DIRECT ADAPTIVE ALGORITHMS FOR CT RECONSTRUCTION

Joseph Shtok, Michael Elad, and Michael Zibulevsky

Computer Science department
Technion - Israel Institute of Technology
Haifa 32000, Israel

ABSTRACT

This work concerns with linear and spatially-adaptive direct reconstruction algorithms for 2-D parallel-beam transmission tomography, extending the Filtered Back-Projection (FBP). The standard apodized Ram-Lak filter kernel is replaced with a bank of statistically trained 2-D convolution kernels, leading to improved reconstruction results. Two types of filter training procedures are considered. The first deals with reconstruction from noisy and truncated projections in a predefined region of interest, for images from a known family. In the second algorithm, termed SPADES, the training aims at improving the impulse response properties of the overall projection-reconstruction scheme. In this algorithm, the degree of smoothing applied to the reconstructed image is spatially controlled by a switch rule. Both methods are shown by simulations to operate well and lead to substantially improved reconstruction results.

Index Terms— Computed Tomography (CT), statistical training, spatial adaptivity, Filtered-Back-Projection (FBP).

1. INTRODUCTION

Filtered Back-Projection (FBP) is a basic direct reconstruction algorithm applied in Computed Tomography (CT). It is fast, simple, cheap, and despite its drawbacks, it is widely used in clinical CT scanners. While being theoretically exact in the continuous domain, FBP fails to account for the numerous physical phenomena present in the data acquisition process and especially for discretization errors. Moreover, it lacks the flexibility required to manage partial input data, like truncated projections of a Region Of Interest (ROI) (desired for reduction of X-ray dosage and hardware cost).

These challenges are met to some extent by more elaborate statistically-based algorithms [1]. Here, the objective function can incorporate the statistical acquisition model and account for partial projections data. Unfortunately, prohibitive computational complexity of such algorithms forces the practitioners to use these ideas in a restricted form. For instance, an iterative restoration of the sinogram (the projection data), followed by the standard FBP, is proposed in [2]. The method developed in [3] employs an exemplar-based classification of the sinogram data patches, combined with training of local 2-D projection filters.

In most works that consider FBP-based algorithms, the choice of the projection filter is restricted to the Ram-Lak (ramp) kernel, with a low-pass window applied in order to eliminate high-frequency noise (naturally amplified by the ideal ramp). Such apodized filter requires parameters tuning and leaves place for improvements. In this paper we propose to replace the ramp filter with a bank of statistically trained 2-D convolution kernels, leading to improved reconstruction results. These filters lead to non-linear, spatially-adaptive reconstruction algorithms that are based on FBP-like linear operators.

As a first step towards an adaptive reconstruction procedure, we develop a generalized version of FBP, in which the analytical 1-D projections filter is replaced by one based on a statistically-trained bank of 2-D kernels, and an additional 2-D filter as post-processing after the back-projection stage. The training objective for these kernels can be designed to address different problems in the tomographic reconstruction, including a reconstruction of a certain family of interest, using projections contaminated with Poisson noise, truncation to a ROI, and more. The filters are computed as minimizers of the Mean-Squared-Error (MSE) between the original training images and those reconstructed by the proposed scheme. More details on this algorithm can be found in [4].

Another approach for filter design is to train kernels that would produce an impulse response of the projection-reconstruction pair with controlled measure of blur and low off-focal perturbations. This tool developed here allows us to compute a sequence of reconstructed images with a varying degree of blur, all obtained from the same sinogram. Due to a natural tradeoff between sharpness and noise reduction, in each such result a different region corresponding to certain smoothness is successfully recovered. Thus, a spatially-adaptive local fusion of these images can (and does) result in a high-quality CT reconstruction.

In this work we present such a Spatially-Adaptive Estimation (termed SPADES). We consider two possible fusion rules for it, the first based on a learning procedure (using Support Vector Regression (SVR) or Neural Network), and the second based on a theoretical consideration. SPADES concludes in executing a small number of FBP-like algorithms in parallel, and combining their outputs by a non-linear fusion operation. This leads to a fast and spatially-adaptive reconstruction algorithm that implicitly employs local smoothness information to achieve high image quality.

This paper is organized as follows. In Section 2 we present the AFBP paradigm and how the objective function is defined for designing the filters of interest. Section 3 describes the design procedure for filters that give a controlled Point Spread-Function (PSF) in the projection-reconstruction cycle. In Section 4 we discuss the SPADES algorithm, which fuses these filters' results, getting spatially adaptive yet non-iterative reconstruction result.

2. AFBP: LINEAR RECONSTRUCTION SCHEME

In 2-D transmission tomography, the attenuation image $f(x)$ is projected along straight lines by means of the 2-D Radon transform: $g_\theta(s) = (\mathbf{R}f)_\ell = \int_\ell f(x)dl$, where $\ell = \ell(\theta, s)$ is the line which makes the angle θ with the x axis and passes at distance s from the origin. The Radon transform of $f(x)$ is sampled at large number of fixed angles (evenly covering the range $[0, \pi]$) and fixed signed distances s (bins). The adjoint $\mathbf{R}^*$, also known as the Back-Projection transform, is defined by $\mathbf{R}^* g(x) = \int_\theta g_\theta([\cos(\theta), \sin(\theta)] \cdot x) d\theta$. We let $\mathcal{A} \subset \mathbb{R}^2$ stand for the domain of an attenuation image, and $\mathcal{P} \subset \mathbb{R}^2$ denote the domain of projections.

In reality, the exact Radon operation is replaced by random Poisson measurements that account for the number y_ℓ (for each projection line ℓ as specified above) of photons received in a unit of time by the corresponding detectors. y_ℓ is modeled as a realization of the random variable $Y_\ell \sim$ Poisson $\left(I_0 e^{-(\mathbf{R}f)_\ell}\right)$. The source X-ray intensity I_0 determines the scale of the parameter of the Poisson distribution, thus influencing the noise level. In sequel we denote $g_\theta(s) = -log\frac{y_\ell}{I_0}$, the Maximum Likelihood estimate of the $\mathbf{R}f$.

The linear reconstruction scheme, labeled as *Adaptive Filtered Back-Projection* (AFBP), is defined to be the transform

$$\mathbf{T}_\kappa = \mathbf{F}_{\kappa^\mathcal{A}} \circ \mathbf{R}^* \circ \mathbf{F}_{\kappa^\mathcal{P}} : \mathcal{P} \to \mathcal{A} \quad (1)$$

with parameter set $\kappa = \{\kappa^\mathcal{P}, \kappa^\mathcal{A}\}$ (a corresponding block diagram is presented in Figure 1).

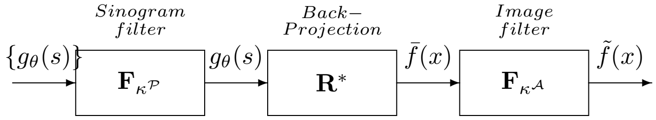

Figure 1: AFBPt reconstruction scheme

Assume that the projections are truncated to the central disk of radius D. We partition this range of $s \in [-D, D]$ into d disjoint sub-segments $[-D, D] = \bigcup_{i=1}^{d} D_i$, and consider a bank $\kappa^\mathcal{P} = \{\kappa_i^\mathcal{P}\}_{i=1}^{d}$ of 2-D convolution kernels, each corresponding to a different segment. The sinogram filter is applied by convolving the neighborhood of the segment D_i of each projection with the kernel $\kappa_i^\mathcal{P}$ (we refer here to the 2-D neighborhood of the projection in the sinogram matrix). In the image domain, a single 2-D convolution kernel $\kappa^\mathcal{A}$ is applied to the Back-Projection output $\bar{f}(x)$.

By replacing the classical 1-D projection filter with 2-D filters, we allow contributions from the neighboring projections in the filtering process. The motivation for using a spatially-variant sinogram filter (a different kernel for each projection segment) comes from considerations regarding reconstruction with missing data. When the projections are truncated to a region of interest, the central part of the remaining projections should be filtered using a symmetric kernel with small spatial support in order to reduce the truncation error. For projection pixels residing near the edges of this region, the information should be gathered in a non-symmetric way, only from the non-truncated part of the projection. In addition, the tuned high-dimensional parameter set (values of the convolution kernels) of $\mathbf{T}_\kappa$ allows to absorb various imperfections in the complex projection-reconstruction process, such as phenomena in acquisition process, discretization errors and minor effects that are not modeled into the reconstruction scheme.

Different training procedures for derivation of the kernel bank κ yield different reconstruction algorithms, as detailed below.

We now turn to describe the methodology developed in [4]. This training objective is designed to account for the following considerations: (1) The attenuation image is expected to come from a certain family of images, represented by an available training set; (2) The projections are contaminated with Poisson noise, stemming from low photon count, with known source intensity; (3) The projections are truncated to a disk containing the ROI plus a small margin. Given a set $\mathcal{F}_{tr}$ of representative images, we build the training set $\mathcal{G}_{tr}$ consisting of noisy truncated sinograms by

$$\mathcal{G}_{tr} = \{g_f^k = (\mathbf{R}f + \xi_f^k)_{|ROI} \mid f \in \mathcal{F}_{tr},\ k = 1...K\}, \quad (2)$$

where $\{g_f^k\}_{k=1}^K$ are generated from $f \in \mathcal{F}_{tr}$ by applying the Radon transform and generating K instances of the Poisson noise.

Using the above training sets, the kernel bank κ is then computed as an optimizer of the objective function

$$(\kappa^*) = \arg\min_\kappa \sum_{g_f^k \in \mathcal{G}_{tr}} \|(\mathbf{T}_\kappa g_f^k - f)_{|ROI}\|_2^2. \quad (3)$$

This function is quadratic in each one of the kernels in $\kappa^\mathcal{P}$ and the $\kappa^\mathcal{A}$. Thus for each of the two sets of variables we can solve the optimization problem using the Conjugate Gradient method, applied to the corresponding linear equation. The training is then carried out in turns, fixing one set of variables and updating the other. Numerical results on a testing image of size 256×256 with reconstructed ROI of radius 30 pixels are presented in Figure 2. As can be seen, the Signal to Noise Ratio (SNR) in the ROI is substantially improved. This algorithm is labeled as AFBPt.

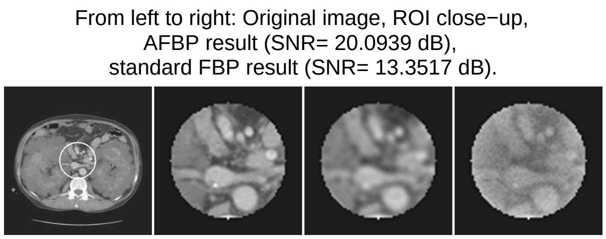

Figure 2: ROI reconstruction from truncated projections

3. AFBPS: PSF-DRIVEN TRAINING

Any linear, spatially-invariant operator $\mathbf{L}$ is completely characterized by its impulse response, also called the Point Spread Function (PSF). This is defined as the image of the Dirac delta function under $\mathbf{L}$. Let $\mathbf{T}$ stand for any linear reconstruction operator $\mathbf{T} : \mathcal{P} \to \mathcal{A}$ and consider the PSF of the composition $\mathbf{T} \circ \mathbf{R}$ (which is - ideally - a linear, spatially invariant operator on $\mathcal{A}$). When $\mathbf{T}$ represents the FBP algorithm, the aforementioned PSF deviates from the ideal one (Dirac delta function) in a number of ways: the central spike is a spread spot, it lacks radial symmetry, and it exhibits a spatial variance. Moreover, it has off-central distortions of global nature (due to the fact 2-D Radon is a global transform).

We propose to train the parameter set κ of the AFBP reconstruction operator $\mathbf{T}_\kappa$ such that the PSF of $\mathbf{T}_\kappa \circ \mathbf{R}$ has a controlled amount of blur and diminished off-focal energy. To that end, κ is computed as the minimizer of an objective function, whose structure is detailed below. For any pixel p, let $f_p(x)$ be the image defined by $f_p(p) = 1$ and zeros elsewhere. Also, let $g_p = \mathbf{R} f_p$, be the projection set created by such an image. The objective function focused on p is

$$\begin{aligned} \Phi(\kappa, p) \quad &= \quad \|(\mathbf{T}_\kappa g_p - 1)_{|R_p}\|_2^2 \\ &\quad + \lambda \|(\mathbf{T}_\kappa g_p)_{|R_p^C}\|_2^2 + \mu \|\kappa^{\mathcal{P}}\|_2^2. \end{aligned} \qquad (4)$$

Here R_p is a small disk centered at p, and R_p^C is its complement in the domain. This objective penalizes the off-focal energy of the PSF $\mathbf{T}_\kappa g_p$ while keeping constant its integral over its central disk. The last component serves as a regularizer, diminishing the energy of the convolution kernels in the projection domain, which reduces noise amplification by the filter. The radius of R_p controls the amount of blur present in the PSF. Since the reconstructed image $\tilde{f}(x)$ can approximately be described as the original $f(x)$ convolved with the PSF, this blur is just the degree of smoothing applied to the original image in the reconstruction process.

Since the operator $\mathbf{T} \circ \mathbf{R}$ is not truly spatially-invariant, we use an objective function summarizing the penalty $\Phi(\kappa, p)$ over some set Ω of image locations p. Our choice for Ω is a coarse rectangular grid, covering the entire image domain and making an acute angle with its coordinate system. To summarize, the kernel bank κ is computed via $\kappa^* = arg \min_\kappa \sum_{p \in \Omega} \Phi(\kappa, p)$. The resulting algorithm is labeled as AFBPs.

The degree of blur introduced by a given reconstruction operator is reflected in its noise tolerance properties: the tradeoff between the X-ray dosage and the reconstruction quality gradually changes as the measure smoothness grows. Graphs of some typical cases, displayed in Figure 3, illustrate this phenomenon.

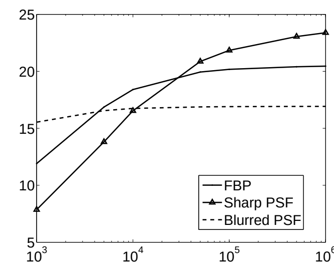

Figure 3: Performance of different kernels on CT images: SNR versus the X-ray power.

4. SPADES: SPATIALLY-ADAPTIVE ESTIMATOR

Our goal is to bridge the gap between the direct and the iterative algorithms by an FBP-based, spatially-adaptive reconstruction scheme. The motivation comes from locally adaptive denoising techniques that apply a smoothing operator with varying degree of blur, dependent on local smoothness of the image (see [5] and references therein). An analogue to denoising with a fixed convolution kernel μ is the AFBPs algorithm trained to produce the PSF which has the shape of μ. The degree of smoothing can be adapted locally in the following way: train a sequence of AFBPs algorithms with gradually increasing spread of the PSF, compute the corresponding sequence of reconstructed images, and apply a local switch rule that would chose the appropriate PSF to be used for each location.

We have designed such a switch rule based on the model in [5]. It addresses the problem of recovering a signal x from observations $y = x + \xi$, where ξ is a homogeneous Gaussian noise. At each location p in the signal domain, a sequence of estimates $\tilde{x}_i$ is computed by fitting to the data a least-squares polynomial of degree m in a disk with growing radius δ_i. Then a rationale based on confidence intervals is employed to define a local switch rule, which chooses one of the estimate values $\tilde{x}_i(p)$.

We have modified and applied this technique to CT reconstruction: different estimates $\tilde{x}_i$ of the original signal are replaced by reconstructed images $\tilde{f}_i = \mathbf{T}_i(g)$. Here $\{\mathbf{T}_i\}_{i=1}^m$ is the sequence of AFBPs transforms with a growing degree of PSF smoothness. While the noise, present in the reconstructed image is neither Gaussian nor independent of data, its second-order statistics can be estimated in order to compute the confidence intervals for $\tilde{f}_i(p)$ at each location p. Then a switch rule, similar to one developed in [5], is applied to choose the most appropriate from the available values $\{\tilde{f}_i(p)\}_{i=1}^m$.

Our experiments show that while this algorithm gives improved image quality over the AFBPt output, the gain is small. This may be explained by insufficient tuning of various parameters, and the assumptions our method relies on. This leads us to a learning-based alternative, described below.

To build a powerful local fusion rule we resort to Support Vector Regression (SVR) or a Neural Network (NN). These are trained on a set $\mathcal{F}_{tr}$ of representative images. For every $f \in \mathcal{F}_{tr}$ we compute the projections $g_\theta(s) = \mathbf{R}f + \xi$ where ξ is the Poisson noise. A sequence of AFBPs reconstruction operators, $\{\mathbf{T}_i\}_{i=1}^m$, described earlier, is then applied to compute the versions $\tilde{f}_i$ of the image. In addition, we use the AFBPt operator $\mathbf{T}_t$, also trained on the set $\mathcal{F}_{tr}$ (see Section 2 for details).

For each image location p, we extract a vector of features containing the following values: (1) The sequence $\tilde{f}_i(p)$, $i = 1, ..., m$; (2) The value $\tilde{f}_t(p)$; and (3) The values of $\tilde{f}_t$ in the 8-neighborhood of p. At the training stage this vector of features, along with the true value $f(p)$, is passed to the learning procedure (this data is extracted from evenly distributed locations in training images). Notice that the information provided is local, and therefore, after the learning process, it is expected to successfully resolve the pixel values for any CT images.

The resulting SPADES reconstruction scheme is given in the Figure 4.

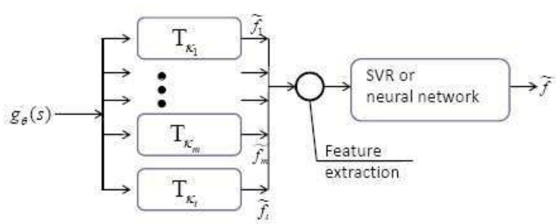

Figure 4: SPADES reconstruction scheme

In our numerical experiments, both SVR and NN have succeeded to improve the reconstruction quality beyond the performance of our best linear estimator $\mathbf{T}_t$ on the testing set compatible to $\mathcal{F}_{tr}$. Results presented here are for a single-layer Neural Network. Its output function is defined as $y(x, w, v) = \sum_{j=1}^N v_j \sigma(\sum_{i=1}^M w_{i,j} x_j + w_{M+1,j})$, where N is the number of neurons, M is the size of the input feature vector, $w_{i,j}$ is the weight on edge connecting i-th input to j-th neuron, and $\sigma(x) = x/(1 + |x|)$ is the sigmoid function.

Figure 5 displays (negatives of) difference images from numerical experiments with 256×256 axial CT thorax obtained from a clinical radiology website MedPix images[1]. A small neural network of 50 neurons was trained and applied to produce the reconstruction. The object background was eliminated in SNR measurements.

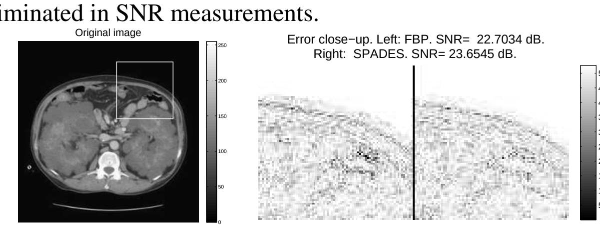

Figure 5: FBP vs. SPADES

5. SUMMARY

In this work we present direct reconstruction algorithms that enjoy the computational complexity of the standard FBP, while leading to improved performance. These algorithms can be easily incorporated in an existing FBP-based reconstruction framework and serve a wide range of tasks in Computed Tomography.

6. REFERENCES

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

[2] J. Bian P.J. La-Riviere and P.A. Vargas, "Penalized-likelihood sinogram restoration for computed tomography," *IEEE Trans. on Medical Imaging*, vol. 25, no. 8, pp. 1022–36, 2006.

[3] K.D. Sauer B.I. Anda and C.A. Bouman, "Nonlinear back-projection for tomographic reconstruction," *IEEE Trans. on Nuclear Science*, vol. 49, no. 1, pp. 61–68, 2002.

[4] M. Zibulevsky J. Shtok, M. Elad, "Adaptive filtered-back-projection for computed tomography," in *IEEEI*, 2008, pp. 528–532.

[5] A. Goldenshluger and A. Nemirovsky, "On spatial adaptive estimation of nonparametric regression," *Mathematical Methods of Statistics*, vol. 6, pp. 135–170, 1997.

[1]courtesy of Prof. James Smirniotopoulos, Uniformed Services University of the Health Sciences