A new framework for sparse regularization in limited angle x-ray tomography

Jürgen Frikel

Image Diagnost International GmbH, Balanstr. 57, D-81541 München, Germany

M6 - Mathematische Modellbildung, Technische Universität München, Zentrum Mathematik, Boltzmannstr. 3, D-85748 Garching bei München, Germany

Abstract

We propose a new framework for limited angle tomographic reconstruction. Our approach is based on the observation that for a given acquisition geometry only a few (visible) structures of the object can be reconstructed reliably using a limited angle data set. By formulating this problem in the curvelet domain, we can characterize those curvelet coefficients which correspond to visible structures in the image domain. The integration of this information into the formulation of the reconstruction problem leads to a considerable dimensionality reduction and yields a speedup of the corresponding reconstruction algorithms.

Keywords: Limited angle tomography, dimensionality reduction, sparse regularization, curvelets, wavefront set

1. Introduction

We consider the problem of limited angle x-ray tomography where only a few projections (measurements) are taken at a small angular range. Such problems arise for example in digital breast tomosynthesis. The goal is to reconstruct the attenuation coefficient f from the noisy measurements

$$y = Rf + \eta, \tag{1}$$

where $\eta \in \mathbb{R}^M$ denotes the noise and $R \in \mathbb{R}^{M\times n}$ is the system matrix of the x-ray transform.

In such problems, the acquired tomographic data y is highly incomplete and, thus, the corresponding reconstruction problem is severely ill-posed [7]. It is well-known that traditional reconstruction methods such as filtered backprojection (FBP) do not perform well in such situations [7].

Due to the ill-posed nature, regularization has to be performed in order to make the reconstruction insensitive to noise. For this reason, additional prior knowledge about the unknown object has to be integrated into the reconstruction.

The classical approach is the so-called Tikhonov regularization, [4], where the regularized solution f_α is obtained as a minimizer of the *Tikhonov functional*

$$T_\alpha(f) = \|Rf - y\|_2^2 + \alpha \|f\|_2^2 . \tag{2}$$

The first term in (2) is the data fidelity term, which controls the reconstruction error. The second term is the so-called penalty or prior term, which controls the smoothness of the solution.

The disadvantage of Tikhonov regularization is that it tends to oversmooth the solution in some situations. This is not wanted in many medical imaging applications. For this reason, other penalties have been proposed in the literature. Prominent examples are the total variation prior (TV) and the Besov norm prior which both enforce smoothness [5, 10].

Recently, the regularization with sparsity constraints has become popular [3]. Here, one seeks for a solution f^* which is 'sparse' in a given dictionary $\Psi = \{\psi_1, \dots, \psi_N\}$, i.e., a solution $f^* = \sum_k x_k \psi_k =: \Psi x$ for which only a small fraction of the coefficients $x_1, \dots, x_N$ is non-zero. It is well-known that ℓ^1-norm prefers sparse solutions [3, 1]. Therefore, in sparse regularization one is interested in minimization of the *ℓ^1-penalized Tikhonov functional*

$$T_\alpha(x) = \|Kx - y\|_2^2 + \alpha \|x\|_1 , \tag{3}$$

where $K = R\Psi \in \mathbb{R}^{M \times N}$.

From the theory of compressed sensing it is known that the integration of sparsity assumptions into the reconstruction has the ability to resolve undersampling [1].

We propose the use of a sparse regularization technique (3) for limited angle problems with angular undersampling. The assumption that the object is sparse in an appropriate basis means that the energy of the object is concentrated in very few (large) coefficients. This implies that the effective dimension is much lower than it pretends to be. Unfortunately, the location of these relevant coefficients is not known a priori.

In this work, we show that by formulating the reconstruction problem in the curvelet domain it is possible to predict the location of the relevant coefficients. The location of these curvelet coefficients depends only on the acquisition geometry. This information is therefore available prior to the acquisition. By exploiting this knowledge, we perform a significant dimensionality reduction in the curvelet domain.

In Section 2, we give a brief description of mathematical principles that we use in order to perform the dimensionality reduction in the curvelet domain. In Section 3, we describe how the location of the relevant (visible) curvelet coefficients can be predicted and formulate the reconstruction problem in the reduced dimension.

2. Mathematical principles

2.1. *Wavefront sets*

We will characterize objects f by means of its singularities (edges). For this reason, we introduce the notion of the wavefront set of f because it is a powerful concept which describes both, location x_0 and direction θ_0 of singularities.

We say that f is smooth near x_0 if there is a cut-off function $\varphi \in C_c^\infty$, $\varphi(x_0) \neq 0$, such that the Fourier transform of φf decays rapidly, i.e., $\widehat{\varphi f}(\xi) = \mathcal{O}(|\xi|^{-N})$, $|\xi| \to \infty$, for all $N > 0$. In what follows, we say that f has a singularity in x_0 if for all cut-off functions φ the Fourier transform of φf is not of rapid decay. The set of all singularities of f is called singular support and is denoted by $\operatorname{sing\,supp}(f)$.

In order to define the orientation of the singularity $x_0 \in \operatorname{sing\,supp}(f)$, we search for directions along which the localized Fourier transform $\widehat{\varphi f}$ does not decay rapidly. We define the *wavefront set* of a function f, $\mathrm{WF}f$, as the set of all tuples (x_0, θ_0) with location $x_0 \in \operatorname{sing\,supp}(f)$ and direction θ_0 such that for all cut-off functions $\varphi \in C_c^\infty$, $\varphi(x_0) \neq 0$, the localized Fourier transform $\widehat{\varphi f}(\xi)$ does not decay rapidly in any polar wedge $W_\delta = \{(r, \omega) : \; |\omega - \theta_0| < \delta\}$. Here, (r, ω) denote polar coordinates in the frequency domain.

The direction θ_0 of a singularity $x_0 \in \operatorname{sing\,supp}(f)$ can be considered as the direction of maximum change of f at x_0. For example, if we consider a function which has a jump singularity along a smooth curve γ, then any $x_0 \in \gamma$ is a singularity with direction θ_0, which is normal to the curve γ in x_0, i.e., the wavefront set is given by the set $\{(x_0, \theta_0) : \; x_0 \in \gamma, \; \theta_0 \text{ normal to } \gamma \text{ in } x_0\}$. However, in general, a singularity can have several distinct directions.

The introduced concept of singularity implies that the essential information of any object f is coded by its wavefront set because f is infinitely smooth in the neighborhood of all $x \notin \operatorname{sing\,supp}(f)$. Thus, our goal is to characterize those singularities (x_0, θ_0) of f which can be reconstructed using only a limited angle data set.

2.2. *X-ray transform and detection of singularities*

We now give a characterization of sigularities which can be reconstructed from the given data set. To this end we consider the continuous 2D x-ray transform $\mathcal{R}f$ of a function f which is given by

$$\mathcal{R}f(\theta, s) = \int_{L(\theta,s)} f(x) \, \mathrm{d}S(x), \tag{4}$$

where $L(\theta, s) = \left\{x \in \mathbb{R}^2 : \; x_1 \cos\theta + x_2 \sin\theta = s\right\}$ is the line with normal direction θ and signed distance from the origin s.

The following characterization was given by E.T. Quinto, [8]: Given the x-ray transform data $\mathcal{R}f(\theta, s)$ for (θ, s) arbitrarily near (θ_0, s_0), one can reconstruct only those singularities of f stably with location $x \in L(\theta_0, s_0)$ and direction θ_0. This means that a singularity (x_0, θ_0) of f is *visible*, [9], from a limited angle data set if the line $L(\theta_0, x_0 \cdot \theta_0)$ is in the data set.

This result provides precise information about the direction of singularities of a function f which can be reconstructed reliably. In particular, this information is available prior to the acquisition process because it depends only on the acquisition geometry. Therefore, given the acquisition geometry we can predict the orientation of those singularities which can be reconstructed reliably.

For example, consider a function which has a jump singularity along a smooth curve γ, then only those singularities (x_0, θ_0) are visible from a limited angle data set if there is a line in the data set which is tangent to γ at x_0.

2.3. Curvelets

Our goal is to integrate the above characterization of visible singularities into the reconstruction process. To this end, we have to separate visible and invisible structures of an object. We use curvelets as a tool for analyzing the wavefront set of a function f because they resolve wavefront sets, i.e., they provide information about both location and direction of singularities [2]. Here, we give only a brief description of the continuous curvelet transform and show how the separation of visible singularities can be performed by analyzing the curvelet coefficients of a function f.

The generating curvelet $\psi_{a,0,0}$ at scale a is defined in the frequency domain using polar coordinates (r, ω):

$$\widehat{\psi}_{a,0,0}(r, \omega) = a^{3/4} W(ar) V(\omega/\sqrt{a}), \tag{5}$$

where $W(r)$ is a radial window and $V(\omega)$ is an angular window. Both are compactly supported smooth functions that satisfy some admissibility conditions, [2]. The curvelets are then rotated using a 2D rotation matrix R_θ, $\theta \in [0, 2\pi)$, and translated by $b \in \mathbb{R}^2$,

$$\psi_{a,b,\theta}(x) = \psi_{a,0,0}(R_\theta(x - b)). \tag{6}$$

The continuous curvelet transform of f, $\Gamma_f(a, b, \theta)$, is defined using the oriented curvelets (6) as function of *scale, location and orientation* by

$$\Gamma_f(a, b, \theta) = \langle f, \psi_{a,b,\theta} \rangle . \tag{7}$$

Due to the parabolic scaling law in equation (5), the *width* of the effective support of these elements is proportional to the *squared length*. Therefore, curvelets $\psi_{a,b,\theta}$ are highly oriented at fine scales, which is the reason for the directional sensitivity of the curvelet transform. Directional sensitivity means that curvelet coefficients $\Gamma_f(a, b, \theta)$ are large at all scales if $\psi_{a,b,\theta}$ matches the location and orientation of a singularity of f and are small otherwise. This implies that it is possible to resolve the location and orientation of a singularity by analyzing the decay properties of curvelet coefficients.

In order to make the above qualitative considerations precise, we need to describe the decay of curvelet coefficients quantitatively. Therefore, we say that Γ_f *decays rapidly* near (x_0, θ_0) if $|\Gamma_f(a, b, \theta)| = \mathcal{O}(a^N)$, $a \to 0$, for all $N > 0$

and uniform over $(b, \theta) \in \mathcal{N}$ for some neighborhood $\mathcal{N}$ of (x_0, θ_0). Now, the wavefront set of a function f can be characterized as follows, [2]: (x_0, θ_0) is in the wavefront set, $(x_0, \theta_0) \in \mathrm{WF}f$ if and only if Γ_f does not decay rapidly near (x_0, θ_0). This implies that a function f is well-approximated by the curvelet coefficients $\Gamma_f(a, b, \theta)$ with $(b, \theta) \in \mathrm{WF}f$.

3. Dimensionality reduction in the curvelet domain

We consider the reconstruction problem (1) in the curvelet domain, i.e., we assume that f is given as a finite linear combination

$$f = \sum_{n=1}^{N} x_n \psi_n \tag{8}$$

with respect to the curvelet dictionary $\Psi = [\psi_1, \ldots, \psi_N]$, where we have used an enumeration $n = n(a, b, \theta)$.

Assume that the acquisition geometry is given by the line parameters (θ_m, s_m). Then, using the decomposition (8), the corresponding noisy measurements are given by

$$\begin{aligned} y_m &= \mathcal{R}f(\theta_m, s_m) + \eta_m \\ &= \sum_{n=1}^{N} x_n \mathcal{R}\psi_n(\theta_m, s_m) + \eta_m, \end{aligned}$$

where η_m denotes the noise. This can be written in the form

$$y = Kx + \eta$$

where K denotes the system matrix which is given by $k_{m,n} = \mathcal{R}\psi_n(\theta_m, s_m)$.

In order to compute a solution of this system of equations, we propose sparse regularization. The curvelet coefficients x_α of the regularized solution f_α are given by

$$x_\alpha = \operatorname*{arg\,min}_{x \in \mathbb{R}^N} \left\{ \|Kx - y\|_2^2 + \alpha \|x\|_1 \right\}. \tag{9}$$

Following the analysis of Section 2.2, we know that only visible singularities of f can be reconstructed. The directions of these singularities are given by the line orientations $\{\theta_m\}_m$. This implies that the directions of singularities of the reconstruction f_α are contained in the set

$$\Theta = \operatorname{conv}\{\theta_m : \ m = 1, \ldots, M\} \subsetneq [0, \pi), \tag{10}$$

where $\operatorname{conv}\{A\}$ denotes the convex hull of a set A. We call call Θ the *set of visible directions.*

Since f_α contains only those singularities with directions in the set Θ, we conclude from Section 2.3 that f_α is well-approximated using curvelet coefficients corresponding to the visible directions Θ – the *visible curvelet coefficients*

x_Θ. The relevant information of f_α is therefore coded by these visible coefficients x_Θ.

We perform the dimensionality reduction in the curvelet domain by computing the visible curvelet coefficients only. To this end, we define the reduced system matrix $\widehat{K}$ as a submatrix of K by $\widehat{k}_{m,n} = \mathcal{R}\psi_n(\theta_m, s_m)$ where $n = n(a, b, \theta)$ with $\theta \in \Theta$. We propose to solve the reduced problem $y = \widehat{K}x + \eta$ by computing a sparse regularization

$$\widehat{x}_\alpha = \arg\min_{x\in\mathbb{R}^n} \left\{ \left\| \widehat{K}x - y \right\|_2^2 + \alpha \left\| x \right\|_1 \right\}. \tag{11}$$

Note that depending on the acquisition geometry the reduced dimension n can be significantly lower than the full dimension N in the curvelet domain. In particular, this is the case when the angular range is very limited. The degree of dimensionality reduction depends on the angular range only. The order of magnitude of n can be estimated via $n \approx \nu \cdot N$, where $\nu = |\Theta| / \pi$ and $|\Theta|$ is the length of the interval Θ.

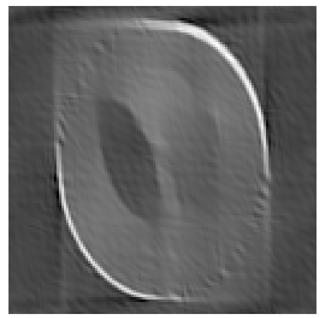
4. Results

We implemented our approach (11) in Matlab using the gradient projection method with step length selection [6]. Limited angle reconstructions f_{full} and f_{red} of a phantom 1(a), (128×128), were computed at an angular range of $90°$ using sparse regularization in the full and reduced curvelet domain. The number of curvelet coefficients used for the computation of f_{full} and f_{red} was $N = 29305$ and $n = 16985$, respectively. The results after 50 iterations are shown in Figure 1(b) and 1(c). The reconstruction error of f_{full} and f_{red} was found to be almost identical at each iteration, Figure 1. For example, the MSE's for f_{full} and f_{red} after 50 iterations are $1.8599 \cdot 10^{-2}$ and $1.8465 \cdot 10^{-2}$, respectively. However, the main advantage of our approach is the significantly lower computational time. In order to compute f_{full} 182 seconds were needed, whereas the computation of f_{red} took only 113 seconds. An additional speedup can be expected using a more sophisticated implementation.

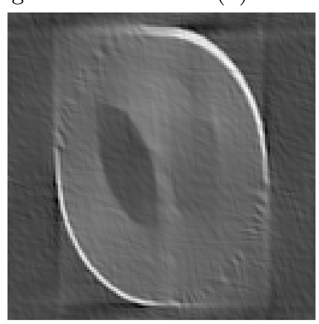
5. Conclusion

We have introduced a new framework for sparse regularization in limited angle x-ray tomography. To this end, we presented a characterization of visible singularities can be predicted. This can be used in order to predict the location of significant (visible) curvelet coefficients. Using this analysis, we performed a dimensionality reduction in the curvelet domain which comes with a significant speedup of corresponding reconstruction algorithms while preserving the image quality. We have demonstrated this fact on a phantom reconstruction.

(a) Original



(b) Full dimensional



(c) Our approach

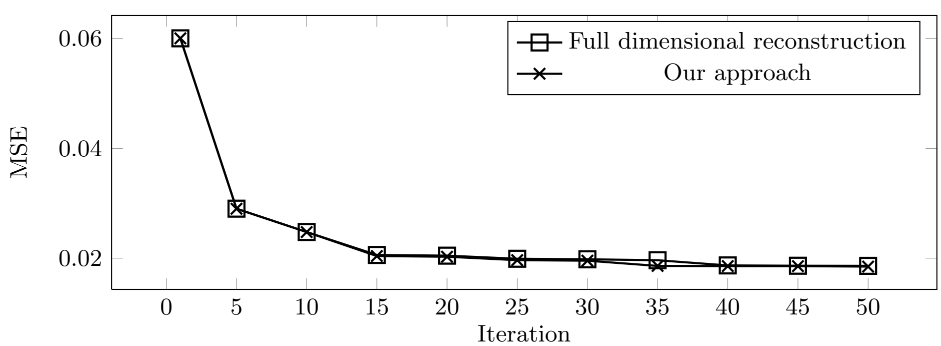


Figure 1: Tomographic reconstruction from limited angular range $1^\circ \leq \theta \leq 90^\circ$, $\Delta\theta = 1^\circ$, using ℓ^1-norm regularization of curvelet coefficients. The image quality obtained after dimensionality reduction (c) is as good as in the case of full dimensional reconstruction (b). However, the computational time for 50 iterations was significantly lower using our approach.

References

[1] Candès, E. J., 2008. The restricted isometry property and its implications for compressed sensing. C. R. Math. Acad. Sci. Paris 346 (9-10), 589–592.

[2] Candès, E. J., Donoho, D. L., 2005. Continuous curvelet transform. I. Resolution of the wavefront set. Appl. Comput. Harmon. Anal. 19, 162–197.

[3] Daubechies, I., et al., 2004. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Comm. Pure Appl. Math. 57 (11), 1413–1457.

[4] Engl, H. W., Hanke, M., Neubauer, A., 1996. Regularization of inverse problems. Vol. 375 of Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht.

[5] Kolehmainen, V., et al., 2003. Statistical inversion for medical x-ray tomography with view radiographs: II. Application to dental radiology. Phys. Med. Biol. 48, 1465–1490.

[6] Loris, I., et al., 2009. Accelerating gradient projection methods for ℓ_1-constrained signal recovery by steplength selection rules. Appl. Comput. Harmon. Anal. 27, 247–254, accepted.
URL `http://arxiv.org/abs/0902.4424`

[7] Natterer, F., 1986. The mathematics of computerized tomography. B. G. Teubner, Stuttgart.

[8] Quinto, E. T., 1993. Singularities of the X-ray transform and limited data tomography in $\mathbb{R}^2$ and $\mathbb{R}^3$. SIAM J. Math. Anal. 24 (5), 1215–1225.

[9] Quinto, E. T., 2006. An introduction to X-ray tomography and Radon transforms. In: The Radon transform, inverse problems, and tomography. Vol. 63 of Proc. Sympos. Appl. Math. Amer. Math. Soc., Providence, RI, pp. 1–23.

[10] Rantala, M., et al., February 2006. Wavelet-based reconstruction for limited angle x-ray tomography. IEEE Transactions on Medical Imaging 25 (2), 210–217.