Tomographic Reconstruction with Adaptive Sparsifying Transforms

Luke Pfister & Yoram Bresler
Computed Tomography

- Linear Measurements: $y = Ax$
- Reconstruction: Filtered Back Projection (FBP)
Computed Tomography

- Linear Measurements: \( y = Ax \)
- Reconstruction: Filtered Back Projection (FBP)
Low Dose Computed Tomography
Model-Based Image Reconstruction

- Three Ingredients
  - System Model
  - Noise Model
  - Signal Model

- Tie together into an optimization problem
Penalized Weighted Least-Squares

$$\min_{x} \frac{1}{2} \| y - Ax \|_W^2 + \lambda J(x)$$
Penalized Weighted Least-Squares

\[
\min_x \frac{1}{2} \| y - Ax \|_W^2 + \lambda J(x)
\]

**System Model**

- \( y \in \mathbb{R}^M \): Log of CT data
- \( A \in \mathbb{R}^{M \times N} \): System matrix
- \( x \in \mathbb{R}^N \): Image estimate
Penalized Weighted Least-Squares

\[
\min_x \frac{1}{2} \|y - Ax\|_W^2 + \lambda J(x)
\]

- Noise Model
  - \( W = \text{diag}\{w_i\} \)
  - \( w_i \) are statistical weights
  - \( W \) is very poorly conditioned
Penalized Weighted Least-Squares

\[
\min_x \frac{1}{2} \| y - Ax \|_W^2 + \lambda J(x)
\]

- Signal Model
  - Regularizer \( J(x) : \mathbb{R}^N \to \mathbb{R} \)
Our Contribution

Propose fast, data-driven regularization with adaptive sparsifying transforms
Signal Models
Signal Models

- Better model $\implies$ better reconstruction
- Data-adaptive sparse representations: sparse signal models adapted for a **particular signal instance**
  - Usually patch based
Patch-based Signal Models
Sparse Signal Models

- Synthesis sparsity
- Transform sparsity
Synthesis Sparsity

- \( x = Da \), \( a \) is sparse

\[
\begin{array}{c}
\begin{array}{c}
x \\
D \\
a
\end{array}
\end{array}
\]

- **Dictionary Learning:** Given \( \{x_j\}_{j=1}^P \), find \( D \) and \( \{a_j\}_{j=1}^P \)
  - Applied to low-dose and limited-angle CT
  - Scales poorly with data size
Transform Sparsity

- $\Phi x = z + e$, $z$ is sparse.
- $e$ captures deviation from sparsity in transform domain

**Transform Learning**: Given $\{x_j\}_{j=1}^P$, find $\Phi$ and $\{z_j\}_{j=1}^P$

- Scales more gracefully with data size
Problem Formulation
Regularization with sparsifying transforms

\[ J(x) = \min_{z, \Phi} \sum_j \frac{\lambda}{2} \| \Phi E_j x - z_j \|_2^2 + \gamma \| z_j \|_0 + \alpha (\| \Phi \|_F^2 - \log \det \Phi) \]
Regularization with sparsifying transforms

\[ J(x) = \min_{z, \Phi} \sum_j \frac{\lambda}{2} \| \Phi E_j x - z_j \|_2^2 + \gamma \| z_j \|_0 + \alpha (\| \Phi \|_F^2 - \log \det \Phi) \]
Regularization with sparsifying transforms

\[ J(x) = \min_{z, \Phi} \sum \frac{\lambda}{2} \| \Phi E_j x - z_j \|_2^2 + \gamma \| z_j \|_0 + \alpha(\| \Phi \|_F^2 - \log \det \Phi) \]
Regularization with sparsifying transforms

\[ J(x) = \min_{z, \Phi} \sum_{j} \frac{\lambda}{2} \| \Phi E_j x - z_j \|_2^2 + \gamma \| z_j \|_0 + \alpha(\| \Phi \|_F^2 - \log \det \Phi) \]
Regularization with sparsifying transforms

\[ J(x) = \min_{z, \Phi} \sum_j \frac{\lambda}{2} \| \Phi E_j x - z_j \|_2^2 + \gamma \| z_j \|_0 + \alpha (\| \Phi \|_F^2 - \log \det \Phi) \]
Reconstruction Problem

\[
\min_{x, \Phi, z_j} \frac{1}{2} \| y - Ax \|_W^2 + \lambda \sum_j \frac{1}{2} \| \Phi E_j x - z_j \|_2^2 + \lambda \gamma \| z_j \|_0 \\
+ \lambda \alpha (\| \Phi \|_F^2 - \log \det \Phi))
\]
\[
\min_x \frac{1}{2} \| y - Ax \|_W^2 + \frac{\lambda}{2} \sum_j \| \Phi E_j x - z_j \|_2^2
\]

**Image Update**

\[
\min_{\Phi} \sum_j \| \Phi E_j x - z_j \|_2^2 + F(\Phi)
\]

**Transform Update**

\[
\min_{z_j} \sum_j \| \Phi E_j x - z_j \|_2^2 + \gamma \| z_j \|_0
\]

**Sparse Code Update**
Regularizer Update

- $\Phi$ update

$$
\Phi^{k+1} = \arg\min_{\Phi} \sum_j \frac{1}{2} \|\Phi E_j x - z_j\|_2^2 + \alpha \left( \|\Phi\|_F^2 - \log \det \Phi \right)
$$

- Closed form solution! [Ravishankar, 2012]

- Requires three matrix products of size $p \times N$ by $N \times p$, and one SVD of size $p \times p$
Regularizer Update

- $z_j$ update

$$z_j^{k+1} = \arg \min_{z_j} \frac{1}{2} \| \Phi E_j x - z_j \|_2^2 + \gamma \| z_j \|_0$$

- Closed-form solution using hard thresholding: $z_j^{k+1} = T_\gamma (\Phi E_j x)$

$$T_\gamma (a) = \begin{cases} 0, & |a| \leq \sqrt{\gamma} \\ a, & \text{else} \end{cases}$$
Image Update

\[
\min_x \frac{1}{2} \|y - Ax\|_W^2 + \sum_j \frac{\lambda}{2} \|\Phi E_j x - z_j\|_2^2
\]

- Weighted least-squares problem in \(x\)

\[
H = A^T W A + \lambda \sum_j E_j^T \Phi^T \Phi E_j
\]
Solution using ADMM [Ramani, 2012]

Big Idea: Use variable splitting to untangle $A$ and $W$

$$\min_x \frac{1}{2} \| y - v \|^2_W + \sum_j \frac{\lambda}{2} \| \Phi E_j x - z_j \|^2_2$$

subject to $v = Ax$
Solution using ADMM [Ramani, 2012]

- **Augmented Lagrangian**

\[
\mathcal{L}(x, v, \eta) = \frac{1}{2} \| y - v \|_W^2 + \sum_j \frac{\lambda}{2} \| \Phi E_j x - z_j \|_2^2 + \frac{\mu}{2} \| v - A x - \eta \|_2^2 - \frac{\mu}{2} \| \eta \|_2^2
\]

- **Alternate between**
  - Minimization over \( x \)
  - Minimization over \( v \)
  - Maximization over \( \eta \)
x-update

- Solve:
  \[
  \left( \mu A^T A + \sum_j E_j^T \Phi^T \Phi E_j \right) x^{k+1} = \mu A^T (v^k - \eta^k) + \sum_j E_j^T \Phi^T z_j
  \]

- Linear **unweighted** least-squares in \( x \)

- Hessian \( H = \mu A^T A + \sum_j E_j^T \Phi^T \Phi E_j \) is approximately shift-invariant

- Solve using Preconditioned Conjugate-Gradient (PCG) with circulant preconditioner
\( v^{k+1} = (W + \mu I)^{-1}(Wy + \mu(Ax^{k+1} + \eta^k)) \)
$\eta^{k+1} = \eta^k - \nu^{k+1} + Ax^{k+1}$
Overall Algorithm (AST-CT)

1: repeat
2:   repeat
3:     Update $\Phi$
4:     $z_j^k \leftarrow T_{\gamma} \Phi E_j x \forall j$
5:   until Halting condition
6: $i \leftarrow 0$, $u^0 \leftarrow Ax^k$, $v^0 \leftarrow \vec{0}$
7: repeat
8:   Use PCG to find approximate solution
9:   of $H \tilde{x}^{i+1} = \mu A^T (u^i - \eta^i) + \lambda \sum_j E_j^T \Phi^T z_j^i$
10: $u^{i+1} \leftarrow (W + \mu I)^{-1} (Wy + \mu (A \tilde{x}^{i+1} + v^i))$
11: $\eta^{i+1} \leftarrow \eta^i - (u^{i+1} - A \tilde{x}^{i+1})$
12: $i \leftarrow i + 1$
13: until Halting condition
14: $x^{k+1} \leftarrow \tilde{x}^{i+1}$
15: until Halting Condition
Experiments
Experiments

- Low-dose data synthesize from clinical image
- Total-variation (TV)
  - $J(x) = \|x\|_{TV}$
  - Apply variable splitting to data fidelity and regularizer
- Dictionary learning (DL):
  - $J(x) = \min_{D,a} \sum_j \| E_j x - Da_j \|_2^2 + \gamma \|a_j\|_0$
  - Solve with orthogonal matching pursuit and K-SVD
## Experiments

<table>
<thead>
<tr>
<th></th>
<th>$D/\Phi$ Update</th>
<th>$a/z$ Update</th>
<th>Image Update</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBP</td>
<td>0</td>
<td>0</td>
<td>2.3</td>
<td>2.3</td>
</tr>
<tr>
<td>TV-CT</td>
<td>0</td>
<td>0</td>
<td>91.3</td>
<td>91.3</td>
</tr>
<tr>
<td>DL-CT</td>
<td>87.5</td>
<td>60.3</td>
<td>85.4</td>
<td>233.3</td>
</tr>
<tr>
<td>AST-CT</td>
<td>4.4</td>
<td>0.2</td>
<td>88.4</td>
<td>93.0</td>
</tr>
</tbody>
</table>
Conclusions
Conclusions

- Proposed the use of adaptive sparsifying transform regularization for low-dose CT reconstruction
- Performs as well as synthesis dictionary learning regularization at the speed of TV regularization
Thanks!