Ultrasound volume reconstruction from 2D Freehand acquisitions using neural implicit representations

François Gaits, Nicolas Mellado, Adrian Basarab

ULTRASOUND VOLUME RECONSTRUCTION FROM 2D FREEHAND ACQUISITIONS USING NEURAL IMPLICIT REPRESENTATIONS

François Gaits[1], *Nicolas Mellado*[1], *Adrian Basarab*[2]

[1]Institut de Recherche en Informatique de Toulouse, UMR CNRS 5505, Université de Toulouse, France

[2]Université de Lyon, Université Claude Bernard Lyon 1, INSA-Lyon, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR 5220, U1206, Villeurbanne, France

ABSTRACT

3D ultrasound reconstruction allows physicians to explore a region of interest (ROI) in 3D while leveraging the advantages of 2D ultrasound imaging: simple, low cost and non-ionizing. It may assist many clinical tasks, such as organ measurement, procedure control or visualization of tissues difficult to interpret through 2D visualization. Recently, new deep learning techniques in the field of novel view synthesis, based on a continuous description of the 3D field, showed promising results in terms of 3D model estimation, robustness to noise and uncertainty, and efficiency. Inspired by these approaches, the objective of this work is to propose a 3D ultrasound reconstruction method based on neural implicit representations. Results on simulated and experimental data show the superiority of the proposed method compared to state-of-the-art voxel-based reconstruction.

***Index Terms*—** Ultrasound reconstruction, neural implicit representation, 3D ultrasound imaging, unsupervised deep learning

1. INTRODUCTION

Ultrasound imaging is one of the most used medical imaging modalities, due to its relatively low-cost, ease-of-use and non-ionizing nature. The standard procedure in most clinical applications is to acquire temporal series of 2D images, i.e., slices, of the examined tissues. This considerably limits the ability of ultrasound imaging to represent the 3D geometry of organs, and motivates an intensive literature on 3D ultrasound imaging. 3D ultrasound volumes could represent a powerful tool to assist physicians in number of applications such as, for example, organ volume measurement [1], medical procedure assistance [2] or fetal examination. 3D ultrasound imaging can be obtained through several techniques: the most usual are based on mechanically-driven moving 1D arrays, or 2D arrays. However, the first option results into low-rate volumes and low-functionality, while the latter suffers from technological limitations related to the high channel count and limited field-of-views.

An alternative to the above-mentioned solutions is to reconstruct 3D ultrasound volumes from a collection of 2D slices acquired by moving manually a 1D ultrasound probe. To perform such a reconstruction, the position of each 2D image in space needs to be known. Three main approaches are usually employed to obtain these positions: mechanical guiding of the probe [3], computation of the relative position of the slices [4] or tracking the probe using six degrees of freedom (6DOF) position sensors, e.g., optical or magnetic trackers [5]. The particular interest of the latter has been shown in a number of applications (e.g. [6–8]).

In this work, we consider the challenging problem of US volume reconstruction from freehand 2D slices acquired by a 1D probe whose position is tracked by a magnetic 6DOF sensor attached to it. Since the seminal works on 3D ultrasound reconstruction, voxel grids have been adopted as the basis of most reconstruction methods [5, 9]. They can be classified in three categories: pixel-based (PBM), voxel-based (VBM) and function-based methods (FBM). Their main challenges are related to their real-time capabilities [9], the difficulty to fill large gaps in data, or to overcome ultrasound imaging limitations such as low spatial resolution and signal-to-noise ratio, or variable depth penetration of the ultrasound waves [6, 10]. PBM usually consist in two steps: bin-filling and hole-filling [11]. In the bin-filling step, the pixel values of each input 2D image are assigned to one or multiple voxels of the 3D volume, based on their position. In the hole-filling step, the voxels with no value assigned are filled using various methods such as fast marching [12] or olympic filling [11].

In contrast to PBM, VBM directly assign a value to each voxel of the volume of interest (VOI) based on the neighboring input images. This approach has the advantage of creating a fully filled 3D grid in one step and is the most common method in 3D ultrasound reconstruction [5]. However, VBM are usually sensitive to noise and spatial inconsistency, i.e., widely separated inputs or region imbalance. The most standard VBM algorithm is the voxel-nearest neighbor (VNN) [13], which assigns to each voxel the value of the closest pixel in the dataset. Complementarily, the distance-weighted (DW) [13, 14] method aggregates the values of multiple near pixels and weight them using their distance, allowing a smoother and more robust reconstruction.

Finally, FBM use the input dataset to estimate functions that describe the values of each voxel in the VOI. Commonly used functions range from polynomial to radial basis functions and Bézier spline [15]. Less commonly used in ultrasound imaging, FBM suffer from large computational time [9], but benefit from strong properties in regard to robustness to noise or missing data.

The objective of this paper is to propose an unsupervised deep learning approach for 3D ultrasound reconstruction. We took inspiration from the neural implicit representations (NIR), a family of approaches that learn volumetric functions from 3D samples [16, 17]. To the best of our knowledge, only two approaches have been proposed to adapt this concept to 3D ultrasound reconstruction [18, 19]. They are both based on a specific variant of NIR called neural radiance field (NeRF). NeRF-based networks learn a field of radiance that is used by a rendering engine (integration along the view directions) to generate novel views. Li *et al.* compared the results of a direct application of NeRF to the standard VNN algorithm for the creation of a spine volume to evaluate spine curvature [18]. Wysocki *et al.* used the output of a NeRF network as multiple parameters to render an ultrasound slice using a simulation process, and used those learned parameters to re-slice the volume [19].

This paper is organized as follows. Section 2 summarizes the concept of NIR. Section 3 describes the proposed method and the process of collecting the data. Section 4 regroups the simulated and experimental results, before reporting perspectives and conclusions.

2. NEURAL IMPLICIT REPRESENTATIONS

In the field of computer graphics, novel view synthesis consists in rendering a new points of view of a scene known only through other images. Recent works, based on the concept of implicit neural representations [16] and sinusoidal representation networks (SIRENs) [17], use deep learning to associate a set of parameters to any point of the scene through the example images. Using an algorithm such as ray marching [20], it becomes then possible to render an image from a new point of view using those parameters.

The forefront work in this field is neural radiance fields [21] (NeRF), which learns color and density associated with the position and view angle of any point in the scene. The density learned can then be used as a side product to estimate the underlying objects present in the scene. Moreover, some works extended this idea and pushed towards learning parameters used for volume reconstruction, such as neural implicit surface [22] (NeuS), which learns a signed distance field to better fit the objects of the scene.

One of the drawbacks of NeRF-based networks is their computational load, representing tens of hours of learning before convergence. However, with the introduction of instant neural graphic primitives [23] (instant NGP), which present a new data structure to accelerate learning, the computational time can be reduced by several orders of magnitude, i.e., from hours to seconds for comparable results [24].

3. MATERIAL AND METHOD

3.1. Overview

Inspired by neural implicit representations, we propose a formulation of the ultrasound volume reconstruction problem as the optimization of a 3D function (see Eq. (2)), and leverage deep-learning mechanisms and previous work on neural representation to approach the optimum. In the case of the freehand ultrasound volume reconstruction, different challenges compared to novel view synthesis arise, such as high speckle noise in the input data or shadowing due to ultrasound wave attenuation in tissues. The proposed framework aims at mitigating these challenges through the choice of the network and learning scheme.

In the seminal works on novel view synthesis methods using deep-learning, three main stage are usually identifiable: network architecture and training, output usage and data embedding. In this work, we focus on volume reconstruction and thus ignore for now the view angle component typical to NeRF networks and focus on the input data (i.e. specification, embedding) that represents the main difference between the targeted application and the existing literature.

3.2. Data acquisition and structure

To show the potential of the proposed method, simulated and experimental images are used. The simulated images are obtained by generating synthetic ultrasound images from a volume for different positions of the simulated ultrasound probe. The volumes considered are synthetic objects with perfectly known geometry, or 3D *in vivo* magnetic resonance volumes. The generation of the scatterers contributing to the simulation of each 2D slice follows the computationally efficient strategy that we have recently proposed in [25]. Data simulated in this way has the advantage of being perfectly controlled, i.e., the underlying medium used can serve as ground truth, the uncertainty of the simulated probe can be controlled, as well as the amount of noise in the images.

The first simulated dataset corresponded to a synthetic medium consisting of a highly echogenic $1.25^3 cm^3$ cube in a low echogenic homogenous medium. The ultrasound slices were generated by mimicking a sectorial sweep (the only movement is a probe tilt). The second dataset was simulated from a pelvic T2 MRI volume, by moving the simulated ultrasound probe in a double orthogonal sweep. The simulated sweeps are constituted respectively of 180 and 240 images. To mimic freehand acquisitions, positional and angular perturbations corresponding to zero-mean Gaussian distributions with standard deviations of 0.1 mm and respectively 0.03 rad, were generated.

The experimental data was acquired with an Ula-Op 64 scanner [26], with a 3.5-MHz LA523 probe from Esaote, equipped with a calibrated magnetic tracker that provides its 3D position in real-time. A synchronization system has been used to synchronize the 2D ultrasound image acquisition with the probe 3D position. Freehand acquisitions have been performed on a phantom mimicking a peripheral nerve block (CAE blue phantom from CAE Healthcare).

For both simulated and experimental cases, the raw data is a list of 2D images, associated with their dimensions and 3D position consisting in the 3D translation and rotation of each image center. Thus, each pixel in any 2D image can be associated with its position in the 3D real world. A collection of the pairs $(position; pixelvalue)$ represents the abstract base of the used data.

3.3. Network and data usage

Inspired by the seminal NeRF [21] and related works [22, 23], we adopt an architecture consisting of eight layers of 256 fully-connected perceptrons with ReLU activation to represent a function Φ_θ such that

$$\Phi_\theta(\gamma(\mathbf{x})) = d_x, \tag{1}$$

where $\mathbf{x}$ is a 3D vector whose elements are the 3D spatial coordinates and d_x is the value of an ultrasound slice at position $\mathbf{x}$. In other words, the aim of the network is to learn, through the iterative update of its parameter set θ, a continuous function Φ_θ that provides the value of the reconstructed volume at any 3D position, while respecting the input data, i.e., is equal to the available ultrasound slice pixel values. γ is a position encoding function that has an important role in the fidelity of the network and its convergence rate. Different choices will be detailed in the next subsection, and supporting results will be presented to highlight its impact on ultrasound volume reconstruction.

At each learning step, the network's results are compared against the expected value on a subset of known locations corresponding to the pixels of one of the ultrasound slices. The loss function is defined as the mean squared error between the network output and ultrasound slice pixel value, yielding to the following optimization process:

$$\arg\min_\theta ||\Phi_\theta(\gamma(\mathbf{x})) - d_x||_2^2. \tag{2}$$

3.4. Positional encoding

As explained previously, one of the key points in NeRF-like networks is the encoding of the input through the function γ in (1). This encoding phase is the first, non-learnable, layer of the network. Its role is to increase the dimension of the position vector $\mathbf{x}$, allowing the network to represent more finely spatial variations. In this work, we evaluate three encoding schemes, described hereafter.

3.4.1. Frequency encoding

The first encoding, proposed in the seminal work on NeRF [21], is a frequency encoding (also referred to as Fourier encoding) and is inspired by the conclusion of Rahaman *et al.* [27] that neural networks are spectrally biased towards low frequencies. Each component of

the input vector **x**, denoted by x in (3), is mapped from $\mathbb{R}$ to $\mathbb{R}^{2L}$ (with L being a meta-parameter) following a function $\gamma(x)$:

$$\gamma(x) = [sin(2^0\pi x), cos(2^0\pi x), \cdots, sin(2^{L-1}\pi x), cos(2^{L-1}\pi x)] \quad (3)$$

3.4.2. Hash grid encoding

Extending on the notion of positional encoding, Müller *et al.* [23] proposed a new approach to expend the input vector to more meaningful values through a process combining hierarchical acceleration structures and deep learning. A 3D coordinate is associated with a set of cells coming from 3D grids spanning from coarser to finer resolution. The final value vector is composed of the corner indexes, arbitrarily associated and hashed by a deep learning process to optimize their expressiveness. In the general case, this approach tremendously speeds up the learning process and allows finer details.

3.4.3. No encoding

As further emphasized and explained in the NeRF followup article [28], encoding aims at enhancing the high frequencies in the learned medium. While this property is of high interest in optical applications, it is not necessarily the case in 3D ultrasound reconstruction, because of the high level of speckle noise that affects the slices. Based on this observation, we also evaluate in this work the reconstruction of 3D ultrasound volumes without any encoding. In this case, the γ function reduces to the identity function, and thus no input data dimensionality augmentation is performed.

3.5. Volume generation

Training our network on a given ultrasound dataset results into a space weighing function Φ_θ, as described in 1. This can be seen in theory as a volume of infinite resolution (within the spatial bounds of the dataset) [27]. From Φ_θ, there are three possible approaches to obtain an exploitable ultrasound volume: (i) use it directly through a volume rendering process such as ray-marching [20], (ii) discretize a given isovalue using marching cube [29] or a similar algorithm, (iii) that discretizes the function in an arbitrarily sized voxel grid. To facilitate the comparison with existing voxel-based reconstruction methods, the latter is adopted in this work.

4. RESULTS

4.1. Impact of encoding

Before evaluating the accuracy of the proposed reconstruction method, we start by analyzing the impact of the encoding in the particular case of ultrasound applications, through the simulated dataset from 3D MRI. Fig. 1 shows an original MRI slice, the corresponding simulated ultrasound 2D image, and the output of our network with different or no encoding. To ensure a fair comparison, the same amount of reconstruction time was imposed for the three reconstruction schemes, instead of fixing the same number of iterations. One may observe that hash grid encoding is the best in learning the high-frequency content of ultrasound images. However, this is not a desirable effect in 3D ultrasound reconstruction, because of the high level of speckle. Frequency encoding shows good properties in removing the speckle, but introduces some undesirable artifacts. Finally, the result with no encoding forces the network to summarize the shape of the medium, yielding smoother results and creating a simplified volume. In our specific scenario, this may be highly desirable. This behavior could be linked with the deep image prior [30], i.e., stopping the network at the right time becomes central to satisfactory reconstruction. Based on this observation, only reconstruction results without encoding are reported hereafter.

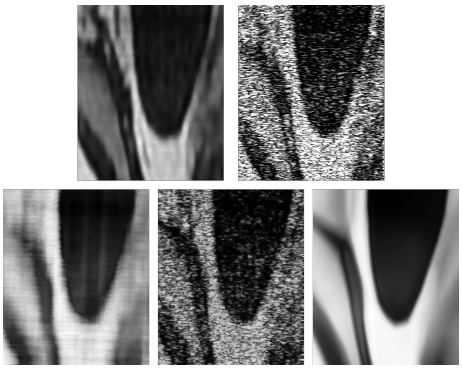

Fig. 1. Comparison of the reproduction of an ultrasound slice by a network trained on the same dataset for 10 minutes (reproduced slice not present in the learning set). Top: MRI slice, simulated ultrasound image. Bottom: reconstructed slice with frequency encoding, hash grid encoding, and without encoding.

4.2. 3D reconstruction results

This section regroups 3D ultrasound reconstruction results with the proposed network without encoding, in comparison with DW method typically used in ultrasound (see, e.g., [13]).

Fig. 2 shows the results on the first simulated dataset of a highly echogenic cube in a low echogenic homogenous medium. It shows the ability of the proposed network to remove noise from the volume and recreate a coherent geometry, while DW reconstruction struggles to efficiently eliminate speckle, resulting into a less-well defined cube. This result highlights the ability of our method to create a less noisy volume compared to DW, focusing on the shape of interest and recreating it more sharply.

Fig. 3 shows results on simulated ultrasound images from an MRI volume in a double orthogonal sweep. Using the larger amount of data at each spatial point due to crossing slices, the proposed method is clearly able to extract the shape of the medium. However, this generalization may come with a drawback, as the details in the ultrasound images tend to be smoothed. This results in gaps being closed and bumps erased in 3D reconstruction, but also leads to a easier-to-interpret reconstruction. Such result may be of high interest in some ultrasound reconstruction applications related to 3D localization or surgery assistance.

Fig. 4 shows the results of the experimental dataset. The proposed network is able to remove most of the noise in the acquisitions

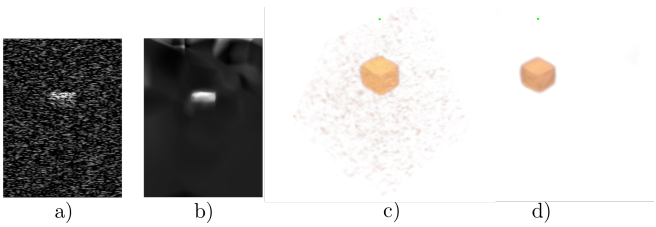

Fig. 2. Reconstruction based on a simulated sectorial sweep of a cube, with probe trajectory (green dot, top): a) ultrasound image given as an input to the network, b) network reconstruction of the image in a), c) DW reconstruction, d) 3-D reconstruction.

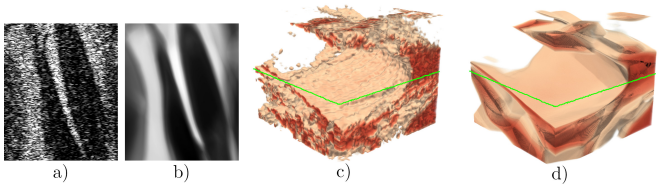

Fig. 3. Reconstruction based on a simulated double sweep (one horizontal, one vertical) of a T2 pelvic MRI based medium, with the probe trajectory shown in green. a) Ultrasound image given as an input to the network, b) reconstruction of the image in a) with the proposed method, c) DW reconstruction, d) 3D reconstruction with the proposed method.

and still display a faithful reconstruction, while some gaps in nerves have been closed and the overall shape is smoother.

Finally, Tab. 1 regroups numerical results of signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) for the three reconstruction results. Both measures have been realized on blocks of voxels extracted from inside (for SNR and CNR) and outside (for CNR) a region of interest (the hyper-echogenic cube, the uterus and one nerve) and computed using the formulas: $SNR = 20 \times \log_{10}(\frac{\mu_{in}}{\sigma_{in}})$ and $CNR = 20 \times \log_{10}(\frac{|\mu_{in} - \mu_{out}|}{\sqrt{\sigma_{in}^2 + \sigma_{out}^2}})$ with μ and σ being respectively the mean and the standard deviation of the regions. These quantitative results confirm the visual impression and strengthen our argument in favor of the proposed network reconstruction interest.

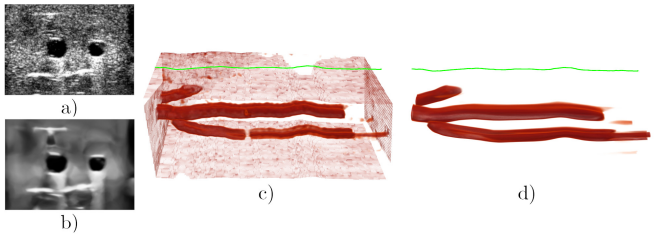

Fig. 4. Reconstruction based on a single horizontal freehand sweep acquired on a nerve phantom with probe trajectory (green). a) Ultrasound image given as an input to the network, b) Network reconstruction of the image in a), c) DW reconstruction, d) 3-D reconstruction.

5. DISCUSSION

5.1. Advantages and drawbacks of neural representation

Using neural representation has many advantages, namely, it can represent a continuous function, acts as a compact representation of the data [21], gives rise to desirable properties and takes advantage of the dynamic field it has originated. But this does not completely alleviate the inherent flaws of deep learning-based methods, such as for example their explainability on *in vivo* data. While the proposed method has a considerably short training time (of the order of a few minutes, on a single computer), it still tends to be outperformed in this respect by more baseline and massively parallel state-of-the-art methods.

5.2. Future work

5.2.1. View angle

Among the existing Neural Implicit Representations, one of the key aspects of the NeRF-like networks is their ability to take into account the view angle when evaluating a certain point in space. This

	Cube		MRI simulation		Nerve phantom	
	Ours	DW	Ours	DW	Ours	DW
SNR	**31.8**	26.1	**28.3**	25.5	**14.4**	12.6
CNR	**30.4**	22.4	**25.8**	20.5	**14.2**	12.3

Table 1. Signal-to-Noise Ratio (SNR) (dB) and Contrast-to-Noise Ratio (CNR) (dB) for the three experiments with DW and proposed methods.

represents a natural extension of the proposed work. The challenges imposed by ultrasound imaging such as depth attenuation (the same point may appear more or less bright depending on the distance from the probe) or shadowing (a bright object may not appear that way when positioned behind a strong impedance change) could be thus alleviated. Another important challenge is to consider the tissue deformation while scanning the medium with the ultrasound probe.

5.2.2. Data input in the learning process

In the current framework, the data points $(position; pixelvalue)$ are tightly tied to the slice notion (every point belongs to a slice, and an iteration covers one slice). However, there is no particular reason to do so, and may even strongly bias the network. Completely decorrelating the point representation could allow multiple improvements to the learning scheme of the network. One could for instance add a slight noise to point position each time they are presented to the network to enforce better generalization. Furthermore, designing an adapted encoding function to the reconstruction problem could accelerate the learning as the previously established functions, while maintaining the denoising and smoothing potential of the results presented in this article.

6. CONCLUSION

This work showed that the 3D freehand ultrasound reconstruction problem can be efficiently addressed using NIR and unsupervised deep learning, and showed encouraging results on denoising, shape generalization and visualization. Furthermore, this application can benefit from the advances of the NIR field while having its own set of specific challenges to consider. Many avenues of research remain to be explored, to enhance the current results that look already very promising.

7. COMPLIANCE WITH ETHICAL STANDARDS

This is a numerical simulation and phantom study, for which no ethical approval was required.

8. AKNOWLEDGEMENTS

No funding was received for conducting this study. The authors have no relevant financial or non-financial interests to disclose.

9. REFERENCES

[1] Aaron Fenster and Dónal B Downey, "Three-dimensional ultrasound imaging and its use in quantifying organ and pathology volumes," *Analytical and bioanalytical chemistry*, vol. 377, pp. 982–989, 2003.

[2] Shyam Natarajan, Leonard S. Marks, Daniel J.A. Margolis, Jiaoti Huang, Maria Luz Macairan, Patricia Lieu, and Aaron Fenster, "Clinical application of a 3d ultrasound-guided prostate biopsy system," *Urologic Oncology: Seminars and*

Original Investigations, vol. 29, no. 3, pp. 334–342, 2011, New Technology in Urology: Balancing Risk and Reward.

[3] Jeffrey Bax, Derek Cool, Lori Gardi, Kerry Knight, David Smith, Jacques Montreuil, Shi Sherebrin, Cesare Romagnoli, and Aaron Fenster, "Mechanically assisted 3d ultrasound guided prostate biopsy system," *Medical physics*, vol. 35, no. 12, pp. 5397–5410, 2008.

[4] Richard Housden, Graham Treece, Andrew Gee, and Richard Prager, "Calibration of an orientation sensor for freehand 3d ultrasound and its use in a hybrid acquisition system," *Biomedical engineering online*, vol. 7, pp. 5, 02 2008.

[5] Richard W Prager, AH Gee, and Laurence Berman, "Stradx: real-time acquisition and visualisation of freehand 3d ultrasound," 1998.

[6] Pierrick Coupé, Pierre Hellier, Noura Azzabou, and Christian Barillot, "3d freehand ultrasound reconstruction based on probe trajectory," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005*, James S. Duncan and Guido Gerig, Eds., Berlin, Heidelberg, 2005, pp. 597–604, Springer Berlin Heidelberg.

[7] Qinghua Huang, Zhaozheng Zeng, et al., "A review on real-time 3d ultrasound imaging technology," *BioMed research international*, vol. 2017, 2017.

[8] Mohammad I. Daoud, Abdel-Latif Alshalalfah, Falah Awwad, and Mahasen Al-Najar, "Freehand 3d ultrasound imaging system using electromagnetic tracking," in *2015 International Conference on Open Source Software Computing (OSSCOM)*, 2015, pp. 1–5.

[9] Farhan Mohamed and Chan Vei Siang, "A survey on 3d ultrasound reconstruction techniques," in *Artificial Intelligence*, Marco Antonio Aceves-Fernandez, Ed., chapter 4. IntechOpen, Rijeka, 2019.

[10] Hyungil Moon, Geonhwan Ju, Seyoun Park, and Hayong Shin, "3d freehand ultrasound reconstruction using a piecewise smooth markov random field," *Computer Vision and Image Understanding*, vol. 151, pp. 101–113, 2016, Probabilistic Models for Biomedical Image Analysis.

[11] DEO Dewi, MHF Wilkinson, TLR Mengko, IKE Purnama, PMA Van Ooijen, AG Veldhuizen, NM Maurits, and GJ Verkerke, "3d ultrasound reconstruction of spinal images using an improved olympic hole-filling method," in *International Conference on Instrumentation, Communication, Information Technology, and Biomedical Engineering 2009*. IEEE, 2009, pp. 1–5.

[12] Tiexiang Wen, Qingsong Zhu, Wenjian Qin, Ling Li, Fan Yang, Yaoqin Xie, and Jia Gu, "An accurate and effective fmm-based approach for freehand 3d ultrasound reconstruction," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 645–656, 2013.

[13] Robert Rohling, Andrew Gee, and Laurence Berman, "A comparison of freehand three-dimensional ultrasound reconstruction techniques," *Medical Image Analysis*, vol. 3, no. 4, pp. 339–359, 1999.

[14] C.D. Barry, C.P. Allott, N.W. John, P.M. Mellor, P.A. Arundel, D.S. Thomson, and J.C. Waterton, "Three-dimensional freehand ultrasound: Image reconstruction and volume analysis," *Ultrasound in Medicine & Biology*, vol. 23, no. 8, pp. 1209–1224, 1997.

[15] Qinghua Huang, Yanping Huang, Wei Hu, and Xuelong Li, "Bezier interpolation for 3-d freehand ultrasound," *IEEE Transactions on Human-Machine Systems*, vol. 45, pp. 385–392, 06 2015.

[16] Lars M. Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger, "Occupancy networks: Learning 3d reconstruction in function space," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4455–4465, 2018.

[17] Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein, "Implicit neural representations with periodic activation functions," 2020.

[18] 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)*, 2021, pp. 1–4.

[19] Magdalena Wysocki, Mohammad Farid Azampour, Christine Eilers, Benjamin Busam, Mehrdad Salehi, and Nassir Navab, "Ultra-nerf: Neural radiance fields for ultrasound imaging," 2023.

[20] K. Perlin and E. M. Hoffert, "Hypertexture," *SIGGRAPH Comput. Graph.*, vol. 23, no. 3, pp. 253–262, jul 1989.

[21] 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," *CoRR*, vol. abs/2003.08934, 2020.

[22] 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," 2023.

[23] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *ACM Trans. Graph.*, vol. 41, no. 4, pp. 102:1–102:15, July 2022.

[24] Yiming Wang, Qin Han, Marc Habermann, Kostas Daniilidis, Christian Theobalt, and Lingjie Liu, "Neus2: Fast learning of neural implicit surfaces for multi-view reconstruction," 2022.

[25] François Gaits, Nicolas Mellado, Gauthier Bouyjou, Damien Garcia, and Adrian Basarab, "Efficient stratified 3d scatterer sampling for freehand ultrasound simulation," 2023.

[26] Piero Tortoli, Luca Bassi, Enrico Boni, Alessandro Dallai, Francesco Guidi, and Stefano Ricci, "Ula-op: an advanced open platform for ultrasound research," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 10, pp. 2207–2216, 2009.

[27] Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred A. Hamprecht, Yoshua Bengio, and Aaron Courville, "On the spectral bias of neural networks," 2019.

[28] Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng, "Fourier features let networks learn high frequency functions in low dimensional domains," 2020.

[29] William E. Lorensen and Harvey E. Cline, "Marching cubes: A high resolution 3d surface construction algorithm," vol. 21, no. 4, pp. 163–169, aug 1987.

[30] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky, "Deep image prior," *International Journal of Computer Vision*, vol. 128, no. 7, pp. 1867–1888, mar 2020.