Neural rendering enables dynamic tomography

Ivan Grega,* **Vikram S. Deshpande**
University of Cambridge

William F. Whitney
Google DeepMind

Abstract

X-ray computed tomography (X-CT) is a prominent method to observe the internal composition of objects. It requires solving the inverse problem of reconstructing the 3d object from 2d X-ray projections. Thousands of projections are typically needed to obtain a detailed reconstruction. Therefore, it has never been possible to reconstruct material deformation under dynamic high-speed loading, which has impeded our understanding of high-speed deformation processes. In this work, we address this limitation by combining high-fidelity X-CT with differentiable neural rendering. In our two-stage approach, we first reconstruct the canonical volume and then use a neural network to predict a temporal deformation field with a cubic spline output parametrization. We demonstrate the reconstruction of deforming objects from very few (two) projections which enables a paradigm shift in reconstruction of dynamic experiments.

1 Introduction

X-ray computed tomography (X-CT) is an established method in medical imaging and materials research. [1, 22] To carry out X-CT, it is necessary to rotate the object or detector to obtain thousands of projections of an object from different angles. In recent years ML methods have helped reduce the number of required projections and the associated radiation dose. [7] However, while the number of required projections can be reduced to a few hundred, 3d reconstruction of high-speed dynamic experiments is still not possible due to the prohibitively high rotation rates which would be required.

Therefore, while much has been discovered about the behaviour of materials (bulk, composites, architected) at *quasi-static* conditions using interrupted tomography, whereby deformation sequence is interrupted at discrete steps [16, 21], 3d characterization of *dynamic* deformations has never been possible. Here we address this limitation. As key contributions, we:

*ig348@cam.ac.uk

Machine Learning and the Physical Sciences Workshop, NeurIPS 2024.

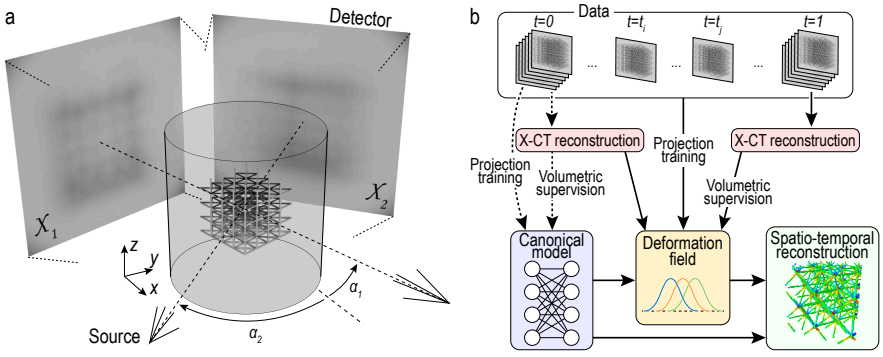

Figure 1: *a)* Projection geometry. *b)* Overview of the method.

- adapt neural radiance fields (NeRF) as a canonical model to capture volumetric density under the assumptions of x-ray attenuation (Beer-Lambert law);
- introduce cubic spline-based deformation field with continuous time parametrization;
- develop a hybrid framework where the high-fidelity X-CT information is combined with sparse projection views;
- and demonstrate experimentally that these techniques enable the reconstruction of dynamic deformations from only 2 projections per intermediate timeframe. [2]

2 Background and related work

Few-projection reconstruction Projection-based digital volume correlation was introduced by Leclerc et al. [8] and Taillandier-Thomas et al. [18] where the authors used a canonical reconstruction of the initial volume and few projections during deformation in combination with a finite element mesh and elastic regularization. The downside of elastic regularization is that a material model has to be assumed *a priori*, which defeats the purpose of DVC if the goal is to find out the material model. The identification of material model from projection was done by Jailin et al. [6]. However, their sample was a thin quasi-two-dimensional dogbone specimen of low complexity.

NeRF-based approaches The use of neural rendering techniques has been demonstrated in X-ray imaging, however most of the approaches rely on pre-training the network on a dataset of objects. [17, 2, 24] This is normally needed because of the ill-posedness of the problem and high number of the degrees of freedom. However, we wish to circumvent this limitation. Our goal is to reconstruct dynamic experiments without pre-training the model on pre-existing database and without assuming a material model *a priori*.

Deformable NeRFs Neural radiance approaches for the reconstruction of dynamic scenes can be broadly classified on the spectrum between full spatio-temporal networks on the one side and methods which decouple canonical volume from deformation field on the other side of the spectrum. Full spatio-temporal models are more difficult to optimize and many types of loss are needed, but they are quite flexible in allowing the topology of the scene to change over time.[23, 3, 9] On the other hand, two-component frameworks with a canonical field and a deformation field are computationally more efficient. However, they are limited to scenes in which temporal changes in the scene are topology-preserving. [12, 5, 20, 10, 13, 14] Other hybrid approaches exist such as low-rank decomposition of the parameters of spatio-temporal field [11, 4] or higher-dimensional hyperspace [13].

3 Methods

Overview The geometry of the X-ray projection setup is illustrated in Figure 1a. X-rays are emitted from the source as a conical beam and pass through the sample, after which the attenuated X-ray intensity is captured at the detector. Multiple projections $\mathcal{X}_i$ are captured at corresponding angles α_i. The data consists of projections recorded at 50 deformation timesteps ($0 \leq t \leq 1$). At steps $t = \{0, 1\}$, we record 1024 projections which enables full X-CT reconstruction, while at the other timesteps only 2 projections separated by 90° are captured. For NeRF reconstruction we use a two-component framework based on a canonical volume and a deformation field similar to the work by Park et al. [12] (Fig. 1b). The architecture of the canonical volume is based on the default *nerfacto* method in `nerfstudio` framework [19]. Unlike in traditional scene reconstruction, there is no color output and appearance does not depend on viewing direction. Therefore, the neural radiance field in our case is a mapping from coordinates to density $\Theta(X_1, X_2, X_3) \rightarrow \sigma$. Rendering with the Beer-Lambert law leads to the equation for transmittance along ray: $T = A_0 \exp\left(-\int \sigma(s)ds\right)$.

Hybrid projection and volumetric reconstruction We develop a hybrid method in which we combine projection loss and volumetric loss (with respect to high-fidelity X-CT reconstructions which are available at $t = 0$ and $t = 1$. The canonical volume is fitted first for $t = 0$ using a combination of 128 equispaced projections at $t = 0$ and high-fidelity X-CT reconstructions (dashed lines in Fig. 1b). The deformation field is then trained (solid lines) using all available data (128 projections at

[2] Further visualizations of the results can be found at `https://neural-xray.github.io/nerfxray/`.

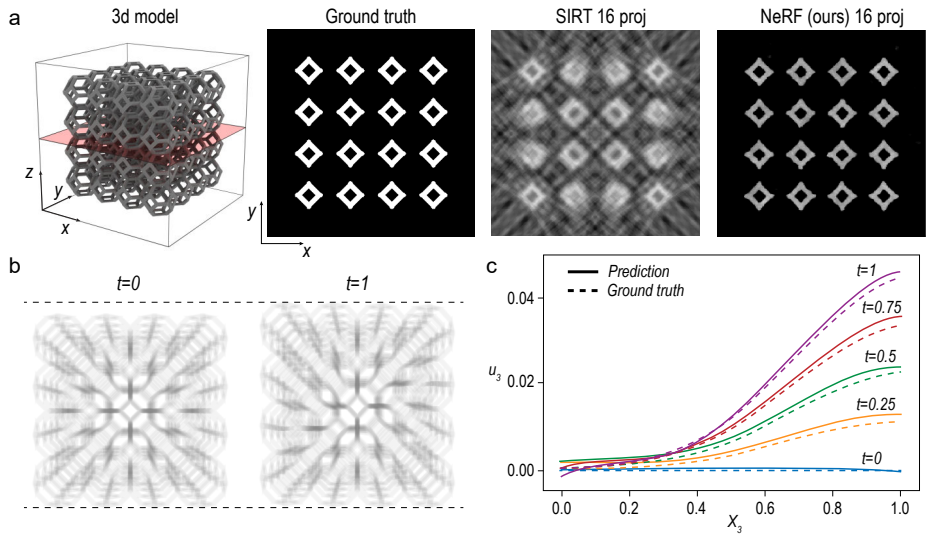

Figure 2: Simulated experiment vertically stretching a Kelvin lattice. *a)* 3d model of lattice and corresponding cross-sections at the highlighted plane. *b)* Simulated X-ray projections used as training data. *c)* Reconstructed displacement field u_3 at various timesteps t. Locations near the top of the object are stretched farther.

$t = 0$, 128 projections at $t = 1$, 2 projections per timestep at $0 < t < 1$, and high-fidelity X-CT reconstructions at $t \in \{0, 1\}$).

Spline-based deformation field Deformation can be expressed as three-component displacement field u_i which is a mapping between the reference (canonical) configuration, X_i and the deformed configuration, x_i, such that $x_i = X_i + u_i$ with $i = 1, 2, 3$ representing x, y, z axes. We model the displacement field using cubic B-spline interpolation based on the work by Rueckert et al. [15]:

$$u(x_i, t) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(\tilde{x}_1) B_m(\tilde{x}_2) B_n(\tilde{x}_3) w(t)_{i_1+l, i_2+m, i_3+n}$$

where $B_i(x)$ is fixed cubic B-spline function, i_1, i_2, i_3 are indices of the relevant grid points, $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ are modulo-subtracted coordinates, and $w(t)_{i_1,i_2,i_3}$ is a trainable weight of the corresponding spline control point. We modified the standard B-spline field to obtain a time-dependent deformation field $u(x_i, t)$ by extracting the weights from an MLP $\Omega(t) \rightarrow w(t)_{i_1,i_2,i_3}$. Enabling continuous parametrization of time has practical implications; for instance, it relaxes the requirements for temporal synchronization of multiple x-ray sources and detectors. While optimization of the deformation field is done jointly on all available timesteps, it was found desirable to prioritize early deformation timesteps during early training steps. Furthermore, the B-spline grid is gradually refined during training.

4 Results

Data efficiency of neural rendering We first establish that NeRFs are an efficient representation for X-ray data using simulated projections of a lattice (Fig. 2a). Cross-sectional slices and normalized correlation metric in Table 1 reveal that when relatively few projections are available, NeRF reconstruction significantly outperforms traditional X-CT method SIRT.

# proj	3	9	16
NeRF (ours)	**0.40**	**0.82**	**0.91**
SIRT	0.29	0.48	0.61

Table 1: Normalized correlation metric for reconstruction of simulated lattice projections.

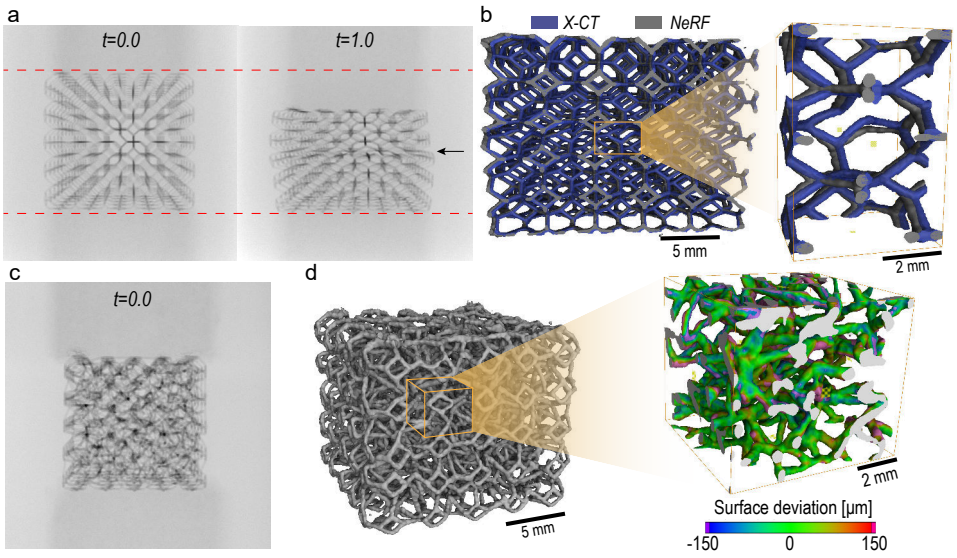

Figure 3: Real-world experiments compressing lattices with a plane of defects *(a,b)* and randomized microstructure *(c,d)*. *a)* X-ray projections with arrow indicating localized deformation. *b)* NeRF and X-CT reconstructions at $t = 0.6$. *c)* X-ray projection of randomized lattice. *d)* NeRF reconstruction at $t = 0.6$. Surface deviations compared to withheld X-CT data shown as inset.

Verification with simulated data First we verified the optimization of the deformation field with simulated data where the deformation field was known. In Fig. 2b we show simulated projections at $t = 0$ and $t = 1$. The fitted displacement u_3 matched the ground truth closely not only at $t = \{0, 1\}$ where high-fidelity volumetric data was available, but also at intermediate timesteps where only 2 projections were available (Fig. 2c). Ablation of the optimization method revealed that the hybrid method with high-fidelity X-CT data at $t \in \{0, 1\}$ and gradual refining of the B-spline grid are important for the reliable convergence of the deformation to the ground truth.

Reconstruction of deformation experiments We applied the framework to two real-world deformation experiments: Kelvin lattice with a plane of defects (Fig. 3a,b) and a lattice with randomized microstructure (Fig. 3c,d); both undergoing compression. Due to the plane of defects in the first lattice, deformation is localized to a thin band (Fig. 3a). In Fig. 3b we compare the reconstructed NeRF with withheld X-CT data at an intermediate timestep $t = 0.6$. Small deviations are perceptible in the reconstructed surface but the overall match is excellent. The second lattice with randomized microstructure is chosen to probe the capacity of the framework to discriminate between density along occluded ray paths (cf. the lack of periodicity between Fig. 3a and c). NeRF reconstruction at intermediate timestep $t = 0.6$ is shown in Fig. 3d. A small cube from the centre of the specimen is isolated and shown as a zoomed inset. The color map corresponds to the surface deviations between the NeRF reconstruction and withheld high-fidelity X-CT. Considering the scale bar and that the resolution of the X-CT is $76\,\mu\text{m}$, the match between the two reconstructions is remarkable.

5 Limitations and conclusions

We adapted neural rendering techniques to the X-ray modality and introduced a two-component method with NeRF canonical model and cubic spline deformation field with continuous time parametrization. We make use of the high-fidelity reconstruction at the beginning and end of the experiment to train the deformation field using only two projections at intermediate stages of deformation. The framework is demonstrated with one simulated and two experimental datasets.

The two-component framework based on the canonical model and deformation field is limited to capturing topology-preserving deformations. While it is not a problem for many experimental systems, a more expressive framework should be used if damage and fracture are present in the deformation.

Broader impact statement

The presented framework has the potential to achieve a paradigm shift in dynamic in-situ experiments visualized by X-ray. Up until now, it was only possible to study materials in 3d under quasi-static conditions using interrupted in-situ X-ray tomography in which X-CT reconstruction with thousands of projections was acquired at every deformation step. Here we develop a framework in which few (two) projections at intermediate stages of deformation are sufficient to obtain a 3d reconstruction. For the first time, it will become possible to develop experiments where high-speed dynamic behaviour of materials is studied in 3d using simultaneous acquisition from few detectors.

References

[1] Simone Carmignato, Wim Dewulf, and Richard Leach, editors. *Industrial X-Ray Computed Tomography*. Springer International Publishing, 2018. ISBN 9783319595719. doi: 10.1007/978-3-319-59573-3. URL `http://dx.doi.org/10.1007/978-3-319-59573-3`.

[2] 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.

[3] Yilun Du, Yinan Zhang, Hong-Xing Yu, Joshua B Tenenbaum, and Jiajun Wu. Neural radiance flow for 4d view synthesis and video processing. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 14304–14314. IEEE Computer Society, 2021.

[4] Sara Fridovich-Keil, Giacomo Meanti, Frederik Rahbæk Warburg, Benjamin Recht, and Angjoo Kanazawa. K-planes: Explicit radiance fields in space, time, and appearance. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12479–12488, 2023.

[5] Xiang Guo, Guanying Chen, Yuchao Dai, Xiaoqing Ye, Jiadai Sun, Xiao Tan, and Errui Ding. Neural deformable voxel grid for fast optimization of dynamic view synthesis. In *Proceedings of the Asian Conference on Computer Vision*, pages 3757–3775, 2022.

[6] Clément Jailin, Ante Buljac, Amine Bouterf, François Hild, and Stéphane Roux. Fast four-dimensional tensile test monitored via x-ray computed tomography: Elastoplastic identification from radiographs. *The Journal of Strain Analysis for Engineering Design*, 54(1):44–53, 2019.

[7] Hyojin Kim, Rushil Anirudh, K Aditya Mohan, and Kyle Champley. Extreme few-view ct reconstruction using deep inference. *arXiv preprint arXiv:1910.05375*, 2019.

[8] Hugo Leclerc, Stéphane Roux, and François Hild. Projection savings in ct-based digital volume correlation. *Experimental Mechanics*, 55(1):275–287, 2015.

[9] Zhengqi Li, Simon Niklaus, Noah Snavely, and Oliver Wang. Neural scene flow fields for space-time view synthesis of dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6498–6508, 2021.

[10] Jia-Wei Liu, Yan-Pei Cao, Weijia Mao, Wenqiao Zhang, David Junhao Zhang, Jussi Keppo, Ying Shan, Xiaohu Qie, and Mike Zheng Shou. Devrf: Fast deformable voxel radiance fields for dynamic scenes. *Advances in Neural Information Processing Systems*, 35:36762–36775, 2022.

[11] Marko Mihajlovic, Sergey Prokudin, Marc Pollefeys, and Siyu Tang. Resfields: Residual neural fields for spatiotemporal signals. *arXiv preprint arXiv:2309.03160*, 2023.

[12] Keunhong Park, Utkarsh Sinha, Jonathan T Barron, Sofien Bouaziz, Dan B Goldman, Steven M Seitz, and Ricardo Martin-Brualla. Nerfies: Deformable neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5865–5874, 2021.

[13] Keunhong Park, Utkarsh Sinha, Peter Hedman, Jonathan T Barron, Sofien Bouaziz, Dan B Goldman, Ricardo Martin-Brualla, and Steven M Seitz. Hypernerf: A higher-dimensional representation for topologically varying neural radiance fields. *arXiv preprint arXiv:2106.13228*, 2021.

[14] Albert Pumarola, Enric Corona, Gerard Pons-Moll, and Francesc Moreno-Noguer. D-nerf: Neural radiance fields for dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10318–10327, 2021.

[15] Daniel Rueckert, Luke I Sonoda, Carmel Hayes, Derek LG Hill, Martin O Leach, and David J Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18(8):712–721, 1999.

[16] Angkur Jyoti Dipanka Shaikeea, Huachen Cui, Mark O'Masta, Xiaoyu Rayne Zheng, and Vikram Sudhir Deshpande. The toughness of mechanical metamaterials. *Nature materials*, 21(3):297–304, 2022.

[17] Mengcheng Sun, Yu Zhu, Hangyu Li, Jiongyao Ye, and Nan Li. Acnerf: enhancement of neural radiance field by alignment and correction of pose to reconstruct new views from a single x-ray. *Physics in Medicine & Biology*, 69(4):045016, 2024.

[18] Thibault Taillandier-Thomas, Clément Jailin, Stéphane Roux, and François Hild. Measurement of 3d displacement fields from few tomographic projections. In *Optics, photonics and digital technologies for imaging applications IV*, volume 9896, pages 99–110. SPIE, 2016.

[19] Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, Abhik Ahuja, et al. Nerfstudio: A modular framework for neural radiance field development. In *ACM SIGGRAPH 2023 Conference Proceedings*, pages 1–12, 2023.

[20] Edgar Tretschk, Ayush Tewari, Vladislav Golyanik, Michael Zollhöfer, Christoph Lassner, and Christian Theobalt. Non-rigid neural radiance fields: Reconstruction and novel view synthesis of a dynamic scene from monocular video. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12959–12970, 2021.

[21] Zifan Wang, Shuvrangsu Das, Akshay Joshi, Angkur JD Shaikeea, and Vikram S Deshpande. 3d observations provide striking findings in rubber elasticity. *Proceedings of the National Academy of Sciences*, 121 (24):e2404205121, 2024.

[22] Philip J. Withers, Charles Bouman, Simone Carmignato, Veerle Cnudde, David Grimaldi, Charlotte K. Hagen, Eric Maire, Marena Manley, Anton Du Plessis, and Stuart R. Stock. X-ray computed tomography. *Nature Reviews Methods Primers*, 1(1), 2 2021. ISSN 2662-8449. doi: 10.1038/s43586-021-00015-4. URL `http://dx.doi.org/10.1038/s43586-021-00015-4`.

[23] Wenqi Xian, Jia-Bin Huang, Johannes Kopf, and Changil Kim. Space-time neural irradiance fields for free-viewpoint video. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9421–9431, 2021.

[24] Yanjie Zheng and Kelsey B Hatzell. Ultrasparse view x-ray computed tomography for 4d imaging. *ACS Applied Materials & Interfaces*, 15(29):35024–35033, 2023.