SPARSITY-BASED SINOGRAM DENOISING FOR LOW-DOSE COMPUTED TOMOGRAPHY

J.Shtok, M. Elad, and M. Zibulevsky

Computer Science department, Technion IIT, Israel

ABSTRACT

We propose a sinogram restoration method which consists of a patch-wise non-linear processing, based on a sparsity prior in terms of a learned dictionary. An off-line learning process uses a statistical model of the sinogram noise and minimizes an error measure in the image domain over the training set. The error measure is designed to preserve low-contrast edges for visibility of soft tissues. Our numerical study shows that the algorithm improves on the performance of the standard Filtered Back-Projection algorithm and effectively allows to halve the radiation dose for the same image quality.

Index Terms— Computed Tomography, sinogram restoration, Sparse-Land paradigm.

1. INTRODUCTION

One of the current challenges in Computed Tomography is to obtain a high-quality image from a low-dose scan, which brings less damage to the patient. The measurements in such a scan are corrupted by a number of physical phenomena [1], while the leading factor is the Poisson noise resulting from poor photon count statistics.

A good way to address this problem is to build an elaborate statistical model of the acquisition process, which accounts for all the sources of data deterioration, and to solve a corresponding (usually, complex) optimization problem using an iterative method [1, 2]. However, high computational cost and the need for an accurate and explicit modeling of the system encourage researchers to seek for other solutions.

An approach alternative to explicit data modeling is to use adaptive processing tools that can automatically adjust and implicitly learn the data statistics. When the learning procedure is accomplished in an off-line training, the resulting algorithm can benefit from both low computational cost of the non-iterative data processing and an improved performance due to the training.

In this work we apply the *Sparse-Land* paradigm to sinogram restoration from a low-count noise. The core idea is to impose a sparsity prior patch-wise on the 2D sinogram matrix, in terms of a *dictionary* $\mathbf{D}$ - a linear transform with the property that every patch of the signal can be well approximated by a linear combination of just a few columns from $\mathbf{D}$. The transform is custom built for a specific family of images via a training set. Aharon and Elad proposed the K-SVD algorithm for *training* of the dictionary $\mathbf{D}$ [3], which is employed in a powerful signal de-noising method [4].

CT reconstruction based on the K-SVD was proposed by Liao and Sapiro in [5]. Utilizing the noise reduction scenario of [4], the algorithm conducts an on-line learning procedure for the dictionary and the coefficients, encoding the reconstructed image patch-wise. The image is then computed by solving a quadratic optimization problem, which balances a data fidelity term and a sparsity prior. This algorithm shows very promising results on phantoms with explicit geometrical structure in some missing data scenarios. However, there are few important aspects of the algorithm which are problematic; in this work we propose a different learning and reconstruction scheme, resolving those issues.

In contrast to the approach described in [5], our algorithm focuses on the sinogram processing, later followed by a standard Filtered Back-Projection (FBP) algorithm. The sparsity prior is applied in the sinogram domain where the noise statistics is known and accounted for. For optimal results, the training objective for the dictionary $\mathbf{D}$ considers the image domain as well as the sinogram domain; we minimize an error measure, applied to the difference between the reference image and the reconstructed version, over a training set. This implies that, as opposed to [5], our work has an off-line part that trains on examples of high-quality images and their sinograms, while the amount of real-time computations is low. The error functional is a weighted L_2 norm, where data-dependent weights are designed to preserve low-contrast edges of the attenuation images. A numerical study shows that the proposed algorithm improves the variance-resolution trade-off of the standard FBP method, effectively allowing to reduce the X-ray dose by 50% for the same image quality (for the specific dose level used in the experiments).

2. THE ALGORITHM

Let f be the sought attenuation image. A CT scan produces photons counts y_ℓ, corresponding to straight lines ℓ through f. Ideally, the measured data comprises of a 2-D Radon transform $\bar{g} = \mathbf{R}f$ (the sinogram) of this image, where $\bar{g}_\ell = \int_\ell f(x)d\ell$ is the line integral along the corresponding line. The relation to the true photon counts is $\bar{y}_\ell = I_0 e^{-\bar{g}_\ell}$, where I_0 is the blank-scan count.

The algorithm in [5] recovers the image f from the noisy sinogram $\hat{g}$ by solving the following optimization problem:

$$\begin{aligned}\{\hat{\alpha}, \hat{\mathbf{D}}, \hat{f}\} = \arg\min_{\alpha,\mathbf{D},f}\{\gamma\|\mathbf{R}f - \hat{g}\|_2^2 + \\ + \textstyle\sum_j \mu_j\|\alpha_j\|_0 + \sum_j \|\mathbf{D}\alpha_j - \mathbf{E}_j f\|_2^2\}.\end{aligned} \tag{1}$$

Here $\mathbf{D}$ is a matrix with unit-norm columns, $\mathbf{E}_j$ extracts the j-th square patch from the image. This patch is sparsely encoded by $\mathbf{D}\alpha_j$. $\|\alpha\|_0$ is the number of non-zero elements in α.

In the training stage, the K-SVD algorithm is employed to optimize (1) for the dictionary $\mathbf{D}$ and the representations

This research was supported by the Israel Science Foundation (ISF) grant number 1031/08.

$\{\alpha_j\}$ simultaneously, using some preliminary image f. These parameters are then kept constant and the objective becomes quadratic optimization problem with respect to the image f. Its solution is the output of the algorithm.

There are a number of drawbacks to the described approach. First, the data fidelity term $\|\mathbf{R}f-\hat{g}\|_2^2$ is formulated in the Radon domain, while the goal is to perform a reconstruction in the image domain. Furthermore, the noise variance is assumed uniform over the sinogram, which leads to the L_2 norm error measure in this term; however, for low-dose measurements, it is important to account for the data-dependent noise statistics. Second, correct weights μ_j are essential for the performance of the algorithm but are hard to compute in this setting, since they depend on the unknown noise statistics in the image domain. These weights are rendered as free parameters. Finally, the algorithm output depends on the preliminary reconstructed image used in the training stage.

We now describe our alternative application of the Sparse-Land approach in CT. In the common statistical model of the CT scan, the measurements y_ℓ are rendered as instances of the random variables $Y_\ell \sim Poisson(\lambda_\ell)$, where $\lambda_\ell = I_0 e^{-\bar{g}_\ell}$.

Using a second-order approximation of the log-likelihood expression $log(\mathbf{P}(Y_\ell = y_\ell|f))$ for this model, one deduces that the variance of the deviation between the measured sinogram value $\hat{g}_\ell = -log(y_\ell/I_0)$, and the ideal value $\bar{g}_\ell$ equals $1/\lambda_\ell$ [6, 7]. In practice, this value is well approximated by $1/y_\ell$. Thus, the goal of sinogram restoration is to minimize the weighted L_2 norm $\|g-\hat{g}\|_{2,\mathbf{W}}$ w.r.t. g, with a diagonal weight matrix $\mathbf{W}$ given by $\mathbf{W}(\ell) = y_\ell$.

Our proposed reconstruction algorithm uses an off-line training procedure, in which two dictionaries are learned. Then, given a noisy data, the sinogram restoration is performed, followed by the standard FBP algorithm.

Off-line procedure: The training requires a set of high-quality reference images and corresponding sinograms contaminated with Poisson noise. This data can be obtained by scanning a number of physical phantoms (or a cadaver) twice, with high and low X-ray doses. Training of the first dictionary is conducted fully in the sinogram domain and uses only the noisy sinograms $\hat{g}$. In the following objective function α_j is a representation vector of the patch $\mathbf{E}_j g$ in terms of $\mathbf{D}$, and T is the number of pixels in a patch. We impose a sparsity prior on the representations α_j of the sinogram patches. In practice, the sparse coding of the patches is substantially better when each patch has zero mean value; therefore mean values of the patches are subtracted before the processing and stored for the sinogram re-assembly stage. This step is not reflected in the equations, in order not to clutter the exposition.

$$\{\mathbf{D}_1, \alpha^*\} = \arg\min_{\mathbf{D},\alpha} \textstyle\sum_j \|\alpha_j\|_0 \;\; s.t. \\ \forall j, \quad |\mathbf{D}\alpha_j - s_j\|_{2,\mathbf{W}}^2 \le T, \quad s_j = \mathbf{E}_j\hat{g}. \tag{2}$$

According to the statistical model, in each sinogram location ℓ the weighted error $\sqrt{y_\ell}(\bar{g}_\ell - \hat{g}_\ell)$ has unit variance and is independent of other locations. Therefore the measure $\|\cdot\|_{2,\mathbf{W}}$ of the difference between the noisy sinogram patch s_j and its sparsely encoded version $D\alpha_j$ (assumed to approximate the true sinogram) has the expected value of T.

The numerical solution of this optimization problem is obtained by a weighted version of the K-SVD algorithm [3], which is generalized for the weighted L_2 norm by using a criss-cross regression [8]. Its iterations consist of two main steps:

- (**Sparse coding**) Solve for each j,
$$\alpha_j^* = \arg\min \|\alpha_j\|_0 \quad s.t. \; \|\mathbf{D}\alpha_j - \mathbf{E}_j\hat{g}\|_{2,\mathbf{W}}^2 \le T. \tag{3}$$
This is done using the Orthogonal Matching Pursuit (OMP) algorithm [9].
- (**dictionary update**) For each column d_i of the dictionary, minimize the error
$$d_i^* = \arg\min_{d_i,\{\alpha_j\}} \sum_{j\in\Omega_i} \|\mathbf{D}\alpha_j - \mathbf{E}_j\hat{g}\|_{2,\mathbf{W}}^2, \tag{4}$$
where Ω_i is the index set of all patches s_j which support includes d_i.

Given the dictionary $\mathbf{D}_1$, we compute for each training image f patch-wise representations α_j^f of the noisy sinogram corresponding to f. Then we perform an additional optimization of the dictionary while keeping the representations fixed. This time the objective function involves the reconstruction transform $\mathbf{T}$ (which is the FBP operation in our case), and the goal is to minimize an error functional in the image domain, applied to the difference between the reference image and the reconstruction:

$$\mathbf{D}_2 = \arg\min_{\mathbf{D}} \sum_{f\in\mathcal{TS}} \|\mathbf{T}\mathbf{M}_E^{-1} \sum_j \mathbf{E}_j^\top \mathbf{D}\alpha_j^f - f\|_{2,\mathbf{Q}}. \tag{5}$$

Here $\mathbf{M}_E = \sum_j \mathbf{E}_j^\top \mathbf{E}_j$ is the correction factor for patches overlapping. The optimizer of (5) is a dictionary close to $\mathbf{D}_1$ which absorbs the effects of applying the reconstruction operator $\mathbf{T}$ to the processed sinograms.

The error functional $\|\cdot\|_{2,\mathbf{Q}}$ is the weighted L_2 norm with a diagonal weight function $\mathbf{Q}$. It is designed to improve the ability of the reconstruction algorithm to recover low-contrast contours: an edge separating two areas of similar intensity is more important than the sharp bone-tissue transition (which is less sensitive to noise). Thus the weight map $\mathbf{Q}$ built for each $f \in \mathcal{TS}$ is inversely proportional[1] to its gradient norm.

On-line procedure: Given the dictionaries $\mathbf{D}_1, \mathbf{D}_2$, the processing of a new noisy sinogram is consists of computing the patch-wise representations $\{\alpha_j\}$ n terms of $\mathbf{D}_1$ and recovering the sinogram as the optimizer of the following objective:

$$g_r = \arg\min_g \lambda\|g-\hat{g}\|_{2,\mathbf{W}}^2 + \sum_j \|\mathbf{D}_2\alpha_j - \mathbf{E}_j g\|_2^2 \tag{6}$$

Here λ controls the trade-off between the data fidelity and a sparsity-promoting term. It is calibrated off-line on the training set, for the best visual quality. The closed-form solution of this problem is

$$g_r = \left(\sum_j \mathbf{E}_j^\top \mathbf{E}_j + \lambda\mathbf{W}\right)^{-1} \left(\sum_j \mathbf{E}_j^\top \mathbf{D}_2\alpha_j + \lambda\mathbf{W}\hat{g}\right). \tag{7}$$

[1] The regions where the gradient values are below a threshold get a small constant weight. The threshold is tuned manually to separate gradient values representing the visually smooth regions from those corresponding to contours in the image.

Finally, the FBP algorithm with non-apodized Ram-Lak kernel is applied to g_r to produce the output image. The method is labeled as ***S**parsity-based **S**inogram **D**enoising* (SSD).

The computational cost of this sinogram pre-processing is linear in the image size, since the operation is local. The sparse-coding OMP step uses only a few atoms from the dictionary, therefore in practice the processing time is low. Experimentally, the on-line part of SSD takes roughly the same time as the following FBP algorithm.

3. NUMERICAL STUDY

In the numerical experiments the sinogram $\bar{g}$ is simulated by applying a pixel-driven implementation of the discrete Radon transform to a discrete reference image. In image of size $n \times n$ pixels we sample $4n$ views (projections) of $\sqrt{2}n$ bins each, evenly distributed over the angle range $[0, \pi]$ (these views are the columns of the sinogram matrix). The noisy photon counts y_ℓ are generated according to the stated model. The X-ray dose is controlled by the value of the blank-scan I_0. It is set to 700 photons, which corresponds to a visibly strong noise in the reconstructed images.

In the experiments we use geometric phantoms, as well as clinical CT images. All images are of dimensions 256×256. Each phantom is built of a large ellipse with boundary, filled with many smaller ellipses with randomly chosen centers, radii and intensity levels. The clinical images are courtesy of Visible Human Project[2], they represent a set of axial sections of the male abdomen. The training is conducted separately for each type of images.

Each sinogram is decomposed into a set of 8×8 overlapping patches, extracted with a 1-pixel shift. The dictionaries $\mathbf{D}_1, \mathbf{D}_2$ of dimensions 64×128 are trained using 20 phantoms. Training data for $\mathbf{D}_1$ training consists of $3 \cdot 10^5$ patches from the noisy sinograms. For phantoms, the average number of atoms required to represent each patch (up to the predicted error) in terms of $\mathbf{D}_1$ is 2.3 atoms when the dictionary is initiated by a subset of the patches and is 1.3 atoms in terms of the trained dictionary. Corresponding numbers for CT images are 5.8 and 2.8 atoms per patch.

One of the quality measures we use is the Signal-to Noise Ratio (SNR), defined for the ideal signal x and a deteriorated version $\hat{x}$ by $SNR(\hat{x}) = -20log_{10}(\|x - \hat{x}\|_2 / \|x\|_2)$. The SNR of the noisy sinogram of a phantom increases during the restoration process (using $\mathbf{D}_1$) from 26 dB to 35 dB. However, a denoising action is also implicitly performed in the standard FBP algorithm by the low-pass apodization filter, therefore the difference in the image domain is not so drastic. The restored sinogram is computed by expression (7) with the value $\lambda = 0.1$. Then the FBP with the basic Ram-Lak filter is applied to produce the output image $\tilde{f}_{SSD}$.

For comparison, we implement the FBP algorithm with an ideal or an apodized[3] Ram-Lak filter, which are labeled $\tilde{f}_{F1}$ and $\tilde{f}_{F2}$ respectively. $\tilde{f}_{F1}$ is the sharpest (and most noisy) image available by FBP. $\tilde{f}_{F2}$ is computed by manually tuning the parameters of the smoothing filter for the highest SNR value on the training set. To appreciate the effective dose reduction, we also produce an image $\tilde{f}_{F2d}$ applying the apodized FBP to a sinogram simulated with a double X-ray dose.

In Figure 1 a visual comparison of SSD and FBP is performed. In the phantom images (upper row) it can be observed that the SSD image displays more faithful image texture than the low-dose FBP images (no spikes and granularity are visible), as well as high detail resolution (the ellipse boundaries are better restored). It is visually comparable to the high-dose $\tilde{f}_{F2d}$ image. In clinical images (lower row) the SSD image features a reduced noise level and better restored details comparing to the $\tilde{f}_{F2}$.

A quantitative comparison of the results is given in Table 3. The standard deviation of the noise is computed in an image region with constant density. The resolution is quantified by approximating the effective point spread function (PSF), applied to f through the projection-reconstruction process in order to obtain the estimate $\tilde{f}$. Technically, we compute the Full-Width-Half-Maximum (FWHM) of the rotation-invariant 2D Gaussian kernel G, which minimizes the MSE $\|\tilde{f} - G * f\|_2$. Of course, it would be desirable to investigate the image quality from the clinical point of view - this could be done via an inspection by a radiologist. Hopefully, in a more realistic setup, such comparison will be conducted in a future work.

It can be observed from the table that the SSD image has a low noise level while by the resolution FWHM measure the image $\tilde{f}_{SSD}$ is second after the sharpest FBP image $\tilde{f}_{F1}$ which is much more noisy. Images $\tilde{f}_{F2d}$ and $\tilde{f}_{SSD}$ display close values in all the measures; SSD image is sharper but has lower SNR, and the behavior of other measures depends on the image type. Overall, these two images are quantitatively comparable, which support our claim on the dose reduction possibility.

Phantoms:

Name	std.dev. of noise	resolution (FWHM)	SNR (dB)	Measure $\|\cdot\|_{2,\mathbf{Q}}$
Phantoms:	$\times 10^{-2}$			$\times 10^{-2}$
$\tilde{f}_{F1}$	12.0	1.35	12.2	1.45
$\tilde{f}_{F2}$	6.2	1.81	16.9	0.63
$\tilde{f}_{F2d}$	5.4	1.66	18.0	0.49
$\tilde{f}_{SSD}$	5.8	1.49	17.6	0.55
Clinical	images:			$\times 10^{7}$
$\tilde{f}_{F1}$	78.37	1.33	21.0	5.24
$\tilde{f}_{F2}$	63.34	1.40	22.48	3.86
$\tilde{f}_{F2d}$	52.33	1.42	23.93	2.79
$\tilde{f}_{SSD}$	39.26	1.27	25.42	1.95

Unfortunately, there is no possibility to directly compare the proposed algorithm with one described in [5], since the latter was presented with the assumption of uniform Gaussian noise model and with no method of choosing the sparse-coding error ϵ.

4. CONCLUSIONS

The introduced algorithm employs the statistical model of the sinogram noise for noise reduction based on sparsity prior in terms of a learned dictionary. The dictionary training op-

[2] Look at $www.nlm.nih.gov/research/visible/visible\_human.html$.

[3] Apodization is performed by applying a low-pass filter, implemented as multiplication of the Ram-Lak filter by a Butterworth window in the Fourer domain.

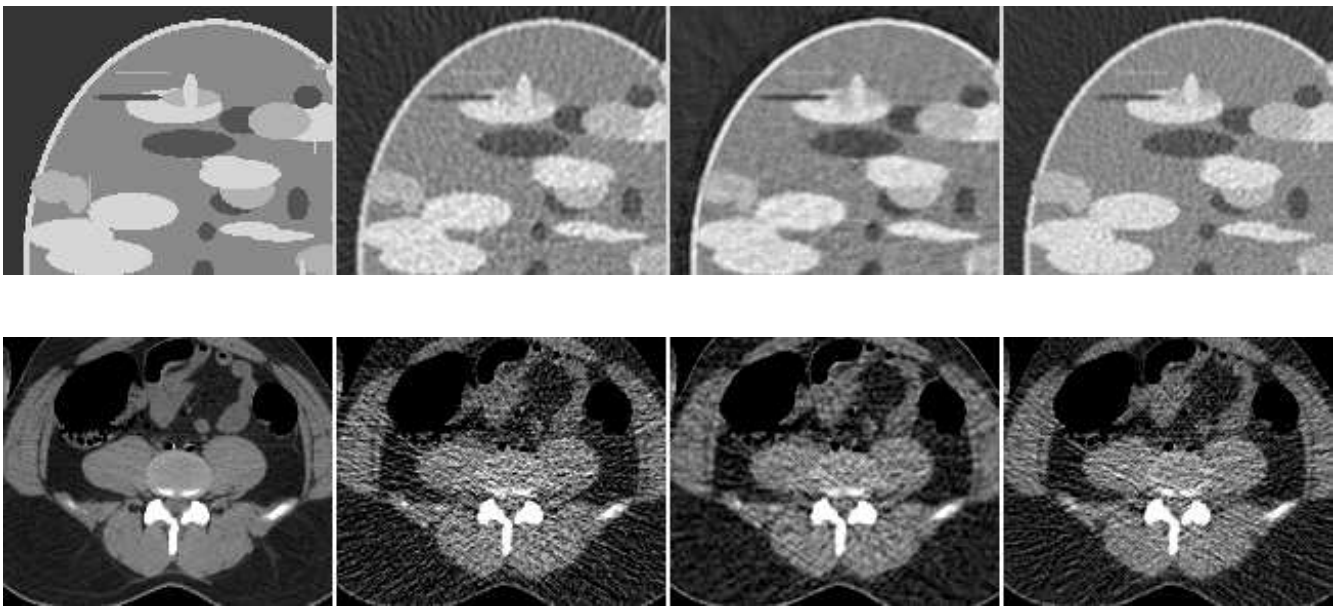

Fig. 1. Upper row: experiment with ellipse phantoms. Lower row: experiment with clinical images, displayed in HU window [-150,270]. Left to right: Reference image, FBP output, SSD output, double-dose FBP result.

timizes the algorithm performance with respect to an edge-preserving error functional in the image domain, expressed in the form of a weighted L_2 norm (notice also that any error measure in the form of the weighed L_2 norm can be easily incorporated in the algorithm instead of the proposed one). The resulting tool for pre-processing of the noisy data reduces the sinogram noise caused by low X-ray dose and effectively allows to maintain the image quality available from higher dose scan with the standard FBP algorithm.

The presented numerical experiments are admittedly synthetic; we use the discrete Radon transform operator to generate the sinogram data from discrete images. Thus, no real-life data distortions, like beam hardening or Compton scatter were confronted. The idea is to present a comparative advantage over the standard algorithm implemented in the same setup, and display a treatment for the isolated (and major) corruption cause which is the Poisson noise on photon counts.

As with any algorithm, based on learning from examples, there is a concern of whether the medical anomalies and special objects will be faithfully recovered. While there is no direct evidence to the robustness of the proposed algorithm, we point to the fact the sparse encoding is performed locally (8×8 squares) and in the sinogram domain; thus, the dictionary atoms will not capture any geometric properties of CT images, but rather the statistical properties of the underlying photon counts.

Due to local nature of the processing, the method can be readily extended to the 3-D scan with a general source trajectory. Another advantage is that the algorithm has a low computational cost for online processing and can easily be incorporated into an existing software of a CT scanner.

5. REFERENCES

[1] J. Bian P.J. La Rivire and P.A. Vargas, "Penalized-likelihood sinogram restoration for computed tomography," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 1022–36, Aug. 2006.

[2] H. Lu Z. Liang J. Wang, T. Li, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose x-ray computed tomography," *IEEE Trans Med Imaging*, vol. 25, no. 10, pp. 1272–1283, 2006.

[3] A. Bruckstein M. Aharon, M. Elad, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal. Proc.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[4] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. on Image Processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[5] H.Y. Liao and G. Sapiro, "Sparse representations for limited data tomography," *ISBI 2008*, , no. 2181, pp. 1375–1378, November 2007.

[6] I.A. Elbakri and J.A. Fessler, "Statistical image reconstruction for polyenergetic x-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 89–99, Feb. 2002.

[7] A. Macovski, *Medical Imaging Systems*, Prentice-Hall, 1983.

[8] S. Zamir K. Ruben Gabriel, "Lower rank approximation of matrices by least squares with any choice of weights," *Technometrics*, vol. 21, no. 4, pp. 489–498, 1979.

[9] P.S. Krishnaprasad Y.C. Pati, R. Rezaiifar, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Twenty-Seventh Asilomar Conference on Signals, Systems and Computers)*, Nov 1993, vol. 1, pp. 40–44.