Projection Data Upsampling for Sparse and Low-Dose CT Scout Scans

Frank Bergner[1], **Rolf Bippus**[1], **Felix K. Kopp**[2], **Peter B. Noël**[2]
[1] Philips Research Röntgenstraße 24-26, 22335 Hamburg, Germany,
frank.bergner@philips.com, rolf.bippus@philips.com
[2] Technische Universiät München, Munich, Germany

Abstract

In Computed Tomography (CT) scout scans are a first initial acquisition that determine scan parameters and final imaging regions. Current state-of-the art scout scans are 2D X-ray images acquired with the stationary CT scanner gantry and the moving table. A 3D scout volume based on a low-dose, sparse-view helical acquisition would be better for the scan planing, but image volumes usually suffer from high noise and especially angular under-sampling artifacts. We therefore evaluate the usage of a convolutional neuronal network to perform a non-linear upsampling and denoising of the input projection data. We compare our approach to a least-squares optimal linear interpolation and evaluate both algorithms with scout scan data synthesized from real patient datasets.

1 Introduction

In CT new technologies like fast tube-current switching could open new applications like fast 3D scout scans for better scan planing [1]. In order to keep the 3D scout dose as low as possible, image acquisition needs to focus the total photon flux in fewer projections, lifting the detected photons over the detector noise, which would otherwise dominate the data. Still the image quality of a reconstructed 3D helical scout is limited by the sparse and noisy projection data, which have characteristic, angular undersampling artifacts as well as noise streaks.

Reconstructions from sparse-view projection data have been studied in the past, for example with iterative approaches like compressed sensing [2] or recently with convolutional neuronal networks (CNN) [3]. Other groups apply CNNs to denoise low-dose images [4]. While image-based approaches have the possibility to use efficient image priors, their drawback is that they have to deal with long-range artifacts, i.e. streaks that can traverse the whole image slices in the worst-case. Furthermore, streaks are highly directional, which implies that networks need to learn all possible directions. Projection domain approaches in contrast lack good priors as structures are moving and overlapping. Yet, they do not have the streak problem as the image based approaches.

In our work we will focus mainly on reducing artifacts by upsampling the projection data along the angular axis in the projection domain. This introduces artificial projections, which do not contain new information but help to reduce the image noise and artifacts. In contrast to previous works we additionally deal with an increased noise level due to an even lower tube-current. The network is therefore required to learn an efficient denoising in addition to interpolation.

2 Materials and Methods

Considering simple 2D parallel-beam models and the underlying Fourier slice theorem it is empirically clear that X-ray projections from close angular positions share low-frequent information. For a sparse-view projection dataset it is therefore possible to calculate a least-mean-squares (LMS) optimal predictor filter that will predict an intermediate view, which was not measured, from a limited

1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

number of acquired views in an angular neighborhood. However, the larger the angular gaps the less high frequent information will be available in the predicted intermediate view. The linear predictor (LP) in our study uses e.g. a neighborhood of 4 views, 3 rows and 17 columns to predict the value of the central detector pixel in original fan beam detector geometry. The LMS-optimal predictor itself is not capable of introducing sharp edges in the interpolated data. Yet, when examining sinogram data with a human eye it appears obvious where sharp sinogram traces should be located. Therefore, as a second approach we evaluate whether a convolutional neuronal network is capable of synthesizing more sharp edges, which would further reduce artifacts caused by angular subsampling. Assuming that the linear predictor covers the low frequency components quite well, we train an additional CNN to generate the high frequency components, its output being added to that of a linear predictor in parallel that is trained along with the CNN. For the CNN we use 6 convolution layers. The input layer is chosen same size as the LP filter, while the other layers are chosen to filter within the row only. The number of filters in each layer is 32, and 1 in the output layer. A leaky-ReLU activation is chosen in all layer except the output layer.

For the training of the given approaches we use the same weighted-least-squares cost function, comparing to real projection from the corresponding full-dose scan. The squared error in each detector pixel is thereby weighted with the inverse variance of estimated noise in the individual pixels. We found this weighting to be mandatory as the predictor would otherwise try to recover mainly the highly attenuated projection regions.

Synthetic scout scan data was generated from subsampling 10 patient scans from a Philips Brilliance iCT Scanner. The projection size was 128 rows and 672 columns with 150 views per turn, which equals a 16-fold sub-sampling of the original data. The relative pitch of the helix was 0.9. In addition, realistic noise was added to the data to simulate a reduction of the tube current down to an equivalent of 3 mAs in the center of rotation [5]. Images were reconstructed on a 0.75 mm grid in order to evaluate the de-streaking performance. 6 datasets were used for the training, 2 for validation during the training, and 2 were kept for a final testing. The training was performed on a GPU using Microsoft CNTK.

3 Results and Discussion

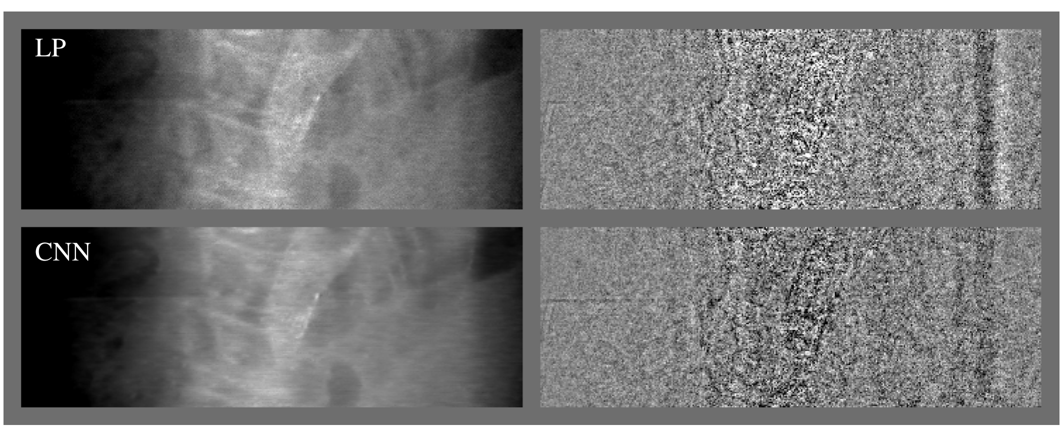

Figure 1: Example of interpolated views from both methods. The right images show the difference with respect to the high-dose, full dataset

The results shown here are from one of the two validation datasets. Fig. 1 shows an interpolated view from both upsampling methods as well as the difference image with respect to the full-dose patient scan. The results from the CNN approach have a better noise appearance and have also a slightly better prediction, visible from less structures in the difference images. In Fig. 2 a reconstructed slice from both datasets is shown from the hip region. The left image contains the slice reconstructed from the simulated sparse view data. The center column shows results from both methods, while the right column shows difference images with respect to the full-dose dataset. The CNN approach shows slightly superior results at the center of rotation as well as object boundaries towards air. Both

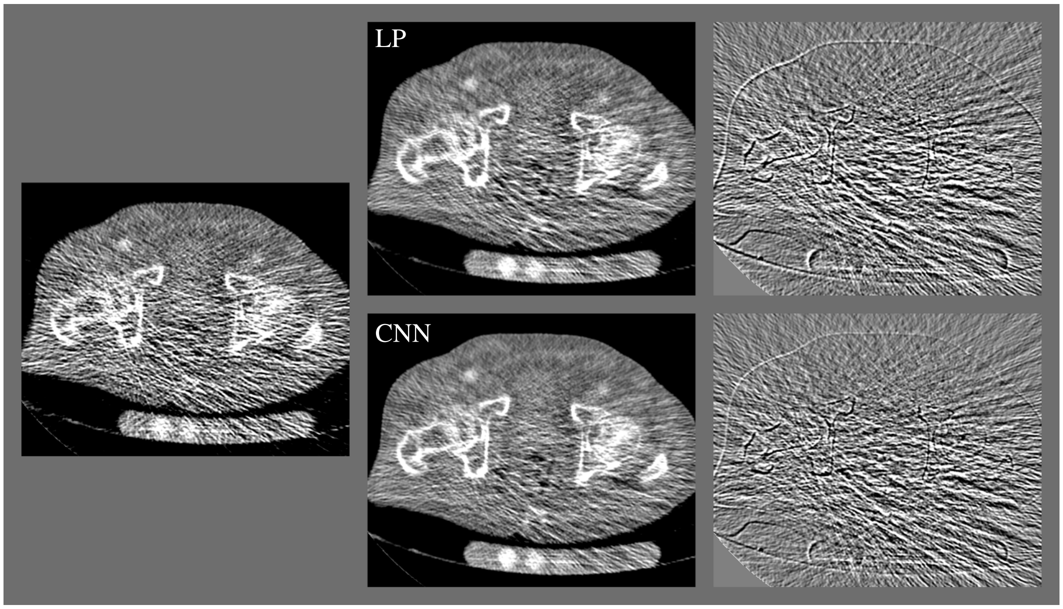

Figure 2: Reconstruction results without upsampling (left), with upsampling (center) and difference with respect to the full-dose data reconstruction (right)

methods show a drop of image noise by about 50%, which is a direct result of the low noise level in the interpolated views.

Overall, the initial results are encouraging, but the limiting factors need to be further studied. First of all the question is whether a 2-fold upsampling is sufficient or 4-fold is necessary to reduce the remaining streaks. The size of the CNN will need to be increased to cover a larger perceptive field across viewing angles and detector rows to enable a better identification of predominant structures. The large angular gaps combined with a high helical pitch can cause structures to move in the order of one detector row from view to view. Currently this is only captured using a row neighborhood in the input layer of the CNN, ignoring row neighborhoods in the remaining layers.

4 Conclusion and Further Steps

In this first study we could show that a CNN is capable of further reducing artifacts and noise better than a purely linear prediction approach. Future experiments will consist of using deeper networks as well as other state-of-the art network topologies like ResNet or U-Net adopted to this problem.

Acknowledgement We acknowledge support through the German Department of Education and Research (BMBF) under grant IMEDO (13GW0072C).

References

[1] Z. Yin, Y. Yao, A. Montillo, M. Wu, P. M. Edic, M. Kalra, and B. D. Man, "Acquisition, preprocessing, and reconstruction of ultralow dose volumetric ct scout for organ-based ct scan planning," *Med. Phys.*, 2015.

[2] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, April 2006.

[3] S. C. Hoyeon Lee, Jongha Lee, "View-interpolation of sparsely sampled sinogram using convolutional neural network," 2017.

[4] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose ct with a residual encoder-decoder convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 36, pp. 2524–2535, Dec 2017.

[5] S. Zabic, Q. Wang, T. Morton, and K. M. Brown, "A low dose simulation tool for ct systems with energy integrating detectors," *Med. Phys.*, vol. 40, no. 3, p. 031102, 2013.