Sparse X-ray tomography

Samuli Siltanen
(Talk delivered by Matti Lassas)

Department of Mathematics and Statistics
University of Helsinki, Finland
samuli.siltanen@helsinki.fi
www.siltanen-research.net

Applied Inverse Problems Conference
Hangzhou, China, May 30, 2017
Finnish Centre of Excellence in Inverse Problems Research

http://wiki.helsinki.fi/display/inverse/Home
Outline

X-ray imaging

Tomographic imaging with sparse data

Hospital case study: diagnosing osteoarthritis

Limited angle tomography

Industrial case study: low-dose 3D dental X-ray imaging

Dynamic tomography
Godfrey Hounsfield and Allan McLeod Cormack developed X-ray tomography.

Hounsfield (top) and Cormack received Nobel prizes in 1979.
Contrast-enhanced CT of abdomen, showing liver metastases
Rotating around the object allows us to form the so-called *sinogram*

https://www.youtube.com/watch?v=5Vyc1TzmNI8
Reconstruction of a function from its line integrals was first invented by Johann Radon in 1917.

\[ f(P) = -\frac{1}{\pi} \int_0^\infty \frac{d\overline{F_P}(q)}{q} \]
Outline

X-ray imaging

Tomographic imaging with sparse data

Hospital case study: diagnosing osteoarthritis

Limited angle tomography

Industrial case study: low-dose 3D dental X-ray imaging

Dynamic tomography
The continuous tomographic model needs to be approximated using a discrete model.

**Continuous model:**

In this schematic setup we have 5 projection directions and a 10-pixel detector. Therefore, the number of data points is 50.

**Discrete model:**
The resolution of the discrete model can be freely chosen according to computational resources

Continuous model:

In this schematic setup we have 5 projection directions and a 10-pixel detector. Therefore, the number of data points is 50.

The number of degrees of freedom in the three discrete models below are 16, 64 and 256, respectively.

Discrete models:
Discretize the unknown by dividing it into pixels

Target (unknown)  

32×32 pixel grid
System matrix $A$, given by the grid and X-rays

735×1024 system matrix $A$, only nonzero elements shown

32×32 pixel grid
System matrix $A$, given by the grid and X-rays

$735 \times 1024$ system matrix $A$, only nonzero elements shown

$32 \times 32$ pixel grid
System matrix $A$, given by the grid and X-rays

735×1024 system matrix $A$, only nonzero elements shown

32×32 pixel grid
System matrix $A$, given by the grid and X-rays

735×1024 system matrix $A$, only nonzero elements shown

32×32 pixel grid
System matrix $A$, given by the grid and X-rays

735×1024 system matrix $A$, only nonzero elements shown

32×32 pixel grid
System matrix $A$, given by the grid and X-rays

735\times1024$ system matrix $A$, only nonzero elements shown

32\times32$ pixel grid
System matrix $A$, given by the grid and X-rays

735×1024 system matrix $A$, only nonzero elements shown

32×32 pixel grid
System matrix $A$, given by the grid and X-rays

735×1024 system matrix $A$, only nonzero elements shown

32×32 pixel grid
What can we expect to see from sparse data?

Object \[ \xrightarrow{A} \] Sinogram

**Theorem 4.2.** A finite set of radiographs tells nothing at all.

For some reason this theorem provokes merriment. It is so plainly one of those mathematical ideals untainted by any possibility of practical application.

[Cormack 1963], [Smith, Solmon & Wagner 1977, Theorem 4.2]
Singular value decomposition $A = U^T D V$

735×1024 system matrix $A$, only nonzero elements shown

Singular values of $A$ (diagonal of $D$)
Illustration of the ill-posedness of sparse tomography

\[ A \rightarrow \text{Difference 0.00992} \]
Illustration of the ill-posedness of sparse tomography
We make a real-world phantom out of cheese

Digital target

Real-world target (cheese phantom)
Constructing and imaging the cheese phantom

https://www.youtube.com/watch?v=l3Astgo_3x0&feature=share

Video: thanks to Tatiana Bubba, Markus Juvonen, Maximilian März, Alexander Meaney and Zenith Purisha
Filtered backprojection

Original phantom

Reconstruction by FBP
Constrained Tikhonov regularization

\[
\arg \min_{f \in \mathbb{R}^n_+} \left\{ \|Af - m\|_2^2 + \alpha \|f\|_2^2 \right\}
\]
Constrained total variation (TV) regularization

\[
\arg \min_{f \in \mathbb{R}^n_+} \left\{ \|Af - m\|_2^2 + \alpha \|\nabla f\|_1 \right\}
\]
TV tomography: \( \arg \min_{f \in \mathbb{R}^n} \{ \| Af - m \|_2^2 + \alpha \| \nabla f \|_1 \} \)

1992 Rudin, Osher & Fatemi: denoise images by taking \( A = I \)
1998 Delaney & Bresler
2001 Persson, Bone & Elmqvist
2003 Kolehmainen, S, Järvenpää, Kaipio, Koistinen, Lassas, Pirttilä & Somersalo (first TV work with measured X-ray data)
2006 Kolehmainen, Vanne, S, Järvenpää, Kaipio, Lassas & Kalke
2006 Sidky, Kao & Pan
2008 Liao & Sapiro
2008 Sidky & Pan
2008 Herman & Davidi
2009 Tang, Nett & Chen
2009 Duan, Zhang, Xing, Chen & Cheng
2010 Bian, Han, Sidky, Cao, Lu, Zhou & Pan
2011 Jensen, Jørgensen, Hansen & Jensen
2011 Tian, Jia, Yuan, Pan & Jiang
2012–present: hundreds of articles indicated by Google Scholar
Bayesian inversion: discretization-invariance and wavelet-based Besov space priors

TV prior: \( c \exp(-\alpha \| \nabla f \|_{L^1}) \).

Besov prior: \( c \exp(-\alpha \| f \|_{B^{1}_{1,1}}) = c \exp(-\alpha \sum_{j,k,\ell} |\langle f, \psi_{j,k,\ell} \rangle|) \).

2004 Lassas and S (TV prior not discretization-invariant)
2009 Lassas, Saksman & S (Besov priors are discretization-invariant)
2010 Stuart
2011 Kolehmainen, Lassas, Niinimäki & S
2012 Dashti, Harris & Stuart
2013 Hämäläinen et al.
2014 Burger & Lucka
2015 Bui-Thanh & Ghattas
2016 Sullivan
Definition of discrete and continuous regularization functionals

Let $D$ be the square $[0, 1]^2 \subset \mathbb{R}^2$. Use anisotropic $BV(D)$ norm

$$\|u\|_{BV} = \|u\|_{L^1} + V(u) = \|u\|_{L^1} + \int_D \left( \left| \frac{\partial u(x)}{\partial x_1} \right| + \left| \frac{\partial u(x)}{\partial x_2} \right| \right) dx.$$ 

Define $S_\infty : BV(D) \to \mathbb{R}$ and $S_j : BV(D) \to \mathbb{R} \cup \{\infty\}$ by

$$S_\infty(u) = \|Au - m\|^2_{L^2(\Omega)} + \alpha_1 \|u\|_{L^1(D)} + \alpha V(u)$$

with positive regularization parameters $\alpha_1 > 0$ and $\alpha > 0$, and

$$S_j(u) = \begin{cases} S_\infty(u), & \text{for } u \in \text{Range}(T_j), \\ \infty, & \text{for } u \not\in \text{Range}(T_j). \end{cases}$$

Linear operator $T_j$ projects to functions that are piecewise constant on a regular $2^j \times 2^j$ square pixel grid.
Our main theorem ensures the convergence of regularized solutions as resolution grows

- There exists a minimizer $\tilde{u}_j \in \text{arg min}(S_j)$ for all $j = 1, 2, 3, \ldots$
- There exists a minimizer $\tilde{u}_\infty \in \text{arg min}(S_\infty)$.
- Any sequence $\tilde{u}_j \in \text{arg min}(S_j)$ of minimizers has a subsequence $\tilde{u}_{j(\ell)}$ that converges weakly in $BV(D)$ to some limit $w \in BV(D)$. Furthermore, $\lim_{\ell \to \infty} V(\tilde{u}_{j(\ell)}) = V(w)$.
- The limit $w$ is a minimizer: $w \in \text{arg min}(S_\infty)$.

[Niinimäki, Lassas, Hämäläinen, Kallonen, Kolehmainen, Niemi & S 2016]
There are some related results in the literature

1992 Vainikko: *On the discretization and regularization of ill-posed problems with noncompact operators*

We use geometric arguments similar to those here:


These works consider TV functionals and $\Gamma$-convergence when discretization is refined, but without a measurement operator:

2009 Chambolle, Giacomini & Lussardi
2012 Gennip & Bertozzi
2013 Bellettini, Chambolle & Goldman
2013 Trillos & Slepčev

This paper achieves a result analogous with ours, using wavelet frames in the finite-dimensional functionals:

2012 Cai, Dong, Osher & Shen
Illustration of the Haar wavelet transform
Daubechies, Defrise and de Mol introduced a revolutionary inversion method in 2004.

Consider the sparsity-promoting variational regularization

$$\arg \min_{f \in \mathbb{R}^n} \left\{ \| Af - m \|_2^2 + \mu \| Wf \|_1 \right\},$$

where $W$ is an orthonormal wavelet transform. The minimizer can be computed using the iteration

$$f_{j+1} = W^{-1} S_\mu W \left( f_j + A^T (m - Af_j) \right),$$

where the soft-thresholding operation

$$S_\mu(x) = \begin{cases} 
  x + \frac{\mu}{2} & \text{if } x \leq -\frac{\mu}{2}, \\
  0 & \text{if } |x| < \frac{\mu}{2}, \\
  x - \frac{\mu}{2} & \text{if } x \geq \frac{\mu}{2},
\end{cases}$$

is applied to each wavelet coefficient separately.
How to choose the thresholding parameter $\mu$?
Here $\mu$ is too small.
How to choose the thresholding parameter $\mu$?
Here $\mu$ is too large.
Automatic parameter choice using controlled wavelet-domain sparsity (CWDS)

Assume given the \emph{a priori} sparsity level \(0 \leq C_{pr} \leq 1\). Denote by \(C_j\) the sparsity of the \(j\)th iterate \(f_j \in \mathbb{R}^n\):

\[
C_j = \frac{\text{(number of nonzero elements in } \mathcal{W}f_j \in \mathbb{R}^n)}{n}.
\]

The CWDS iteration is based on proportional-integral-derivative (PID) controllers:

\[
\begin{align*}
e_j & = C_j - C_{pr} \\
\mu_{j+1} & = \max\{0, \mu_j + \beta e_j\}, \\
f_{j+1} & = \mathcal{W}^{-1} S_{\mu_{j+1}} \mathcal{W} \left( f_j + A^T(m - Af_j) \right), \\
f_{j+1} & = \max\{0, f_{j+1}\}.
\end{align*}
\]

CWDS choice of the thresholding parameter $\mu$
CWDS choice of the thresholding parameter $\mu$
The shearlet transform gives multi-resolution and orientation-aware building blocks for image data.

Wavelets can only detect the singular support of a function. Shearlets provide more information as they can resolve the wavefront set.

Left: Schematic diagram of the frequency plane tiling of several elements of a 2D shearlet system, for different values of dilation and shearing parameters.
Loris and Verhoeven introduced an algorithm applicable to shearlet-sparsifying inversion

[Loris & Verhoeven 2011] The minimization of

$$
\arg\min_{f \in \mathbb{R}^n} \{ \|Af - m\|_2^2 + \mu\|Tf\|_1 \}
$$

can be computed using this iteration:

$$
\begin{align*}
    g_{j+1} &= f_j + A^T(m - Af_j), \\
    w_{j+1} &= P_{\mu} \left( w_j + T(g_{j+1} - T^Tw_j) \right), \\
    f_{j+1} &= g_{j+1} - T^Tw_{j+1},
\end{align*}
$$

where $P_{\mu} = I - S_{\mu}$.
Outline

X-ray imaging

Tomographic imaging with sparse data

Hospital case study: diagnosing osteoarthritis

Limited angle tomography

Industrial case study: low-dose 3D dental X-ray imaging

Dynamic tomography
This is a joint work with

**Tatiana Bubba**, University of Helsinki, Finland

**Sakari Karhula**, Oulu University Hospital, Finland

**Juuso Ketola**, Oulu University Hospital, Finland

**Maximilian März**, TU Berlin

**Miika T. Nieminen**, University of Oulu, Finland

**Zenith Purisha**, University of Helsinki, Finland

**Juho Rimpeläinen**, University of Helsinki, Finland

**Simo Saarakkala**, Oulu University Hospital, Finland
We consider small specimens of human bone imaged using microtomography.

Slice of 3D reconstruction by FDK based on **596 angles**

Three-dimensional structure
We pick out a smaller region of interest for osteoarthritis analysis.

Slice of 3D reconstruction by FDK based on **596 angles**.

Slice of 3D region of interest after binary thresholding.
We use two numerical quality measures applied to segmented three-dimensional bone structure.

Trabecular thickness

Trabecular separation

[Bouxsein, Boyd, Christiansen, Guldberg, Jepsen, & Müller 2010]
The goal is to reduce measurement time by recording fewer radiographs.

3D FDK reconstruction based on 40 angles

3D shearlet-sparsity reconstruction based on 40 angles
Bone quality parameters from ground truth

Healthy bone

- Thickness: 8.1
- Separation: 15.3

Osteoarthritic bone

- Thickness: 5.9
- Separation: 21

[Purisha, Karhula, Rimpeläinen, Nieminen, Saarakkala & S, submitted]
Results from FDK reconstructions

<table>
<thead>
<tr>
<th>Projections</th>
<th>Thickness</th>
<th>Separation</th>
</tr>
</thead>
<tbody>
<tr>
<td>596</td>
<td>8.1</td>
<td>15.3</td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[Purisha, Karhula, Rimpeläinen, Nieminen, Saarakkala & S, submitted]
Results from 3D shearlet-sparsity reconstructions

<table>
<thead>
<tr>
<th>Projections</th>
<th>Thickness</th>
<th>Separation</th>
</tr>
</thead>
<tbody>
<tr>
<td>596</td>
<td>8.1</td>
<td>15.3</td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>596</td>
<td>5.9</td>
<td>21</td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[Purisha, Karhula, Rimpeläinen, Nieminen, Saarakkala & S, submitted]
Outline

X-ray imaging

Tomographic imaging with sparse data

Hospital case study: diagnosing osteoarthritis

Limited angle tomography

Industrial case study: low-dose 3D dental X-ray imaging

Dynamic tomography
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram

0°  90°  180°
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Construction of limited-angle sinogram
Singular value decomposition $A = U^T D V$

$735 \times 1024$ system matrix $A$, only nonzero elements shown

Singular values of $A$ (diagonal of $D$)
Limited data gives only part of the wavefront set

Stable part of wavefront set

Unstable part of wavefront set

See [Greenleaf & Uhlmann 1989], [Quinto 1993], and [Frikel & Quinto 2013]
Filtered backprojection

Stable part of WF set

Reconstruction by FBP
Constrained total variation (TV) regularization

\[
\arg \min_{f \in \mathbb{R}^n_+} \left\{ \|Af - m\|_2^2 + \alpha \|\nabla f\|_1 \right\}
\]

Stable part of WF set

TV regularized reconstruction
Outline

X-ray imaging

Tomographic imaging with sparse data

Hospital case study: diagnosing osteoarthritis

Limited angle tomography

Industrial case study: low-dose 3D dental X-ray imaging

Dynamic tomography
The VT device was developed in 2001–2012 by

Nuutti Hyvönen
Seppo Järvenpää
Jari Kaipio
Martti Kalke
Petri Koistinen
Ville Kolehmainen
Matti Lassas
Jan Moberg
Kati Niinimäki
Juha Pirttilä
Maaria Rantala
Eero Saksman
Henri Setälä
Erkki Somersalo
Antti Vanne
Simopekka Vänskä
Richard L. Webber
Application: dental implant planning, where a missing tooth is replaced with an implant.
This is the classical imaging procedure of the panoramic X-ray device

https://www.youtube.com/watch?v=QFTXegPxC4U
The resulting image shows a sharp layer positioned inside the dental arc.
Nowadays, a digital panoramic imaging device is standard equipment at dental clinics.

A panoramic dental image offers a general overview showing all teeth and other structures simultaneously.

Panoramic images are not suitable for dental implant planning because of unavoidable geometric distortion.
We reprogram the panoramic X-ray device so that it collects projection data by scanning.

Number of projection images: 11

Angle of view: 40 degrees

Image size: $1000 \times 1000$ pixels

The unknown vector $f$ has 7,000,000 elements.
CBCT imaging gives 100 times more radiation than VT reconstruction


Outline

X-ray imaging

Tomographic imaging with sparse data

Hospital case study: diagnosing osteoarthritis

Limited angle tomography

Industrial case study: low-dose 3D dental X-ray imaging

Dynamic tomography
There exists at least one minimizer for our generalized level set functional

**Theorem:** Let \( \mathcal{A} \) be an operator modeling 2D Radon transforms measured at several times. If \( \alpha > 0 \) satisfies an upper bound involving the signal-to-noise ratio, then the nonlinear functional

\[
F_n(\phi) := \frac{1}{2} \| \mathcal{A} g(\phi) - m \|_2^2 + \frac{\alpha}{2} \sum_{1 \leq |\beta| \leq n} \| D^\beta \phi \|_2^2
\]

has a **global minimizer**. The minimizer is unique for \( n = 1 \).

We model the X-ray attenuation function as \( g(\Phi(x, y)) \), where

\[
g(\tau) = \begin{cases} 
\tau, & \text{if } \tau \geq 0 \\
0, & \text{if } \tau < 0.
\end{cases}
\]

[Niemi, Lassas, Kallonen, Harhanen, Hämäläinen and S 2015]
The optical flow constraint

Assuming a constant image intensity of $f(x, t)$ along a trajectory $x(t)$ with a vector field $\frac{dx}{dt} = \mathbf{v}(x, t)$, we get by using the chain-rule

$$0 = \frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{i=1}^{2} \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} = \partial_t f + \nabla f \cdot \mathbf{v}.$$

In short the optical flow constraint is given by

$$\partial_t f + \nabla f \cdot \mathbf{v} = 0.$$
Optical flow functional

We have one equation for the flow field \( \mathbf{v} = (v^1, v^2)^T \) and hence the problem is underdetermined.

To obtain an approximation we are minimizing

\[
J_{\text{flow}}(f, \mathbf{v}) = \| \partial_t f + \nabla f \cdot \mathbf{v} \|_1 + \beta \sum_{i=1}^{2} \| v^i \|_{TV}.
\]

The \( L^1 \)-norm is more robust with respect to outliers for this model. [Zach, Pock, and Bischof, 2007]
An alternating algorithm for optical flow regularization for dynamic tomography

The joint model can be transformed into an iterative two-step method. Given $v^k$ we compute

$$f^{k+1} = \arg\min_f \int_0^T \frac{1}{p} \|Af - m\|_p^p + \alpha \|f\|_{TV} + \gamma \left\| \partial_t f + \nabla f \cdot v^k \right\|_1 dt$$

$$v^{k+1} = \arg\min_v \int_0^T \left\| \partial_t f^{k+1} + \nabla f^{k+1} \cdot v \right\|_1 + \frac{\beta}{\gamma} \sum_{j=1}^2 \|v_j\|_{TV} dt.$$ 

These subproblems are linear and convex.

[Burger, Dirks, Frerking, Hauptmann, Helin & S, submitted]
Demonstration of dynamic tomography methods

https://www.youtube.com/watch?v=JTdVAQTFKxI&feature=youtu.be
Thank you for your attention!

www.siltanen-research.net
All Matlab codes freely available at this site!

Part I: Linear Inverse Problems
1 Introduction
2 Naïve reconstructions and inverse crimes
3 Ill-Posedness in Inverse Problems
4 Truncated singular value decomposition
5 Tikhonov regularization
6 Total variation regularization
7 Besov space regularization using wavelets
8 Discretization-invariance
9 Practical X-ray tomography with limited data
10 Projects

Part II: Nonlinear Inverse Problems
11 Nonlinear inversion
12 Electrical impedance tomography
13 Simulation of noisy EIT data
14 Complex geometrical optics solutions
15 A regularized D-bar method for direct EIT
16 Other direct solution methods for EIT
17 Projects
Another great resource is Per Christian Hansen’s 3D tomography toolbox TVreg

**TVreg**: Software for 3D Total Variation Regularization (for Matlab Version 7.5 or later), developed by Tobias Lindstrøm Jensen, Jakob Heide Jørgensen, Per Christian Hansen, and Søren Holdt Jensen.

Website: http://www2.imm.dtu.dk/ pcha/TVReg/
These books are recommended for learning the mathematics of practical X-ray tomography

1983 Deans: The Radon Transform and Some of Its Applications
1986 Natterer: The mathematics of computerized tomography
1988 Kak & Slaney: Principles of computerized tomographic imaging
1996 Engl, Hanke & Neubauer: Regularization of inverse problems
1998 Hansen: Rank-deficient and discrete ill-posed problems
2001 Natterer & Wübbeling: Mathematical Methods in Image Reconstruction
2008 Buzug: Computed Tomography: From Photon Statistics to Modern Cone-Beam CT
2008 Epstein: Introduction to the mathematics of medical imaging
2010 Hansen: Discrete inverse problems
2012 Mueller & S: Linear and Nonlinear Inverse Problems with Practical Applications
2014 Kuchment: The Radon Transform and Medical Imaging
Assumptions on the linear forward map $\mathcal{A}$

Assume either (A) or (B) about the linear operator $\mathcal{A}$:

(A) $\mathcal{A} : L^2(D) \to L^2(\Omega)$ is compact and $\mathcal{A} : L^1(D) \to D'(\Omega)$ is continuous with some open and bounded set $\Omega \subset \mathbb{R}^2$. This covers the case of classical Radon transform with image domain $D$ and sinogram domain $\Omega$. We denote the set of distributions by $D'(\Omega)$.

(B) $\mathcal{A} : L^1(D) \to \mathbb{R}^M$ is bounded. This covers the practically important discrete pencil beam model of tomographic measurements.