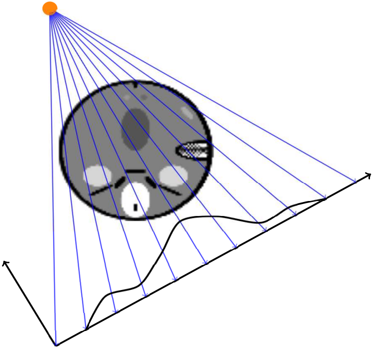
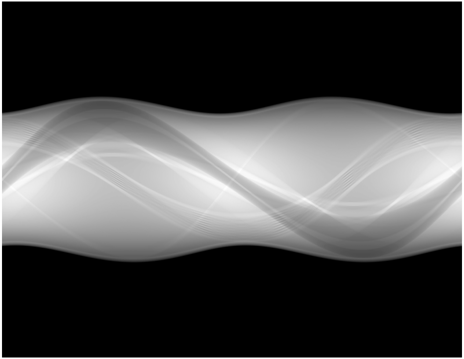
Adaptive Sparsifying Transforms for Iterative Tomographic Reconstruction

Luke Pfister & Yoram Bresler

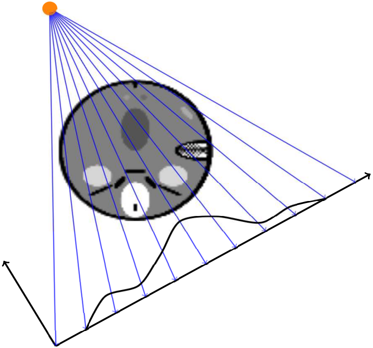
Computed Tomography

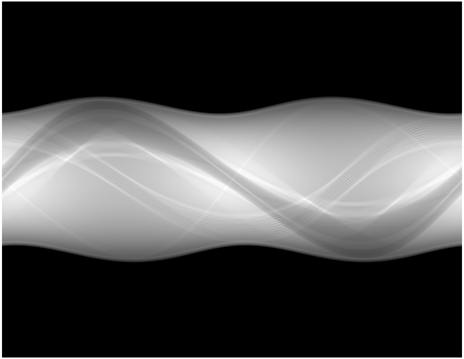




- Linear Measurements: $y = Ax$
- Reconstruction: Filtered Back Projection

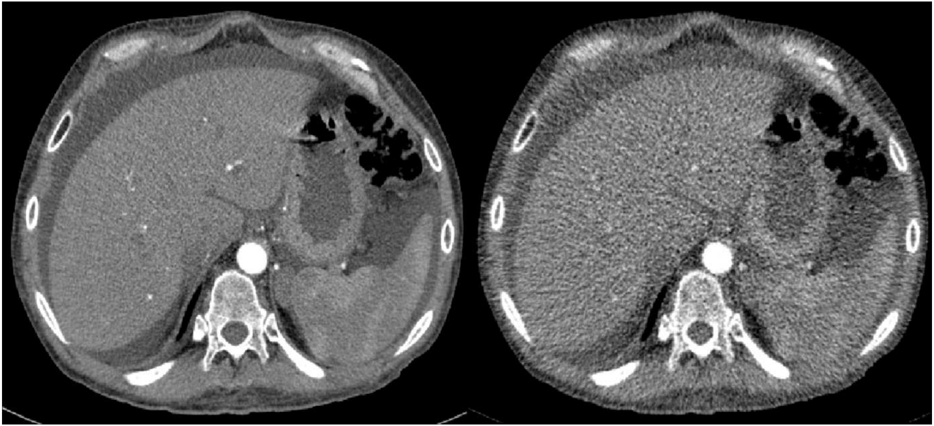
Computed Tomography





- Linear Measurements: $y = Ax$
- Reconstruction: Filtered Back Projection

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
  - $W = \operatorname{diag}\{w_i\}$
  - 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 Contributions

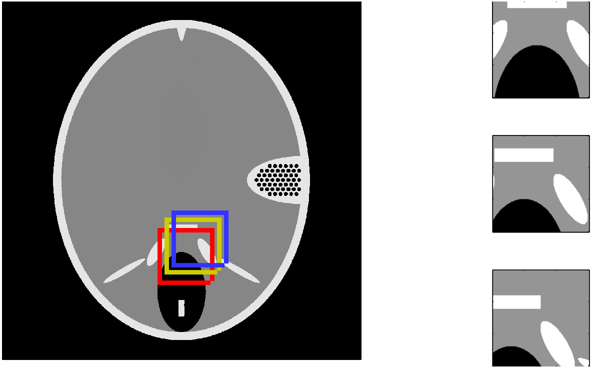
- Propose fast, data-driven regularization with adaptive sparsifying transforms
- Accelerate reconstruction using linearized ADMM

Signal Models

Signal Models

- Better model $\implies$ better reconstruction
- Data-adative sparse representations: sparse signal models adapted for a **particular signal instance**
  - Usually patch based

Patch-based Signal Models

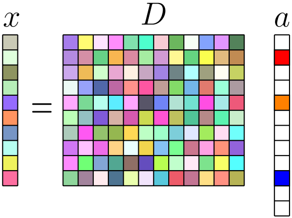


Sparse Signal Models

- Synthesis sparsity
- Transform sparsity

Synthesis Sparsity

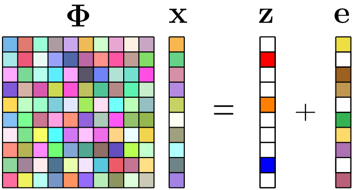
- $x = Da,\ a$ is sparse



- **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

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



- **Transform Learning**: Given $\{x_j\}_{j=1}^{P}$, find Φ 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

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

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

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

Reconstruction Problem

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

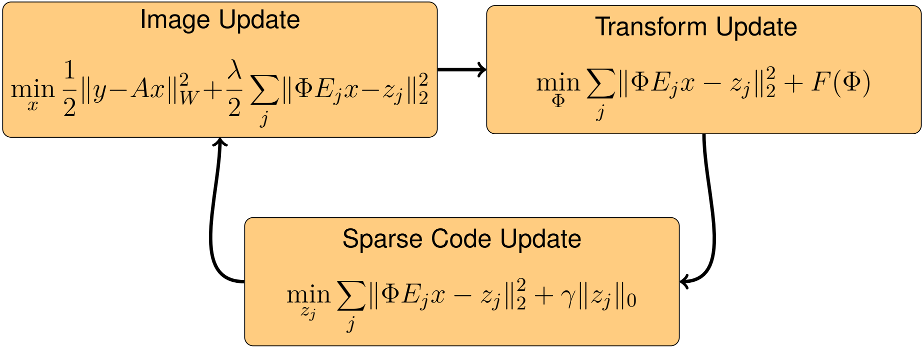


Image Update
$\min_x \frac{1}{2}\|y - Ax\|_W^2 + \frac{\lambda}{2}\sum_j \|\Phi E_j x - z_j\|_2^2$
Transform Update
$\min_\Phi \sum_j \|\Phi E_j x - z_j\|_2^2 + F(\Phi)$
Sparse Code Update
$\min_{z_j} \sum_j \|\Phi E_j x - z_j\|_2^2 + \gamma \|z_j\|_0$

Regularizer Update

- Φ 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*: $z_j^{k+1} = \mathcal{T}_\gamma\left(\Phi E_j x\right)$

$$\mathcal{T}_\gamma(a) = \begin{cases} 0, & |a| \le \sqrt{\gamma} \\ a, & \text{else} \end{cases}$$

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

$$\min_{x} \frac{1}{2}\|y - v\|_W^2 + \sum_j \frac{\lambda}{2}\|\Phi E_j x - z_j\|_2^2$$
$$\text{s.t. } v = Ax$$

Solution using ADMM [Ramani, 2012]

- Augmented Lagrangian

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

- Alternate between
  - Minimization over x
  - Minimization over v
  - Maximization over η

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-update

$$v^{k+1} = (W + \mu I)^{-1}(Wy + \mu(Ax^{k+1} + \eta^k))$$

η-update

$$\eta^{k+1} = \eta^k - v^{k+1} + Ax^{k+1}$$

Overall Algorithm (AST-CT)

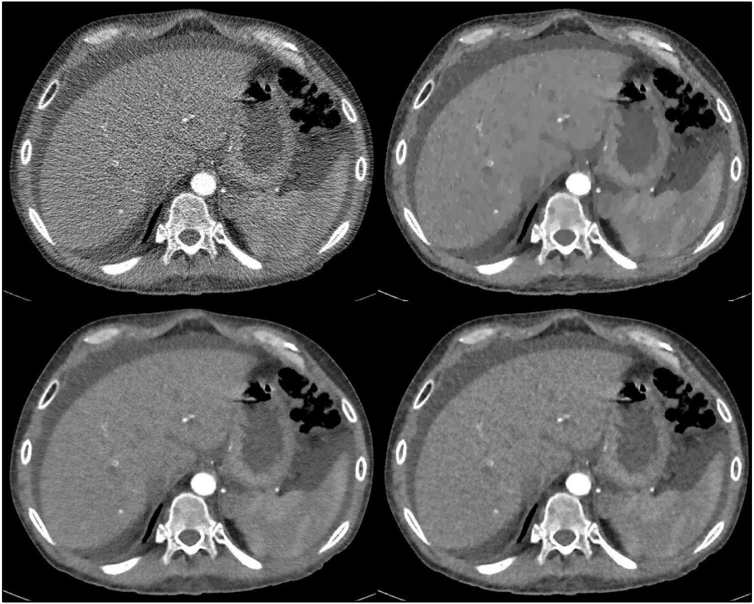
repeat
  repeat
    Update Φ
    $z_j^k \leftarrow \mathcal{T}_\gamma \Phi E_j x \; \forall j$
  until Halting condition
  $i \leftarrow 0$, $u^0 \leftarrow Ax^k$, $v^0 \leftarrow \vec{0}$
  repeat
    Use PCG to find approximate solution
    of $H\tilde{x}^{i+1} = \mu A^T(u^i - \eta^i) + \lambda \sum_j E_j^T \Phi^T z_j^i$
    $u^{i+1} \leftarrow (W + \mu I)^{-1} \left(Wy + \mu(A\tilde{x}^{i+1} + v^i)\right)$
    $\eta^{i+1} \leftarrow \eta^i - \left(u^{i+1} - A\tilde{x}^{i+1}\right)$
    $i \leftarrow i + 1$
  until Halting condition
  $x^{k+1} \leftarrow \tilde{x}^{i+1}$
until Halting Condidition

Experiments

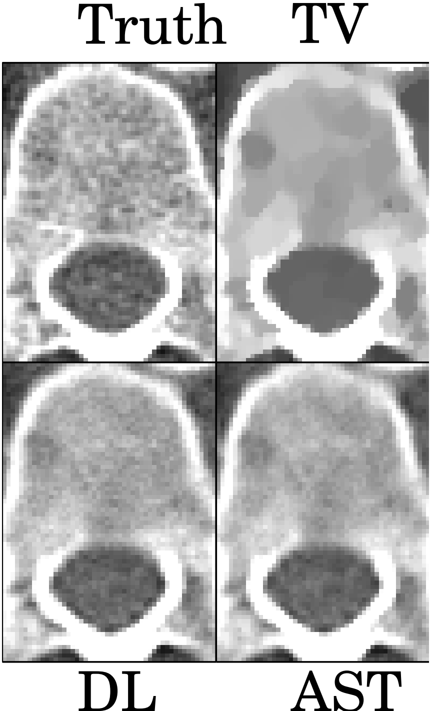
Experiments

- 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):
  - $J(x) = \min_{D,a_j} \sum_j \|E_j x - D a_j\|_2^2 + \gamma \|a_j\|_0$
  - Solve with orthogonal matching pursuit and K-SVD

FBP TV



DL AST



Truth
TV
DL
AST

Experiments

	D/Φ Update	a/z Update	Image Update	Total
FBP	0	0	2.3	2.3
TV-CT	0	0	91.3	91.3
DL-CT	87.5	60.3	85.4	233.3
AST-CT	4.4	0.2	88.4	93.0

Linearized ADMM

Linearized ADMM

- Also called the split inexact Uzawa method
- Many possible derivations- We follow the approach of Zhang & Osher[1]

[1] *A Unified Primal Dual Framework Based on Bregman Iteration*, 2011

Linearized ADMM

$$
\begin{aligned}
\mathcal{L}(x, v, \eta) =& \frac{1}{2}\|y - v\|_W^2 + \sum_j \frac{\lambda}{2}\|\Phi E_j x - z_j\|_2^2 + \frac{\mu}{2}\|v - Ax - \eta\|_2^2 \\
& - \frac{\mu}{2}\|\eta\|_2^2 + \|x - x^k\|_Q^2
\end{aligned}
$$

- Q: Positive definite matrix
- v and η updates remain unchanged

x update

$$\left(\mu A^T A + \sum_j E_j^T \Phi^T \Phi E_j + Q\right) x^{k+1} = \mu A^T (v^k - \eta^k) + \sum_j E_j \Phi^T z_j + Q x^k$$

x update

- $Q = \delta I - \mu A^T A$
- $\delta > \mu \|A^T A\|_2$

$$\left(\delta I + \sum_j E_j^T \Phi^T \Phi E_j\right) x^{k+1} = \mu A^T (v^k - \eta^k + A x^k) + \sum_j E_j \Phi^T z_j$$

x update

- $Q = \delta I - \mu A^T A$
- $\delta > \mu \|A^T A\|_2$

$$\underbrace{\left(\delta I + \sum_j E_j^T \Phi^T \Phi E_j\right)}_{\text{Circulant!}} x^{k+1} = \mu A^T (v^k - \eta^k + A x^k) + \sum_j E_j \Phi^T z_j$$

- Solve using FFT; no need to precondition!
- Trading slower convergence for faster iterations

x update

- $Q = \delta I - \mu A^T A$
- $\delta > \mu \|A^T A\|_2$

$$\underbrace{\left(\delta I + \sum_j E_j^T \Phi^T \Phi E_j\right)}_{\text{Circulant!}} x^{k+1} = \mu A^T (v^k - \eta^k + A x^k) + \sum_j E_j \Phi^T z_j$$

- Solve using FFT; no need to precondition!
- Trading slower convergence for faster iterations

x update

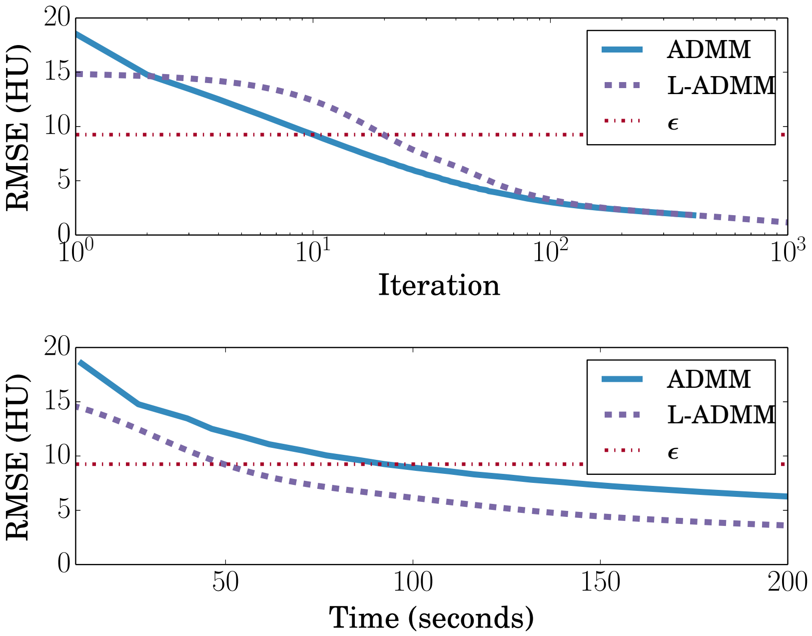
- $Q = \delta I - \mu A^T A$
- $\delta > \mu \|A^T A\|_2$

$$\underbrace{\left(\delta I + \sum_j E_j^T \Phi^T \Phi E_j\right)}_{\text{Circulant!}} x^{k+1} = \mu A^T (v^k - \eta^k + Ax^k) + \sum_j E_j \Phi^T z_j$$

- Solve using FFT; no need to precondition!
- Trading slower convergence for faster iterations

L-ADMM vs ADMM

- Run 4000 ADMM iterations to solve x-update; call result x^*.
- Compare L-ADMM vs ADMM (with circulant preconditioner)
- Compare RMSE between iterates and x^*
- $\epsilon =$ RMSE of ADMM after 10 iterations



RMSE (HU)
Iteration
ADMM
L-ADMM
ϵ
20
15
10
5
0
10^0
10^1
10^2
10^3
RMSE (HU)
Time (seconds)
ADMM
L-ADMM
ϵ
20
15
10
5
0
50
100
150
200

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
- Linearized ADMM: Improves speed of image update subproblem
  - No preconditioner!
  - Use for other regularizers?
  - Similar technique without using ADMM? (Moreau-Yosida regularization)

Thanks!

Moreau-Yosida Regularization

- Let f be a convex function over $\mathbb{R}^n$
- Moreau-Yosida regularization:

$$g(x) \triangleq \min_{y \in \mathbb{R}^n} f(y) + \frac{1}{2}\|y - x\|_2^2$$

- f and g share the same set of minimizers