A Deep Learning Architecture for Limited-Angle Computed Tomography Reconstruction

Kerstin Hammernik[1], Tobias Würfl[2], Thomas Pock[1,3], Andreas Maier[2]

[1]Institute of Computer Graphics and Vision, Graz University of Technology
[2]Pattern Recognition Lab, Friedrich-Alexander-University
[3]Digital Safety and Security Department, AIT Austrian Institute of Technology
{hammernik}@icg.tugraz.at

Abstract. Limited-angle computed tomography suffers from missing data in the projection domain, which results in intensity inhomogeneities and streaking artifacts in the image domain. We address both challenges by a two-step deep learning architecture: First, we learn compensation weights that account for the missing data in the projection domain and correct for intensity changes. Second, we formulate an image restoration problem as a variational network to eliminate coherent streaking artifacts. We perform our experiments on realistic data and we achieve superior results for destreaking compared to state-of-the-art non-linear filtering methods in literature. We show that our approach eliminates the need for manual tuning and enables joint optimization of both correction schemes.

1 Introduction

Computed Tomography (CT) is a clinical routine imaging modality that is used to diagnose certain diseases and trauma. In some applications, CT data cannot be acquired over the full angular range which is known as limited-angle CT. Examples for such setups are robot assisted scanners in medicine or scanning of very large objects in industrial CT. As limited-angle CT does not acquire data over the full angular range, the projection data is incomplete which results in intensity inhomogeneities as well as streaking artifacts in the image domain. Further sources for streaking artifacts are the non-linear attenuation of polychromatic X-rays or inelastic scattering of photons. All these artifacts are corrected with specialized heuristic compensation procedures that tune each step independently.

Many specialized iterative algorithms exist which clearly improve the image quality [1,2]. A disadvantage of iterative techniques is their high runtime requirement. In contrast, analytical algorithms are less demanding, but typically suffer from intensity inhomogeneities and streaking artifacts in the image domain due to missing projections. To correct for intensity inhomogeneities, Riess et al. [3] use a heuristic scheme to estimate compensation weights. Würfl et al. [4] reformulate filtered back-projection as a neural network and learn compensation weights for limited-angle CT reconstruction. However, their approach cannot account for the remaining streaking artifacts due to the missing non-linear filtering step.

To correct for remaining streaking artifacts, Riess et al. [3] apply a bilateral filter [5] after the compensation of missing projection data. Although there exists a number of other non-linear filtering methods such as BM3D [6] and Total (Generalized) Variation (T(G)V) [7,8], they can mainly correct for unstructured Gaussian noise. These models cannot describe the complex image content as they make assumptions on the image statistics such as piece-wise constancy in the case of TV. This motivates the use of deep learning approaches that learn the statistics of images [9], and thus can account for coherent noise artifacts. Recently, Hammernik et al. [10] proposed a deep learning approach to remove coherent backfolding artifacts in accelerated magnetic resonance image reconstruction.

In this paper, we propose a deep learning architecture for limited-angle CT reconstruction. In a first step, we estimate compensation weights in the projection domain [4]. The focus of our work is the second step: We propose to learn a non-linear filtering method inspired by variational image restoration problems [9] to remove streaking artifacts in the image domain.

2 Materials and Methods

The basic network architecture for artifact compensation in limited-angle CT is illustrated in Fig. 1. To account for missing projection data due to the limited-angle CT acquisition, a neural network architecture is used to estimate the compensation weights in the first step. In a second step, we eliminate streaking artifacts using a variational network architecture.

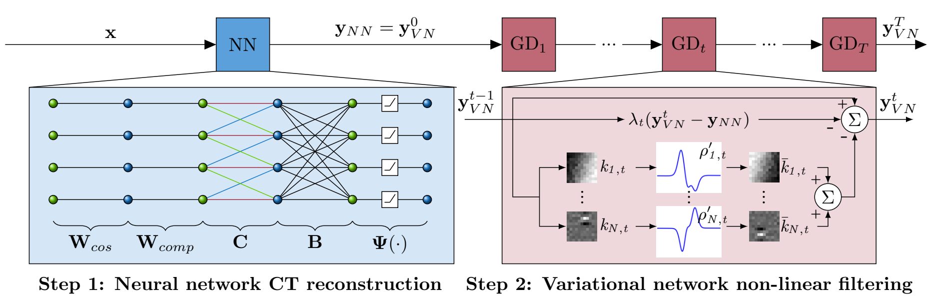

Fig. 1. Deep learning architecture for limited-angle CT reconstruction. The first neural network (blue) models filtered backprojection and corrects the intensity inhomogeneities in the image domain by learning compensation weights $\mathbf{W}_{\text{comp}}$ in the projection domain. The second variational network (red) formulates non-linear filtering as T unrolled gradient descent steps (GD). In each step t, the filters $k_{i,t}$, derivative of potential functions $\rho'_{i,t}$ and the regularization parameter λ_t are learned to remove the remaining streaking artifacts.

2.1 Step 1: A neural network to learn compensation weights

To correct for intensity inhomogeneities in limited-angle CT, we use the network architecture of Würfl et al. [4]. The input of the neural network is a sinogram with missing angular data, denoted by $\boldsymbol{x}$. The network reformulates the fan-beam reconstruction as

$$\boldsymbol{y}_{NN} = \boldsymbol{\Psi}\left(\mathbf{B}\mathbf{C}\mathbf{W}_{\text{comp}}\mathbf{W}_{\cos}\boldsymbol{x}\right) \tag{1}$$

where $\mathbf{B}$ denotes the backprojection operator, $\mathbf{C}$ implements filtering with a one-dimensional convolution kernel and the weighting operators $\mathbf{W}_{\cos}$, $\mathbf{W}_{\text{comp}}$ implement elementwise multiplications with cosine weights and compensation weights, respectively. The non-negativity constraint is realized via the operator $\boldsymbol{\Psi}$. While the operators $\mathbf{B}, \mathbf{C}$ and $\mathbf{W}_{\cos}$ are fixed, the compensation weights in $\mathbf{W}_{\text{comp}}$ are learned. After training, the network can be applied to a new sinogram and yields the intensity corrected reconstruction $\boldsymbol{y}_{NN}$. This output defines the input for the following variational network that focuses on destreaking.

2.2 Step 2: A variational network to remove streaking artifacts

To remove streaking artifacts in the neural network reconstruction $\boldsymbol{y}_{NN}$, we learn a non-linear filtering method. We seek an optimal image with eliminated streaking artifacts $\boldsymbol{y}_{VN}$. Based on the theory of [9], we formulate a network for non-linear filtering as a fixed number of T unrolled gradient descent steps

$$\boldsymbol{y}_{VN}^{t} = \boldsymbol{y}_{VN}^{t-1} - g^{t}(\boldsymbol{y}_{VN}^{t-1}). \tag{2}$$

The gradient of these steps is set to the gradient of a variational model

$$g^{t}(\boldsymbol{y}_{VN}^{t-1}) = \nabla_{\boldsymbol{y}} E(\boldsymbol{y})|_{\boldsymbol{y}=\boldsymbol{y}_{VN}^{t-1}}. \tag{3}$$

The variational image restoration problem is given as

$$E(\boldsymbol{y}) = \frac{\lambda}{2}\|\boldsymbol{y}_{VN} - \boldsymbol{y}_{NN}\|_2^2 + \sum_{i=1}^{N_k} \rho_i(\mathbf{K}_i \boldsymbol{y}_{VN}) \tag{4}$$

where the first term is a data fidelity term that measures the similarity to the network input $\boldsymbol{y}_{NN}$ and the second term is the regularization term that imposes prior knowledge on the image $\boldsymbol{y}_{VN}$. The impact of both terms is regulated by a parameter λ. For the regularization term, we apply N_k convolution operators $\mathbf{K}_i$, followed by non-linear functions $\rho_i : \mathbb{R}^N \mapsto \mathbb{R}$ to $\boldsymbol{y}_{VN}$. Note that applying the convolution operator $\mathbf{K}_i\boldsymbol{y}$ equals to a convolution with filter kernels $k_i * \boldsymbol{y}$. Plugging the gradient of the variational model into Eq. 2 yields

$$\boldsymbol{y}_{VN}^{t} = \boldsymbol{y}_{VN}^{t-1} - \sum_{i=1}^{N_k} \mathbf{K}_{i,t}^{T} \rho_{i,t}'(\mathbf{K}_{i,t}\boldsymbol{y}_{VN}^{t-1}) - \lambda_t(\boldsymbol{y}_{VN}^{t-1} - \boldsymbol{y}_{NN})\,. \tag{5}$$

This allows the parameters to adapt in every gradient descent step. In the gradient calculation, we additionally introduce the derivative of potential functions

$\rho'_{i,t} : \mathbb{R}^N \mapsto \mathbb{R}^N$ and transpose convolution operators $\mathbf{K}_{i,t}^T$. The vector $\rho'_{i,t}$ is understood in a point-wise manner. After a training procedure, we obtain the convolution kernels $k_{i,t}$, non-linear derivatives of potential functions $\rho'_{i,t}$ and the regularization parameter λ_t for each of the T gradient steps by minimizing the the mean-squared error (MSE)

$$L_{MSE} = \frac{1}{2S} \sum_{s=1}^{S} \| \boldsymbol{y}_{VN,s}^T - \boldsymbol{z}_s \|_2^2 \tag{6}$$

where S is the number of training samples and $\boldsymbol{z}$ defines the full scan reference.

2.3 Experimental Setup

To obtain training data, we simulated 450 fan-beam projections of size 512×512 from volumetric datasets of ten different patients. For evaluation, we performed a 5-fold cross validation and split the dataset into 80% training data and 20% validation data. As we focused on the evaluation of destreaking, we refer the interested reader to [4] for more details on the estimation of compensation weights. For our variational network architecture, we report results for different kernel sizes $k \in \{5, 7, 9, 11, 13\}$ and fixed the number of filter kernels $N_k = 24$ and gradient steps $T = 5$ empirically. For training, we used the L-BFGS-B optimizer and run 1200 iterations in total, i.e. 5×100 iterations pre-training of each gradient step and 700 iterations joint training of all gradient steps [9].

3 Results

We compared our variational network results to bilateral filtering, BM3D, TV and TGV quantitatively and to BM3D qualitatively. Table 1 shows the mean values and standard deviations for Peak Signal-to-Noise Ratio (PSNR) and Structured Similarity Index (SSIM). In order to perform a fair comparison, the parameters for all methods were estimated by grid search such that the PSNR of the validation data was maximized. Figure 2 shows the qualitative comparison of different methods and illustrates that the variational network result has less streaking artifacts and appears more natural compared to BM3D, which is like T(G)V and bilateral filtering not well suited for structured noise. Our deep learning architecture outperforms all methods qualitatively and quantitatively. The best results were achieved for a filter kernel size of 13.

4 Discussion

We propose a two-step deep learning architecture to correct for imperfections in limited-angle CT reconstruction due to missing projection data. In a first step, we correct intensity inhomogeneities in the image domain by learning compensation weights in the projection domain. In a second step, we train a variational network to learn regularization to remove structured streaking artifacts. Our results

Method	PSNR	SSIM
Neural Network	34.66 ± 2.07	0.908 ± 0.015
Bilateral Filtering ($\sigma_s = 0.5, \sigma_c = 0.1$)	29.93 ± 3.61	0.907 ± 0.021
BM3D ($\sigma = 1.5$)	34.75 ± 2.09	0.911 ± 0.015
TV ($\lambda = 300$)	34.82 ± 2.10	0.914 ± 0.014
TGV ($\lambda = 2, \alpha_0 = 0.01, \alpha_1 = 0.02$)	34.80 ± 2.09	0.914 ± 0.014
Variational Network ($k = 5$)	36.16 ± 2.13	0.930 ± 0.010
Variational Network ($k = 7$)	36.86 ± 2.01	0.938 ± 0.010
Variational Network ($k = 9$)	38.14 ± 2.27	0.947 ± 0.009
Variational Network ($k = 11$)	37.87 ± 1.98	0.949 ± 0.009
Variational Network ($k = 13$)	*38.23 ± 2.06*	*0.952 ± 0.010*

Table 1. Quantitative comparison of non-linear filtering methods along with the used parameter settings. The comparison is performed in terms of PSNR and SSIM (mean ± standard deviation) in the field-of-view. The intensity corrected neural network reconstruction defines the input to all methods. Our variational network results outperform all reference methods significantly.

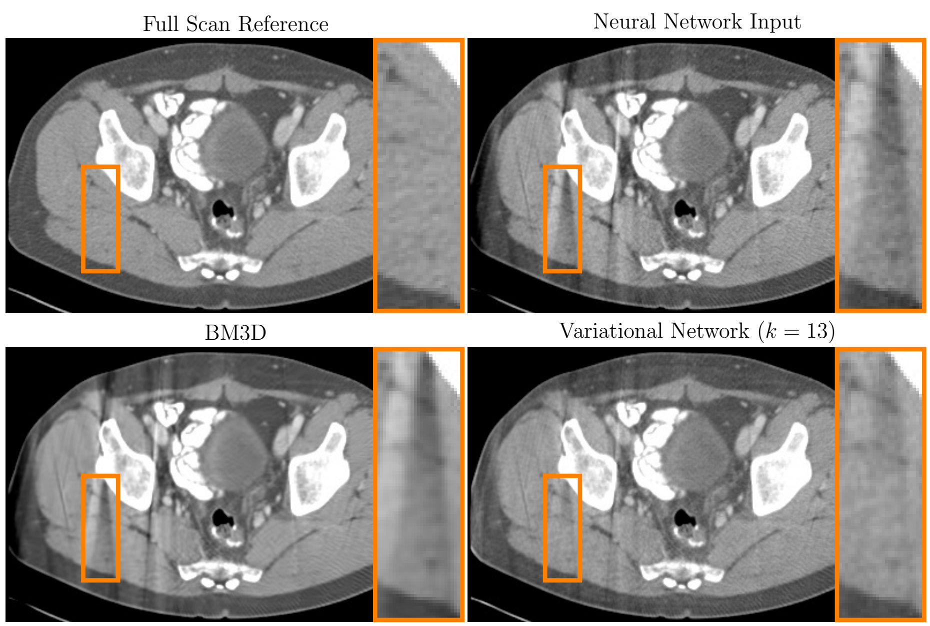

Fig. 2. Qualitative comparison of different non-linear filtering methods to the full scan reference. The neural network result is the intensity corrected output of a first correction step and defines the input to all methods. The variational network reconstruction with kernel size $k = 13$ shows significantly reduced streaking artifacts compared to BM3D.

reduce streaking artifacts significantly and outperform current state-of-the-art non-linear filtering approaches that can mainly deal with unstructured noise. The strength of our proposed model is that it eliminates the need for manual tuning and replaces heuristic compensation steps by data-driven optimization. In the future, we want to explore further extensions to our network architecture that account for more physical effects and train both networks jointly.

Acknowledgement We acknowledge grant support from the Austrian Science Fund (FWF) under the START project BIVISION, No. Y729. The authors would like to thank Dr. Cynthia McCollough, the Mayo Clinic, the American Association of Physicists in Medicine funded by grants EB017095 and EB017185 from the National Institute of Biomedical Imaging and Bioengineering for providing the used data.

References

1. Huang Y, Taubmann O, Huang X, Haase V, Lauritsch G, Maier A. A New Weighted Anisotropic Total Variation Algorithm For Limited Angle Tomography. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro; 2016. p. 585–588.
2. Huang Y, Taubmann O, Huang X, Haase V, Lauritsch G, Maier A. A New Scale Space Total Variation Algorithm for Limited Angle Tomography. In: Proceedings of the 4th International Meeting on Image Formation in X-Ray Computed Tomography; 2016. p. 149–152.
3. Riess C, Berger M, Wu H, Manhart M, Fahrig R, Maier A. TV or not TV? That is the question. In: The 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine; 2013. p. 341–344.
4. Würfl T, Ghesu FC, Christlein V, Maier A. Deep Learning Computed Tomography. In: Medical Image Computing and Computer-Assisted Intervention; 2016. p. 432–440.
5. Tomasi C, Manduchi R. Bilateral Filtering for Gray and Color Images. In: International Conference on Computer Vision; 1998. p. 839–846.
6. Dabov K, Foi A, Katkovnik V, Egiazarian KO. Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Transactions on Image Processing. 2007; p. 2080–2095.
7. Rudin LI, Osher S, Fatemi E. Nonlinear Total Variation Based Noise Removal Algorithms. Physica D: Nonlinear Phenomena. 1992;60(1-4):259–268.
8. Bredies K, Kunisch K, Pock T. Total Generalized Variation. SIAM Journal on Imaging Sciences. 2010;3(3):492–526.
9. Chen Y, Yu W, Pock T. On learning optimized reaction diffusion processes for effective image restoration. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. vol. 07; 2015. p. 5261–5269.
10. Hammernik K, Knoll F, Sodickson DK, Pock T. Learning a Variational Model for Compressed Sensing MRI Reconstruction. In: Proceedings of the International Society of Magnetic Resonance in Medicine. 24; 2016. p. 1088.