Regularization in Tomographic Reconstruction Using Thresholding Estimators

Jérôme Kalifa*, Andrew Laine, and Peter D. Esser

***Abstract*—In tomographic medical devices such as single photon emission computed tomography or positron emission tomography 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 (WP) decompositions, is studied. This approach is based on the fact that the decompositions provide a near-diagonalization of the inverse Radon transform and of prior information in medical images. A WP decomposition is adaptively chosen for the specific image to be restored. Corresponding algorithms have been developed for both two-dimensional and full three-dimensional reconstruction. These procedures are fast, noniterative, and flexible. Numerical results suggest that they outperform filtered back-projection and iterative procedures such as ordered- subset-expectation-maximization.**

***Index Terms*—Dyadic wavelet transform, PET, SPECT, tomographic reconstruction, wavelet packets.**

I. 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 single photon emission computed tomography (SPECT) and positron emission tomography (PET) cameras.

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

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

where $\{f[n_1, n_2]\}_{0 \le n_1 < N_1,\, 0 \le n_2 < N_2}$ is an 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]

$$(\mathcal{R}_c f_c)(t,\alpha) = \int_{\mathbb{R}}\int_{\mathbb{R}} f_c(x_1,x_2)\delta(x_1\cos\alpha + x_2\sin\alpha - t)dx_1dx_2, \tag{2}$$

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 $t - 1/2 \le n_1 \cos\alpha + n_2 \sin\alpha < t + 1/2$.

When three-dimensional (3-D) data is processed, we treat it as a series of tomographic projections of N_3 translated 2-D slices of the observed object. When necessary, the tomographic projections are 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 image acquisition [3]. Thus, the 3-D dataset is written as

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

The noise W is usually modeled as a Gaussian white noise W, which is independent of f, or as Poisson noise, whose intensity at each pixel depends on the intensity of $\mathcal{R}(f)$.

A tomographic reconstruction procedure incorporates the following steps.

- *Filtered-Back-Projection*: The basis for tomographic reconstruction is the identity, in the continuous case

$$f_c(x_1,x_2) = \mathcal{R}_c^* K_c((\mathcal{R}_c f_c)(t,\alpha)) = \mathcal{R}_c^*(q_c \star (\mathcal{R}_c f_c)(t,\alpha)) \tag{4}$$

  where $\star$ denotes a convolution, q_c is the one-dimensional (1-D) ramp filter whose Fourier transform satisfies $\widehat{q_c}(\omega) = |\omega|$, and the back-projection operator $\mathcal{R}_c^*$ is the adjoint of $\mathcal{R}_c$

$$(\mathcal{R}_c^* p)(\vec{x}) = \int_0^{2\pi} p(t,\alpha)d\alpha$$

  with $\vec{x} = (x_1, x_2)$, $t = \vec{x}^T \cdot \vec{\alpha}$, and $\vec{\alpha} = (\cos\alpha, \sin\alpha)^T$. The filtered back-projection (FBP) algorithm is the application of a discrete operator $\mathcal{R}^{\text{inv}}$ which is the discretization of the operator $\mathcal{R}_c^* K_c$. It can be directly computed with a radial interpolation and a deconvolution by a 1-D filter q which is the discretized version of q_c. The application of the filter q amplifies the high-frequency components of the tomographic projections Y in the direction of t.
- *Regularization*: The deconvolution comes from the fact that the Radon transform is a smoothing transform. Consequently, back-projecting in the presence of additive noise is an ill-posed inverse problem: numerically speaking, a direct computation of $\mathcal{R}^{\text{inv}} Y$ is contaminated by a large additive noise $Z = \mathcal{R}^{\text{inv}} W$, which means that a regularization has to be incorporated in the reconstruction procedure.

Manuscript received October 24, 2001; revised November 22, 2002. This work was supported in part by Siemens Medical Systems and the Whitaker Foundation. *Asterisk indicates corresponding author.*

*J. Kalifa is with the Department of Biomedical Engineering, Columbia University, New York, NY 10027 USA.

A. Laine is with the Department of Biomedical Engineering, Columbia University, New York, NY 10027 USA.

P. D. Esser is with the Department of Radiology, Columbia-Presbyterian Medical Center, New York, NY 10027 USA.

Digital Object Identifier 10.1109/TMI.2003.809691

Current approaches for regularization in tomographic reconstruction can be classified into two families:

1) Regularized FBP (RFBP) is a linear filtering technique in the Fourier space, in which the Fourier transform $\hat{q}[k]$ of the filter q is replaced by a filter $\hat{q}[k]\hat{r}[k]$ where r is a low-pass filter which attenuates the amplification of high frequencies. RFBP 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. This has been proven by Donoho [4], who has showed the sub-optimality of RFBP to recover piece-wise regular signals, such as medical images.
2) Iterative statistical model-based techniques are designed to implement expectation-maximization (EM) and maximum *a posteriori* (MAP) estimators [5], [6]. In some cases, these approaches can provide an improvement over RFBP, but these estimators suffer from the following drawbacks:
   - *Computation time.* Almost all the corresponding algorithms are too computer-intensive for clinical applications, with the exception of ordered-subset-expectation-maximization (OS-EM) [7], which is an accelerated implementation of an EM estimator. In MAP methods, useful priors usually give local maxima, but the computational cost of relaxation methods remains prohibitive.
   - *Theoretical understanding and justification.* EM estimation lacks a theoretical foundation to understand and characterize the estimation error. The theoretical properties of MAP estimators have been more thoroughly studied and are 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 have to be stopped after a limited number of iterations. Beyond this critical number, the noise may be magnified, and EM and OS-EM converge to a nonmaximum-likelihood solution. The number of iterations must be chosen by the user.

In this paper, a new family of estimation procedures is studied to address these limitations. These techniques are based on a thresholding procedure in a time-frequency decomposition, namely a wavelet or wavelet packet (WP) transform.

Section II introduces thresholding estimators in time-frequency decompositions and their application to tomographic reconstruction. Section III explains how the best WP transform is chosen among the variety of possible wavelet and WP representations, using a statistical estimation of the final error. Section IV describes the corresponding fast noniterative tomographic reconstruction algorithm to recover 2-D images. Section V describes how the reconstruction algorithm can be adapted to 3-D data to take advantage of the spatial correlations of the data in the transaxial direction (z axis). Section VI presents sample numerical results on SPECT and PET data. These numerical results are then compared with the results obtained with state of the art procedures currently used in existing medical devices, namely RFBP and OS-EM.

Wavelets have been previously introduced in tomography by a large number of researchers. The most popular application of wavelets in tomography is local reconstruction [8]–[15]. Delaney and Bressler [16] as well as Blanc-Féraud *et al.* [17] used wavelet transforms to obtain accelerated implementations of a standard FBP. Bhatia, Karl, and Willsky [18], [19] combine wavelets with a MAP model to derive sparse formulations of the problem. Other authors have used wavelet methods to implement a postfiltering of a reconstructed image after it was reconstructed by a standard algorithm [20]. Sahiner and Yagle [21] use wavelet transforms to derive constraints on an iterative reconstruction algorithm. Finally, the wavelet-vaguelette decomposition (WVD) [4], [22]–[24], which is related to the work presented here, will be discussed in this paper.

A. Notation

Upper cases are used to represent signals which are the results of statistical processes.

II. Thresholding Estimators

The operator $\mathcal{R}^{\text{inv}}$ is considered as an approximate discrete inverse Radon transform operator. Let f_i be

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

The difference image $f - f_i$ is the radial interpolation error, and is in general very low compared with 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. Spline-based interpolation techniques are currently the most popular for tomographic reconstruction [25], [26].

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

$$X = f_i + Z \tag{5}$$

where $X = \mathcal{R}^{\text{inv}}Y$ and $Z = \mathcal{R}^{\text{inv}}W$. If the noise Z was Gaussian white, Donoho and Johnstone have established [27] 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 tomographic medical images. A thresholding estimator $\tilde{F}$ of f_i in $\mathcal{B}$ is defined as

$$\tilde{F} = \sum_m \rho_m \left(\langle X, g_m \rangle \right) g_m^* \tag{6}$$

where ρ_m is a thresholding operator. Typical simple thresholding rules include hard thresholding

$$\rho_m(x) = \begin{cases} x, & \text{if } |x| > T_m \\ 0, & \text{if } |x| \le T_m \end{cases} \tag{7}$$

and soft thresholding

$$\rho_m(x) = \begin{cases} x - T_m, & \text{if } x \ge T_m \\ x + T_m, & \text{if } x \le -T_m \\ 0, & \text{if } |x| \le T_m \end{cases}. \tag{8}$$

The threshold T_m is chosen to be proportional to the standard deviation σ_m of the transform coefficient $\langle Z, g_m \rangle$ of the back-projected noise, which is a random variable.

A. Wavelet-Vaguelette Decomposition

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 back-projected noise Z, whose behavior is very specific due to the fact that it has been distorted by back-projection and deconvolution processes. 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 in the image f_i.

The image f_i is a spatially inhomogeneous, piece-wise regular signal, which is compactly represented in a wavelet decomposition. When the noise W is Gaussian white, then the noise $Z = \mathcal{R}^{\text{inv}}W$ remains Gaussian because $\mathcal{R}^{\text{inv}}$ is linear. To obtain a diagonal representation of the noise Z, one must find a decomposition in which the covariance of Z, and hence $\mathcal{R}^{\text{inv}}$, is nearly diagonal. Since the inverse Radon transform is a Calderon–Zygmund operator [28], it is also nearly-diagonal in a wavelet basis.

These two properties of wavelet bases led Donoho [4] to suggest the use of thresholding estimators in wavelet bases for several linear inverse problems, including the inversion of the Radon transform. Such an estimator is given by (6), where the basis $\mathcal{B} = \{g_m, g_m^*\}_{0 \le m \le N_1 * N_2 - 1}$ is an orthogonal or a bi-orthogonal wavelet basis. Donoho established the minimax optimality of this approach, called a WVD, and showed its superiority with respect to other approaches such as FBP, for the recovery of piece-wise regular signals.

However, the WVD as studied by Donoho was developed for a continuous model of the back-projection operator, and assumes that the additive noise W is always Gaussian white. Moreover, the asymptotic optimality results establish the performance of a WVD estimator for high resolution data, which is not the case for PET and SPECT medical images. This means that, unfortunately, despite numerical implementations and refinements by other researchers [22]–[24], the theoretical interest of the WVD is not matched with a significant gain of performance when compared with other techniques such as RFBP, when applied to real clinical PET and SPECT data. The purpose of this paper is to build estimators which share the same theoretical properties as WVD, but also provide an important additional flexibility and adaptivity which are essential to improve the numerical performances and image quality of the resulting algorithms.

The minimax optimality properties of the Wavelet-Vaguelette Decomposition can only be established when the additive noise W is a Gaussian white noise. When W is a Poisson noise, the coefficients of its decomposition in a wavelet transform are not independent, and the minimax optimality properties cannot be verified. In practice, however, the strategy of finding a decomposition in which $\mathcal{R}^{\text{inv}}$ is nearly diagonal remains valid, and guarantees that the numerical values of the transform coefficients of filtered back-projected noise $Z = \mathcal{R}^{\text{inv}}W$ will be nearly independent, even if W is a Poisson noise. Section III-C explains how the algorithm is adapted depending on the Gaussian or Poisson nature of W.

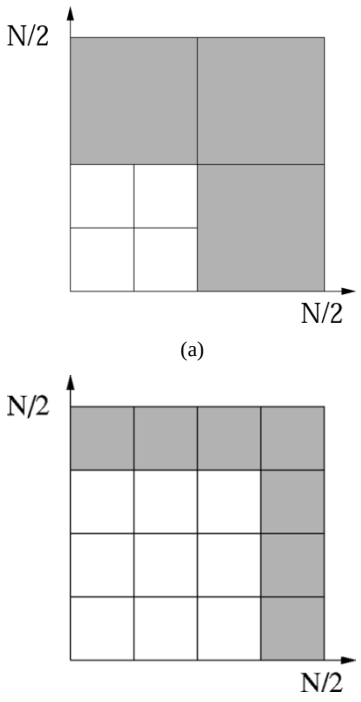


Fig. 1. This figure illustrates the 2-D discrete Fourier domain for positive frequencies. (a) Segmentation induced by a wavelet transform. The grey areas correspond to the wavelet coefficients which are always put to zero by the thresholding operator because these coefficients have been contaminated by the numerical explosion of the back-projected noise at high frequencies. (b) Segmentation induced by a particular WP transform. The highest frequencies in which the information is completely dominated by the back-projected can be isolated more accurately, and some information at intermediate frequencies are recovered by the thresholding operator.

B. Wavelet Packets

A major problem of the WVD comes from the relatively poor resolution in frequency of the wavelet transform. Fig. 1(a) illustrates the partitioning of the 2-D discrete Fourier domain induced by an orthogonal wavelet basis. At the finest scale of the wavelet transform, which corresponds to frequencies higher than $N/4$ in the horizontal and vertical directions, all the wavelet coefficients are contaminated by the numerical explosion of the back-projected noise Z. These coefficients are put to zero by the thresholding operator, because the threshold values T_m depend on the standard deviation σ_m of the back-projected noise Z, which is very large at highest frequencies. Alternatively, if the threshold T_m is chosen at a lower value, the noise remaining in the reconstructed image is too important.

Kalifa and Mallat [29] have generalized Donoho's approach to adapt it to other types of decompositions, including WP bases. WP bases are decompositions which can provide a compact representation of an observed image f, as well as a more accurate segmentation of the frequency domain than a wavelet basis, to improve the near-diagonalization of the noise Z. It is shown in [29] that the thresholding estimation risk $r(\tilde{F}) = \mathcal{E}(\|\tilde{F} - f_i\|^2)$ is of the same order, up to a $\log(N)$ factor, of the decision risk

$$r_d(f_i) = \sum_m \min(\sigma_m^2, |\langle f_i, g_m \rangle|^2) . \tag{9}$$

To minimize (9), we need to concentrate the energy $\mathbf{E}\{||Z||^2\}$ over few vectors g_m which produce coefficients σ_m larger than $|\langle f_i, g_m\rangle|$, and among the remaining vectors g_p concentrate the energy $||f_i||^2$ over few large coefficients $|\langle f_i, g_p\rangle|$ that are above the noise level σ_p.

Fig. 1(b) gives an example: a more accurate segmentation of the Fourier domain as compared with a wavelet transform enables isolation of the highest frequencies, in which each coefficient $\langle f_i, g_m\rangle$ of the information is below the standard deviation σ_m of the coefficients of the back-projected noise. Because this WP decomposition, as opposed to a Fourier transform, provides a compact representation of information, thresholding allows the recovery of most of the information in the rest of the Fourier domain.

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 back-projected noise Z and the representation of the data f_i to be recovered. There is no single time-frequency decomposition (such as a Fourier basis, a wavelet basis, or a specific WP basis) which fits all applications of SPECT and PET imaging. However, a WP basis can be adaptively chosen from a dictionary of different WP bases. This enables us to optimize the choice of the WP transform for a specific type of observed image and for the specific nature of the back-projected noise Z. This additional adaptivity brings a significant improvement of numerical performances with respect to a Wavelet-Vaguelette estimator.

III. Choice of Wavelet Packet Decomposition

A WP dictionary is a rapidly constructible set of distinct and numerous orthogonal bases $\{\mathcal{B}^\gamma\}_\gamma$. It is possible, within this dictionary, 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 [30], 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 $\mathbf{l}^2$ estimation error (risk)

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

The quadratic estimation error can also be replaced with other measures or error at the same computational cost. For example, the $\mathbf{l}^1$ error $\mathcal{E}(||\tilde{F} - f_i||)$ is sometimes considered in the image processing community as a better measure to assess the perceptual quality of reconstructed images. For the method presented in Section III-A, we have experienced the use of both the $\mathbf{l}^1$ and $\mathbf{l}^2$ estimation errors with similar results. We will use the $\mathbf{l}^2$ estimation error in this presentation since its theoretical properties are easier to manipulate and because the quadratic estimation error is used to compute the PSNRs of reconstructed images.

Two alternatives are proposed to compute the choice of the WP decomposition. 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.

A. Use of Phantom Images

Phantom images are synthetic images modeling observed organs or anatomical structures, without any noise or artifacts. 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, (6) becomes

$$\tilde{F} = \sum_m a_m \langle X, g_m\rangle g_m^* \tag{10}$$

where a_m is either zero or one. 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 back-projected 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 WP basis $\mathcal{B}^\gamma = \{g_m^\gamma, g_m^{\gamma,*}\}_{0 \le m \le N_1 * N_2 - 1}$ is

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

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 WP basis $\mathcal{B}^\gamma$ such that $\mathcal{C}(\mathcal{B}^\gamma)$ is minimal.

B. Use of The Stein Unbiased Risk Estimator (SURE)

The Stein Unbiased Risk Estimator [31] is an estimator of the risk when ρ is a soft thresholding operator. For a WP basis $\mathcal{B}^\gamma$, it is given by

$$\tilde{r}^\gamma(f_i) = \sum_m \Phi_m\left(|\langle X, g_m^\gamma\rangle|^2\right) \tag{11}$$

with

$$\Phi_m(u) = \begin{cases} |u - \sigma_m^2|, & \text{if } u \le T_m^2 \\ \sigma_m^2 + T_m^2, & \text{if } u > T_m^2 \end{cases} \tag{12}$$

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

$$\tilde{r}^{\gamma_1}(f_i) = \min_\gamma \tilde{r}^\gamma(f_i)\ . \tag{13}$$

The estimated risk is calculated in (11) as an additive cost function over the noisy coefficients.

The SURE-based approach to compute the best basis is in general the most efficient 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.

C. Model of the Noise

The cost functions used to compute the best basis algorithm depend on the back-projected noise Z. To generate a model of

the noise Z, it is necessary to first generate a model of the additive noise W observed in the sinograms. The model of the noise Z is obtained by back-projecting 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 assumed to be either Gaussian white noise or Poisson noise.

- When W is assumed to be a white Gaussian noise, the problem is to estimate its standard deviation. Donoho and Johnstone [27] 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 W is assumed to be a Poisson noise, the sinograms Y are roughly denoised using the Poisson intensity estimation method by Fryzlewicz and Nason [32]. The resulting denoised sinograms Y_{den} cannot be back-projected 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.

IV. Reconstruction Algorithm

The tomographic reconstruction algorithm is carried out by the following steps.

1). FBP without regularization of the tomographic projections Y to obtain the back-projected image $X = f_i + Z$.
2). (Optional) Computation of the best WP basis $\mathcal{B}^{\gamma_1}$ optimized for a specific image to be restored, using one of the two methods presented in **Section III**. The best basis can be recomputed for each image, or can be computed once and stored in advance, to save computation time. However the computation of the best basis algorithm is very fast.
3). WP transform of the back-projected image X in the best basis $\mathcal{B}^{\gamma_1}$ to obtain the WP coefficients $\{\langle X, g_m \rangle\}_m$.
4). Thresholding of the WP coefficients.
5). Inverse WP transformation of the thresholded coefficients to obtain the estimate image $\tilde{F}$.

The thresholding operator is preferably a soft thresholding, with a threshold value $T_m = \lambda_m \sigma_m$ proportional to the standard deviation σ_m of the noise coefficients. λ_m is typically chosen between 1.5 and 3. It can either be a constant $\lambda_m = \lambda$ or depend on the WP vector g_m: in this case, it is smaller when the support of the Fourier transform of g_m is concentrated in lower frequencies. This imposes that the remaining noise on the reconstructed image will have a nearly flat power spectrum, because the coefficients of the noise which correspond to higher frequencies will be more attenuated, and the remaining noise will practically behave like a white Gaussian noise. This is useful when combined with a supplemental thresholding in another decomposition, as explained in Section V. Finally, note that the soft thresholding, which attenuates on the whole image the intensity of the remaining noise, guarantees that the reconstructed data will be sufficiently regular and free of strong artifacts.

The WP transform and its inverse are computed with fast filter bank algorithms of complexity $O(N)$ for signals of N samples [33]. Numerical results are improved if the WP transform and its inverse are undecimated, i.e., translation-invariant, in which case the filter bank algorithm is equivalent to the "à trous" algorithm [33].

V. Extension to 3-D Reconstruction

So far, the WP reconstruction has been presented for 2-D reconstruction of slices. We now consider (3), where we have 3-D data in the form of a series of tomographic projections of N_3 translated 2-D slices of an 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 information and noise. In this case, a regularization is computed on the whole 3-D data, but the back-projections are still computed slice by slice.

The FBP operator $\mathcal{R}^{\text{inv}}$ is still a 2-D operator; assuming that the power spectrum of the additive noise W is constant in every direction, the power spectrum of the filtered back-projected noise $Z = \mathcal{R}^{\text{inv}} W$ will remain constant in the transaxial direction and will not depend on the position on the z axis, contrary to the n_1 and n_2 axis. As a consequence, there is no need to use a decomposition with a good resolution in the Fourier domain along the z axis. The best decomposition must only provide a compact representation of spatially inhomogeneous data, which means that a wavelet decomposition is the most appropriate. The best results are obtained with a combination of a slice-by-slice 2-D regularization in a WP decomposition, using the algorithm of Section IV, 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 slice z, an estimation $\tilde{F}[.,.,z]$ is first computed with a thresholding of the WP 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 full 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 perpendicular 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. It should be mentioned that other types of transform are currently being designed to provide geometrical adaptivity for denoising and inverse problems, see in particular [34] and [35].

A 3-D dyadic wavelet transform is computed with a family of wavelets which are the discretized translations and dilatations of

three wavelets ψ^1, ψ^2 and ψ^3 that are the partial derivatives of a smoothing function θ

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

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

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

and

$$\psi_{j,l_1,l_2,v}^k[n_1,n_2,z]=\frac{1}{2^{3j/2}}\psi^k\left[\frac{n_1-l_1}{2^j},\frac{n_2-l_2}{2^j},\frac{z-v}{2^j}\right]. \tag{14}$$

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

$$T_j^k\tilde{F}[l_1,l_2,v]=<\tilde{F},\psi_{j,l_1,l_2,v}^k>,\ k=1,2,3\,. \tag{15}$$

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 the smoothed $\tilde{F}$ 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

$$M_j\tilde{F}=\sqrt{|T_j^1\tilde{F}[l_1,l_2,v]|^2+|T_j^2\tilde{F}[l_1,l_2,v]|^2+|T_j^3\tilde{F}[l_1,l_2,v]|^2}. \tag{16}$$

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 each 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 [33]. The 3-D tomographic reconstruction algorithm is decomposed in the following steps.

1). For all $0\le z<N_3$, computation of the regularized 2-D image $\{\tilde{F}[.,.,z]\}$ using the algorithm described in **Section IV**.
2). Three–dimensional 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 **(16)**.
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 $\{\tilde{F}_r[n_1,n_2,z]\}_{n_1,n_2,z}$.

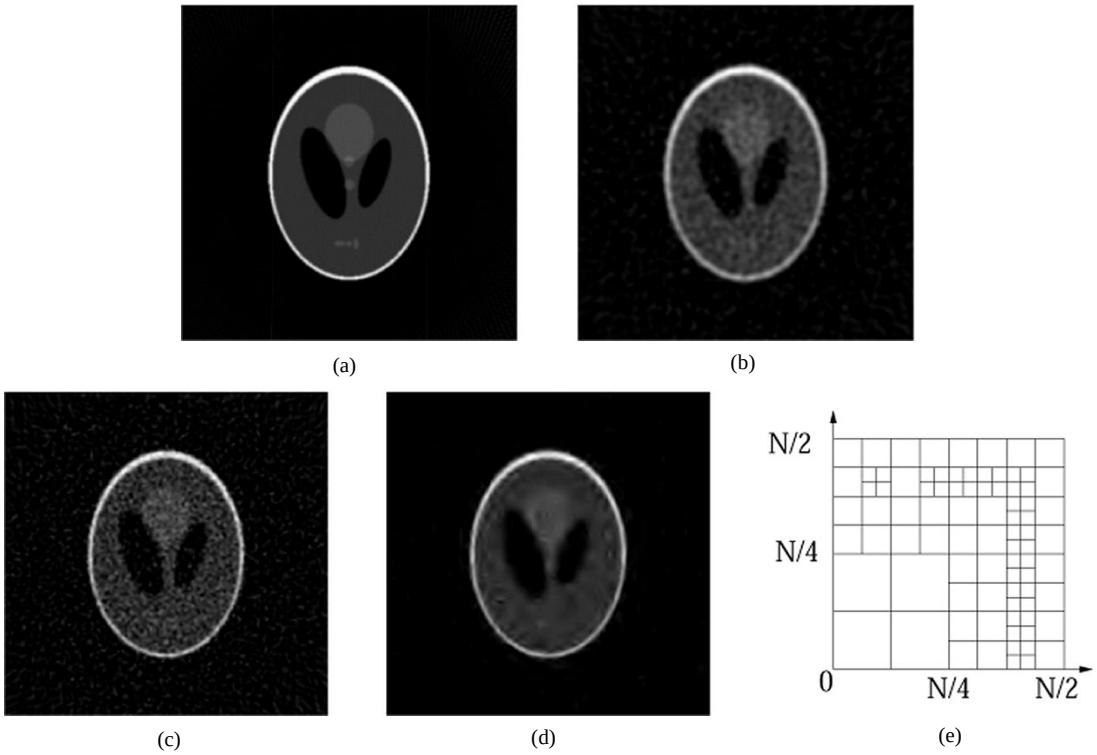
The thresholding operator on the modulus coefficients is a hard thresholding, since a supplemental soft thresholding would add an unnecessary smoothing on the reconstructed data.

VI. Numerical Results

Numerical results are provided for synthetic phantom images to demonstrate the metrical performances, in terms of signal-to-noise ratio (PSNR), of the WP-based reconstruction with respect to RFBP and WVD. Numerical results on real clinical SPECT and PET data are also provided to demonstrate the perceptual performance of the WP reconstruction. Examples include bone, brain and Jaszak phantom images, which have very different dynamics and properties.

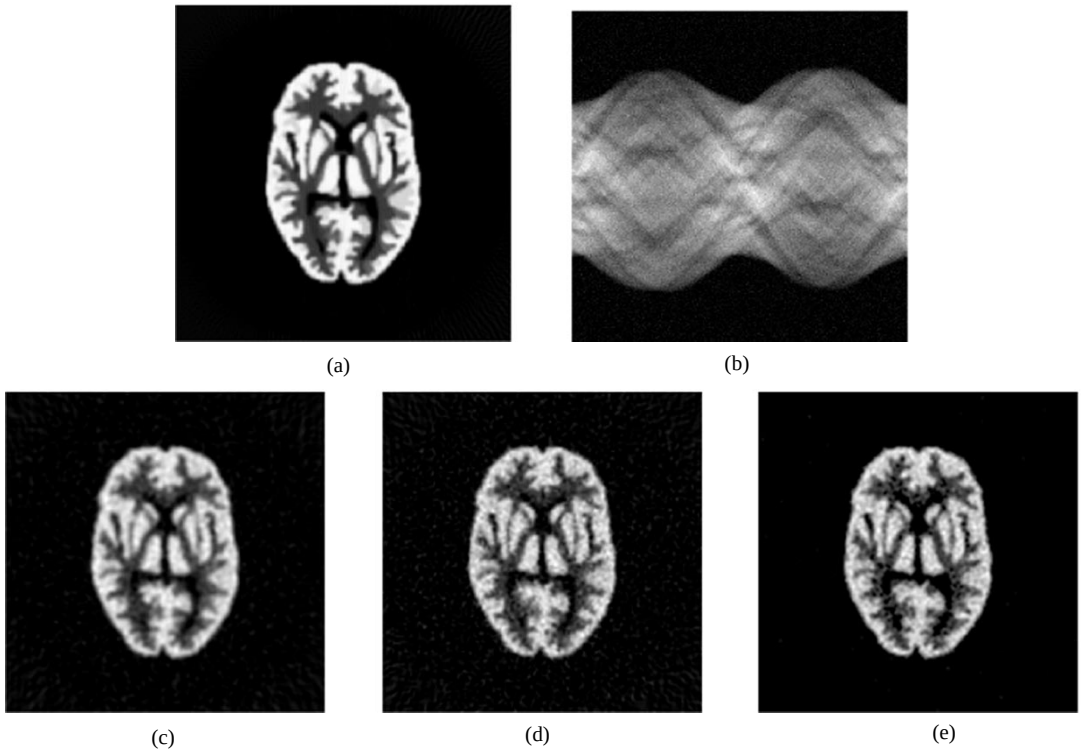
Fig. 2 compares the reconstructions of a RFBP reconstruction, a WVD reconstruction, and a WP reconstruction on a synthetic 256×256 Shepp-Logan phantom, starting from 192 projections, as in a standard PET device. Since an ideal reference image is available, this example enables us to compare the metrical performance directly. The filter used for RFBP has been optimized to obtain the best possible PSNR. Experiences have been conducted with different shapes of filters, including cosine, Hamming, Hann, and Shepp-Logan filters. The best filter in this example is a Hann filter. The WVD algorithm includes a translation-invariant decomposition [23] and its thresholding strategy has been carefully optimized to provide the highest possible PSNR, using scale-dependent thresholds. In all the experiments, the same filter bank has been used for the WVD and for the WP algorithm, namely *symmlets* with four null moments, and the WP tree has been selected using the SURE-based procedure of Section III-B. As mentioned in Section IV, for both WP and WVD, the threshold values $T_m=\lambda_m\,\sigma_m$ are proportional to the estimated standard deviation σ_m of the noise coefficients, where λ_m is scale-dependent for WVD and tree node-dependent for WP, and is typically between 1.5 and 3. The PSNRs of the RFBP-reconstructed image and of the WVD-reconstructed image are respectively 19.5 dB and 18.2 dB while the PSNR of the WP-reconstructed image is 23.5 dB. Fig. 2(e) illustrates the frequency segmentation roughly induced by the WP basis chosen by the best basis algorithm for this image. The best WP basis is usually composed of wavelets in low frequencies, but performs a moderately finer segmentation in higher frequencies.

Fig. 3 provides another comparison with a synthetic phantom of a brain. Both the parameters of the RFBP and of the WVD have been modified to be specifically optimized for this new image, while the same default parameters have been kept for the WP-reconstructed image to assert the robustness of the algo-



(a) (b) (c) (d) (e)

Fig. 2. (a) Reference image of a Shepp-Logan Phantom. (b) Reconstructed image with RFBP. ($\mathrm{PSNR} = 19.5$ dB) (c) Reconstructed image with a WVD ($\mathrm{PSNR} = 18.2$ dB). (d) Reconstructed image with a WP thresholding ($\mathrm{PSNR} = 23.5$ dB). (e) Segmentation of frequencies induced by the selected WP basis. At low frequencies, for $k < N/4$, the selected vectors are regular wavelets, while the WP basis performs a finer segmentation in higher frequencies.



(a) (b) (c) (d) (e)

Fig. 3. (a) Reference image of a 256×256 Brain Phantom. (b) Sinograms simulated with 192 angular positions and 192 detectors, as in standard PET devices. (c) Image reconstructed with an optimized RFBP ($\mathrm{PSNR} = 19.8$ dB). (d) Image reconstructed with a WVD estimator ($\mathrm{PSNR} = 18.7$ dB). (e) Image reconstructed with a WP thresholding ($\mathrm{PSNR} = 21.6$ dB).

rithm. The PSNR of the RFBP-reconstructed image is 19.8 dB, the PSNR of the WVD-reconstructed image is 18.7 dB, while the PSNR of the WP-reconstructed image is 21.6 dB. For both images, the WVD is slightly inferior in terms of metrical performances to RFBP, while the WP reconstruction algorithm clearly outperforms both algorithms by a substantial margin.

Fig. 4 compares numerical results computed on SPECT clinical data of 128×128 bone images, using an OS-EM

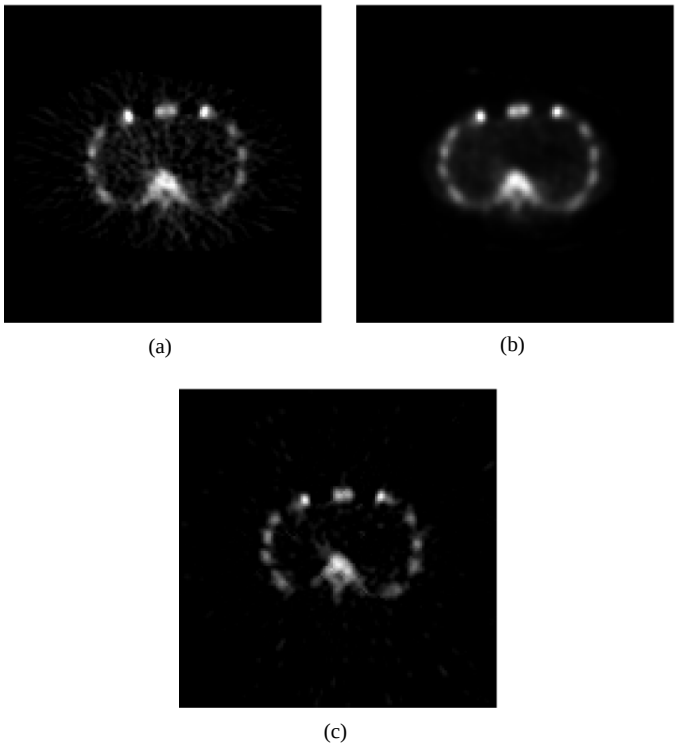


Fig. 4. A 128×128 SPECT image of bone reconstructed with (a) RFBP, (b) OS-EM, and (c) thresholding in a WP basis, starting from sinograms with 128 projections.

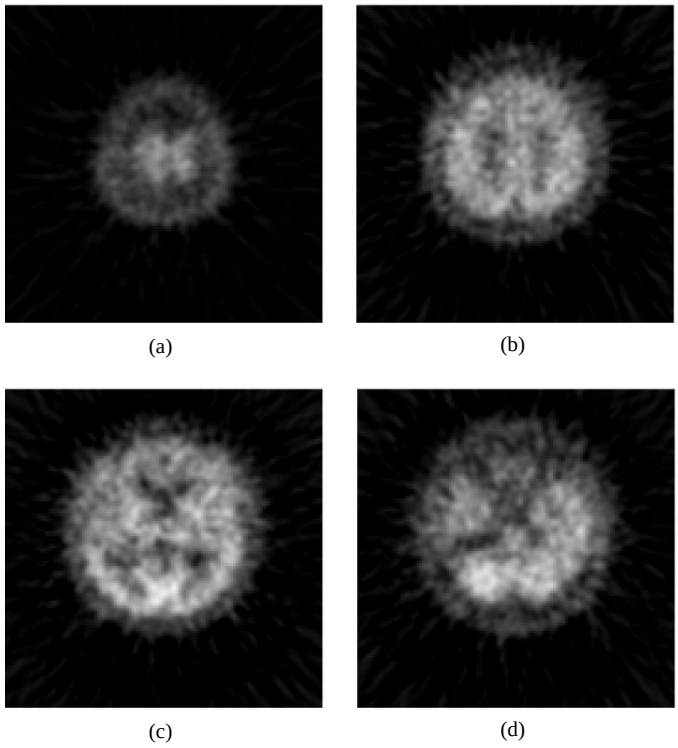


Fig. 5. Slices of a 128×128×90 3-D SPECT Brain scan reconstructed with FBP reconstruction, starting from sinograms with 128 projections. To compare with Fig. 6.

reconstruction, a RFBP and a WP-based reconstruction, and starting from 128 projections. Again, the filter used for RFBP has been optimized to provide the best perceptual results. The parameters used for the OS-EM reconstruction, such as the number of iterations, are the parameters used by physicians in clinical conditions for an image of this type. 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. Here the OS-EM algorithm was stopped after eight iterations. On the other hand, the RFBP-reconstructed image is corrupted by a significant amount of noise and artifacts, which cannot be reduced unless the reconstructed image becomes extremely smoothed. With the WP reconstruction algorithm, the amount of smoothness of a reconstructed image can be controlled precisely, while the noise is reduced significantly as compared with an image reconstructed with RFBP or OS-EM.

Figs. 5 and 6 exhibit two series of sections of a 3-D SPECT-acquired brain volume reconstructed with RFBP, and with the combination of the 2-D WP and 3-D dyadic wavelet regularization presented in Section V. Once again, the RFBP-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 WPs and dyadic wavelets.

Finally, Fig. 7 compares numerical results on PET images of a Jaszak phantom, starting from 192 projections, using a FBP and a WP-based reconstruction. Fig. 7(a) and (b) exhibits a single slice of the reconstructed volume, while Fig. 7(c) and (d) exhibits the resulting images after summing together five adjacent slices, as is often done in practice. The RFBP-reconstructed images are corrupted by a significant noise and their smaller structures are more difficult to detect than in the WP-reconstructed images.

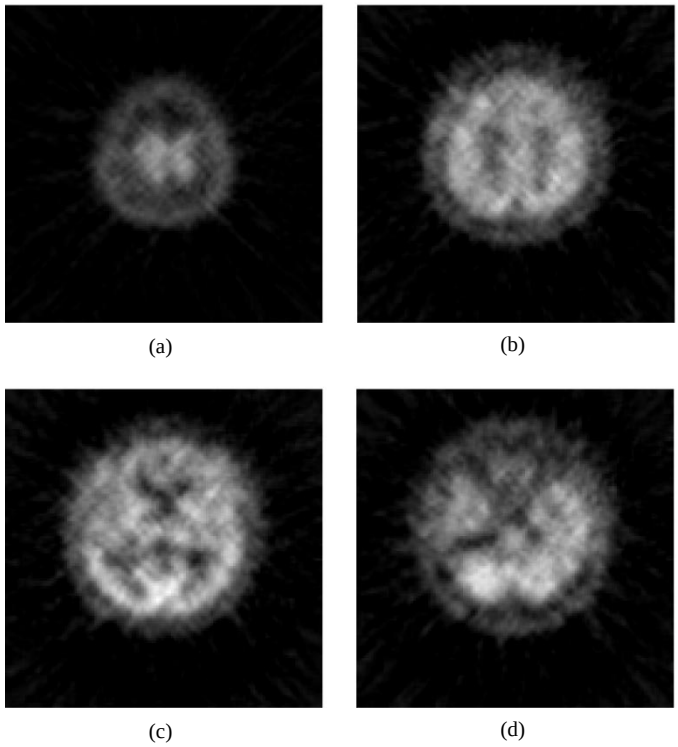


Fig. 6. Slices of a 128×128×90 3-D SPECT Brain scan reconstructed with a combined thresholding in 2-D WP and 3-D dyadic wavelet decompositions. To compare with Fig. 5.

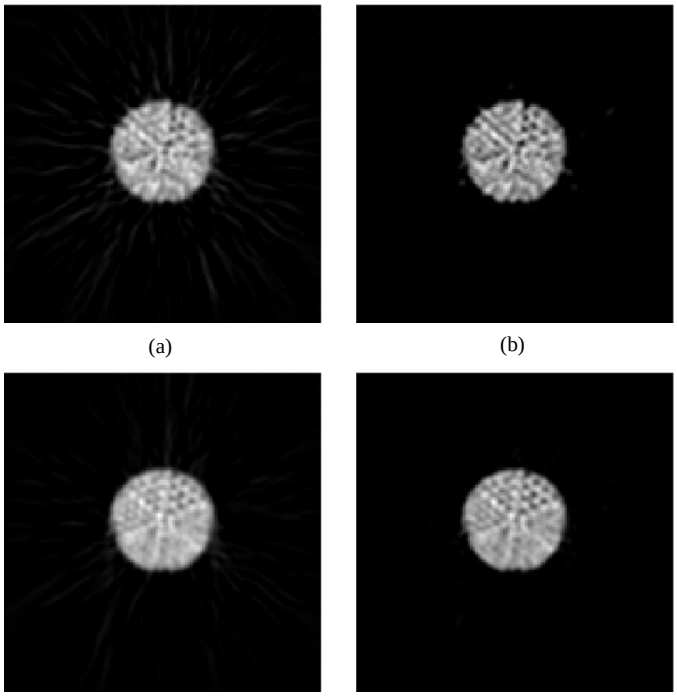


Fig. 7. (a) 192×192 PET image of a single slice of an RFBP-reconstructed Jaszak phantom, starting from sinograms with 192 projections. (b) PET image of a single slice of a WP-reconstructed Jaszak phantom. (c) PET image of a sum of 5 adjacent slices of an RFBP-reconstructed Jaszak phantom. (d) PET image of a sum of 5 adjacent slices of a WP-reconstructed Jaszak phantom.

VII. Summary

A new family of tomographic reconstruction algorithms based on a thresholding in wavelet and WP decompositions has been developed for the recovery of PET and SPECT images. This approach is based on the fact that wavelet and WP decompositions are adapted to both the nature of the medical images to be recovered and the properties of the inverse Radon transform. Corresponding algorithms are fast, and available for 2-D and 3-D data.

Numerical results on a variety of PET and SPECT images demonstrate a significant improvement, perceptually and metrically, with respect to state of the art methods such as FBP and OS-EM, currently used in existing medical devices.

Acknowledgment

The authors would like to thank F. Gonczi and S. Guedon for helping us implement some of the algorithms.

References

[1] S. Deans, *The Radon Transform and Some of its Applications*. New York: Wiley, 1983.
[2] P. Toft, "The Radon transform – theory and implementation," Ph.D. dissertation, Tech. Univ. Denmark, Dept. Math. Modeling, Lyngby, Denmark, 1996.
[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. Med. Imag.*, vol. 16, pp. 145–158, Apr. 1997.
[4] D. Donoho, "Nonlinear solution of linear inverse problems by wavelet-vaguelette decompositions," *J. Appl. Comput. Harmonic Anal.*, vol. 2, no. 2, pp. 101–126, 1995.
[5] L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imag.*, vol. MI-1, pp. 113–122, 1982.
[6] G. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York: Wiley, 1997.
[7] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imag.*, vol. 13, pp. 601–609, Dec. 1994.
[8] C. Berestein and D. Walnut, *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, Eds. Boca Raton, FL: CRC, 1996.
[9] T. Olson and J. DeStefano, "Wavelet localization of the radon transform," *IEEE Trans. Image Processing*, vol. 42, pp. 2055–2067, Aug. 1994.
[10] T. Olson, "Optimal time-frequency projections for localized tomography," *Ann. Biomed. Eng.*, vol. 23, pp. 622–836, 1995.
[11] F. Rashid-Farrokhi, K. Liu, C. Berenstein, and D. Walnut, "Wavelet-based multiresolution local tomography," *IEEE Trans. Image Processing*, vol. 22, pp. 1412–1430, Oct. 1997.
[12] S. Zhao, G. Wang, and J. Hsieh, "Wavelet sampling and localization schemes for the radon transform in two dimensions," *SIAM J. Appl. Math.*, vol. 57, pp. 1749–1762, 1997.
[13] M. Bottema, B. Moran, and S. Suorova, "An application of wavelets in tomography," *Digital Signal Processing*, vol. 8, pp. 244–254, 1998.
[14] W. Maldych, "Tomography, approximate reconstructions, and continuous wavelet transforms," *J. Appl. Comput. Harmonic Anal.*, vol. 7, pp. 54–100, 1999.
[15] S. Zhao, "Wavelet filtering for filtered backprojection in computer tomography," *J. Appl. Comput. Harmonic Anal.*, vol. 6, pp. 346–373, 1999.
[16] A. Delaney and Y. Bresler, "Multiresolution tomographic reconstruction using wavelets," *IEEE Trans. Image Processing*, vol. 4, pp. 799–813, June 1995.
[17] L. Blanc-Féraud, P. Charbonnier, P. Lobel, and M. Barlaud, "A fast tomographic reconstruction algorithm in the 2-d wavelet transform domain," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, vol. 5, Apr. 1994, pp. V/305–V/308.
[18] M. Bhatia, W. Karl, and A. Willsky, "A wavelet-based method for multiscale tomographic reconstruction," *IEEE Trans. Med. Imag.*, vol. 15, pp. 92–101, Feb. 1996.
[19] ——, "Tomographic reconstruction and estimation based on multiscale natural-pixel bases," *IEEE Trans. Image Proc.*, vol. 6, pp. 463–478, 1997.
[20] J.-W. Lin, A. F. Laine, and S. R. Bergmann, "Improving pet-based methods using the wavelet transform for positron emission tomography," *IEEE Trans Biomed. Eng.*, vol. 48, pp. 202–212, Feb. 2001.
[21] B. Sahiner and A. Yagle, *Time-Frequency and Wavelets in Biomedical Signal Processing*, M. Akay, Ed. Piscataway, NJ: IEEE Press, 1997, pp. 473–498.
[22] E. Kolaczyk, "A wavelet shrinkage approach to tomographic image reconstruction," *J. Amer. Statist. Assoc.*, vol. 91, pp. 1079–1090, 1996.
[23] N. Lee and B. Lucier, "Wavelet methods for inverting the radon transform with noisy data," *IEEE Trans. Image Processing*, vol. 10, Jan. 2001.
[24] Y. Choi, J. Y. Koo, and N. Y. Lee, "Image reconstruction using the wavelet transform for positron emission tomography," *EEE Trans Med Imag.*, vol. 20, pp. 1188–1193, Nov. 2001.
[25] J. Guédon and Y. Bizais, "Bandlimited and haar filtered back-projection reconstruction," *IEEE Trans. Med. Imag.*, vol. 12, pp. 430–440, Sept. 1994.
[26] M. Lautsch, "A spline inversion formula for the radon transform," *SIAM J. Numer. Anal.*, vol. 26, no. 2, pp. 456–467, 1989.
[27] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, Dec. 1994.
[28] Y. Meyer and R. Coifman, *Wavelets – Calderón–Zygmund and Multilinear Operators*, ser. Cambridge Studies in Advanced Mathematics. Cambridge, U.K.: Cambridge Univ. Press, 1997.
[29] J. Kalifa and S. Mallat, "Thresholding estimators for linear inverse problems and deconvolution," *Ann. Statist.*.
[30] R. Coifman and M. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inform. Theory*, vol. 38, pp. 713–718, Mar. 1992.
[31] C. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Statist.*, vol. 9, pp. 1135–1151, 1981.
[32] P. Fryzlewicz and G. Nason, "Poisson intensity estimation using wavelets and the Fisz transformation," Univ. Bristol, Dept. Math., Bristol, U.K., Tech. Rep., 2001.
[33] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. New York: Academic, 1999.
[34] E. Candes and D. Donoho, "Recovering edges in ill-posed linear inverse problems: Optimality of curvelet frame," *Ann. Statist.*, 2000.
[35] M. N. Do and M. Vetterli, "The finite ridgelet transform for image representation," IEEE Trans. Image Processing, vol. 12, pp. 16–28, Jan. 2003, to be published.