



HAL
open science

Neural fields for sparse-view 3D X-ray imaging: preliminary work

Romain Vo, Julie Escoda, Caroline Vienne, Adrien Stolidi, Etienne Decencière

► **To cite this version:**

Romain Vo, Julie Escoda, Caroline Vienne, Adrien Stolidi, Etienne Decencière. Neural fields for sparse-view 3D X-ray imaging: preliminary work. ORASIS 2023 - journées francophones des jeunes chercheurs en vision par ordinateur, Laboratoire LIS, UMR 7020, May 2023, Carqueiranne, France. hal-04219565

HAL Id: hal-04219565

<https://hal.science/hal-04219565v1>

Submitted on 27 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Neural Fields for Sparse-View 3D X-ray Imaging: preliminary work

Romain Vo^{1,2} Julie Escoda¹ Caroline Vienne¹ Adrien Stolidi¹ Étienne Decencière²

¹ Université Paris-Saclay, CEA, List, F-91190, Palaiseau, France

² Mines Paris, PSL University, Centre for mathematical morphology (CMM), 77300 Fontainebleau, France

romain.vo@cea.fr

Résumé

La tomographie par rayons X est de plus en plus utilisée dans de nombreux domaines industriels pour sa capacité unique à contrôler l'intégrité et la conformité dimensionnelle des pièces. Pourtant, elle ne parvient pas à être adoptée comme technique standard de contrôle en ligne en raison de son coût excessif en temps d'acquisition. Dans ce travail, nous développons une méthodologie hybride, qui combine différentes techniques d'apprentissage profond pour effectuer une reconstruction tomographique à partir de peu de vues sur des données 3D à faisceau conique.

Mots Clef

Reconstruction Tomographique, Rayons X, Deep learning, Problème inverse, Régularisation par débruitage, Champ de neurones.

Abstract

X-ray Computed Tomography has been increasingly used in many industrial domains for its unique capability of controlling both the integrity and dimensional conformity of parts. Still, it fails to be adopted as a standard technique for on-line monitoring due to its excessive cost in terms of acquisition time. In this work we develop a hybrid methodology, which combines state of the art techniques to perform CT reconstruction from few views using deep learning on 3D cone-beam data.

Keywords

Computed Tomography, X-ray, Deep learning, Inverse problem, Regularization by Denoising, Neural fields.

1. Introduction

X-ray CT is a non-destructive evaluation tool used to inspect and fully visualize the inner structure of a 3D physical object. It is widely used in medical imaging, industrial inspection and material science for the visualization of fine

details it provides. For industrial inspection, CT fails to be adopted as a standard technique for on-line monitoring due to its excessive cost in terms of acquisition time. In this context, the reconstruction of CT images is equivalent to solving an *ill-posed* inverse problem, where the linear system is underdetermined.

Common strategies to solve *ill-posed* inverse problems rely on the concept of variational regularization. The objective is to combine knowledge of the forward operator in the form of a data fidelity term to minimize, and regularize the solution by using a penalty term. Recently, guided by the effectiveness of deep learning (DL) and convolutional neural networks (CNN) on computer vision tasks there has been a growing interest for *data-driven* techniques to solve inverse problems, including the reconstruction of X-ray CT [1].

2. Background

Inverse problems

The CT image reconstruction problem can be formulated as an ill-posed linear system:

$$b = Af^\dagger + \epsilon, \quad (1)$$

where one seeks to recover the true underlying object $f^\dagger \in \mathbb{R}^n$ from a sparse set of measurements $b \in \mathbb{R}^m$, where $m \ll n$. In CT we consider the *Radon Transform* as the forward operator $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ that maps the object to the measurements. The noise ϵ is a random variable that models the measurement noise.

A standard approach for CT reconstruction is to solve a composite optimization problem by minimizing an objective function composed of a data fidelity term $\mathcal{L}(f) = \frac{1}{2} \|Af - b\|_2^2$ and a regularization term $\mathcal{R}(f)$ which encodes *a priori* information on f and penalizes unlikely solutions:

$$\hat{f} = \arg \min_{f \in \mathbb{R}^n} \mathcal{L}(f) + \lambda \mathcal{R}(f). \quad (2)$$

Neural fields

Neural fields are a recent class of deep learning models that have seen widespread success in 3D shape modelling, novel view synthesis or surface reconstructions [2, 3].

The main idea behind neural fields is to represent the solution f to the inverse problem as a *field*, a quantity defined for all spatial coordinates, and to parametrize this field by a neural network F_Φ with learned weights Φ . In 3D CT we consider the attenuation field which maps a 3D coordinate $\mathbf{x} \in \mathbb{R}^3$ to a scalar value. Considering the optimization problem in (2), the neural field parameters Φ can be recovered by solving:

$$\arg \min_{\Phi} \mathcal{L}(F_\Phi) + \lambda \mathcal{R}(F_\Phi). \quad (3)$$

Regularization by Denoising (RED)

Regularization by Denoising [4] is a recent framework for regularized imaging inverse problems which allows seamless integration of many *denoising engine* in the regularization functional. The regularization term ρ proposed by Romano *et al.* [4] is defined as follows:

$$\rho(f) = f^T (f - D(f)), \quad (4)$$

here $D : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a denoising engine that takes as input the noisy image f and outputs a denoised version of f . Given some assumptions on the denoiser D , Romano *et al.* further show that the gradient of the regularization term ρ w.r.t f can be obtained without the need of explicitly differentiating f . When the function D is differentiable, has a symmetric Jacobian and is locally homogeneous, the gradient of ρ can be computed exactly as:

$$\nabla \rho(f) = f - D(f), \quad (5)$$

3. Proposed methodology

In this work we propose to merge the two approaches to efficiently optimize a neural field for the reconstruction of CT volumes from a sparse set of views. The two methods are complementary, neural fields reduce the memory requirements for reconstruction when working with high-resolution data. The RED framework allows to decouple the regularization and the data-fidelity and allows to use *off-the-shelf* artefact removal networks.

To evaluate the approach we work with a dataset of natural cork samples of size $(1024 \times 1024 \times 1024)$ and SVCT reconstruction using only 50 views out of 720 available. For the artefact removal engine D_θ we trained a classic UNet [5] to map a sparse reconstruction slice to a dense reconstruction slice. Then we trained the neural field F_Φ using Eq. (3) and show promising performances on our dataset.

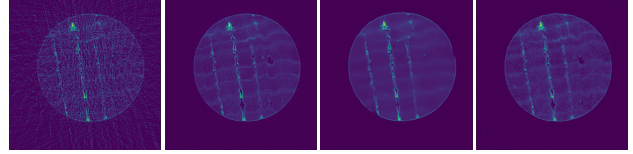


Figure 1. Illustration of natural cork samples reconstructions. From left to right : sparse and dense reconstructions then post-processing of the sparse reconstruction using only D_θ and reconstruction using the regularized neural field F_Φ .

4. Conclusion and perspectives

In this paper we tackle the challenge of reconstructing highly-resolved 3D CT volumes from a sparse set of views acquired via a cone-beam geometry setup. The following part of work will focus on thoroughly evaluating the approach on our dataset. Finally we would like to validate the method on public benchmarks in the field of CT reconstruction.s

References

- [1] J. Leuschner, M. Schmidt, P. S. Ganguly, V. Andriashen, S. B. Coban, A. Denker, D. Bauer, A. Hadjifaradji, K. J. Batenburg, P. Maass, and M. van Eijnatten, “Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications,” *Journal of Imaging*, vol. 7, p. 44, Mar. 2021. 1
- [2] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, “NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis,” in *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I*, (Berlin, Heidelberg), pp. 405–421, Springer-Verlag, Aug. 2020. 2
- [3] Y. Xie, T. Takikawa, S. Saito, O. Litany, S. Yan, N. Khan, F. Tombari, J. Tompkin, V. Sitzmann, and S. Sridhar, “Neural Fields in Visual Computing and Beyond,” *Computer Graphics Forum*, vol. 41, pp. 641–676, May 2022. 2
- [4] Y. Romano, M. Elad, and P. Milanfar, “The Little Engine That Could: Regularization by Denoising (RED),” *SIAM Journal on Imaging Sciences*, vol. 10, pp. 1804–1844, Jan. 2017. 2
- [5] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds.), Lecture Notes in Computer Science, (Cham), pp. 234–241, Springer International Publishing, 2015. 2