Hybrid Reconstruction Using Shearlets and Deep Learning for Sparse X-Ray Computed Tomography

Andi Braimllari[a,b], Theodor Cheslerean-Boghiu[a,b], and Tobias Lasser[a,b]

[a]Computational Imaging and Inverse Problems, Department of Informatics, Technical University of Munich, Germany
[b]Munich Institute for Biomedical Engineering, Technical University of Munich, Germany

ABSTRACT

In sparse X-ray Computed Tomography, the radiation dose to the patient is lowered by measuring fewer projection views compared to a standard protocol. In this work we investigate a hybrid approach combining shearlet representation with deep learning for reconstruction of sparse-view X-ray computed tomography. The proposed method is hybrid in that it reconstructs the parts that can provably be retrieved by utilizing a model-based approach, and it in-paints the parts that provably cannot through a learning-based approach. In doing so, we attempt to benefit from the best aspects of model- and learning-based methods. We demonstrate first promising results on publicly available data.

Keywords: sparse-view X-ray computed tomography, shearlets, deep learning.

1. INTRODUCTION

X-Ray Computed Tomography (CT) is an essential technique that provides deep insight of a patient or an object of interest in a non-invasive manner. The forward model can be formulated as

$$Rf = y + \eta \tag{1}$$

where R represents the X-ray transform, f the quantity to be reconstructed (the absorption coefficients), while y represents the X-ray measurements and η some noise.

An important aspect of medical X-ray CT is the radiation exposure of the patients. One technique to lower the radiation dose is by lowering the number of projection views.[1] This is referred to as sparse-view or sparse X-Ray CT. The reconstructions of such sparse measurements tend to feature streak artifacts near edges tangent to the acquired X-rays.[1] With increasing sparsity, and thus with an increasing lack of measurement data, the severity of these streak artifacts increases as well. Hence a reconstruction approach alleviating the impact of these streak artifacts is highly desirable.

The visibility principle[2] tells us in essence that the visible part of an object is comprised of the set of edges tangent to the acquired X-rays, and the invisible part is comprised of the edges non-tangent to these X-rays. Moreover, which edges can or cannot be reconstructed is dependent on the acquisition geometry, therefore known before acquisition.

Following[3] we leverage this principle by employing shearlets to resolve the wavefront set of such a signal. This enables us to properly reconstruct the visible edges using ℓ_1-regularization and to in-paint the invisible ones using deep learning. Using only model-based methods, by the visibility principle, we cannot retrieve the invisible information. On the other hand, using only deep learning on such an ill-posed problem, we might get satisfactory results up to a point, but we will not be able to certainly assert as to how much the original signal has changed.[3]

Therefore, our proposed approach, which we will term **SDLX** for the remainder of this work, is a hybrid method that benefits from the best aspects of both model-based and learning-based approaches for sparse-view

Send correspondence to Andi Braimllari (andi.braimllari@gmail.com) or Tobias Lasser (lasser@in.tum.de)

7th International Conference on Image Formation in X-Ray Computed Tomography,
edited by Joseph Webster Stayman, Proc. of SPIE Vol. 12304, 123040N
 · 0277-786X · doi: 10.1117/12.2646385

X-ray CT reconstruction. The SDLX method is very closely related to,[3] which was developed for limited-angle X-ray CT. In the following we will present the SDLX method in detail, along with first results on a publicly available data set.

2. METHODS

2.1 SHEARLETS

Shearlets are a mathematical concept building on top of existing wavelet-theory components with distinct advantages. They represent a multi-scale framework that provides optimally sparse approximations of multivariate, anisotropic data. The approximation rate of shearlets is of $O(N^{-2})$, comparatively better than the $O(N^{-1})$ of wavelets. Shearlets are constructed by applying three operations, translation, dilation, and shearing, see[4] for details. They are applied to a single generating function ψ, resulting in a shearlet system

$$\psi_{a,s,t} = |\det M_{as}|^{1/2} \psi M_{as}(-t) \tag{2}$$

Here, $a \in \mathbb{R}_+$ dictates the dilation matrix A_a, $s \in \mathbb{R}$ dictates the shearing matrix S_s, while $t \in \mathbb{R}^2$ represents the translations. The composite matrix M_{as} is then defined as $M_{as} = A_a^{-1} S_s^{-1}$. The **continuous shearlet transform** is then

$$SH(f) = \langle f, \psi_{a,s,t} \rangle \tag{3}$$

For the **discerete shearlet transform**, we sample the parameter space $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^2$ at discrete points. This defines the **regular discrete shearlet system** as

$$SH(\psi) = \{\psi_{j,k,m} 2^{3j/4} \psi(S_k A_{2^j}), \forall (j,k,m) \in \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}^2\} \tag{4}$$

The **discrete shearlet transform** is defined similarly to the continuous scenario.

Given the directional bias exhibited by regular shearlets,[5] we will instead be using the cone-adapted shearlets as they provide a remedy to it. For a visual representation of the tiling that is generated in the Fourier domain, see Figure 1. Additionally, we want to explicitly specify and emphasize, that the **cone-adapted discrete shearlet systems** as mentioned above, under mild assumptions, form a Parseval frame.[6] Based on this fact and the above statements, we know that

$$f = SH^T(SH(f)) \tag{5}$$

This equation represents a powerful reconstruction formula of the **discrete shearlet transform**, which is essential to our hybrid approach.

By far the most important property of the shearlet systems to the hybrid approach is their ability to resolve the wavefront set of the signals at hand.[3] This allows us to differentiate the visible and invisible boundaries, and is accomplished by distinguishing different decay rates of the **shearlet transform**.

It is worth pointing out that we utilize classical shearlet systems, which are band-limited (compact support in the frequency domain[4]). Our implementation of such a **discrete cone-adapted band-limited shearlet transform** is based on,[7] to which we refer the reader to for further details. We also utilize the **shearlet transform** available in[8] as an intermediary operation after running the ℓ_1-regularization.

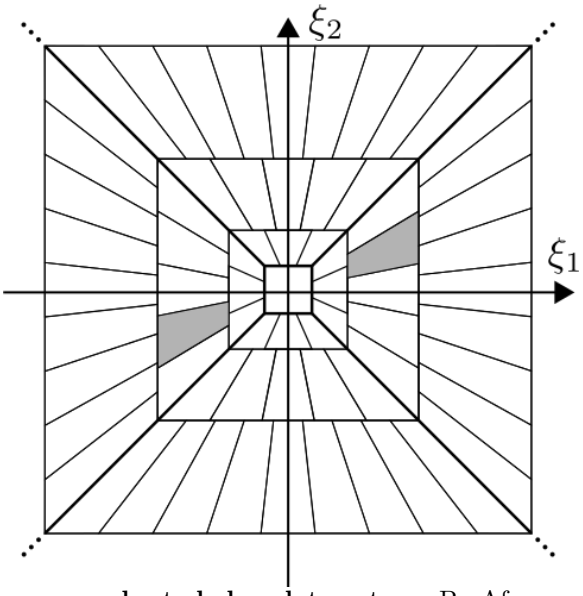


Figure 1: Frequency tiling of the **cone-adapted shearlet system**. By Afg genzel - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=27761187

2.2 SPARSE REGULARIZATION WITH ADMM

Sparse regularization attempts to leverage the assumption that the output of a problem can be described by a fewer number of inputs, or put differently that for every output there exists a sparsifying representation system.[3] More specifically for low-dose CT, it has been shown that such methods enable more accurate reconstructions given very few tomographic measurements.[3] Therefore, such a paradigm is of interest to us for tackling sparse-view X-ray CT.

Alternating Direction Method of Multipliers (ADMM) is a general algorithm that works quite well in splitting the minimization of the sparsity-promoting ℓ_1-regularization term and the data fidelity term. Details on **ADMM** can be found in.[9]

Based on (1), we are now able to construct and utilize shearlet-based sparse regularization. Explicitly expressing the reconstruction problem built so far, we write

$$\operatorname{argmin}_{f \geq 0} \frac{1}{2}\|Rf + y\| + \|SH(f)\|_{1,w} \tag{6}$$

in which $w \in \mathbb{R}_{\geq 0}^{W \times H \times L}$ represents the weights to the regularization term. This allows us to split the wavefront set of the signal into the visible and invisible parts, as described in the visibility principle. In this equation we have also explicitly specified a constraint for non-negative solutions, as it leads to better reconstructions.[3]

We solve this minimization problem using **ADMM**.

2.3 RECOVERING THE INVISIBLE USING DEEP LEARNING

Deep Learning is one of the most influential paradigms of the last few decades with impressive results in a plethora of fields. In the recent years, considerable attempts have also been made towards the field of medical imaging as well. Many of the current model architectures and techniques pre-process the measurements or post-process the reconstructions, which can produce impressive results. However, it is not always immediately obvious if data fidelity has been preserved. In a medical setting this is not something that can be brushed aside easily, as accuracy is crucial.

In the hybrid approach that we are working on, the influence of deep learning is kept to a minimum. It is only used for in-painting missing information that can provably not be retrieved through classical model-based approaches. The architecture that we are using is **PhantomNet**, as proposed in.[3]

PhantomNet is a fully-convolutional neural network based on one of the most prevalent architectures, **U-Net**. Different from **U-Net**, it is also a multi-channel input and multi-channel output network, based on the fact that it operates on the phase space and works with shearlet coefficients. More specifically, it takes a signal of shape (L, W, H) (e.g. $(61, 512, 512)$ and outputs a signal of the same shape. Here, W and H respectively represent the width and height of the image, while L dictates the number of layers of the shearlet coefficients. We refer the reader to[3] for full details on the architecture.

2.4 THE SDLX METHOD

Using the separate components of the hybrid approach outlined above, we now summarize all the steps that make up the **SDLX** method.

1. **Retrieve the visible coefficients**
   Compute ℓ_1-regularized solutions of the following problem

$$g \in \text{argmin}_{f \geq 0} \ \frac{1}{2}\|Rf + y\| + \|SH(f)\|_{1,w} \tag{7}$$

   by utilizing **ADMM** (or any other appropriate solver), which retrieves the visible coefficients based on the provided measurements. The input to this step are the sparse-view measurements y, while the output g is a reconstruction with sparse-view artifacts.

2. **Estimate the invisible coefficients**
   We apply the **shearlet transform** to all of the images g generated above, which maps them from (W, H) to (L, W, H). The **PhantomNet** uses these shearlet coefficients as input, and it outputs objects of the same shape, which are the in-painted shearlet coefficients. After training **PhantomNet** (**PN**), we use this model to estimate the invisible coefficients. If its weights are well adjusted, the following approximation should hold to a satisfactory threshold,

$$\mathbf{PN}(SH(g)) \approx SH(f)_{inv} \tag{8}$$

3. **Combine the visible and invisible coefficients**
   Up until here we have the retrieved visible coefficients and a decent-enough estimation of the invisible coefficients (output of **PN**). We sum them together and bring the entire output back to the spatial-domain through the **inverse shearlet transform**

$$f_{SDLX} = SH^T(SH(g)_{vis} + PN(SH(g))) \tag{9}$$

   Here, f_{SDLX} is our end-result (of shape (W, H)), which contains the reconstruction of the sparse-view measurements along with the in-painting of the missing information.

The run-time of the proposed **SDLX** method is dominated by ADMM solving the ℓ_1-regularized problem. The implementations of our proposed algorithm are available at.[10,11]

3. EXPERIMENTS AND RESULTS

The dataset used here is the one provided by the Mayo Clinic for the AAPM Low-Dose CT Grand Challenge.[12] It contains human abdomen scans with width and height of 512. We chose the low-dose scans, with pixel intensities in $[0 : 255]$.

For training of the PhantomNet we selected 10 patients, with the IDs of L004, L006, L014, L019, L033, L049, L056, L057, L058, and L064, which comprised a total number of 1525 scans. For testing we chose another patient, with an ID of L071.

To generate training pairs for the **PhantomNet**, we first simulate sparse, 64 projection view sinograms (over an arc of 360 degrees) of the 1525 training images using,[11] adding 1% Gaussian noise. Then we compute

Table 1: Metrics of the 64-view reconstruction results on patient L071 compared to the ground truth. The lower the **RE**, the better. The higher the **PSNR**, **SSIM**, **HaarPSI**, the better.

Metrics				
Method	RE	PSNR	SSIM	HaarPSI
f_{CG}	0.073	21.664	0.221	0.339
f_{ADMM}	0.061	22.409	0.246	0.352
f_{SDLX}	**0.026**	**26.001**	**0.271**	**0.626**

ℓ_1-regularized reconstructions of those sinograms using ADMM as in (7), with 10 iterations of ADMM and 5 inner iterations of the conjugate gradient method on the normal equation. We manually selected the parameters of ADMM as $\rho_1 = 1/2$, $\rho_2 = 1$ (as in[3]), and $w = 0.001$. Afterwards, we apply the **forward shearlet transform** from.[8]

We train the **PhantomNet** for 100 epochs in single-batches (e.g. one $(61, 512, 512)$ object as a batch) on a learning rate of $7e-5$ and weight decay of $1e-7$. The chosen optimizer is Adam. The loss function is the mean squared error loss from torch.nn.MSELoss.

For testing, we use the data from the patient with ID of L071, and simulate sparse-view sinograms with 64 projection views as above, adding 1% Gaussian noise. We execute the full **SDLX** method as in subsection 2.4, using the same ADMM parameters as for training the net. In the last step, we sum together the visible coefficients from the ℓ_1-regularization and the estimated invisible coefficients that the trained **PhantomNet** predicts, and apply the **inverse shearlet transform** to it.

An example result f_{SDLX} of our proposed **SDLX** method is shown in Figure 2 for one of the slices of patient ID L071, which was not seen during training. We also compare with the ground truth and reconstructions using the same ADMM as in the first step of SDLX, f_{ADMM}, and the result of an unregularized CG reconstruction using 10 iterations, f_{CG}.

It is apparent that this hybrid method is capable of in-painting the missing singularities for sparse-view CT. **SDLX** outperforms the other methods, a claim also supported by the metrics, as displayed in Table 1. The metrics utilized are the Relative Error (RE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Haar Wavelet-Based Perceptual Similarity Index (HaarPSI).

We also ran the same testing experiment without adding noise to the simulated sinograms. The trained model (on noisy simulated data) performed just as well on the unseen data, which serves as an indicator that **SDLX** is fairly robust towards noise.

4. DISCUSSION AND CONCLUSION

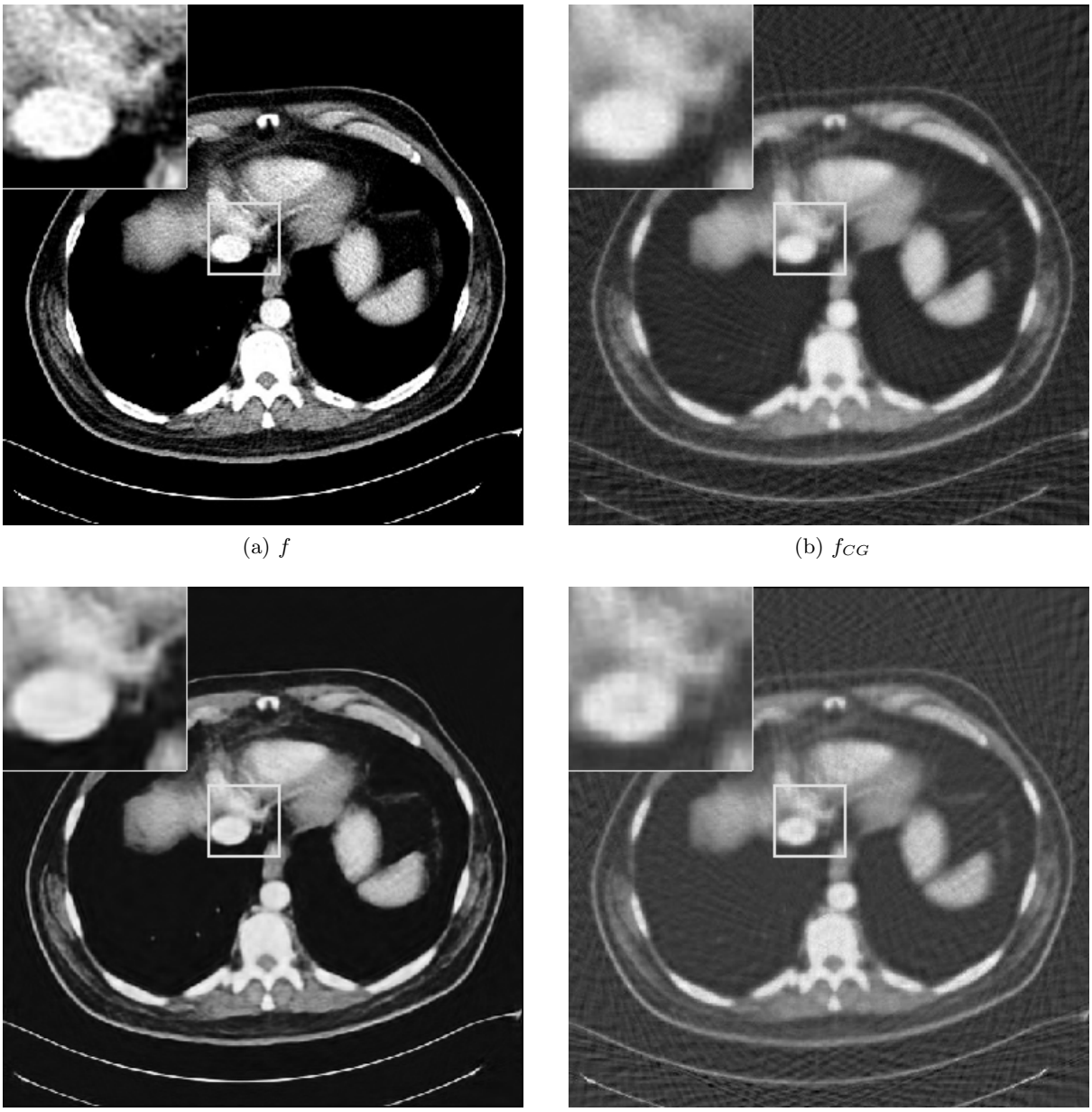
The results from the experiment indicate that the hybrid approach works for sparse-view CT. However, a certain smoothing effect is also visible in the results. First experiments (not shown here) indicate that further tuning of the training of PhantomNet might negate this effect.

One aspect of the SDLX method that might introduce unexplained features is the deep learning step. Fortunately, this element is utilized here in a relatively controlled manner, given that it only handles the inference of the invisible coefficients. Further tuning of the hyper-parameters might be beneficial, as might be the study of more advanced models, such as transformers, instead of the U-net.

In summary, **SDLX** works because shearlets are capable of resolving the wavefront sets of the signals we are dealing with, and these decomposed coefficients adhere to certain rules which we can then learn. Adapting the work for limited-angle X-ray CT in,[3] our first experiments for sparse-view X-ray CT on a publicly available data set show promising results.

ACKNOWLEDGMENTS

We would like to express our gratitude to David Frank, a maintainer of the *elsa* library,[11] for his code reviews.



(a) f (b) f_{CG}

(c) f_{SDLX} (d) f_{ADMM}

Figure 2: 64-view reconstruction results on patient L071. The ground truth is f, while f_{SDLX} is the output of our proposed SDLX algorithm. For comparison, we also show the reconstruction of the first step of the SDLX algorithm (f_{ADMM}) as well as an unregularized iterative CG reconstruction f_{CG}. The pixel intensities lie in $[0:255]$. For further details see section 3.

REFERENCES

[1] Han, Y. and Ye, J. C., "Framing U-Net via Deep Convolutional Framelets: Application to Sparse-view CT," *IEEE Transactions on Medical Imaging* **37**, 1418–1429 (2018).

[2] Quinto, E. T., "Artifacts and Visible Singularities in Limited Data X-Ray Tomography," *Sensing and Imaging* **18** (2017).

[3] Bubba, T. A., Kutyniok, G., Lassas, M., März, M., Samek, W., Siltanen, S., and Srinivasan, V., "Learning

The Invisible: A Hybrid Deep Learning-Shearlet Framework for Limited Angle Computed Tomography," *Inverse Problems* **35**, 064002 (2018).

[4] Kutyniok, G. and Labate, D., "Introduction to Shearlets," *Applied and Numerical Harmonic Analysis* (2012).

[5] Guo, K., Kutyniok, G., and Labate, D., "Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators," (2005).

[6] Christensen, O., "An Introduction to Frames and Riesz Bases," (2003).

[7] Häuser, S. and Steidl, G., "Fast Finite Shearlet Transform," (2014).

[8] Ronchetti, M. and Voigtlaender, F., "Adaptive transform for manifold-valued data," (2016).

[9] Boyd, S., Parikh, N., Peleato, E. C. B., and Eckstein, J., "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends in Machine Learning* **3**, 1–122 (2011).

[10] Braimllari, A., "SDLX, https://github.com/AndiBraimllari/SDLX," (2021).

[11] Lasser, T., Hornung, M., and Frank, D., "elsa - an elegant framework for tomographic reconstruction," in [*15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*], Matej, S. and Metzler, S. D., eds., **11072**, 570 – 573, International Society for Optics and Photonics, SPIE.

[12] McCollough, C., "TU-FG-207A-04: Overview of the Low Dose CT Grand Challenge," *Medical Physics* (2016).