remote sensing, ground surfaces are usually represented by a DSM, *i.e.* a 2D map of the altitude of the ground at each position. For our experiments we propose to use the same formalism to define the surface in the proposed radar field. From $dsm : \mathbb{R}^2 \to \mathbb{R}$, we define

$$\mathcal{F}(\boldsymbol{x}) = z - dsm(x,y) \quad (17)$$

with $\boldsymbol{x} = (x,y,z) \in [0,1]^3$. While this function is not exactly an SDF since its value might not necessarily correspond to the exact distance to the surface, it nonetheless defines a surface similarly to Eq. 6 and therefore the learning process presented in Section 3 can be used to learn the surface. The results shown in Sec. 5 were computed using a rasterized DSM that can be assimilated to an image whose pixel represent the height value of the surface. It is the value of each pixel that is then directly optimized. Subpixel heights are estimated using a bilinear interpolation of the nearest values and normals to the defined surface are computed accordingly. One of the advantage is that the Eikonal loss (8) is not necessary and it is possible define simple surface regulation terms like

$$\sum_{\boldsymbol{x}} \sum_{\boldsymbol{y} \in \mathcal{N}(\boldsymbol{x})} \|dsm(\boldsymbol{x}) - dsm(\boldsymbol{y})\|_2^2, \quad (18)$$

with $\mathcal{N}(\boldsymbol{x})$ the set of neighbors of $\boldsymbol{x}$, that we use for the experiments. Note that the presented concept should remain compatible with the more classic definitions of radiance fields.
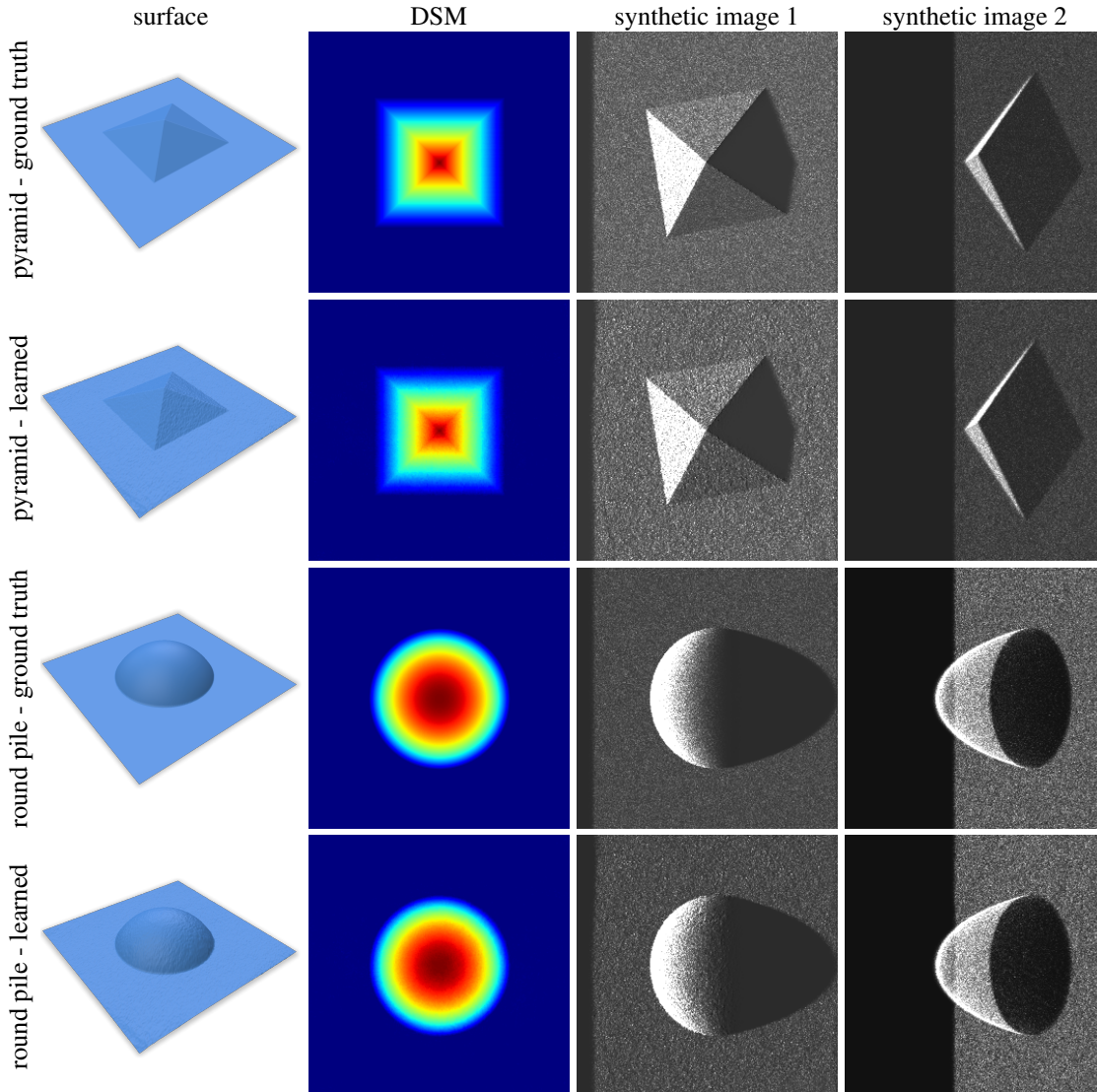
Figure 3. Results of surface learning using five SAR images on two toy examples. The first represents a pyramid (first row is the ground truth and second row the learned model) and the second a round pile (ground truth in the third row and learned model in the last row). From left to right: visualization of the surface model, DSM corresponding to the surface model, and two images corresponding to two views used during training.

5. Experiments

We apply the proposed radar field model to synthetic examples generated from toy surfaces and real DSMs to demonstrate its potential.

Synthetic data generation. We generate the synthetic SAR acquisitions, using the same image formation model described in Section 4. Indeed, given that the ground truth is also represented by a DSM, the radar field synthesis equations can be used to simulate the images corresponding to the query views. The main difference is that noise needs to be added during the process. For this, following Eq. (16), we use instead $\tilde{s}(d) = n \times s(d)$ with $n \sim \Gamma(1,1)$ and $s(d)$

defined by Eq. (13). The optimization process requires only the generated noisy images and the geometric parameters associated to each view.

Parameters. All the experiments presented in this section were done using the same parameters and the images generated corresponds to the same five camera positions. The models were trained during 10000 steps with an initial learning rate of 1. We reduced the learning rate to 0.1 at epoch 5000 and to 10^{-2} at step 8000. Azimuth planes were sampled uniformly from all images and grouped into batches of size 64, and we sampled 256 rays per azimuth plane. Except when specified otherwise, we consider Lambertian surfaces with $\theta = 1$.

surface DSM synthetic image 1 synthetic image 2

Fuji - ground truth

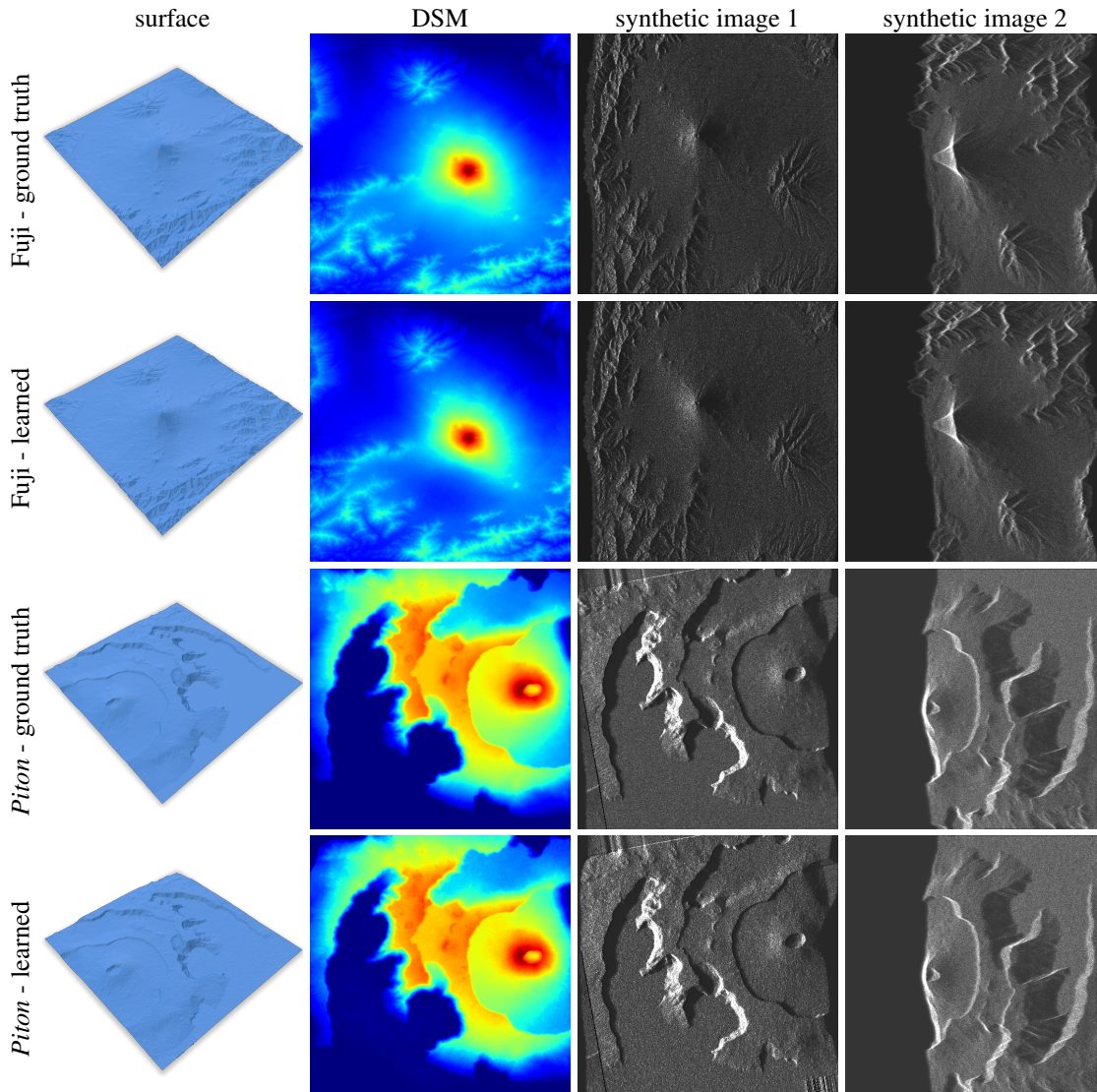Fuji - learned

Piton - ground truth

Piton - learned

Figure 4. Results of surface learning using five SAR images on simulated data from two real DSMs. The first represents mount Fuji (first row is the ground truth and second row the learned model) and the second the *Piton de la Fournaise* (ground truth in the third row and learned model in the last row). From left to right: visualization of the surface model, DSM corresponding to the surface model, and two images corresponding to two views used during training.

Results. Figures 3 and 4 show results both with toy examples (a pyramid and a circular pile) as well as simulations using real DSMs (of the mount Fuji and the *Piton de la Fournaise*). These results show that it is indeed possible to learn the ground surface from a few opportunistic SAR acquisitions using the proposed radar fields. By quantifying the altitude reconstruction errors relative to the ground pixel resolution, we observe errors in the order of 10^{-2} pixels. Lastly, Figure 5 shows an example of reconstruction of a non Lambertian surface, where the specularity map and geometry are both estimated during the optimization. In all these examples, the learned surfaces are very accurate despite the presence of a strong noise in the training data. Ad-

ditional results are available in the supplementary material.

6. Discussion and conclusion

In this section, we discuss the limitations and possible improvements of the model presented in Section 4.

Multi bounce. Depending on the surface configuration, the SAR signal may be subject to multiple bounces. This phenomena occurs when the signal received by the sensor is not the one directly reflected by an ideal Lambertian surface, but one coming from a second or further reflection. Corners are a common source of double bounces. We distinguish between two cases: the first is when bounces occur
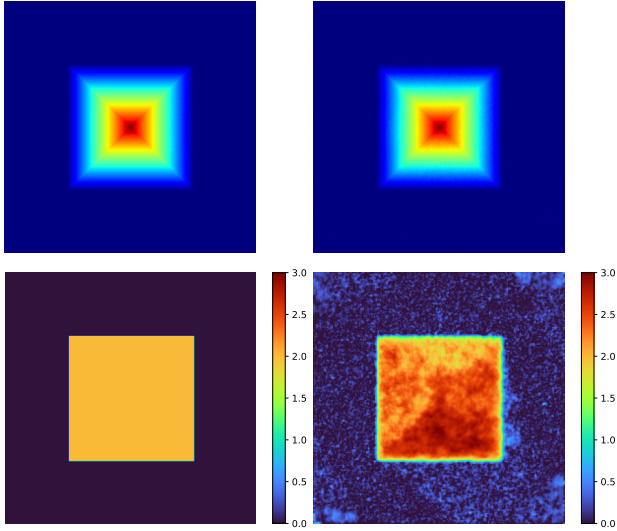
Figure 5. Example showing that it is possible to learn the θ coefficient (see Eq.(12)) characterizing the surface property (bottom) at the same time as the surface model (top). Left is the ground truth while right is the learned model. The learned model was initialized with $\theta = 1$ everywhere.
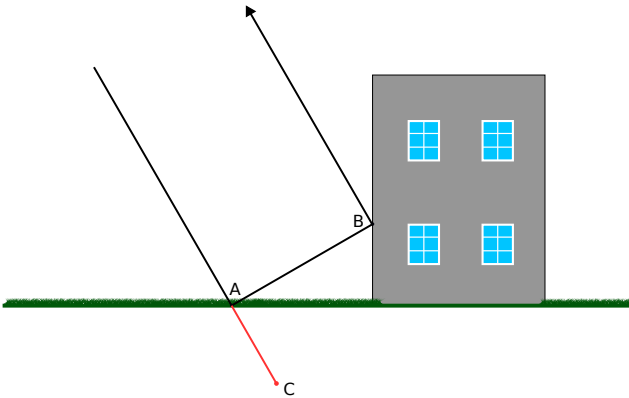

Figure 6. Example of multiple bounces for SAR signal. The signal is partially reflected in A but also bounces and is reflected a second time in B. This second bounce is sensed as a virtual reflection in C.

in a given azimuth plane, while the second is a more generic case where bounces are not restricted to that plane.

Figure 6 illustrates the case of double bounce in a given azimuth plane. Part of the signal is reflected at the first contact with the surface – as modeled so far – but also bounces to another part of the surface before being reflected back to the sensor. This process could be simulated by continuing rays after the first surface intersection, similar to how additional rays are computed in [35] to predict shadows. This increases the complexity of the model by a factor equal to the maximum number of bounces considered. It is important to note that the apparent position of the reflection associated to the double bounce is not the same as the true reflection posi-

tion. Indeed, the acquisition system measures the traveling time of the impulse signal from emission to reception and as such assumes that the reflection happened at half the traveling time. For example, in Figure 6, the signal that bounced from A to B and back to the receiver appears at the same position as if it were reflected by point C.

The generic case is more complex than the latter because the zero-doppler assumption is not true anymore. This means that a ray corresponding to a given azimuth plane can impact the signal of another azimuth plane. Such a generic multi-bounce scenario would require batching multiple azimuth planes together to render each pixel, making the method highly unfeasible unless strong priors or other constraints are exploited in future work.

Phase and polarimetric information. One additional attribute of SAR data that could be incorporated is the phase information. Indeed, SAR images are made of complex numbers containing the intensity and the phase of the signal. While phase is already used during for the focusing preprocessing step, it can also be used in the model. SAR Interferometry (InSAR) exploits the phase of multiple images with adequate acquisition configurations. Multitemporal InSAR images can be used to measure small terrain deformation [36], and across-track InSAR [5] can be used to build digital surface models (DSM) using a pair of images acquired from two antennas (e.g., SRTM), as mentioned in Section 2. Therefore, we see potential in the phase information to improve altitude accuracy in future work. Similarly, polarimetric information could also be modeled. Polarimetric information is widely used to determine the scattering mechanism [28] and as such can be useful when estimating the surface model. It can also be useful when estimating additional scene properties, such as semantic classification.

Conclusion. We presented radar fields, an extension of radiance fields to SAR imagery. While the model presented in this paper is a proof of concept tested on simulated synthetic data, we believe that it opens up new exciting opportunities, especially in remote sensing where high-quality SAR images are becoming increasingly common and the generation of surface models from satellite images is a major research topic. Moreover, it also shows a potential to inspire future hybrid models combining both opportunistic optical and SAR acquisitions in a joint framework.

References

[1] Copernicus DEM handbook, last consulted 10/10/23. 2

[2] Benjamin Attal, Eliot Laidlaw, Aaron Gokaslan, Changil Kim, Christian Richardt, James Tompkin, and Matthew O'Toole. Törf: Time-of-flight radiance fields for dynamic scene view synthesis. *Advances in neural information processing systems*, 34:26289–26301, 2021. 2

[3] Stefan Josef Auer. *3D synthetic aperture radar simulation for interpreting complex urban reflection scenarios*. PhD thesis, Technische Universität München, 2011. 4

[4] Timo Balz and Uwe Stilla. Hybrid GPU-based single-and double-bounce SAR simulation. *IEEE Transactions on Geoscience and Remote Sensing*, 47(10):3519–3529, 2009. 4

[5] Alessandra Budillon, Christoph H. Gierull, Vito Pascazio, and Gilda Schirinzi. Along-track interferometric SAR systems for ground-moving target indication: Achievements, potentials, and outlook. *IEEE Geoscience and Remote Sensing Magazine*, 8(2):46–63, 2020. 8

[6] Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. TensoRF: Tensorial radiance fields. In *Computer Vision – ECCV 2022*, pages 333–350, 2022. 3

[7] Changan Chen, Alexander Richard, Roman Shapovalov, Vamsi Krishna Ithapu, Natalia Neverova, Kristen Grauman, and Andrea Vedaldi. Novel-view acoustic synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6409–6419, 2023. 2

[8] Yu Chen and Gim Hee Lee. DReg-NeRF: Deep registration for neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 22703–22713, 2023. 2

[9] Zixuan Chen, Jianhuang Lai, Lingxiao Yang, and Xiaohua Xie. Cunerf: Cube-based neural radiance field for zero-shot medical image arbitrary-scale super resolution. *arXiv preprint arXiv:2303.16242*, 2023. 2

[10] Abril Corona-Figueroa, Jonathan Frawley, Sam Bond-Taylor, Sarath Bethapudi, Hubert PH Shum, and Chris G Willcocks. Mednerf: Medical neural radiance fields for reconstructing 3D-aware ct-projections from a single x-ray. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 3843–3848. IEEE, 2022. 2

[11] Ian G Cumming and Frank H Wong. *Digital signal processing of synthetic aperture radar data*. Artech House, 2005. 4

[12] Emanuele Dalsasso, Loic Denis, and Florence Tupin. SAR2SAR: A semi-supervised despeckling algorithm for SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:4321–4329, 2021. 5

[13] François Darmon, Bénédicte Bascle, Jean-Clément Devaux, Pascal Monasse, and Mathieu Aubry. Improving neural implicit surfaces geometry with patch warping. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6260–6269, 2022. 3

[14] Kangle Deng, Andrew Liu, Jun-Yan Zhu, and Deva Ramanan. Depth-supervised NeRF: Fewer views and faster training for free. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 3

[15] Dawa Derksen and Dario Izzo. Shadow neural radiance fields for multi-view satellite photogrammetry. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1152–1161, 2021. 2

[16] Thibaud Ehret, Axel Davy, Pablo Arias, and Gabriele Facciolo. Joint demosaicking and denoising by fine-tuning of bursts of raw images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8868–8877, 2019. 5

[17] Gordon Farquharson, William Woods, Craig Stringham, Navneet Sankarambadi, and Lucas Riggi. The capella synthetic aperture radar constellation. In *EUSAR 2018; 12th European Conference on Synthetic Aperture Radar*, pages 1–5. VDE, 2018. 2

[18] Qiancheng Fu, Qingshan Xu, Yew Soon Ong, and Wenbing Tao. Geo-NeuS: Geometry-consistent neural implicit surfaces learning for multi-view reconstruction. *Advances in Neural Information Processing Systems*, 35:3403–3416, 2022. 3

[19] Michael Gableman and Avinash Kak. Incorporating season and solar specificity into renderings made by a NeRF architecture using satellite images. *arXiv preprint arXiv:2308.01262*, 2023. 2

[20] Kyle Gao, Yina Gao, Hongjie He, Denning Lu, Linlin Xu, and Jonathan Li. NeRF: Neural radiance field in 3D vision, a comprehensive review. *arXiv e-prints*, pages arXiv–2210, 2022. 1

[21] J. W. Goodman. Some fundamental properties of speckle∗. *J. Opt. Soc. Am.*, 66(11):1145–1150, 1976. 5

[22] N. R. Goodman. Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *The Annals of Mathematical Statistics*, 34(1):152 – 177, 1963. 5

[23] 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. 2

[24] Inwoo Hwang, Junho Kim, and Young Min Kim. Ev-NeRF: Event based neural radiance field. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 837–847, 2023. 2

[25] Dario Izzo and Pablo Gómez. Geodesy of irregular small bodies via neural density fields: geodesynets. *arXiv preprint arXiv:2105.13031*, 2021. 2

[26] Animesh Karnewar, Tobias Ritschel, Oliver Wang, and Niloy Mitra. Relu fields: The little non-linearity that could. In *ACM SIGGRAPH 2022 Conference Proceedings*, pages 1–9, 2022. 3

[27] Simon Klenk, Lukas Koestler, Davide Scaramuzza, and Daniel Cremers. E-NeRF: Neural radiance fields from a moving event camera. *IEEE Robotics and Automation Letters*, 8(3):1587–1594, 2023. 2

[28] J.S. Lee and Eric Pottier. *Polarimetric Radar Imaging : From basics to applications.* CRC Press, Taylor & Francis Editor,, 2009. 8

[29] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. In *International Conference on Machine Learning*, pages 2965–2974. PMLR, 2018. 5

[30] Aviad Levis, Pratul P Srinivasan, Andrew A Chael, Ren Ng, and Katherine L Bouman. Gravitationally lensed black hole emission tomography. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19841–19850, 2022. 2

[31] Honggen Li, Hongbo Chen, Wenke Jing, Yuwei Li, and Rui Zheng. 3D ultrasound spine imaging with application of neural radiance field method. In *2021 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4. IEEE, 2021. 2

[32] Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H Taylor, Mathias Unberath, Ming-Yu Liu, and Chen-Hsuan Lin. Neuralangelo: High-fidelity neural surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8456–8465, 2023. 3

[33] Weng Fei Low and Gim Hee Lee. Robust e-NeRF: NeRF from sparse & noisy events under non-uniform motion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 18335–18346, 2023. 2

[34] Roger Marí, Gabriele Facciolo, and Thibaud Ehret. Sat-NeRF: Learning multi-view satellite photogrammetry with transient objects and shadow modeling using RPC cameras. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1310–1320, 2022. 2

[35] Roger Marí, Gabriele Facciolo, and Thibaud Ehret. Multi-date earth observation NeRF: The detail is in the shadows. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2034–2044, 2023. 2, 8

[36] Didier Massonnet, Marc Rossi, César Carmona, Frédéric Adragna, Gilles Peltzer, Kurt Feigl, and Thierry Rabaute. The displacement field of the Landers earthquake mapped by radar interferometry. *Nature*, 364(6433):138–142, 1993. 8

[37] 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. In *Computer Vision – ECCV 2020*, pages 405–421, 2020. 1, 3, 5

[38] Wolfgang Pitz and David Miller. The TerraSAR-X satellite. *IEEE Transactions on Geoscience and Remote Sensing*, 48 (2):615–622, 2010. 2

[39] Yingjie Qu and Fei Deng. Sat-mesh: Learning neural implicit surfaces for multi-view satellite reconstruction. *Remote Sensing*, 15(17):4297, 2023. 2

[40] Clément Rambour, Loïc Denis, Florence Tupin, Hélène Oriot, Yue Huang, and Laurent Ferro-Famil. Urban surface reconstruction in SAR tomography by graph-cuts. *Computer Vision and Image Understanding*, 188:102791, 2019. 2

[41] Michael Recla and Michael Schmitt. Deep-learning-based single-image height reconstruction from very-high-resolution SAR intensity data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 183:496–509, 2022. 2

[42] Michael Recla and Michael Schmitt. From relative to absolute heights in SAR-based single-image height prediction. In *2023 Joint Urban Remote Sensing Event (JURSE)*, pages 1–4. IEEE, 2023. 2

[43] Albert Reed, Thomas Blanford, Daniel C Brown, and Suren Jayasuriya. Implicit neural representations for deconvolving sas images. In *OCEANS 2021: San Diego–Porto*, pages 1–7. IEEE, 2021. 2

[44] A. Reigber and A. Moreira. First demonstration of airborne SAR tomography using multibaseline l-band data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5): 2142–2152, 2000. 2

[45] Barbara Roessle, Jonathan T Barron, Ben Mildenhall, Pratul P Srinivasan, and Matthias Nießner. Dense depth priors for neural radiance fields from sparse input views. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 3

[46] Viktor Rudnev, Mohamed Elgharib, Christian Theobalt, and Vladislav Golyanik. EventNeRF: Neural radiance fields from a single colour event camera. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4992–5002, 2023. 2

[47] Siyuan Shen, Zi Wang, Ping Liu, Zhengqing Pan, Ruiqian Li, Tian Gao, Shiying Li, and Jingyi Yu. Non-line-of-sight imaging via neural transient fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(7):2257–2268, 2021. 2

[48] Yilei Shi, Richard Bamler, Yuanyuan Wang, and Xiao Xiang Zhu. SAR tomography at the limit: Building height reconstruction using only 3–5 TanDEM-X bistatic interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 58 (11):8026–8037, 2020. 2

[49] Craig Stringham, Gordon Farquharson, Davide Castelletti, Eric Quist, Lucas Riggi, Duncan Eddy, and Scott Soenen. The capella x-band sar constellation for rapid imaging. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 9248–9251. IEEE, 2019. 2

[50] Tang Tao, Longfei Gao, Guangrun Wang, Peng Chen, Dayang Hao, Xiaodan Liang, Mathieu Salzmann, and Kaicheng Yu. LiDAR-: Novel LiDAR view synthesis via neural radiance fields. *arXiv preprint arXiv:2304.10406*, 2023. 2

[51] Haithem Turki, Deva Ramanan, and Mahadev Satyanarayanan. Mega-NeRF: Scalable construction of large-scale NeRFs for virtual fly-throughs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12922–12931, 2022. 2

[52] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction. In *Advances in Neural Information Processing Systems*, pages 27171–27183. Curran Associates, Inc., 2021. 3

[53] Magdalena Wysocki, Mohammad Farid Azampour, Christine Eilers, Benjamin Busam, Mehrdad Salehi, and Nassir Navab. Ultra-nerf: Neural radiance fields for ultrasound imaging. *arXiv preprint arXiv:2301.10520*, 2023. 2

[54] Songlin Xie, Lei Zhang, Gwanggil Jeon, and Xiaomin Yang. Remote sensing neural radiance fields for multi-view satellite photogrammetry. *Remote Sensing*, 15(15):3808, 2023. 2

[55] Yiheng Xie, Towaki Takikawa, Shunsuke Saito, Or Litany, Shiqin Yan, Numair Khan, Federico Tombari, James Tompkin, Vincent Sitzmann, and Srinath Sridhar. Neural fields in visual computing and beyond. In *Computer Graphics Forum*, pages 641–676. Wiley Online Library, 2022. 1

[56] Linning Xu, Yuanbo Xiangli, Sida Peng, Xingang Pan, Nanxuan Zhao, Christian Theobalt, Bo Dai, and Dahua Lin. Grid-guided neural radiance fields for large urban scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8296–8306, 2023. 2

[57] Lior Yariv, Jiatao Gu, Yoni Kasten, and Yaron Lipman. Volume Rendering of Neural Implicit Surfaces. In *Advances in Neural Information Processing Systems*, pages 4805–4815. Curran Associates, Inc., 2021. 3

[58] Lulin Zhang and Ewelina Rupnik. Sparsesat-NeRF: Dense depth supervised neural radiance fields for sparse satellite images. In *ISPRS Annals 2023*, 2023. 2

[59] Tongtong Zhang and Yuanxiang Li. Fast satellite tensorial radiance field for multi-date satellite imagery of large size. *arXiv preprint arXiv:2309.11767*, 2023. 2

[60] Xiaopeng Zhao, Zhenlin An, Qingrui Pan, and Lei Yang. Nerf2 - neural radio-frequency radiance fields. *arXiv preprint arXiv:2305.06118*, 2023. 2

[61] Manfred Zink, Markus Bachmann, Benjamin Brautigam, Thomas Fritz, Irena Hajnsek, Alberto Moreira, Birgit Wessel, and Gerhard Krieger. TanDEM-X : The new global DEM takes shape. *IEEE Geoscience and Remote Sensing Magazine*, 2(2):8–23, 2014. 2