

Wavelet-based reconstruction for limited-angle X-ray tomography

Maaria Rantala, Simopekka Vänskä, Seppo Järvenpää, Martti Kalke,
Matti Lassas, Jan Moberg and Samuli Siltanen

Abstract—The aim of X-ray tomography is to reconstruct an unknown physical body from a collection of projection images. When the projection images are only available from a limited angle of view, the reconstruction problem is a severely ill-posed inverse problem. Statistical inversion allows stable solution of the limited-angle tomography problem by complementing the measurement data by *a priori* information. In this work, the unknown attenuation distribution inside the body is represented as a wavelet expansion, and a Besov space prior distribution together with positivity constraint is used. The wavelet expansion is thresholded before reconstruction to reduce the dimension of the computational problem. Feasibility of the method is demonstrated by numerical examples using *in vitro* data from mammography and dental radiology.

I. INTRODUCTION

In medical Computerized Tomography (CT), the three-dimensional structure of the inner organs of a patient is reconstructed from an extensive collection of X-ray projection images taken from all around the patient. The radiation dose of a CT scan and the cost of scanners limits the use of CT to investigations of serious diseases. In contrast, we consider low-dose three-dimensional X-ray imaging, where only few projection images are taken. Such imaging finds applications in dental radiology, surgical imaging, thorax imaging and mammography. However, in such situations projection images are typically available only within a limited angle of view, leading to the notoriously ill-posed problem of limited-angle tomography.

It is well-known that traditional backprojection-based reconstruction methods produce significant artifacts when applied to limited-angle data [1], [2]. Statistical inversion (SI) is more suitable for reconstruction since *a priori* information about the unknown body can be efficiently combined with measurement data, leading to improved inversion. In statistical inversion the unknown object is discretized and interpreted as a random variable f . The X-ray measurement is modeled as the *likelihood distribution* $\pi(y|f)$, where y represents the measured image data. All available *a priori* information about f is modeled as the *prior distribution* $\pi(f)$. The solution of the inverse problem is the *posterior distribution* given by the

Bayes formula:

$$\pi_{\text{post}}(f) \propto \pi(y|f)\pi_{\text{pr}}(f).$$

To get a reconstruction of the object we compute a representative from the posterior distribution, for example the *maximum a posteriori* (MAP) estimate

$$f_{\text{MAP}} := \operatorname{argmax} \pi_{\text{post}}(f), \quad (1)$$

or the *conditional mean* (CM) estimate

$$f_{\text{CM}} := \int f \pi_{\text{post}}(f) df.$$

The choice of the prior distribution is the crucial point of statistical inversion. Gaussian smoothness prior is used in [3], more general Gibbs prior in [4], and geometrical prior models e.g. in [5], [6]. For a thorough review of applications of statistical inversion to limited data tomographic problems see [7, section 4.4]. In cases when the function f is *a priori* known to be nonnegative and piecewise smooth, the total variation (TV) prior is used in limited-angle X-ray tomography in [8], [9], [10]. It is found to give edge-preserving reconstructions.

Based on [11], the visible singularities in limited-angle tomography are analyzed in [12]. Roughly speaking, the edges and cracks in the tissue that are tangented by some detected X-rays can be reliably reconstructed. This theoretical result shows up in practice, too: two-dimensional slices through the total variation reconstruction perpendicular to the measured X-rays are very close to the true attenuation coefficient, while slices along the X-rays are not acceptable reconstructions [8]. However, often the diagnostically important information can be found in the good-quality slices.

Despite the success of total variation reconstructions, room for improvement remains. First, it is shown in [13] that the CM estimate with total variation prior loses its edge-preserving nature with very fine discretization. Second, the computational burden of total variation reconstruction is quite high.

In this work we address the above problems by using a wavelet expansion approximation and Besov space *a priori* information to compute a *maximum a posteriori* estimation for an image reconstruction. (See [14], [15] for an introduction to wavelets.) We also introduce a *pre-thresholding method* in which thresholding is applied to the wavelet coefficients prior to the computation of the reconstruction.

The inverse problem is to determine the wavelet coefficients from given projection images and available *a priori* information. We specify *a priori* information in terms of the statistical properties of random wavelet coefficients. Besov spaces are

Manuscript received XXX; revised XXX. This work was supported by National Technology Agency of Finland (TEKES, contract 206/03).

M. Rantala, M. Kalke, J. Moberg and S. Siltanen are with GE Healthcare, P.O. Box 20, FIN-04301 Tuusula, Finland.

S. Vänskä is with Rolf Nevanlinna Institute, P.O. Box 68, FIN-00014 University of Helsinki, Finland.

M. Lassas and S. Järvenpää are with Helsinki University of Technology, P.O. Box 1100, FIN-02015 TKK, Finland.

the natural spaces to work in with wavelets. When a function is represented by a wavelet expansion, one can see in which Besov space the function belongs to by studying the decay of the wavelet coefficients [14]. The Besov space norm can be interpreted as a weighted sum of the wavelet coefficients with scale-dependent weights. In this work, we use Besov spaces as *a priori* information, i.e., we penalize a large Besov norm of the reconstruction.

The locality and the scaling property of wavelets imply that in practical cases usually only a relatively small number of the coefficients is essentially non-zero. It is shown in [16] that if the wavelet coefficients are set zero with certain probabilities, the non-zero coefficients are normally distributed and the wavelet is smooth enough, the function falls to any Besov space by choosing the corresponding hyperparameters. We use a different approach and introduce a straightforward algorithm, which first decreases the set of potentially non-zero coefficients before the coefficients are actually computed. This *pre-thresholding* procedure reduces the computational burden before the reconstruction.

We demonstrate the feasibility of our method by numerical examples using *in vitro* data from dental radiology and mammography. In the dental examples we find that Besov prior reconstructions are edge-preserving and comparable in quality to the total variation reconstructions reported in [8]. Our method gives promising reconstructions also in the case of mammography, while the total variation prior is perhaps not suitable for modeling fine scale structure of breast tissue.

We compare our reconstruction method to (unfiltered) back-projection, also called tomosynthesis [17]. This is because tomosynthesis has been much investigated as a radiological technique both in dentistry and mammography [18]. Also, it is known theoretically that certain parts of boundaries between tissues can be seen in tomosynthetic reconstruction [12]. We do not use filtered backprojection (FBP) in the comparison because FBP is known to perform badly when applied to limited-angle data, see e.g. [1, Fig.VI.1].

Wavelet-based backprojection type methods are studied without using *a priori* information in [19], [20], [21], [22] (full-angle data) and in [23] (limited-angle data). Statistical inversion is used in reconstructing wavelet coefficients of f from full-angle data in [24], [25]. Statistical wavelet based methods (including the use of Besov space priors) are applied to image processing [26], [27], [28], [29], [30] and to geophysics [31]. For other wavelet-based statistical prior models, see [32], [33]. To our best knowledge, this is the first paper to use wavelets for incorporating *a priori* information in statistical solution of limited-angle tomography problems.

This paper is organized as follows. Wavelet expansions and Besov spaces are discussed in Section II. We present the measurement model in Section III. Bayesian inversion with Besov priors is discussed in Section IV. Optimization of computation by pre-thresholding the wavelet expansion is explained in Section V. Numerical results are presented in Section VI, and discussed in Section VII. We conclude our investigation in Section VII.

II. WAVELET EXPANSION AND BESOV SPACES

Let ϕ and ψ be the orthonormal scaling and wavelet functions, respectively, of some one-dimensional multiresolution analysis. Then in two dimensions,

$$\phi(x) = \phi(x_1)\phi(x_2), \quad x = (x_1, x_2)$$

is the scaling function, and

$$\begin{aligned} \psi^1(x) &= \phi(x_1)\psi(x_2), \\ \psi^2(x) &= \psi(x_1)\phi(x_2), \\ \psi^3(x) &= \psi(x_1)\psi(x_2), \end{aligned} \quad (2)$$

are the wavelet functions of three types, see [14]. Denote by $\psi_{j k \ell}(x) = 2^{-\frac{j}{2}} \psi^\ell(2^{-j}x - z_{jk})$, $\phi_{jk}(x) = 2^{-\frac{j}{2}} \phi(2^{-j}x - z_{jk})$, the scaled, dilated and translated functions. Here, index $j \in \mathbb{Z}$ is related to the scale, k to the location z_{jk} in space, and ℓ to the wavelet type, $\ell = 1, 2, 3$. The dilations z_{jk} form an appropriate grid in the computational domain. The wavelet expansion of a function $f \in L^p(\mathbb{R}^2)$, $1 \leq p \leq \infty$, is

$$f = \sum_k c_{0k} \phi_{0k} + \sum_{j \geq 0} \sum_k \sum_{\ell=1}^3 w_{jk\ell} \psi_{jk\ell}. \quad (3)$$

A Besov space is a natural space to operate in when we have wavelet representations for functions. Suppose that the wavelet family is r -regular or C^r -regular, see [14]. A function $f \in L^p(\mathbb{R}^2)$ with the wavelet expansion (3) is in Besov space B_{pq}^s , $s < r$, $1 \leq p, q < \infty$, [14], if the Besov norm

$$\begin{aligned} \|f\|_{B_{pq}^s} &= \|c_0\|_{\ell^p} + \left[\sum_{j=0}^{\infty} \left(2^{j(s+1-\frac{2}{p})} \|w_j\|_{\ell^p} \right)^q \right]^{\frac{1}{q}} \\ &=: \|f\|_p + \|f\|_{\tilde{B}_{pq}^s}, \end{aligned} \quad (4)$$

is finite. Here $c_0 = (c_{0k})$ and $w_j = (w_{jkl})$ are sequences indexed in some way. Note that increasing the parameter s is related to demanding more smoothness from the function. See [28] for more information on statistical use of Besov spaces.

In practice, we are operating with compactly supported functions (images), with compactly supported wavelets (e.g. the Daubechies-wavelets), and with finite scale approximations, i.e., in space

$$V_J = \text{span}\{\phi_{Jk} \mid k = 1, \dots, K_J\}.$$

The scale $J > 0$ and the number of possible translations K_J are determined by the size of the image. Since

$$V_J = V_{J_0} \oplus \prod_{j=J_0}^{J-1} \text{span}\{\psi_{j k \ell}\}_{k \ell}, \quad J_0 < J,$$

we can represent a function $f \in V_J$ either by the scaling functions,

$$f = \sum_{k=1}^{K_J} c_{Jk} \phi_{Jk},$$

or by the wavelet expansion,

$$f = \sum_{k=1}^{K_{J_0}} c_{J_0 k} \phi_{J_0 k} + \sum_{j=J_0}^{J-1} \sum_{k=1}^{K_j} \sum_{\ell=1}^3 w_{jk\ell} \psi_{jk\ell}. \quad (5)$$

By orthonormality, it holds that

$$c_{jk} = \int f \phi_{jk} dx, \quad w_{jkl} = \int f \psi_{jkl} dx,$$

and by the dilation equations, these integrals can be implemented by a discrete filtering procedure, see [34].

III. THE MEASUREMENT MODEL

Here, we describe the pencil beam model for X-ray attenuation (see [7] for more details).

Let the object be a compactly supported density $f = f(x)$. Suppose that an X-ray source is placed on one side of the object and the radiation is detected on the other side at a detector point. Each source point and detector point pair indicates one data value

$$\int_s f d\ell,$$

where s is the line joining the source and the detector point. In practice, the detector is a line grid (in the two-dimensional case) and it depends on the measurement equipment if the detector points or the source point, or both, are moving with respect to the object.

Let the number of the detector points be D and the number of the projection images be I . Then the measurement consists of $M = D \cdot I$ lines and of the corresponding integrals,

$$\int_{s_m} f d\ell = y_m, \quad m = 1, \dots, M,$$

$m = m(i, d)$, with $i = 1, \dots, I$ and $d = 1, \dots, D$. By approximating and representing f with the wavelet expansion (5), we get a $M \times N$ matrix equation

$$Aw = y, \quad (6)$$

where $y = (y_m)$ and

$$w = \begin{pmatrix} (c_{J_0k}) \\ (w_{jkl}) \end{pmatrix}.$$

The entries of the matrix A are

$$A_{mn} = \int_{s_m} \psi_n d\ell,$$

where $\psi_n = \phi_{J_0k}$ or $\psi_n = \psi_{jkl}$ with some indexing $n = 1, \dots, N$.

Note the following relationship between representing f by the pixel values and by the wavelet expansion. Let B be the matrix corresponding to (5) relating the pixel values x of f and the wavelet coefficients w ,

$$x = Bw.$$

The measurement equation for x is

$$\mathcal{A}x = y$$

with

$$\mathcal{A}_{mn} = \int_{s_m \cap \{\text{pixel } n\}} 1 d\ell.$$

Now, it holds that

$$\mathcal{A}B = A.$$

IV. BAYESIAN INVERSION WITH BESOV NORM PRIOR

The idea of Bayesian inversion is to interpret the variables as random variables and determine the probability densities for them.

For the joint probability $\pi(f, y)$, it holds that

$$\pi(f|y)\pi(y) = \pi(f, y) = \pi(y|f)\pi_{\text{pr}}(f). \quad (7)$$

Here the conditional distribution $\pi(f|y)$ is the *posterior distribution* of interest,

$$\pi_{\text{post}}(f) = \pi(f|y) = \frac{\pi(y|f)\pi_{\text{pr}}(f)}{\pi(y)}. \quad (8)$$

As seen from equation (8), the density $\pi(y)$ can be considered as a normalizing constant. We suppose that the noise in the measurement (6) is additive, uncorrelated, and Gaussian,

$$Aw = y + \epsilon, \quad \epsilon \sim N(0, \sigma^2). \quad (9)$$

The feasibility of the noise model (9) is discussed in [7]. The *likelihood distribution* is

$$\pi(y|f) = C e^{-\|Aw - y\|_2^2 / 2\sigma^2}$$

with some normalizing constant C .

Finally, $\pi_{\text{pr}}(f)$ is the *prior density*, which contains the information that we *a priori* have on f . We seek the reconstruction in a finite dimensional space V_J , so we *a priori* assume that

$$f \in V_J$$

with some $J > 0$. We represent *a priori* information by penalizing large Besov norm of f . This is computationally straightforward when we use a wavelet representation for f , and moreover, we can control the smoothness of the reconstruction by tuning the Besov space parameters. Since the object is an X-ray attenuation coefficient, we also know that it is non-negative. Hence, we choose a prior of a Besov norm together with the approximative positivity condition,

$$\pi_{\text{pr}}(f) = C \exp \left(-\alpha_1 \left(\|f\|_p^p + \|f\|_{\bar{B}_{pq}^s}^q \right) - \alpha_2 g_+(f) \right).$$

Here α_j 's are positive control constants and

$$g_+(f) = \sum_{n=1}^M \gamma(f_n),$$

where f_n , $n = 1, \dots, N$, denotes the pixel values of f and

$$\gamma(f_n) = \begin{cases} \frac{1}{2}(f_n)^2, & f_n < 0 \\ 0, & f_n \geq 0 \end{cases}.$$

Now, the posterior density is

$$\pi_{\text{post}}(f) = C \exp \left(-\frac{1}{2\sigma^2} \|Aw - y\|_2^2 - \alpha_1 (\|f\|_p^p + \|f\|_{\bar{B}_{pq}^s}^q) - \alpha_2 g_+(f) \right). \quad (10)$$

This is the statistical solution of the inverse problem. To get a single representative (image) of the posterior distribution, we compute the MAP estimate (1). This is equivalent to minimizing the functional

$$F(f) = \frac{\|Aw - y\|_2^2}{2\sigma^2} + \alpha_1 (\|f\|_p^p + \|f\|_{\bar{B}_{pq}^s}^q) + \alpha_2 g_+(f). \quad (11)$$

Note that minimizing (11) can also be interpreted as solving (6) as a regularized problem with a penalty function containing the Besov norm and positivity.

Here the minimization is done with the Barzilai and Borwein minimization method [35], [36]. We start with an initial guess and take steps in the direction of the negative gradient of the penalty function by computing with the formula

$$f_{k+1} = f_k - s_k \nabla F(f_k),$$

where the step length is given by

$$\frac{1}{s_k} = \frac{(f_k - f_{k-1})^T (\nabla F(f_k) - \nabla F(f_{k-1}))}{(f_k - f_{k-1})^T (f_k - f_{k-1})}.$$

Note that the first step length can not be computed with this formula. An adequate step length to start with can be found e.g. by minimization along a one-dimensional line. A suitable choice in practice is to take s_0 small enough.

V. PRE-THRESHOLDING

The locality and the scaling property of wavelets imply that in practical cases a significant amount of the coefficients of a wavelet expansion is zero, after some thresholding rule is applied. We want to take advantage of this reduction and decrease the size of the minimization problem by dropping out the coefficients that are approximately zero. We introduce a procedure with which one can do this thresholding before actually reconstructing the wavelet coefficients. We call this *pre-thresholding*.

To be more specific, let the wavelet expansion of the correct density function $f \in V_J$ be given by (5). Suppose that we use some thresholding rule, and the coefficients close to zero are set to zero. Denote the indices of the remaining non-zero wavelet coefficients by \mathcal{W}_T . Then the sum

$$\sum_{k=1}^{K_{J_0}} c_{J_0 k} \phi_{J_0 k} + \sum_{(j,k,\ell) \in \mathcal{W}_T} w_{j k \ell} \psi_{j k \ell}$$

gives a good approximation for f . This gives a motivation for doing the reconstruction in two steps. First, determine the set \mathcal{W}_T , and secondly, solve the values for the corresponding coefficients by minimizing (11), as described above.

In practice, we can not reconstruct the set \mathcal{W}_T exactly without knowing f . Instead, we seek a set \mathcal{W} of indices such that the sets

$$\mathcal{W}_n = \mathcal{W}_T \setminus \mathcal{W} = \mathcal{W}_T \cap \mathcal{W}^C, \quad (12)$$

and

$$\mathcal{W}_e = \mathcal{W} \setminus \mathcal{W}_T = \mathcal{W} \cap \mathcal{W}_T^C, \quad (13)$$

are both small. Here, \mathcal{W}_n is the set of coefficients that are non-zero for the true density function f but will not be reconstructed, and \mathcal{W}_e is the set of extra coefficients that are zero for f but are taken into the minimization process.

In obtaining the set \mathcal{W} , we have used the following *back-projection pre-thresholding procedure*. Compute the back-projection reconstruction

$$x_{\text{BP}} = A^T y,$$

and the corresponding wavelet coefficients w_{BP} . Threshold coefficients w_{BP} , and choose the remaining coefficients as \mathcal{W} . We apply the following thresholding rule: Define thresholding parameter τ , $0 \leq \tau \leq 1$. Set the smallest $\tau \cdot 100\%$ of the finest scale coefficients to zero, and similarly for every scale the smallest $\tau \cdot 2^{-\frac{1}{2}(j-1)} \cdot 100\%$ of the j th finest scale coefficients to zero. Note that w_{BP} can be computed directly by

$$w_{\text{BP}} = A^T y,$$

or by computing the wavelet transform of x_{BP} , depending on whether one uses the matrix A or \mathcal{A} .

Since the backprojection x_{BP} is fast to compute, the computation of \mathcal{W} is also fast. In the limited-angle sparse data case, the quality of x_{BP} is not very high and the image has significant artifacts. However, the singularities that are perpendicular to the projection directions appear in x_{BP} (see [12]). Hence, we expect to have the majority of the reconstructable coefficients \mathcal{W}_T of the real singularities within set \mathcal{W} , but also some extra coefficients due to the artifacts.

How to choose the thresholding parameter τ in practice? We determine τ in a test case, and then use the same τ with similar data (e.g. for mammography). We choose a τ that gives a visually optimal reconstruction, by which we mean that if τ is chosen to be larger, we get visually observable artifacts in the reconstruction. This is demonstrated later in Section VI (see Figures 4 and 6).

VI. RESULTS

We test our method with three different data. In all cases the two-dimensional projection images are sliced to one-dimensional projection functions from which we compute two-dimensional reconstructions. Then the three-dimensional reconstructions are built by attaching these two-dimensional reconstruction slices together. All computations are done with a desktop PC (Intel processor, 2.4 GHz CPU, 1 GB RAM) using Matlab 6.5. We use the Daubechies 6 -wavelet (db6) with three scaling levels, and Besov parameters $p = q = 1.5$ and $s = 0.5$.

We consider the wavelet reconstructions with all the coefficients to be good comparison objects for testing the pre-thresholding method. However, in the first simple test case, we also have the sparse full-angle data to compare with.

A. Tooth Phantom Case

The first data is measured from a tooth phantom (a third mandibular molar) with the Focus intraoral X-ray source and the Sigma intraoral sensor (by GE Healthcare). The full-angle data contains 23 projections taken equally spaced from a total viewing angle of 187° . The limited-angle data contains nine of these projections, creating a total viewing angle of 68° .

Figure 1 presents a comparison between reconstructions computed from the two different data sets, and also between reconstructions computed with all wavelet coefficients and with pre-thresholding. The full-angle reconstructions are considered here as the ground truth. The relative errors (in the sense of $(\max_n |f_n - \hat{f}_n|) / \max_n |f_n|$) in the limited-angle reconstructions are computed along the middle horizontal lines.

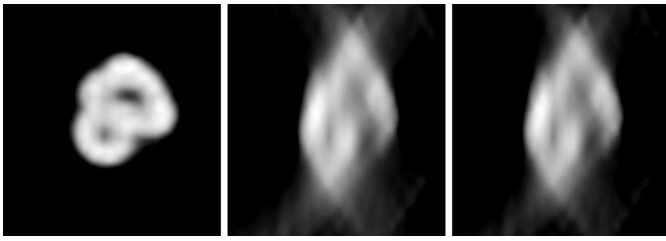


Fig. 1. 2D reconstruction of a tooth phantom. Left: wavelet reconstruction with all coefficients from full-angle data (187° viewing angle) as the 'ground truth'. Middle: wavelet reconstruction with all coefficients from limited-angle data (68° viewing angle), showing 24 % of relative error along the middle horizontal line. Right: wavelet reconstruction with pre-thresholding with $\tau = 0.8$ from limited-angle data, showing 26 % of relative error along the middle horizontal line.

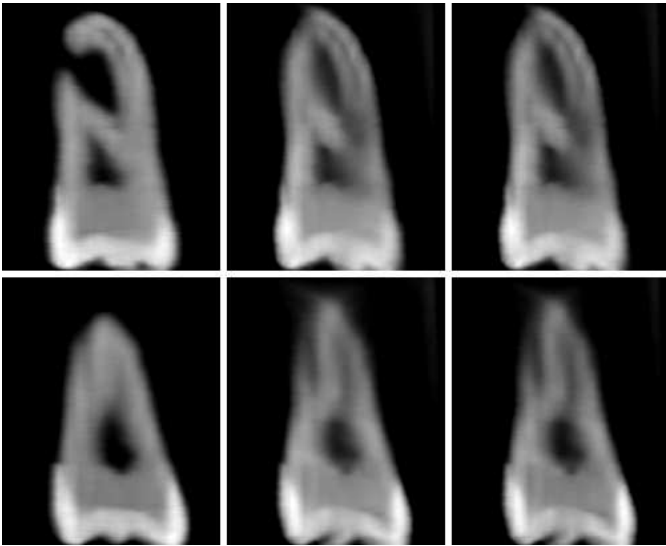


Fig. 2. Slices parallel to the detector from 3D reconstructions of a tooth phantom. Left column: wavelet reconstruction with all coefficients from full-angle data (187° viewing angle) as the 'ground truth'. Middle column: wavelet reconstruction with all coefficients from limited-angle data (68° viewing angle). Right column: wavelet reconstruction with pre-thresholding with $\tau = 0.8$ from limited-angle data.

The errors are 24 % in the reconstruction with all coefficients and 26 % in the reconstruction with pre-thresholding with $\tau = 0.8$. Here the image sizes are 174×164 pixels.

We attach these two-dimensional slices perpendicular to the detector of the X-ray device to form a three-dimensional volume. In practice a volume is usually examined as slices parallel to the detector of the device. Figure 2 presents two examples of such slices. The image sizes are 201×164 pixels.

B. Mammography Case

The second data is measured from a mammographic specimen with a digital full-field mammography equipment (by GE Healthcare). The seven projection images are taken from a viewing angle of 30° .

Figure 3 presents a comparison between the various reconstructions we experiment with. Here the image sizes are 143×468 pixels. The pre-thresholding parameters used are $\tau = 0.7$ and $\tau = 0.9$, which produce total zeroing percentages of 62 % and 79 %, respectively. The whitest detail in the

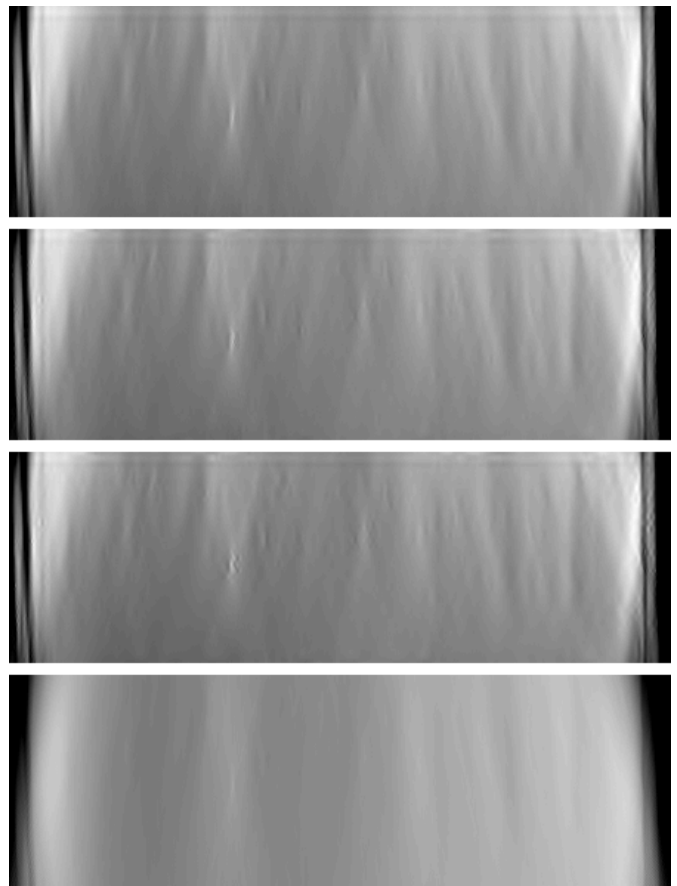


Fig. 3. 2D reconstructions of a mammography specimen. Top: wavelet reconstruction with all coefficients. Second: wavelet reconstruction with pre-thresholding with $\tau = 0.7$. Third: wavelet reconstruction with pre-thresholding with $\tau = 0.9$. Bottom: backprojection reconstruction.

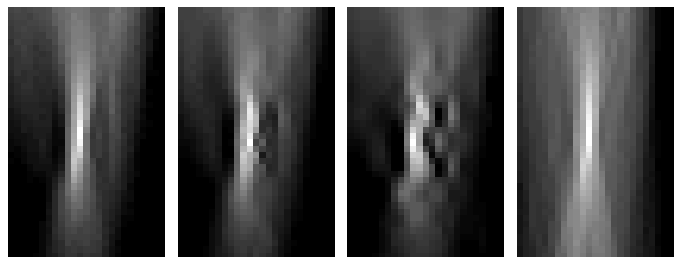


Fig. 4. A detail in 2D reconstructions of a mammography specimen. Left: wavelet reconstruction with all coefficients. Second: wavelet reconstruction with pre-thresholding with $\tau = 0.7$. Third: wavelet reconstruction with pre-thresholding with $\tau = 0.9$. Right: backprojection reconstruction.

images is a calcified vein, and it is presented also in detail in Figure 4, in which the image sizes are 40×25 pixels. The reductions in reconstruction times are 7 % with $\tau = 0.7$ and 12 % with $\tau = 0.9$.

Figure 5 presents slices parallel to the detector from the three-dimensional volume. Here the image sizes are 294×468 pixels. The slice of Figure 5 presents the calcified vein shown also in the perpendicular slices of Figure 3. In Figure 6 the vein is presented in detail (60×40 pixels).

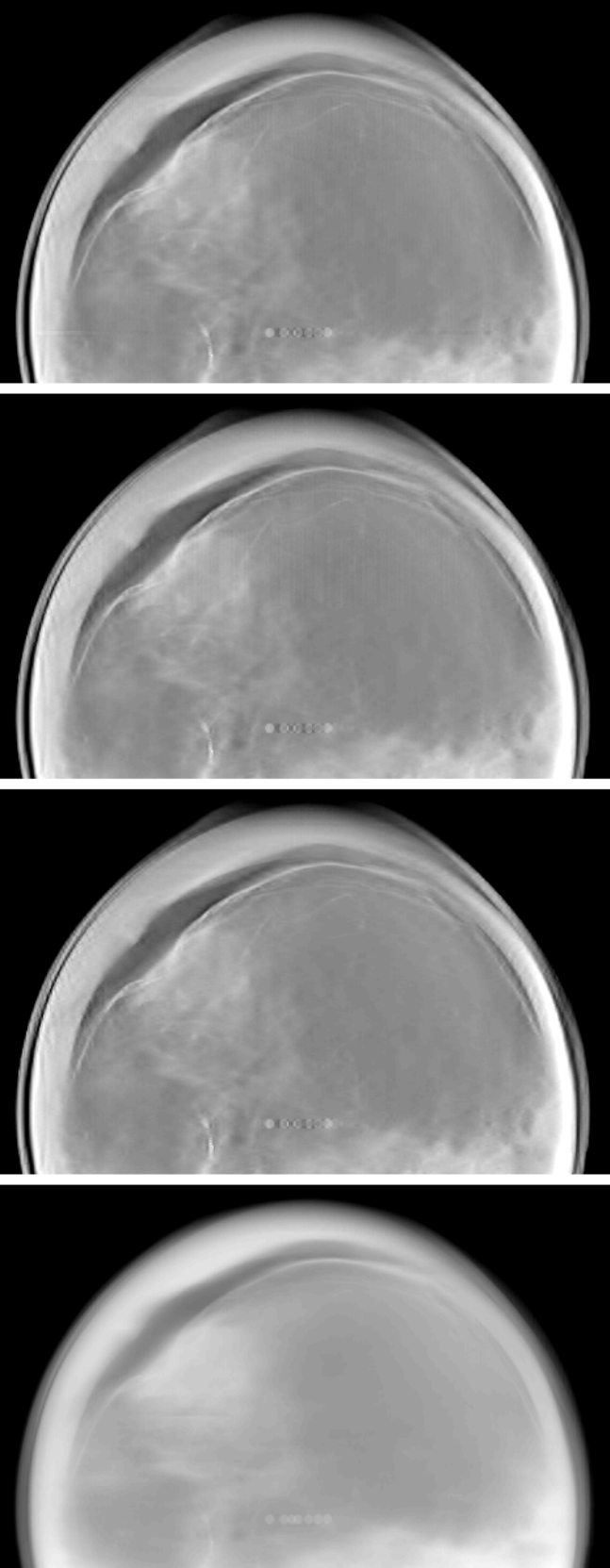


Fig. 5. Slices parallel to the detector from 3D reconstructions of a mammography specimen. Top: wavelet reconstruction with all coefficients. Second: wavelet reconstruction with pre-thresholding with $\tau = 0.9$. Third: wavelet reconstruction with pre-thresholding with $\tau = 0.9$. Bottom: backprojection reconstruction.

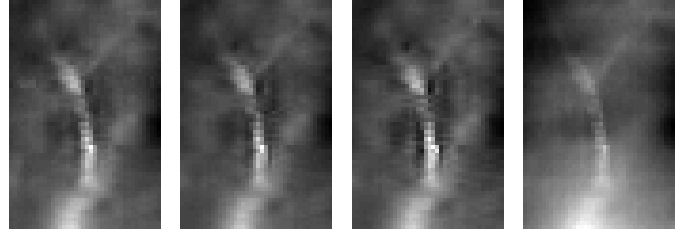


Fig. 6. A detail in slices parallel to the detector from 3D reconstructions of a mammography specimen. Left: wavelet reconstruction with all coefficients. Second: wavelet reconstruction with pre-thresholding with $\tau = 0.7$. Third: wavelet reconstruction with pre-thresholding with $\tau = 0.9$. Right: backprojection reconstruction.

C. Intraoral Dental Imaging Case

The third data we experiment with is measured with the Focus intraoral X-ray source and the Sigma intraoral sensor from a head phantom. Here, the seven projection images are taken from a viewing angle of 60° .

Figure 7 presents reconstruction slices parallel to the detector. The left image column shows one of the three roots of the middle tooth and the right image column the other two. Here the pre-thresholding parameters are $\tau = 0.7$ and $\tau = 0.95$, leading to total zeroing percentages of 61 % and 83 %, respectively. The image sizes are 179×227 pixels. In this example the reductions in reconstruction times are 5 % with $\tau = 0.7$ and 7 % with $\tau = 0.95$.

VII. DISCUSSION

In the test examples, the Besov prior without pre-thresholding produces good images comparing to the backprojection reconstruction. Backprojection, also known as tomosynthesis, is currently a popular technique for reconstruction from limited angle data, see [18]. Further, by using the backprojection pre-thresholding procedure we are able to drop out over 60 % of the unknown coefficients without degrading largely the reconstructed images.

In all three examples, the parameter p is chosen to be a value between 1 and 2 by comparing visually various reconstructions (note that p needs to be greater than 1 to ensure differentiability of the objective function in the minimization problem). The parameter $q = p$ is chosen for simplicity, and s by visual inspection the largest value that does not smoothen up the reconstructions too much. Also, a value of s between 0.3 and 0.6 is found to be appropriate for natural (photographic) images in [37], giving indirect support for our choice $s = 0.5$.

Next we discuss the performance of our method in detail for the three example cases.

A. Tooth Phantom Case

In this example we have ground truth available, namely the reconstruction using full-angle data. As expected from the theory of recoverable singularities in X-ray tomography, we are able to reconstruct slices parallel to the detector (and thus perpendicular to the rays) with reasonable 24 % accuracy. Further, zeroing as much as over 70 % of the coefficients does

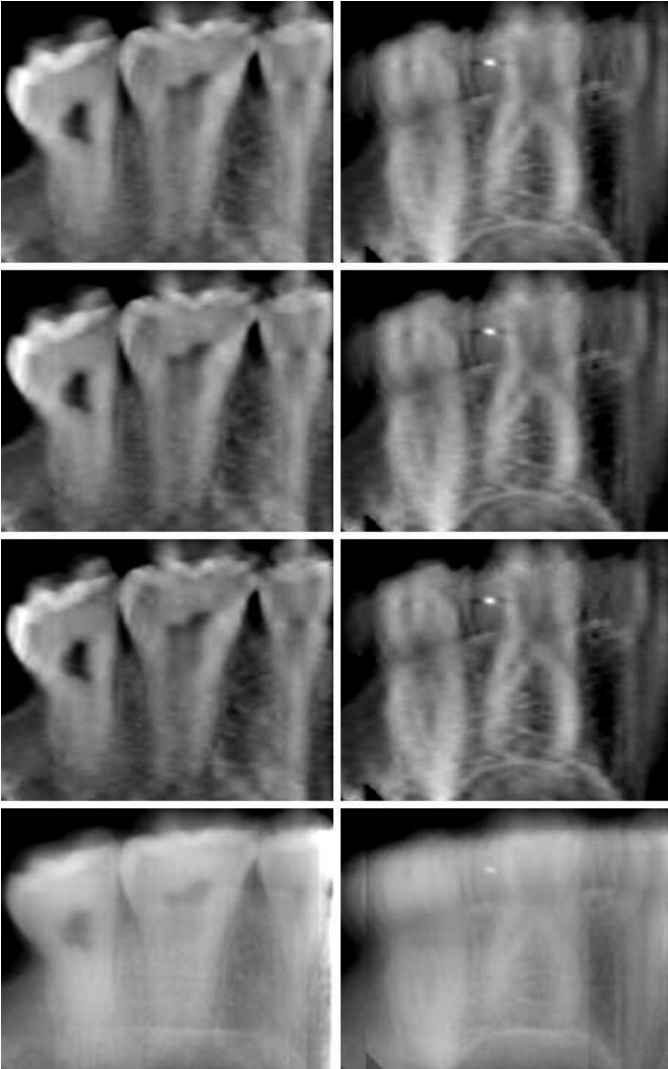


Fig. 7. Slices parallel to the detector from 3D reconstructions of a dental phantom. Top row: wavelet reconstruction with all coefficients. Second row: wavelet reconstruction with pre-thresholding with $\tau = 0.7$. Third row: wavelet reconstruction with pre-thresholding with $\tau = 0.95$. Bottom row: backprojection reconstruction.

not produce significantly more artifacts, and the quantitative error in the slice parallel to the detector grows only moderately (from 24 % to 26 %). See Figure 1.

Further, the three-dimensional reconstruction formed as a stack of two-dimensional slices is shown in Figure 2. The reconstruction from limited-angle data is very close to the ground truth, and there is no significant difference between the full and pre-thresholded reconstructions.

B. Mammography Case

The opening angle of 30° is achievable with an existing mammographic device and is thus realistic for diagnostic mammography. The benefits of three-dimensional mammography over traditional projection mammography include the possibility to recover the shape of clusters of microcalcifications.

Figure 5 demonstrates that the Besov prior approach yields

reconstructions with better sharpness and contrast than back-projection. This is also evident from Figure 6 where we pick out a calcified vein as a diagnostically interesting feature.

Pre-thresholding with parameter $\tau = 0.7$ seems to yield almost identical image quality than full wavelet tree reconstruction, while parameter $\tau = 0.9$ is arguably too large. See Figures 5 and 6. From the computational point of view, the pre-thresholding procedure stabilizes the minimization problem by reducing the amount of the unknowns: fewer iteration steps are needed with the reduced problem. Also, the choice $\tau = 0.7$ leads to 62 % reduction in disc storage space for representing the reconstruction and a moderate 7 % reduction in computation time.

C. Intraoral Dental Imaging Case

The imaging geometry chosen here models the situation at a typical dental clinic. The opening angle of 60° is achievable with standard intraoral imaging equipment. Benefits of three-dimensional intraoral imaging over traditional projection imaging include the possibility to recover the exact position and shape of tooth roots. This knowledge is often needed before removing a tooth.

Figure 7 shows two different slices through the teeth. The three roots of the middle tooth are visible in different slices, one in the left image column and the other two in the right image column. Reconstructions using the Besov prior show significantly better contrast than the backprojection reconstruction.

Compare the top row and third row from top in Figure 7. Pre-thresholding with parameter $\tau = 0.95$ seems to yield almost identical image quality than full wavelet tree reconstruction. This leads to 83 % reduction in disc storage space for representing the reconstruction and a moderate 7 % reduction in computation time. Similarly to the mammography case, fewer iteration steps are needed in the reduced minimization problem.

VIII. CONCLUSION

We consider limited-angle tomography problems arising in medical X-ray imaging and demonstrate the feasibility of statistical wavelet-based Besov prior techniques. Our examples show (using *in vitro* specimen data) that one can significantly improve the clinical quality of reconstructions compared to backprojection techniques, such as tomosynthesis. In particular, dental reconstructions using a Besov prior are of similar quality as the total variation reconstructions in [8]. In contrast to the total variation prior, however, we expect wavelet-based Besov space priors to be discretization invariant in the sense defined in [13].

While we consider here stacks of two-dimensional tomographic problems, it is straightforward to generalize our method to three-dimensional setting.

Further, we introduce a pre-thresholding approach for reducing the number of free parameters in the tomographic problem. Benefits of pre-thresholding include 60–80 % reduction in disc storage requirements, faster and more stable convergence of the optimization method, and slightly reduced computation

times and random access memory needs. We expect the advantage of these benefits to be significantly greater in genuinely three-dimensional settings.

REFERENCES

- [1] F. Natterer, *The mathematics of computerized tomography*, SIAM, 2001.
- [2] R. Rangayyan, A. P. Dhawan and R. Gordon, "Algorithms for limited-view computed tomography: an annotated bibliography and a challenge," *Applied optics*, vol. 24, pp. 4000–4012, 1985.
- [3] K. M. Hanson and G. W. Wecksung, "Bayesian approach to limited-angle reconstruction in computed tomography," *J. Opt. Soc. Am.*, vol. 73, pp. 1501–1509, 1983.
- [4] K. Sauer, J. Sachs Jr. and K. Klifa, "Bayesian estimation of 3-D objects from few radiographs," *IEEE Transactions on Nuclear Science*, vol. 41, pp. 1780–1790, 1994.
- [5] K. M. Hanson, G. S. Cunningham and R. J. McKee, "Uncertainty assessment for reconstructions based on deformable geometry," *Int. J. Imaging Syst. Technol.*, vol. 8, pp. 506–512, 1997.
- [6] K. M. Hanson, G. S. Cunningham and R. J. McKee, "Uncertainties in tomographic reconstructions based on deformable models," In K. M. Hanson, editor, *Medical Imaging: Image processing*, vol. 3034 of *Proc. SPIE*, pp. 276–286, 1997.
- [7] S. Siltanen, V. Kolehmainen, S. Järvenpää, J. P. Kaipio, P. Koistinen, M. Lassas, J. Pirttilä and E. Somersalo, "Statistical inversion for medical X-ray tomography with few radiographs I: General theory," *Phys. Med. Biol.*, vol. 48, pp. 1437–1463, 2003.
- [8] V. Kolehmainen, S. Siltanen, S. Järvenpää, J. P. Kaipio, P. Koistinen, M. Lassas, J. Pirttilä and E. Somersalo, "Statistical inversion for medical X-ray tomography with few radiographs II: Application to dental radiology," *Phys. Med. Biol.*, vol. 48, pp. 1465–1490, 2003.
- [9] M. Persson, D. Bone and H. Elmqvist, "Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography," *Physics in Medicine and Biology*, vol. 46, pp. 853–866, 2001.
- [10] A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography," *IEEE Transactions on Image Processing*, vol. 7, pp. 204–221, 1998.
- [11] V. Guillemin and S. Sternberg, *Geometric asymptotics*, Amer. Math. Soc., Providence, RI, 1977.
- [12] E. T. Quinto, "Singularities of the X-ray transform and limited data tomography in \mathbb{R}^2 and \mathbb{R}^3 ," *SIAM J. Math. Anal.*, vol. 24, pp. 1215–1225, 1993.
- [13] M. Lassas and S. Siltanen, "Can one use total variation prior for edge-preserving Bayesian inversion?," *Inverse Problems*, vol. 20, pp. 1537–1563, 2004.
- [14] Y. Meyer, *Wavelets and operators*, Cambridge Univ. Press, 1992.
- [15] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674–693, 1989.
- [16] F. Abramovich, T. Sapatinas and B. W. Silverman, "Wavelet thresholding via a Bayesian approach," *J.R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 60, pp. 725–749, 1998.
- [17] D. G. Grant, "Tomosynthesis: a three-dimensional radiographic imaging technique," *IEEE Transactions on Biomedical Engineering*, vol. 19, pp. 20–28, 1972.
- [18] J. T. Dobbins III and D. J. Godfrey, "Digital x-ray tomosynthesis: current state of the art and clinical potential," *Physics in Medicine and Biology*, vol. 48, pp. R65–R106, 2003.
- [19] A. H. Delaney and Y. Bresler, "Multiresolution tomographic reconstruction using wavelets," *IEEE Transactions on Image Processing*, vol. 4, pp. 799–813, 1995.
- [20] F. Rashid-Farrokhi, K. J. R. Liu, C. A. Berenstein and D. Walnut, "Wavelet-based multiresolution local tomography," *IEEE Transactions on Image Processing*, vol. 6, pp. 1412–1430, 1997.
- [21] T. Olson and J. DeStefano, "Wavelet localization of the Radon transform," *IEEE Transactions on Image Processing*, vol. 42, pp. 2055–2067, 1994.
- [22] S. Zhao, G. Welland and G. Wang, "Wavelet sampling and localization schemes for the Radon transform in two dimensions," *SIAM J. Appl. Math.*, vol. 57, pp. 1749–1762, 1997.
- [23] B. Sahiner and A. E. Yagle, "Limited angle tomography using wavelets," *Proc. 1993 IEEE Med. Im. Conf.*, pp. 1912–1916, San Francisco CA., 1994.
- [24] T. Frese, C.A. Bouman and K. Sauer, "Adaptive wavelet graph model for bayesian tomographic reconstruction," *IEEE Transactions on Image Processing*, vol. 11, pp. 756–770, 2002.
- [25] M. Bhatia, W.C. Karl and A.S. Willsky, "A wavelet-based method for multiscale tomographic reconstruction," *IEEE Trans. Medical Imaging*, vol. 15, pp. 92–101, 1996.
- [26] R. W. Buccigrossi and E. P. Simoncelli, "Image compression via joint statistical characterization in the wavelet domain," *IEEE Transactions in Image Processing*, vol. 8, pp. 1688–1701, 1999.
- [27] K. Berkner, M. J. Gormish and E. L. Schwartz, "Multiscale sharpening and smoothing in Besov spaces with applications to image enhancement," *Applied and Computational Harmonic Analysis*, vol. 11, pp. 2–31, 2001.
- [28] H. Choi and R.G. Baraniuk, "Multiple Wavelet Basis Image Denoising using Besov Ball Projections," *IEEE Signal Processing Letters*, vol. 11, No. 9, pp. 717–720, 2004.
- [29] M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Transactions on Image Processing*, vol. 12, pp. 906–916, 2003.
- [30] G. Wang, J. Zhang and G.-W. Pan, "Solution of inverse problems in image processing by wavelet expansion," *IEEE Transactions on Image Processing*, vol. 4, pp. 579–593, 1995.
- [31] J. Kane, F. J. Herrmann and M. N. Tököz, "Wavelet domain linear inversion with application to well logging," *Preprint*, 2001.
- [32] P. Müller and B. Vidakovic (eds.), *Bayesian inference in wavelet-based models*, Springer Lecture Notes in Statistics, vol. 141, 1999.
- [33] B. Vidakovic, *Statistical modeling by wavelets*, Wiley Series in Probability and Statistics, 1999.
- [34] E. Aboufadel and S. Schlicker, *Discovering wavelets*, John Wiley & Sons, 1999.
- [35] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA Journal of Numerical Analysis*, vol. 8, pp. 141–148, 1988.
- [36] M. Raydan, "The Barzilai and Borwein method for the large scale unconstrained minimization problem," *SIAM J. Optim.*, vol. 7, pp. 26–33, 1997.
- [37] R. A. DeVore, B. Jawerth and B. J. Lucier, "Image Compression Through Wavelet Transform Coding," *IEEE Transactions on Information Theory*, vol. 38, pp. 719–746, 1992.