Direct Sparse-View CT Reconstruction using LSTM Processing of Stacked Back Projections[1]

Wenrui Li, Gregery T. Buzzard, and Charles A. Bouman[2]
Purdue University

Abstract

Sparse-view CT is important in a wide range of applications because of its potential to reduce acquisition time and dosage. Analytical reconstruction methods perform poorly with sparse views, so until recently the only practical approach to sparse-view reconstruction has been iterative methods such as model-based iterative reconstruction (MBIR). MBIR can produce high quality reconstructions from sparse data [6, 2] and can also incorporate prior models based on deep neural networks (DNNs) [8]. However, MBIR tends to be computationally intensive.

Over the past few years, image reconstruction using DNNs has emerged as a fundamentally new approach with the advantages that a) it can dramatically reduce computation, and b) given sufficient training data, it can be directly trained to incorporate complex prior information. Methods for DNN reconstruction fall into four groups [5]: (i) image domain methods, which post-process the analytical reconstruction using a DNN[10]; (ii) sensor-domain methods, which pre-process the sinogram data [3, 4]; (iii) hybrid-domain learning methods that process in both domains [1]; and (iv) direct, end-to-end DNN reconstruction methods, which go directly from the sinogram data to the reconstructed image. Among these methods, direct reconstruction methods, such as AUTOMAP [9], offer the greatest potential for high-quality sparse view reconstruction, but they are very computationally expensive and difficult to train.

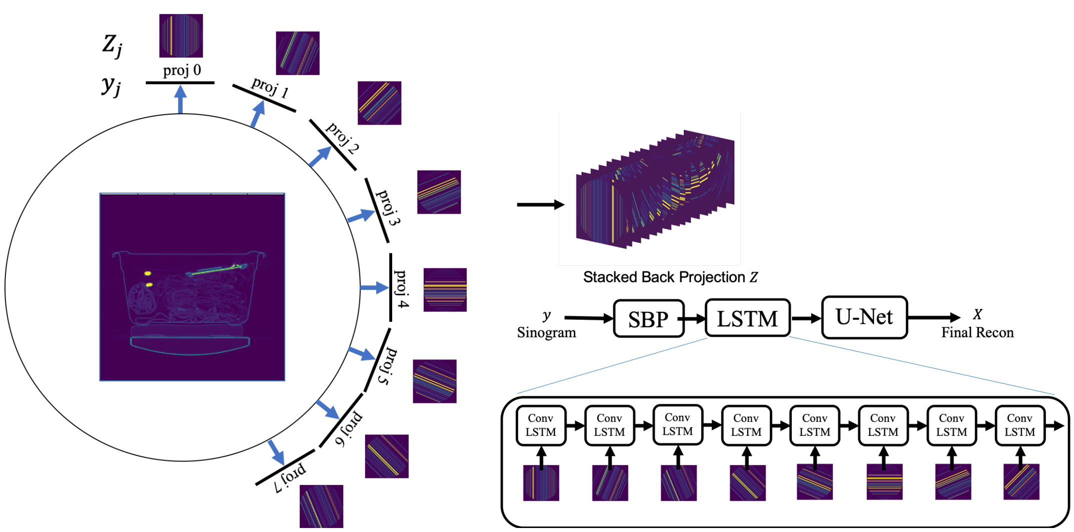


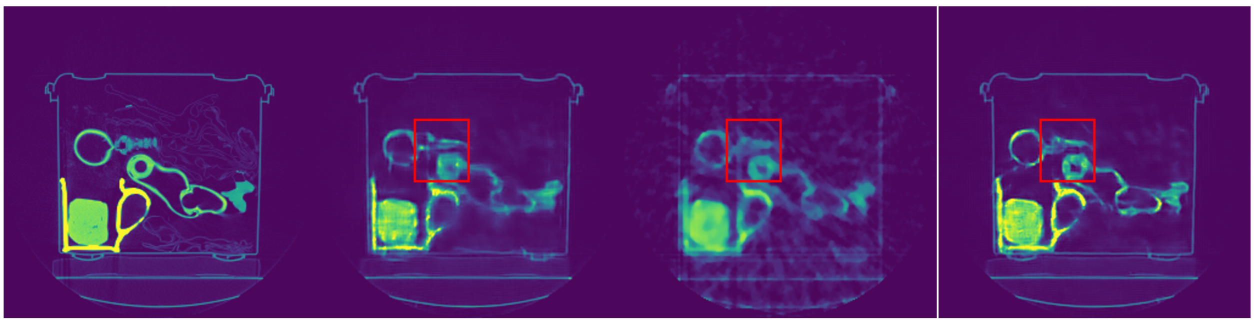
Figure 1: **Conceptual view of stacked backprojection and SBP-LSTM:** Individual, single-view projections are back projected (left), then collected into a tensor called the stacked backprojection (top right) for input to a CNN. At the bottom right, the individual backprojected views are shown as input to the LSTM, prior to further processing by a U-Net.

[1]Presented at the 2020 International Conference on Computational Photography, https://iccp2020.engr.wustl.edu/

[2]Wenrui Li and Charles A. Bouman are with the School of Electrical and Computer Engineering, Purdue University, 465 Northwestern Ave., West Lafayette, IN 47907-2035, USA. Gregery T. Buzzard is with the Department of Mathematics, Purdue University, West Lafayette, IN, USA. Li and Bouman were partially supported by the US Dept. of Homeland Security, S&T Directorate, under Grant Award 2013-ST-061-ED0001. Buzzard was partially supported by the NSF under grant award CCF-1763896. Email: {li3120,buzzard,bouman}@purdue.edu

In this poster, we build on the ideas in [7] and propose the LSTM-based stacked back projection (SBP-LSTM), which allows for computationally efficient direct DNN reconstruction going directly from the sinogram to the image. Illustrated in Figure 1, the key innovation of SBP-LSTM is that by individually back-projecting the views of the full sinogram into the image domain, it is possible to implement full direct reconstruction, while maintaining a computationally efficient structure. We compare several deep neural network structures including CNN, U-Net, and a novel LSTM U-Net architecture. Our experimental results demonstrate that SBP-LSTM using a U-Net structure results in the best overall quality reconstructions with reduced streaking artifacts and modest training data requirements.

Figures 2 and 3 compare the results of SBP-LSTM to alternative sparse-view reconstruction algorithms on both simulated and real CT data. Results indicate that the proposed SBP-LSTM algorithm can sharpen the reconstruction while reducing streaking artifacts.



(a) Ground Truth | (b) FBP+U-Net with MSE 0.045/0.886 | (c) MBIR($q = 1.2, \sigma_x = 0.8$) 0.051/0.816 | (d) SBP+LSTM+U-Net with MSE 0.037/0.912

Figure 2: **16-view Reconstruction on Simulated Sinogram:** Comparisons to ground truth using RMSE/SSIM values. a) Reference full-view MBIR reconstruction; b) Sparse-view MBIR reconstruction; c) FBP with U-Net post processing; d) Stacked Back Projection (SBP) with U-Net processing; d) SBP with LSTM + U-Net processing. Display range is from 0 (air) to 2000 Hounsfield units (HU). SBP+LSTM+U-Net can resolve the edges of the rings better than FBP+U-Net, SBP+U-Net, and MBIR on simulated data.

References

[1] M. U. Ghani and W. C. Karl. Deep Learning Based Sinogram Correction for Metal Artifact Reduction. In *Electronic Imaging, Computational Imaging XVI*, pages 472–1–4728, Jan. 2018. 1

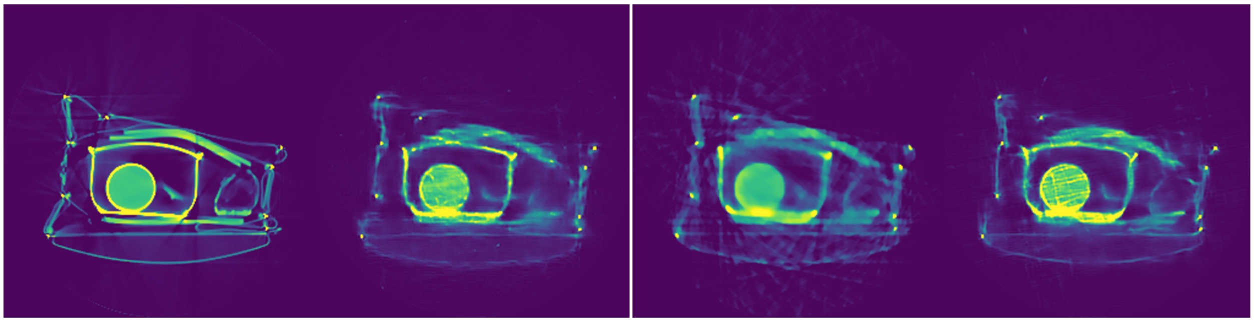
[2] S. J. Kisner, E. Haneda, C. A. Bouman, S. Skatter, M. Kourinny, and S. Bedford. Model-based ct reconstruction from sparse views. In *Second International Conference on Image Formation in X-Ray Computed Tomography*, pages 444–447, June 24-27 2012. 1

[3] H. Lee, J. Lee, and S. Cho. View-Interpolation of Sparsely Sampled Sinogram using Convolutional Neural Network. In *Medical Imaging 2017: Image Processing*, volume 10133, page 1013328. International Society for Optics and Photonics, 2017. 1

[4] H. Lee, J. Lee, H. Kim, B. Cho, and S. Cho. Deep-Neural-Network based Sinogram Synthesis for Sparse-View CT Image Reconstruction. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 3(2):109–119, Mar. 2019. arXiv: 1803.00694. 1

[5] S. Ravishankar, J. C. Ye, and J. A. Fessler. Image Reconstruction: From Sparsity to Data-adaptive Methods and Machine Learning. *arXiv:1904.02816 [cs, eess, stat]*, Apr. 2019. arXiv: 1904.02816. 1

[6] J.-B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh. A three-dimensional statistical approach to improved image quality for multislice helical ct. *Medical Physics*, 34(11):4526–4544, 2007. 1



(a) Full-View MBIR (b) FBP+U-Net with MSE 0.059/0.793 (c) MBIR ($q = 1.2, \sigma_x = 0.8$) 0.055/0.826 (d) SBP+LSTM+U-Net with MSE 0.058/0.805

Figure 3: **16-view Reconstruction on Real Sinogram:** Comparisons to Full-View MBIR reconstruction using RMSE/SSIM values. a) Reference full-view MBIR reconstruction; b) FBP with U-Net post processing; c) Stacked Back Projection (SBP) with U-Net processing; d) SBP with LSTM + U-Net processing. Display range is from 0 (air) to 2000 Hounsfield units (HU).

[7] D. H. Ye, G. T. Buzzard, M. Ruby, and C. A. Bouman. Deep Back Projection for Sparse-View CT Reconstruction. In *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 1–5, Nov. 2018. 2

[8] D. H. Ye, S. Srivastava, J.-B. Thibault, K. Sauer, and C. Bouman. Deep residual learning for model-based iterative ct reconstruction using plug-and-play framework. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6668–72, April 2018. 1

[9] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen. Image Reconstruction by Domain-Rransform Manifold Learning. *Nature*, 555(7697):487–492, Mar. 2018. 1

[10] A. Ziabari, D. H. Ye, S. Srivastava, K. D. Sauer, J.-B. Thibault, and C. A. Bouman. 2.5d deep learning for ct image reconstruction using a multi-gpu implementation. In *52st Asilomar Conference on Signals, Systems and Computers*, pages 2044–2049, Oct. 2018. 1