Invited Paper

Regularization in tomographic reconstruction using thresholding estimators

Jérôme Kalifa[a], Andrew Laine[a] and Peter D. Esser[b]

[a]Department of Biomedical Engineering, Columbia University, New York, NY
[b]Department of Radiology, Columbia-Presbyterian Medical Center, New York, NY

ABSTRACT

In tomographic medical devices such as SPECT or PET cameras, image reconstruction is an unstable inverse problem, due to the presence of additive noise. A new family of regularization methods for reconstruction, based on a thresholding procedure in wavelet and wavelet packet decompositions, is studied. This approach is based on the fact that the decompositions provide a near-diagonalization of the inverse Radon transform and of the prior information on medical images. An optimal wavelet packet decomposition is adaptively chosen for the specific image to be restored. Corresponding algorithms have been developed for both 2-D and full 3-D reconstruction. These procedures are fast, non-iterative, flexible, and their performance outperforms Filtered Back-Projection and iterative procedures such as OS-EM.

Keywords: 2-D and 3-D tomographic reconstruction, PET, SPECT, wavelet packets, dyadic wavelet transform.

1. INTRODUCTION

We are interested in the problem of tomographic reconstruction of images from transmission data, which we call tomographic projections or *sinograms*. Although the work presented here has a wide range of applications for various tomographic devices, we will focus on medical images with SPECT and PET cameras.

A slice of an object observed by a tomographic device is represented by a 2-D discrete image $f[n_1, n_2]$. An estimation of f must be computed with a tomographic reconstruction procedure from the sinograms produced by tomographic devices, denoted $Y[t, \alpha]$, and defined as:

$$Y[t, \alpha] = \mathcal{R}(f[n_1, n_2]) + W[t, \alpha] \tag{1}$$

where $\{f[n_1, n_2]\}_{0 \leq n_1 \leq N_1 - 1,\, 0 \leq n_2 \leq N_2 - 1}$ is the observed image, W is an additive noise, and $\mathcal{R}$ is the discrete Radon transform which models the tomographic projection process. The discrete Radon transform is derived from its continuous version $\mathcal{R}_c$, which is equivalent to the X-ray transform in two dimensions and is defined as[1]

$$(\mathcal{R}_c f_c)(t, \alpha) = \int_{\mathbb{R}} \int_{\mathbb{R}} f_c(x_1, x_2) \delta(x_1 \cos\alpha + x_2 \sin\alpha - t) dx_1 dx_2. \tag{2}$$

where $f_c(x_1, x_2) \in \mathbf{L^2}(\mathbb{R}^2)$, δ is the Dirac mass, $\alpha \in [0, 2\pi)$, and $t \in \mathbb{R}$. There are several different ways to define the discrete Radon transform based on the continuous Radon transform.[2] Typically, a line integral along $x_1 \cos\alpha + x_2 \sin\alpha = t$ is approximated by a summation of the pixel values inside the strip $t - 1/2 \leq n_1 \cos\alpha + n_2 \sin\alpha < t + 1/2$.

When 3-D data is available, we consider that we have a series of tomographic projections of N_3 translated 2-D slices of the observed object. When necessary, the tomographic projections have been transformed via rebinning techniques in order to obtain tomographic projections of 2-D slices: this approach is in general not necessary for SPECT images, but is increasingly common in 3-D PET imaging.[3] The 3-D dataset is written

$$\forall\ 0 \leq z < N_3,\ \ Y[t, \alpha, z] = \mathcal{R}(f[n_1, n_2, z]) + W[t, \alpha, z]. \tag{3}$$

The noise W is usually modelled as a Poisson, or sometimes Gaussian noise. However, since the tomographic projections Y are often processed to incorporate various corrections, such as attenuation correction, scatter correction, resolution correction or geometric correction, the resulting noise is also distorted and does not always comply with such prior statistical models.

Wavelets: Applications in Signal and Image Processing IX, Andrew F. Laine, Michael A. Unser, Akram Aldroubi, Editors, Proceedings of SPIE Vol. 4478 (2001) © 2001 SPIE · 0277-786X/01/$15.00

A tomographic reconstruction procedure must incorporate the following steps: first, a *backprojection* which may be viewed as the application of a discretized inverse Radon transform $\mathcal{R}^{inv}$ on the tomographic projections Y. This can be directly computed with a radial interpolation and a deconvolution to amplify the high frequency components of the tomographic projections Y in the direction of t. This deconvolution comes from the fact that the Radon transform is a smoothing transform. Consequently, backprojecting in the presence of additive noise is an *ill-posed* inverse problem: numerically speaking, a direct computation of $\mathcal{R}^{inv}Y$ is contaminated by a large additive noise $Z = \mathcal{R}^{inv}W$, which means that a *regularization* has to be incorporated in the reconstruction procedure.

Current approaches for regularization in tomographic reconstruction can be summarized as follows:

1. Filtered Back-Projection (FBP) is a linear filtering technique in the Fourier space. FBP suffers from performance limitations due to the fact that the sinusoids of the Fourier basis are not adapted to represent spatially inhomogeneous data as found in medical images.
2. Iterative statistical model-based techniques are designed to implement Expectation-Maximization (EM) and Maximum A Posteriori (MAP) estimators.[4,5] In some cases, these approaches can provide an improvement over FBP, but these estimators suffer from the following drawbacks:
   - *Computation time.* Almost all the corresponding algorithms are too computers-intensive for clinical applications, with the exception of OS-EM[6] (an accelerated implementation of an EM estimator). In MAP methods, useful priors usually give local maxima, and the computational cost of relaxation methods is prohibitive.
   - *Theoretical understanding and justification.* EM estimation lacks a theoretical foundation to understand and characterize the estimation error. Some MAP estimators are in some cases better understood, yet no optimality for a realistic model has been established.
   - *Convergence.* EM estimators are ill-conditioned, in the sense that the corresponding iterative algorithms should be stopped after a limited number of iterations. Beyond this critical number, the noise is magnified, and EM and OS-EM converge to a non-ML (Maximum Likelihood) solution.

This study aims to adress these limitations by building a family of estimation procedures which provide better numerical results, both metrically and perceptually, with a sound theoretical basis and for which the estimation error is understood and characterized. The estimators should also be implemented with fast and flexible algorithms.

2. THRESHOLDING ESTIMATORS

Let f_i be the discrete inverse Radon transform of the discrete Radon transform of f:

$$f_i = \mathcal{R}^{inv}(\mathcal{R}(f)).$$

The difference image $f - f_i$ is the radial interpolation error, and is always very low compared to the estimation error due to the presence of noise. In this paper, our focus is not on interpolation techniques, but on regularization: the image f_i is considered to be our reference (ideal) image.

The estimation problem in (1) is also equivalent to the denoising problem

$$X = f_i + Z \tag{4}$$

where $X = \mathcal{R}^{inv}Y$ and $Z = \mathcal{R}^{inv}W$. If the noise Z was white, Donoho and Johnstone have established[7] that a thresholding estimator in a properly selected vector family $\mathcal{B} = \{g_m, g_m^*\}_{0 \le m \le N_1 * N_2 - 1}$, typically a wavelet basis, would be optimal to recover spatially inhomogeneous data as found in medical images. A thresholding estimator $\tilde{F}$ of f_i in $\mathcal{B}$ is defined as

$$\tilde{F} = \sum_m \rho_m \left(\langle X, g_m \rangle \right) g_m^* \ , \tag{5}$$

where ρ_m is a thresholding operator. The typical and simplest thresholding rules include hard thresholding

$$\rho_m(x) = \begin{cases} x & \text{if } |x| > T_m \\ 0 & \text{if } |x| \le T_m \end{cases} \ , \tag{6}$$

and soft thresholding

$$\rho_m(x) = \begin{cases} x - T_m & \text{if } x \geq T_m \\ x + T_m & \text{if } x \leq -T_m \\ 0 & \text{if } |x| \leq T_m \end{cases} . \quad (7)$$

In our situation, the choice of the decomposition $\mathcal{B}$ does not only depend on the prior information on the object f_i, but also on the backprojected noise Z, whose behavior is very specific due to the fact that it has been distorted by a backprojection. The assumption underlying thresholding estimators is that each coefficient in the decomposition $\mathcal{B}$ can be estimated independently without a loss of performance. As a consequence, such estimators are efficient if the coefficients of the noise and of the object to be recovered are indeed nearly independent in $\mathcal{B}$. This means that $\mathcal{B}$ must provide a near-diagonalization of the noise Z and of the prior information on the image f_i.

The image f_i is a spatially inhomogeneous, piece-wise regular signal, which is compactly represented in a wavelet decomposition. To obtain a diagonal representation of the noise Z, we want to find a decomposition in which the discrete inverse Radon transform (backprojection) is nearly diagonal. Since the inverse Radon transform is a Calderon-Zygmund operator,[8] it is also nearly-diagonal in a wavelet basis.

These two properties of wavelet bases led Donoho[9] to suggest the use of thresholding estimators in wavelet bases for several linear inverse problems, including the inversion of the Radon transform. Donoho established the asymptotic optimality (minimax sense) of this approach, called Wavelet-Vaguelette Decomposition (WVD), and its superiority with respect to other approaches such as Filtered BackProjection. However, the WVD as studied by Donoho is a theoretical concept developped for continuous models, based on the assumption that the additive noise W is Gaussian white. Despite numerical implementations and refinements by other researchers,[10,11] this technique is not as such adapted to the tomographic reconstruction of real medical images. In the meanwhile, Kalifa and Mallat[12] generalized Donoho's approach to adapt it to other types of decompositions, not restricted to wavelet bases.

The choice of the best time-frequency decomposition in which the thresholding estimation is computed is a matter of compromise between the representation of the backprojected noise Z and the representation of the data f_i to be recovered. There is not one "best" time-frequency decomposition (such as a Fourier basis or a wavelet basis) which fits all applications of SPECT and PET imaging. The estimators constructed in this study exhibit the same asymptotic optimality properties as the wavelet-vaguelette estimator of Donoho, but they have been adapted to the specific difficulties arising in medical tomographic imaging, such as the type of noise, the low resolution, the distortions and the perceptual constraints for the diagnosis.

Wavelet packet bases are decompositions which can provide a compact representation of the observed image f, and since a wavelet packet transform provides a more accurate segmentation of the frequency domain than a wavelet transform, this improves the near-diagonalization of the noise Z. Besides, a wavelet packet basis can be adaptively chosen from a dictionary of different wavelet packet bases. We shall see in the next section how this enables us to optimize the choice of the wavelet packet transform to a specific type of observed image and to the nature of the backprojected noise Z.

3. CHOICE OF WAVELET PACKET DECOMPOSITION

A wavelet packet dictionary is a rapidly constructible set of different numerous orthogonal bases $\{\mathcal{B}^\gamma\}_\gamma$. It is possible, within this dictionnary, to search for a "best" basis $\mathcal{B}^{\gamma_1}$ for a specific problem, according to a criterion chosen in advance. This criterion is usually a cost function which is minimal in the best basis. This best basis is computed using the fast best basis algorithm of Coifman and Wickerhauser,[13] with $O(N \log N)$ operations for an image of N samples.

Regularization in tomography is an estimation problem, and the best basis $\mathcal{B}^{\gamma_1}$ for estimating f_i is obtained empirically by minimizing an estimation of the final quadratic estimation error (risk)

$$r(\tilde{F}, f_i) = \mathcal{E}(\|\tilde{F} - f_i\|^2).$$

Two alternatives are proposed. In both cases, the resulting best basis is designed to discriminate the noise Z and the information in the signal, hence a thresholding can remove most of the noise without removing information.

3.1. Use of phantom images

Phantom images are synthetic images of observed organs or anatomical structures, without any noise or artifacts. Figure 1 shows an example of a brain phantom image. A phantom image f_{ph} provides a reasonable representation of how the image f_i of the observed object should appear. When phantom images of the observed organ are available, they can be used for the computation of the best basis, assuming that the phantom image is a mathematical model of the image f_i to be recovered.

If the thresholding operator ρ is a hard thresholding, equation (5) becomes

$$\tilde{F} = \sum_m a_m \langle X, g_m \rangle g_m^* \ , \tag{8}$$

where a_m is either 0 or 1. When $a_m = 1$, the quadratic estimation error on the corresponding coordinate is equal to the variance of the random variable $\langle Z, g_m \rangle$ of the coordinate of the backprojected noise Z. When $a_m = 0$, the quadratic estimation error is the energy $|\langle f_i, g_m \rangle|^2$ of the coordinate of f_i. The optimal choice of the values of a_m depends on the signal f_i which is unknown in practice; however the phantom image f_{ph} can be used as model for f_i, in which case the cost function for a given wavelet packet basis $\mathcal{B}^\gamma = \{g_m^\gamma, g_m^{\gamma,*}\}_{0 \le m \le N_1 * N_2 - 1}$ is

$$\mathcal{C}(\mathcal{B}^\gamma) = \sum_m \min(|\langle f_{ph}, g_m^\gamma \rangle|^2, (\mathcal{E}\langle Z, g_m^\gamma \rangle)^2)$$

which can be computed in practice with a numerical model of the noise Z (see below). The best basis algorithm is used to find the wavelet packet basis $\mathcal{B}^\gamma$ such that $\mathcal{C}(\mathcal{B}^\gamma)$ is minimal.

3.2. Use of the Stein Unbiased Risk Estimator (SURE)

The Stein Unbiased Risk Estimator[14] is an estimator of the risk when ρ is a soft thresholding operator. For a wavelet packet basis $\mathcal{B}^\gamma$, it is given by

$$\tilde{r}^\gamma(f_i) = \sum_m \Phi_m(|\langle X, g_m^\gamma \rangle|^2) \, , \tag{9}$$

with

$$\Phi_m(u) = \begin{cases} u - \sigma_m^2 & \text{if } u \le T_m^2 \\ \sigma_m^2 + T_m^2 & \text{if } u > T_m^2 \end{cases} \ . \tag{10}$$

where σ_m is the standard deviation of the random variable $\langle Z, g_m^\gamma \rangle$.

The empirical best basis $\mathcal{B}^{\gamma_1}$ for estimating f is obtained by minimizing the estimated risk

$$\tilde{r}^{\gamma_1}(f_i) = \min_\gamma \tilde{r}^\gamma(f_i) \ . \tag{11}$$

The estimated risk is calculated in (9) as an additive cost function over the noisy coefficients.

The SURE-based approach to compute the best basis is in general the easiest to implement because it can be difficult to obtain phantoms whose properties, such as dynamics as well as spatial and spectral behaviors, are close to the images to be reconstructed.

3.3. Model of the noise

The cost functions used to compute the best basis algorithm depend on the backprojected noise Z. To generate a model of the noise Z, it is necessary to first generate a model of the additive noise W observed on the sinograms, and the model of the noise Z is obtained by backprojecting the model of the noise W. The goal here is not to find an accurate estimation of the realization of the noise W on the available data, but to evaluate its amplitude as well as its spatial and spectral behavior.

Depending on the type of images, the noise W is usually assumed to be either a Gaussian white noise or a Poisson noise. However, in some cases, the noise W cannot be modelled using a statistical prior. This is the case when the sinograms Y have been distorted by various corrections (such as scatter, attenuation, resolution or geometry), which have also significantly modified the statistical nature of the resulting noise W.

- When the noise W is assumed to be a white Gaussian noise, the problem is to estimate its standard deviation. Donoho and Johnstone[7] showed that an accurate estimator can be calculated from the median of the finest scale wavelet coefficients. Once the standard deviation has been estimated, a numerical model of W is computed using a white noise random generator.

- When the noise W is assumed to be a Poisson noise, the sinograms Y are roughly denoised using a Poisson intensity estimation method by Fryzlewicz and Nason.[15] The resulting denoised sinograms Y_{den} cannot be backprojected and produce tomographic images of good quality, however the difference $Y - Y_{den}$ between the original and the denoised sinograms is a good estimation of the Poisson noise W.

- When there is no available statistical prior for the noise W, a numerical model of W must be generated. Depending on the imaging device, one way to do this is to generate "empty" sinograms, for which no objects were put inside the camera but noise can still be observed, or to use the sinograms at the extremities of 3-D datasets, when available, where the sections have no information but only noise. These techniques work only when the noise is not strongly data-dependent.

Note that in general, it is safe to assume that the noise W can be modelled as a Poisson noise.

4. RECONSTRUCTION ALGORITHM AND NUMERICAL RESULTS

The tomographic reconstruction algorithm is decomposed in the following steps:

1. Backprojection without regularization of the tomographic projections Y to obtain the backprojected image $X = f_i + Z$.
2. (Optional) Computation of the best wavelet packet basis $\mathcal{B}^{\gamma_1}$ optimized for the specific image to be restored, using one of the two methods presented in section 3. The best basis can be recomputed for each image, or can have been computed once and stored in advance, to save computation time. However the computation of the best basis algorithm is very fast.
3. Wavelet packet transform of the backprojected image X in the best basis $\mathcal{B}^{\gamma_1}$ to obtain the wavelet packet coefficients $\{\langle X, g_m \rangle\}_m$.
4. Thresholding on the wavelet packet coefficients.
5. Inverse wavelet packet transform of the thresholded coefficients to obtain the estimated image $\tilde{F}$.

The wavelet packet transform and its inverse are computed with fast filter bank algorithms of complexity $O(N)$ for signals of N samples.[16] Numerical results are improved if the wavelet packet transform and its inverse are undecimated, i.e. translation-invariant, in which case the filter bank algorithm is the "à trous" algorithm.[16]

Figure 2 compares numerical results computed on SPECT clinical data of bone images, using an OS-EM reconstruction, a Filtered Back-Projection and a wavelet-packet based reconstruction. The OS-EM reconstructed image is very smooth because the OS-EM algorithm has to be stopped after a limited number of iterations, otherwise the noise is strongly amplified and the algorithm converges to a noisy reconstructed image. On the other hand, the FBP-reconstructed image is corrupted by a significant noise and artifacts, which cannot be reduced unless the reconstructed image becomes extremely smoothed. With the wavelet packet reconstruction algorithms, the amount of smoothness of a reconstructed image can be controlled precisely, while the noise is reduced significantly as compared to an image reconstructed with FBP or OS-EM. Besides, wavelet packet reconstruction methods are adapted to a broader range of images (different objects, low or high count) than OS-EM and FBP. The superiority of the wavelet-packet based reconstruction method over FBP and OS-EM is currently being established through ROC studies on various SPECT and PET data.[17]

Figure 3 compares numerical results on PET images of a Jaszak phantom, using a Filtered Back-Projection and a wavelet-packet based reconstruction. On the FBP-reconstructed image, only three out of the five holes cannot be confused with noise. The noise can be interpreted as information, and the small variations in amplitude, especially on the three holes at the left of the image, are almost invisible. Besided, the shape of the most visible hole, at the bottom right, is not as round and regular as it should be. The five holes are visible on the WP-reconstructed image, and none of the remaining noise could be misinterpreted as information.

5. EXTENSION TO 3-D RECONSTRUCTION

So far, the wavelet packet reconstruction has been presented for 2-D reconstruction of slices. We now consider equation (3), where we have 3-D data in the form of a series of tomographic projections of N_3 translated 2-D slices of the observed object. It is useful to take advantage of the correlations of the signal in the transaxial direction (z-axis) to obtain a better discrimination between the information and the noise. In this case, a regularization is computed on the whole 3-D data, but the backprojections are still computed slice by slice, since we are in the case where the Radon transform is still a bidimensional operator.

The best results are obtained with a combination of a slice-by-slice 2-D regularization in a wavelet packet decomposition, using the algorithm of section 4, and a supplemental fully 3-D regularization on the whole 3-D volume, using a second thresholding estimator in a 3-D dyadic wavelet decomposition: for each section z, an estimation $\tilde{F}[.,.,z]$ is first computed with a thresholding of the wavelet packet coefficients. The error $\tilde{F} - f$ can be considered as a residual noise. This noise has a power spectrum which is nearly flat at high frequencies, and it is nearly diagonalized in a wavelet decomposition.

To take fully advantage of the 3-D information in the data, we want to apply a 3-D dyadic wavelet transform on the volume, where the wavelets can be adaptively oriented perpendicularly to the singularities of the signal. This directional selectivity enables us to maximize the correlation between the vectors of the wavelet family and the information of the signal. The efficiency of noise removal is thus greatly improved.

A 3-D dyadic wavelet transform is computed with a family of wavelets which are the discretized translations and dilations of three wavelets ψ^1, ψ^2 and ψ^3 that are the partial derivatives of a smoothing function θ:

$$\psi^1(x_1,x_2,x_3) = \frac{\partial\theta(x_1,x_2,x_3)}{\partial x_1} \quad ,\psi^2(x_1,x_2,x_3) = \frac{\partial\theta(x_1,x_2,x_3)}{\partial x_2} \quad ,\psi^3(x_1,x_2,x_3) = \frac{\partial\theta(x_1,x_2,x_3)}{\partial x_3}. \tag{12}$$

For a given j, ψ_j^1, ψ_j^2 and ψ_j^3 have been equally dilated and are translated on the same position, but they have respectively a horizontal, vertical and transaxial direction; let us denote

$$\psi_j^k[n_1,n_2,z] = \frac{1}{2^{3j/2}}\psi^k\left[\frac{n_1}{2^j},\frac{n_2}{2^j},\frac{z}{2^j}\right] \quad \text{for } 1 \le k \le 3,$$

and

$$\psi_{j,l_1,l_2,v}^k[n_1,n_2,z] = \frac{1}{2^{3j/2}}\psi^k\left[\frac{n_1-l_1}{2^j},\frac{n_2-l_2}{2^j},\frac{z-v}{2^j}\right]. \tag{13}$$

For a volume image $\tilde{F}$, the wavelet transform of $\tilde{F}[n_1,n_2,z]$ at a scale 2^j has three components which can be written as frame inner products:

$$T_j^k\tilde{F}[l_1,l_2,v] = <\tilde{F},\psi_{j,l_1,l_2,v}^k>, \quad k=1,2,3\ . \tag{14}$$

Because ψ^1, ψ^2 and ψ^3 are partial derivatives of θ, these three components are proportional to the coordinates of the gradient vector of $\tilde{F}$ smoothed by a dilated version of θ.

From these coordinates, one can compute the angle of the gradient vector, which indicates the direction in which the partial derivation of $\tilde{F}$ smoothed has the largest amplitude. The amplitude of this maximum partial derivative is equal to the modulus of the gradient vector and is therefore proportional to the wavelet modulus

$$M_j\tilde{F} = \sqrt{|T_j^1\tilde{F}[l_1,l_2,v]|^2 + |T_j^2\tilde{F}[l_1,l_2,v]|^2 + |T_j^3\tilde{F}[l_1,l_2,v]|^2}. \tag{15}$$

We do not threshold independently each wavelet transform component $T_j^k\tilde{F}[l_1,l_2,v]$. Instead, we threshold the modulus $M_j\tilde{F}[l_1,l_2,v]$. This is equivalent to selecting first a direction in which the partial derivative is maximum at the scale 2^j, and thresholding the amplitude of the partial derivative in this direction. This can be viewed as an adaptive choice of the wavelet direction in order to best correlate the signal. The coefficients of the dyadic wavelet transform are then computed back from the thresholded modulus and the angle of the gradient vector.

The dyadic wavelet transform is implemented with a fast filter bank "à trous" algorithm.[16] The 3-D tomographic reconstruction algorithm is decomposed in the following steps:

1. For all $0 \leq z < N_3$, computation of the regularized 2-D image $\{\tilde{F}[.,.,z]\}$ using the algorithm of section 4.
2. 3-D dyadic wavelet decomposition of the volume $\{\tilde{F}[n_1,n_2,z]\}_{n_1,n_2,z}$ to obtain 3-D dyadic wavelet coefficients $\{T_j^k \tilde{F}[l_1,l_2,v]\}_{j,k,l_1,l_2,v}$.
3. Computation of the modulus coefficients $\{M_j \tilde{F}[l_1,l_2,v]\}_{j,l_1,l_2,v}$ of the 3-D dyadic wavelet coefficients, following equation (15).
4. Thresholding of the modulus coefficients.
5. Computation of the denoised 3-D dyadic wavelet coefficients from the thresholded modulus coefficients.
6. Inverse 3-D dyadic wavelet transform from the denoised 3-D dyadic wavelet coefficients to obtain the regularized volume $\{\tilde{F}_r[n_1,n_2,z]\}_{n_1,n_2,z}$.

Figure 4 and 5 exhibit two series of sections of a brain volume reconstructed with Filtered Back-Projection, and with the combination of the 2-D wavelet packet and 3-D dyadic wavelet regularization. Once again, the FBP-reconstructed images exhibit noise and artifacts which are likely to be interpreted as information. These patterned artifacts do not appear on the images reconstructed with wavelet packets and dyadic wavelets.

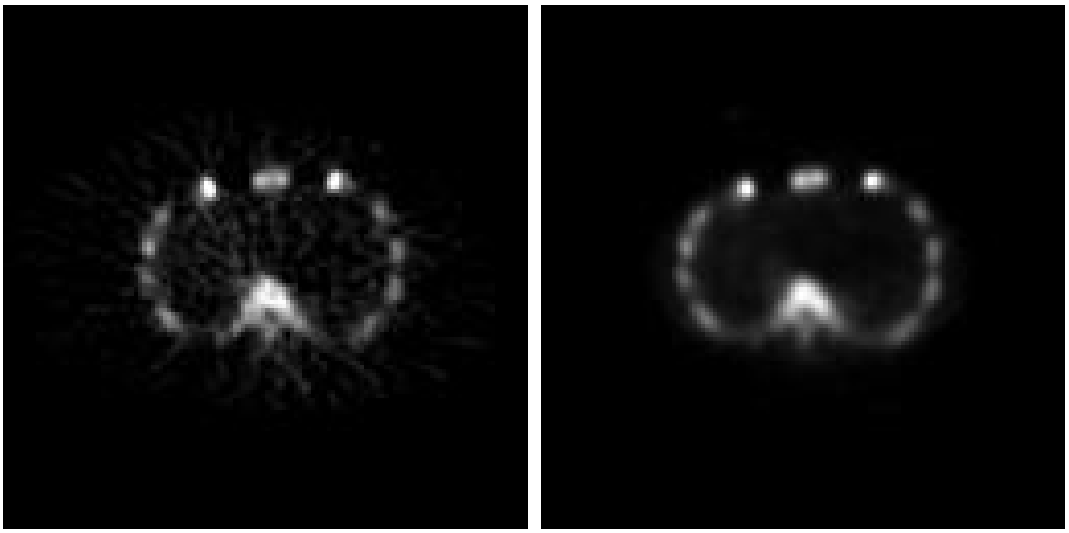
REFERENCES

1. S. Deans, *The Radon transform and some of its applications*, John Wiley and Sons, 1983.
2. P. Toft, *The Radon Transform - Theory and Implementation*. PhD thesis, Department of Mathematical Modelling, Technical University of Denmark, 1996. available at http://www.sslug.dk/~pto/cv.html.
3. M. Defrise, P. Kinahan, D. Townsend, C. Michel, M. Sibomana, and D. Newport, "Exact and approximate rebinning algorithms for 3-d PET data," *IEEE Trans. on Medical Imaging* **16**, pp. 145–158, April 1997.
4. L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *Transactions on Medical Imaging* , pp. 113–122, 1982.
5. G. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, New York:Wiley, 1997.
6. H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Transactions on Medical Imaging* , 1994.
7. D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika* **81**, pp. 425–455, December 1994.
8. Y. Meyer and R. Coifman, *Wavelets - Calderón-Zygmund and multilinear operators*, Cambridge Studies in Advanced Mathematics, 1997.
9. D. Donoho, "Nonlinear solution of linear inverse problems by wavelet-vaguelette decompositions," *J. of Appl. and Comput. Harmonic Analysis* **2**(2), pp. 101–126, 1995.
10. E. Kolaczyk, "A wavelet shrinkage approach to tomographic image reconstruction," *J. Amer. Statist. Assoc.* **91**, pp. 1079–1090, 1996.
11. N. Lee and B. Lucier, "Wavelet methods for inverting the Radon transform with noisy data," *IEEE Trans. on IP, to appear* , 2000.
12. J. Kalifa and S. Mallat, "Thresholding estimators for inverse problems and deconvolution," 1999. Submitted for publication, available at http://www.cmap.polytechnique.fr/~kalifa/.
13. R. Coifman and M. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Info. Theory* **38**, pp. 713–718, March 1992.
14. C. Stein, "Estimation of the mean of a multivariate normal distribution," *Annals of Statistics* **9**, pp. 1135–1151, 1981.
15. P. Fryzlewicz and G. Nason, "Poisson intensity estimation using wavelets and the fisz transformation," Tech. Rep. 01-10, Department of Mathematics, University of Bristol, 2001.
16. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 2nd edition, 1999.
17. J. Kalifa, P. Esser, R. V. Heertum, and A. Laine, "A new family of non iterative tomographic reconstruction methods," in *Society of Nuclear Medicine 48th Annual Meeting*, (Toronto, Canada), 2001.

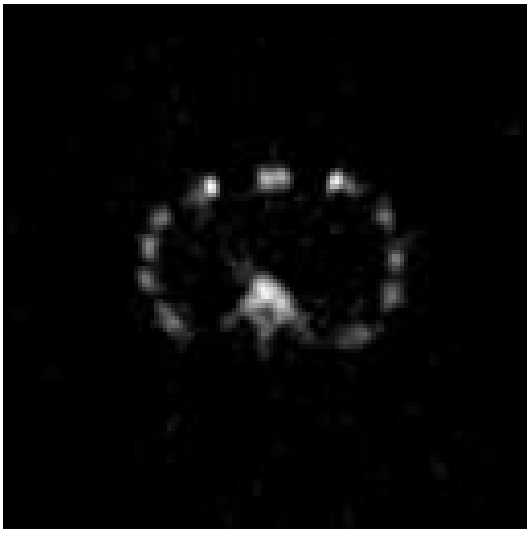
Acknowledgments

This work is supported in part by Siemens Medical Systems and the Whitaker Foundation. The authors would also like to thank François Gonczi and Sylvain Guedon for helping us implement some of the algorithms.

Figure 1. Brain phantom image. This image is completely synthetic, free of any noise or artifacts.



(a) (b)



(c)

Figure 2. SPECT image of a bone reconstructed with (a) Filtered BackProjection (b) OS-EM (c) thresholding in a wavelet packet basis.

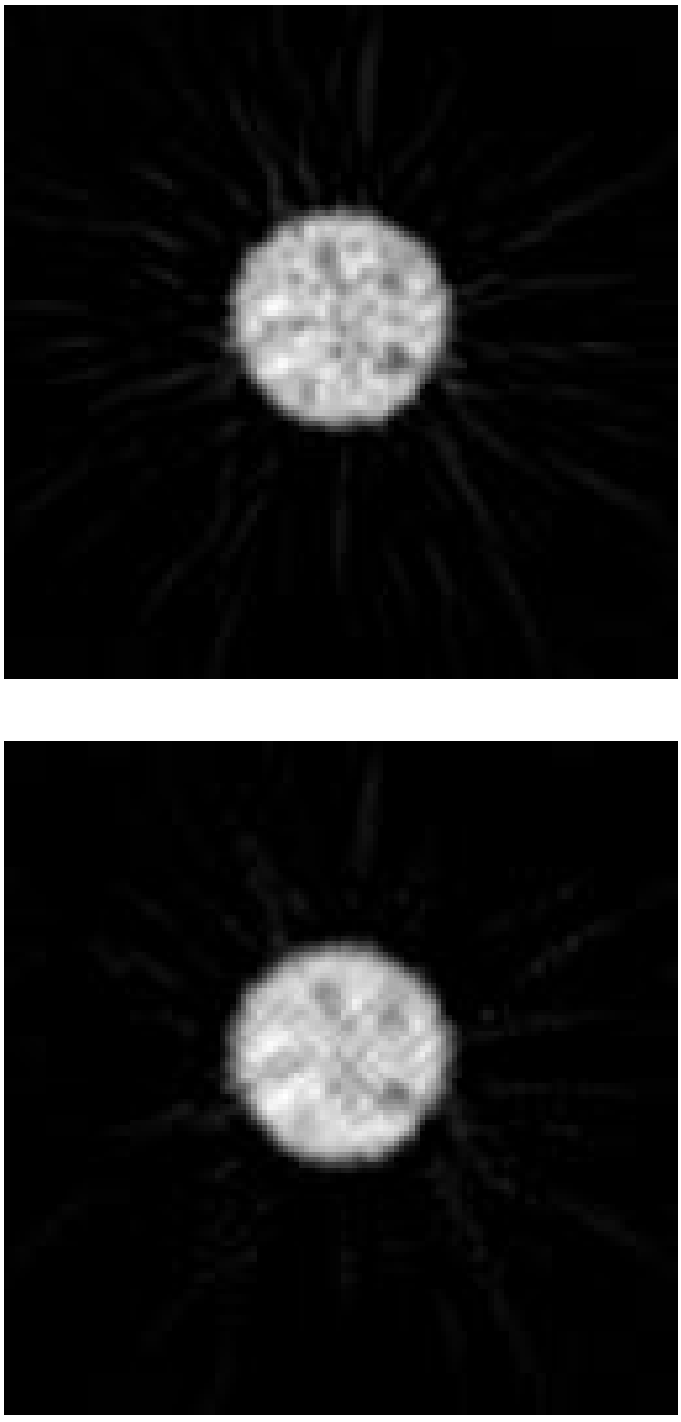


Figure 3. PET image of a Jaszak phantom reconstructed with (a) Filtered Back-Projection and (b) wavelet packets.

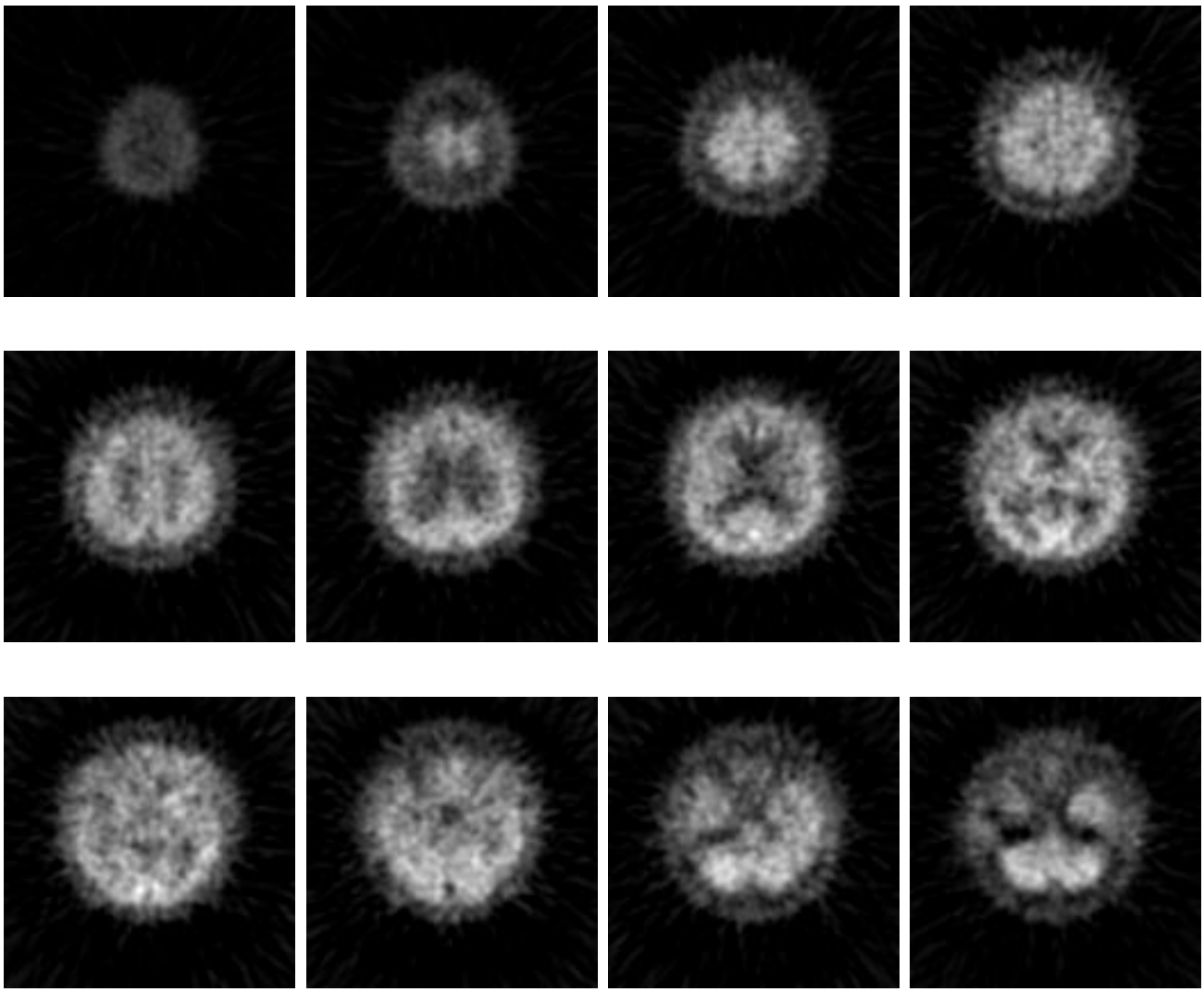


Figure 4. SPECT Brain scanner, FBP reconstruction. Compare with Figure 5.

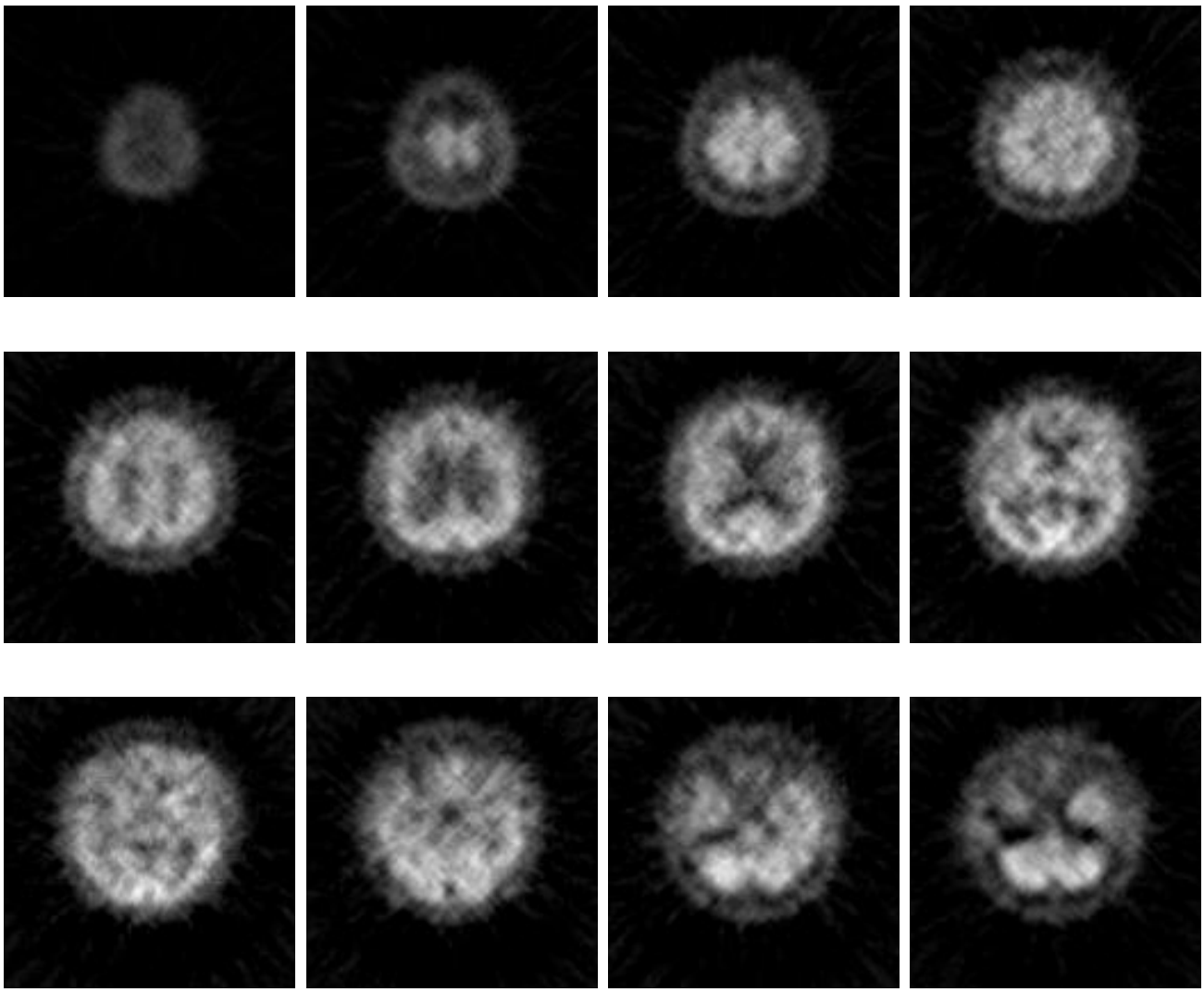


Figure 5. SPECT Brain scanner, 2-D Wavelet Packet (WP) and 3-D dyadic wavelet reconstruction. Compare with Figure 4.