DTh1D.4 OSA Imaging and Applied Optics Congress 2021 (3D, COSI, DH, ISA, pcAOP) © OSA 2021

Limited-view Cone Beam CT reconstruction using 3D Patch-based Supervised and Adversarial Learning

Anish Lahiri[1], Marc L. Klasky[2], Jeffrey A. Fessler[1], and Saiprasad Ravishankar[3]

[1]Dept. of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI 48109, [2]Los Alamos National Laboratory, Los Alamos, NM 87545, [3]Depts. of Computational Math Science & Engineering, and Biomedical Engineering, Michigan State University, East Lansing, MI 48824

Abstract

We present a novel multi-stage algorithm for CBCT reconstruction from very limited projections. Our proposed method uses 3D patch-based supervised and adversarial learning from scarce training data, combined with physics (forward) models and statistical priors.

Introduction

Computed Tomography (CT) has become commonplace in a wide variety of applications ranging from healthcare applications, materials science, global security, inertial confinement fusion, as well as a number of applications of interest to the stockpile stewardship program. However, in order to accurately reconstruct a volume of size $N \times N \times N$, traditional techniques require N sets of projections at angles evenly spaced between 0° and 180° degrees. However, in many applications this is not possible or would require excessive time for data acquisition. Furthermore, in many dynamic tomography applications, only a very limited set of measurements is possible.

In ill-posed settings such as sparse/limited-view tomography or limited-angle tomography, traditional reconstruction techniques, such as filtered back-projection (FBP) or maximum likelihood estimation, are known to produce images with severe streaking artifacts. These problems can be ameliorated with improved system modeling and the inclusion of regularization. However, even these approaches break down as the angular subsampling goes above a factor of 20 [3]. Consequently, they are not sufficient in an many applications of interest such as in dynamic radiography.

In recent years, there have been tremendous advances in limited-view tomographic reconstruction in the medical physics and security arenas. These advances have been largely the result of the introduction of machine learning and deep learning (DL) approaches to limited-view CT reconstructions (see review in [6]). While significant enhancements to limited-view reconstructions may be realized, the training of deep networks typically entails large representative training sets that are often difficult to obtain. Consequently, our objective in this work is to develop machine learning-based reconstruction pipelines that enable highly accurate reconstructions using *extremely limited views as well as very limited training data.*

Algorithm and Methods

Our proposed approach for cone beam CT (CBCT) reconstruction from limited views adopts an iterative strategy (somewhat similar to [1, 5]), combining both physics-based forward models and machine learned components. In each iteration or hyperlayer of our algorithm, we use a combination of deep supervised and adversarial learning-based de-streaking of a corrupted estimate of the object being reconstructed, followed by a data-consistency (i.e., imaging forward and noise model-consistent) and regularized update of the resulting volume. An edge-preserving regularized reconstruction [8] is used to initialize the first iteration of our algorithm. A pictorial depiction of this process is shown in Fig. 1.

Importantly, to enable learning from scarce training data, we adopt a 3D patch-based approach to learning and de-streaking CBCT subvolumes. During both training and testing, the inputs to the deep generative adversarial network (GAN) are smaller sub-volumes of the CBCT image volume. Essentially, this yields several training examples in the form of overlapping (3D) patches from a limited number of image volumes. Moreover, as depicted in Fig. 1, for our generator, we use a deep CNN that maps 3D input volumes (containing a 2D slice that is to be de-streaked and its spatially neighboring slices) to 2D slices [4]. This enables us to use 3D contextual information. The training targets for the 'supervised' generator are set to be the ground truth central slice corresponding to the input sub-volume to the network. The cost for training the generator and discriminator for the kth 'iteration' (or *hyperlayer*) of our method are as follows:

$$\min_{\theta_G} \; -\lambda \mathbb{E}\big[D(G(\mathcal{P}_3 \boldsymbol{x}_{k-1};\theta_G))\big] + \mathbb{E}\big[\|\mathcal{P}_3^{2,\text{mid}} \boldsymbol{x}_{\text{GT}} - G(\mathcal{P}_3 \boldsymbol{x}_{k-1})\|_1\big] \tag{1}$$

$$\min_{\theta_D} \; \mathbb{E}\big[D(G(\mathcal{P}_3 \boldsymbol{x}_{k-1});\theta_D)\big] - \mathbb{E}\big[D(\mathcal{P}_3^{2,\text{mid}} \boldsymbol{x}_{\text{GT}};\theta_D)\big], \tag{2}$$

where D and G denote the discriminator and generator networks respectively and θ_D and θ_G are the corresponding network weights. $\boldsymbol{x}_{\text{GT}}$ is a ground truth image volume, $\boldsymbol{x}_{k-1}$ is the output volume from the $(k-1)$th stage. $\mathcal{P}_3$ is a 3D patch extraction operator, and $\mathcal{P}_3^{2,\text{mid}}$ is a 2D patch extraction operator that extracts the slice in the central position corresponding to the patch extracted by $\mathcal{P}_3$. The parameter λ balances the weighting adversarial vs. supervised learning with the second term in the loss (1) denoting ℓ_1 reconstruction loss (expectation computed over all possible patches from ground truth volumes). During testing, all overlapping subvolumes in an input volume $\boldsymbol{x}_{k-1}$ are processed using the generator network, and then aggregated into a single volume before being passed through a data-consistency unit that optimizes the following cost:

$$\boldsymbol{x}_k = \min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \beta \|\boldsymbol{x} - \boldsymbol{x}_{\text{G}}\|_2^2, \tag{3}$$

OSA Imaging and Applied Optics Congress 2021 (3D, COSI, DH, ISA, pcAOP) © OSA 2021

where $\boldsymbol{A}$ is the CT system matrix, $\boldsymbol{y}$ are the acquired CT sinogram (post-logarithm) measurements, $\boldsymbol{x}_{\mathrm{G}}$ is the patch aggregated image obtained from the generator, and β is a regularization parameter. The generators and discriminators in our proposed algorithm are trained one hyperlayer at a time in a greedy fashion based on outputs from previous layer.

In our experiments, we used $500 \times 500 \times 8$ overlapping patches from one walnut in a walnut CT dataset in [7] for training our patch-based architecture, while a different walnut is used for testing. The MIRT toolbox [2] was used to generate measurements for limited-view CBCT with 8 acquired projections. An FDK reconstruction was used to initialize the edge-preserving regularized [6] reconstruction.

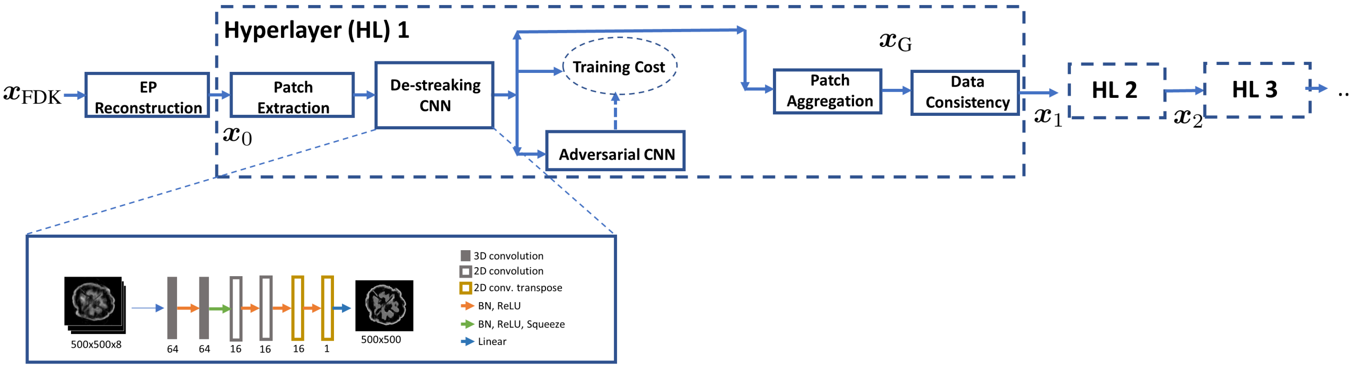

Figure 1: Proposed iterative algorithm for cone Beam CT reconstruction, each stage or *hyperlayer* of our algorithm has the same composition as depicted for Hyperlayer 1.

Results and Discussion

Fig. 2 shows the results of reconstruction of a test walnut volume from 8 projections using our proposed method. The mean absolute error (also depicted in the figure) improves significantly across the stages of our algorithm, until it stabilizes after a few stages. However, we observed that the reconstructions show some edge artifacts that grow prominent after several stages of our algorithm. This may be because training several stages of the proposed algorithm on a single walnut may lead to overfitting.

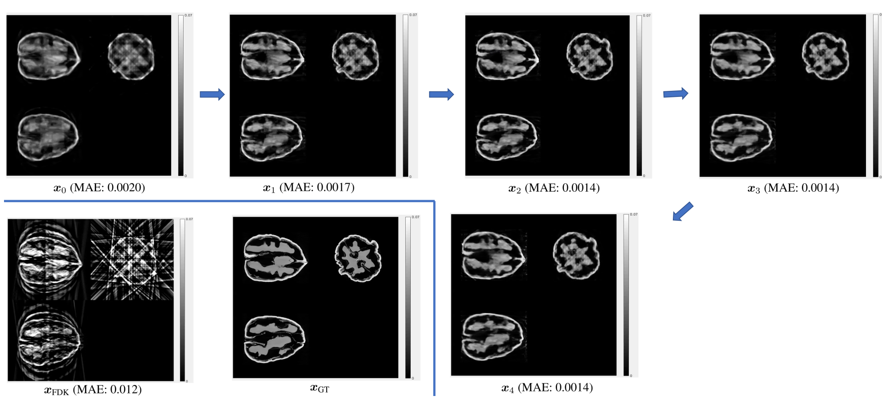

Figure 2: Center-slices along x-y, y-z, and x-z directions from the reconstruction of a test walnut across various stages ($\boldsymbol{x}_1 - \boldsymbol{x}_4$) using the proposed method. The ground truth using a full set of views ($\boldsymbol{x}_{\mathrm{GT}}$) and FDK reconstruction using only 8 views ($\boldsymbol{x}_{\mathrm{FDK}}$) have been shown in the bottom left. The initialization for our method has been depicted on top left ($\boldsymbol{x}_0$). Mean absolute errors (MAE) are shown, wherever applicable.

References

[1] I. Y. Chun and J. A. Fessler. Deep bcd-net using identical encoding-decoding cnn structures for iterative image recovery. In *2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, pages 1–5, 2018.

[2] Jeffrey A Fessler. Michigan image reconstruction toolbox. http://web.eecs.umich.edu/ fessler/irt/irt.

[3] Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522, 2017.

[4] Jiajia Luo, Haneda Eri, Ali Can, Sathish Ramani, Lin Fu, and Bruno De Man. 2.5 d dictionary learning based computed tomography reconstruction. In *Anomaly Detection and Imaging with X-Rays (ADIX)*, volume 9847, page 98470L. International Society for Optics and Photonics, 2016.

[5] S. Ravishankar, A. Lahiri, C. Blocker, and J. A. Fessler. Deep dictionary-transform learning for image reconstruction. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1208–1212, 2018.

[6] S. Ravishankar, J. C. Ye, and J. A. Fessler. Image reconstruction: From sparsity to data-adaptive methods and machine learning. *Proceedings of the IEEE*, 108(1):86–109, 2020.

[7] H. D. Sarkissian, F. Lucka, M. V. Eijnatten, G. Colacicco, S. Coban, and K. J. Batenburg. A cone-beam x-ray ct data collection designed for machine learning. *ArXiv*, abs/1905.04787, 2019.

[8] X. Zheng, S. Ravishankar, Y. Long, and J. A. Fessler. Pwls-ultra: An efficient clustering and learning-based approach for low-dose 3d ct image reconstruction. *IEEE Transactions on Medical Imaging*, 37(6):1498–1510, 2018.