Image Reconstruction in Medical X-ray CT (and PET)

Jeffrey A. Fessler

EECS Dept., BME Dept., Dept. of Radiology
University of Michigan

http://web.eecs.umich.edu/~fessler

ISMRM Sedona Workshop:
Data Sampling & Image Reconstruction

2020-01-27

Disclosure

- Research support from GE Healthcare from 2001-2016
  for low-dose CT image reconstruction
- CT research supported in part by NIH grants R01 HL 098686 & U01 EB018753
- Equipment support from Intel Corporation and NVIDIA
- Gift support from KLA

Acknowledgment:
Thanks to many collaborators and many students and post-docs

Outline

Outline

MR image reconstruction via compressed sensing

Measurement model:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$$

- $\boldsymbol{y}$ k-space data
- $\boldsymbol{A}$ system model (gradient encoding, sensitivity encoding, B0 map, ...)
  (wide matrix for under-sampled data, aka compressed sensing)
- $\boldsymbol{x}$ unknown image to be reconstructed
- $\boldsymbol{\varepsilon}$ complex noise in k-space

MR image reconstruction via compressed sensing

Measurement model:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

- $\boldsymbol{y}$ k-space data
- $\boldsymbol{A}$ system model (gradient encoding, sensitivity encoding, B0 map, ...)
  (wide matrix for under-sampled data, aka compressed sensing)
- $\boldsymbol{x}$ unknown image to be reconstructed
- $\boldsymbol{\varepsilon}$ complex noise in k-space

Regularized image reconstruction formulation: (Lustig, Donoho, Pauly: MRM, 2007) [1]

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \beta_1 \|\boldsymbol{T}\boldsymbol{x}\|_1 + \beta_2 \|\boldsymbol{x}\|_{\mathrm{TV}}$$

- $\boldsymbol{T}$ sparsifying transform such as orthogonal wavelets
- $\|\boldsymbol{x}\|_{\mathrm{TV}}$ total variation (TV) regularizer. In 1D: $\|\boldsymbol{x}\|_{\mathrm{TV}} = \sum_j |x_j - x_{j-1}|$
- β regularization parameters
- arg min : requires iterative methods

FDA approval for clinical use in commercial systems 2017 & 2018 [2] [3] [4]

Outline

History: Statistical reconstruction for PET I

J. Fessler

- Iterative method for emission tomography (Kuhl, 1963)
  (earliest iterative method for medical imaging?)
- FBP for PET (Chesler, 1971)
- Weighted least squares for 3D SPECT (Goitein, NIM, 1972)
- Richardson/Lucy iteration for image restoration (1972, 1974)
- Poisson likelihood (emission) (Rockmore and Macovski, TNS, 1976)

$$\boldsymbol{y} \sim \text{Poisson}\{\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}\} \Longrightarrow L(\boldsymbol{x}) = \mathbf{1}'(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}) - \boldsymbol{y}' \log.(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b})$$

- Expectation-maximization (EM) algorithm (Shepp and Vardi, TMI, 1982)
- Regularized (aka Bayesian) Poisson emission reconstruction (Geman and McClure, ASA, 1985)
- Ordered-subsets EM (OSEM) algorithm (Hudson and Larkin, TMI, 1994)
- Commercial release of OSEM for PET scanners circa 1997

History: Statistical reconstruction for PET II

- Today, most (all?) commercial PET systems include *unregularized* OSEM
- Some pre-clinical PET systems use regularized reconstruction

  Qi and Leahy et al. 1998
- Some clinical PET systems more recently have used edge-preserving regularizers

  Ahn et al. 2015
- Relative difference prior: Nuyts et al. 2002

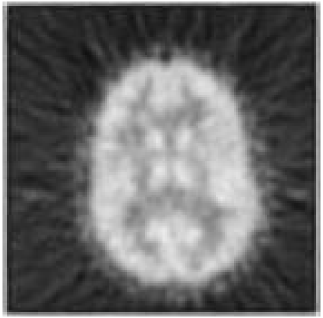
$$\psi(a, b) = \frac{(a-b)^2}{(a+b) + \gamma\,|a-b|} \qquad (\text{ cf TV: } |a-b| \text{ })$$

- 15 years between key EM paper (1982) and commercial adoption (1997)
  (25 years if you count the R/L paper in 1972 that is the same as EM)
- 30 years between early MAP methods and clinical regularized methods

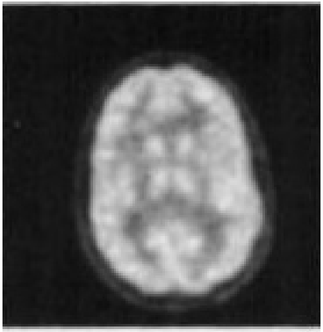
Key factors in PET

- OS algorithm accelerated convergence by order of magnitude
- Computers got faster (but problem size grew too)
- Key clinical validation papers?
- Key numerical observer studies?
- Nuclear medicine physicians grew accustomed to appearance of images reconstructed using statistical methods

FBP:

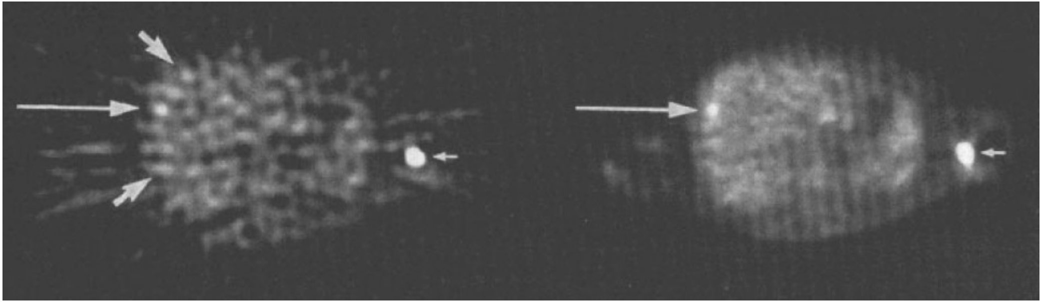


ML-EM:



Llacer et al., 1993

Whole-body PET example



FBP ML-OSEM

Meikle et al., 1994

Key factor in PET: Poisson model for measurement statistics

Current PET image reconstruction research

- MR-guided PET image reconstruction for PET-MR (or MR-PET) systems
- Motion-compensated image reconstruction
- Reduced dose PET image reconstruction
-

Current PET image reconstruction research

- MR-guided PET image reconstruction for PET-MR (or MR-PET) systems
- Motion-compensated image reconstruction
- Reduced dose PET image reconstruction
- Machine learning methods for PET image reconstruction
  - Post-process initial reconstructed image [19]
  - Improve sinogram then apply FBP [20]
  - Unrolled-loop iterative reconstruction [21, 22, 23]
  - Direct from sinogram to image: “learned FBP” (2D only, using CNN!) [24]
  -

Current PET image reconstruction research

- MR-guided PET image reconstruction for PET-MR (or MR-PET) systems
- Motion-compensated image reconstruction
- Reduced dose PET image reconstruction
- Machine learning methods for PET image reconstruction
  - Post-process initial reconstructed image [19]
  - Improve sinogram then apply FBP [20]
  - Unrolled-loop iterative reconstruction [21, 22, 23]
  - Direct from sinogram to image: "learned FBP" (2D only, using CNN!) [24]
  - *cf.* (LSI!) ANN for SPECT image recon, C. Floyd, IEEE-T-MI Sep. 1991 [25]

Outline

X-ray CT scans

CT image reconstruction problem:
Determine unknown attenuation map $\boldsymbol{x}$ given sinogram data $\boldsymbol{y}$ using system matrix $\boldsymbol{A}$.

(No moving parts to animate)

MR image reconstruction problem:
Determine unknown magnetization image $\boldsymbol{x}$ given k-space data $\boldsymbol{y}$ using system matrix $\boldsymbol{A}$

Clinical CT system / instrumentation advances: Classic

- From single slice to multi-slice
  1999 4-slice, 2003 64-slice, ...
  More recently: 256 or 320 detector rows
  $256 \cdot 0.625 = 160$mm axial coverage

Clinical CT system / instrumentation advances: Classic

- From single slice to multi-slice
  1999 4-slice, 2003 64-slice, ...
  More recently: 256 or 320 detector rows
  $256 \cdot 0.625 = 160$mm axial coverage
- From axial scan to helical scans ($\approx$ 1989)

Clinical CT system / instrumentation advances: Classic

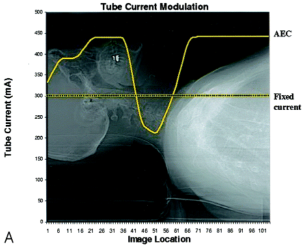
- From single slice to multi-slice
  1999 4-slice, 2003 64-slice, ...
  More recently: 256 or 320 detector rows
  $256 \cdot 0.625 = 160$mm axial coverage
- From axial scan to helical scans ($\approx$ 1989)
- Faster rotation ($\approx$ 0.3 sec?)

Clinical CT system / instrumentation advances: Classic

J. Fessler

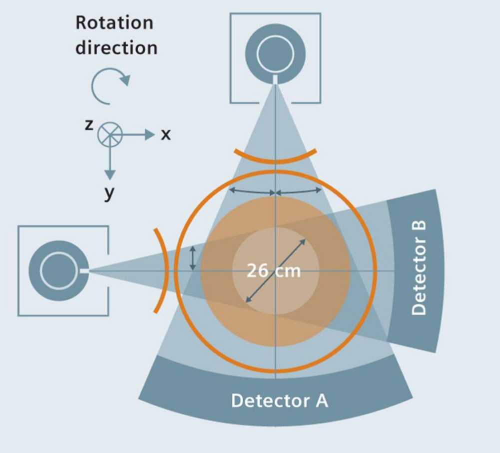
- From single slice to multi-slice
  1999 4-slice, 2003 64-slice, …
  More recently: 256 or 320 detector rows
  $256 \cdot 0.625 = 160$mm axial coverage
- From axial scan to helical scans ($\approx$ 1989)
- Faster rotation ($\approx$ 0.3 sec?)
- Tube current modulation
  to reduce dose in helical scans

`http://www.ajnr.org/content/27/10/2221`



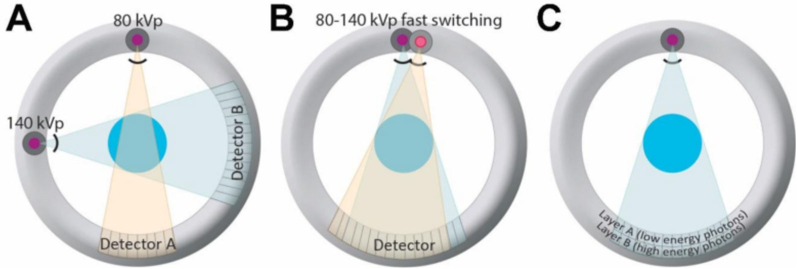
A

- Dual X-ray source / detector systems (2005)



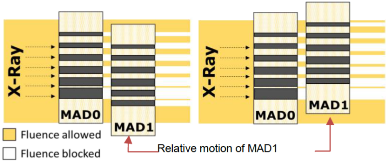
https://www.siemens-healthineers.com/no/computed-tomography/news/mso-back-to-the-future.html

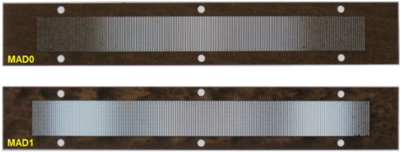
- Dual energy systems (for material separation)
  - Slow kVp switching
  - Dual source/detectors systems
  - Fast kVp switching
  - Dual layer detectors

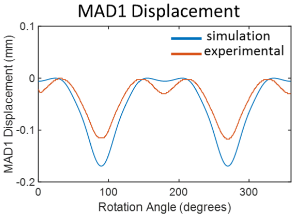


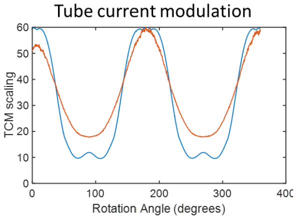
[26]

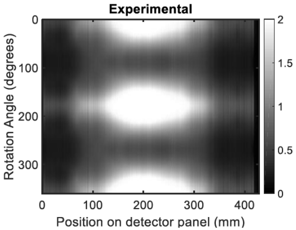
- X-ray fluence modulation [27]











- photon-counting detectors
  - cut electronic noise
  - multi-spectral data
  - possibly with new contrast agents (*e.g.*, gold nanoparticles)

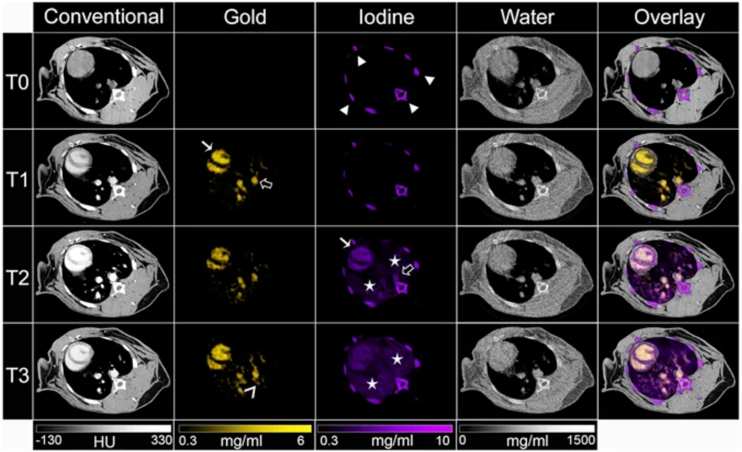
IEEE Transactions on Radiation & Plasma Medical Sciences

Special issue on
Single photon counting spectral x-ray computed tomography imaging

Call for papers

Guest Editors
Katsuyuki Taguchi, Dimitra G. Darambara, Michael Campbell, and Rafael Ballabriga

Spectral CT example



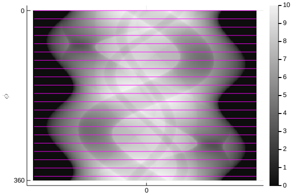
"color CT" [28]

Dose reduction methods

- reduce tube current
- X-ray tube-current modulation
- X-ray fluence modulation
- eliminate electronic noise using photon counting
-

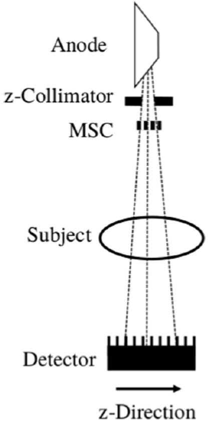
Dose reduction methods

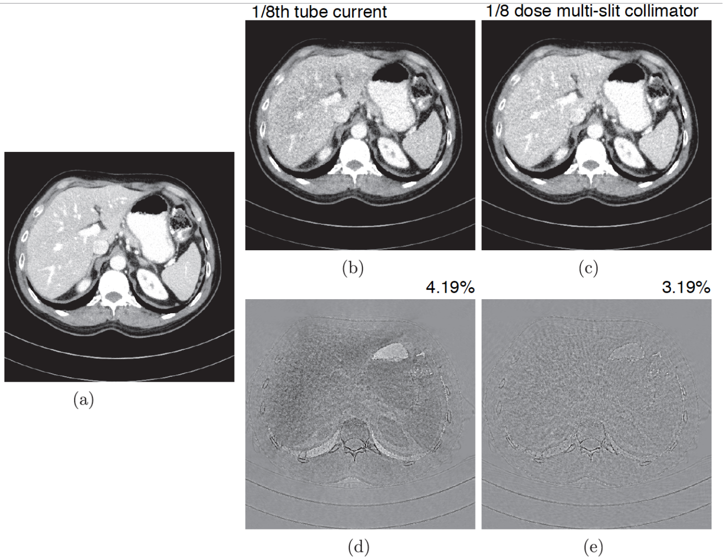
- reduce tube current
- X-ray tube-current modulation
- X-ray fluence modulation
- eliminate electronic noise using photon counting
- sparse view CT (cf radial undersampling in MRI)
  - Easy for slow flat-panel C-arm systems
  - Hard for fast rotating helical systems



Sparse-view CT example

NYU, Muckley et al. [29]
multi-slit collimator





Outline

$$\begin{matrix}\text{Unknown} \\ \text{object} \\ \boldsymbol{x}\end{matrix} \rightarrow \boxed{\begin{matrix}\text{Imaging} \\ \text{system}\end{matrix}} \rightarrow \begin{matrix}\text{Data} \\ \boldsymbol{y}\end{matrix} \rightarrow \boxed{\text{Recon}} \rightarrow \begin{matrix}\text{Image} \\ \hat{\boldsymbol{x}}\end{matrix}$$

How to reconstruct object $\boldsymbol{x}$ from data $\boldsymbol{y}$?

Non-iterative methods:

- analytical / direct
  - Filtered back-projection (FBP) for CT (textbook: Radon transform)
  - Inverse FFT for MRI (textbook: FFT)
- idealized description of the system ("textbook model")
  - geometry / sampling
  - disregards noise and simplifies physics
- typically fast

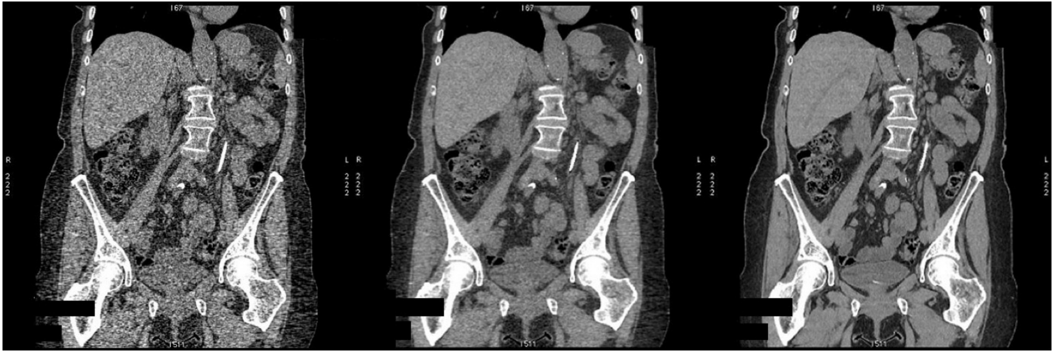
Iterative methods:

- model-based / statistical
- based on "reasonably accurate" models for physics and statistics
- usually much slower

Statistical image reconstruction: CT example

J. Fessler

- A picture is worth 1000 words
- (and perhaps several 1000 seconds of computation?)



Thin-slice FBP	ASIR (denoise)	Statistical
Seconds	A bit longer	Much longer

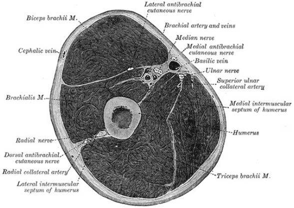
(Same sinogram, so all at same dose)

Why statistical/iterative methods for CT?

- Accurate physics models
  - X-ray spectrum, beam-hardening, scatter, ...
    $\Longrightarrow$ reduced artifacts? quantitative CT?
  - X-ray detector spatial response, focal spot size, ...
    $\Longrightarrow$ improved spatial resolution?
  - detector spectral response (*e.g.*, photon-counting detectors)
    $\Longrightarrow$ improved contrast between distinct material types?

- Nonstandard geometries
  - transaxial truncation (wide patients)
  - long-object problem in helical CT
  - irregular sampling in "next-generation" geometries
  - coarse angular sampling in image-guidance applications
  - limited angular range (tomosynthesis)
  - "missing" data, *e.g.*, bad pixels in flat-panel systems

- Appropriate models of (data dependent) measurement statistics
  - weighting reduces influence of photon-starved rays (*cf.* FBP)
    $\Longrightarrow$ reducing image noise or X-ray dose
- Object constraints / priors
  - nonnegativity
  - object support
  - piecewise smoothness
  - object sparsity (*e.g.*, angiography)
  - sparsity in some basis
  - motion models
  - dynamic models
  - ...



Henry Gray, Anatomy of the Human Body, 1918, Fig. 413.

Constraints may help reduce image artifacts or noise or dose.

Similar motivations/benefits in PET and SPECT.

Disadvantages of iterative methods for CT?

J. Fessler

- Computation time
- Must reconstruct entire FOV
- Complexity of models and software
- Algorithm nonlinearities
  - Difficult to analyze resolution/noise properties (*cf.* FBP)
  - Tuning parameters
  - Challenging to characterize performance / assess IQ

Sub-mSv example

3D helical X-ray CT scan of abdomen/pelvis:
100 kVp, 25-38 mA, 0.4 second rotation, 0.625 mm slice, 0.6 mSv.

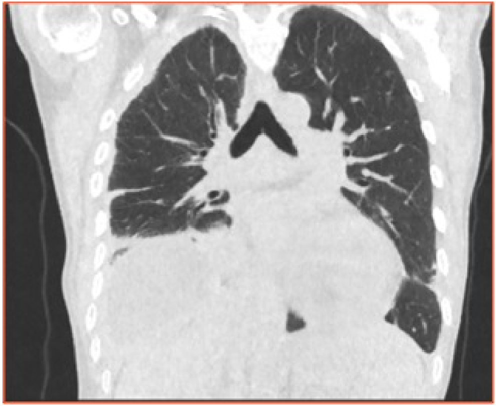
FBP ASIR Statistical

MBIR example: Chest CT

Helical chest CT study with dose = 0.09 mSv.
Typical CXR effective dose is about 0.06 mSv.

(Health Physics Soc.: `http://www.hps.org/publicinformation/ate/q2372.html`)

FBP



MBIR

Veo (MBIR) images courtesy of Jiang Hsieh, GE Healthcare

History: Statistical reconstruction for X-ray CT*

J. Fessler

- Iterative method for X-ray CT (Hounsfield, 1968)
- ART (Kaczmarz) for tomography (Gordon, Bender, Herman, JTB, 1970)
- ...
- Roughness regularized LS for tomography (Kashyap & Mittal, 1975)
- Poisson likelihood (transmission) (Rockmore and Macovski, TNS, 1977)
- EM algorithm for Poisson transmission (Lange and Carson, JCAT, 1984)
- Iterative coordinate descent (ICD) (Sauer and Bouman, T-SP, 1993)
- Ordered-subsets algorithms (Manglos et al., PMB 1995)
  (Kamphuis & Beekman, T-MI, 1998)
  (Erdoğan & Fessler, PMB, 1999)
- ...
- Commercial OS for Philips BrightView SPECT-CT (2010)
- Commercial ICD for GE CT scanners (Veo) (circa 2010)
- FDA 510(k) clearance of Veo (Sep. 2011)
- First Veo installation in USA (at UM) (Jan. 2012)

(* numerous omissions, including many denoising methods)

5 decades of CT image reconstruction research

J. Fessler

1. 70's "Analytical" methods (integral equations): FBP
2. 80's Algebraic methods (as in "linear algebra")
   Solve $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$
3. 90's Statistical methods
   - LS / ML methods
   - Bayesian methods (Markov random fields, ...)
   - regularized methods
4. 00's Compressed sensing methods
   (mathematical sparsity models)
5. 10's Adaptive / data-driven methods
   machine learning, deep learning, ...

Optimization problem formulation:

$$\hat{\boldsymbol{x}} = \underbrace{\operatorname*{arg\,min}_{\boldsymbol{x} \geq \boldsymbol{0}}}_{\substack{\text{optimization} \\ \text{algorithm}}} \Psi(\boldsymbol{x}), \qquad \underbrace{\Psi(\boldsymbol{x})}_{\substack{\text{cost} \\ \text{function}}} \triangleq \underbrace{\frac{1}{2} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_{\boldsymbol{W}}^2}_{\substack{\text{data-fit term} \\ \text{physics \& statistics}}} + \underbrace{\beta \sum_{j=1}^{N} \sum_{k \in \mathcal{N}_j} \psi(x_j - x_k)}_{\substack{\text{regularizer} \\ \text{prior models}}}$$

$\boldsymbol{y}$: measured data (sinogram)
$\boldsymbol{A}$: system matrix (physics / geometry)
$\boldsymbol{W}$: weighting matrix (statistics)
$\boldsymbol{x}$: unknown image (attenuation map)
β : regularization parameter(s)
$\mathcal{N}_j$: neighborhood of jth voxel
ψ : edge-preserving potential function
(piece-wise smoothness / gradient sparsity)

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x} \geq \boldsymbol{0}} \Psi(\boldsymbol{x}), \quad \Psi(\boldsymbol{x}) \triangleq \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_{\boldsymbol{W}}^2 + \sum_j \sum_k \beta_{j,k} \, \psi(x_j - x_k)$$

Apparent topics:

- regularization design / parameter selection ψ, β_{jk}
- statistical modeling $\boldsymbol{W}$, $\|\cdot\|$
- system modeling $\boldsymbol{A}$
- optimization algorithms (arg min)
- assessing IQ of $\hat{\boldsymbol{x}}$

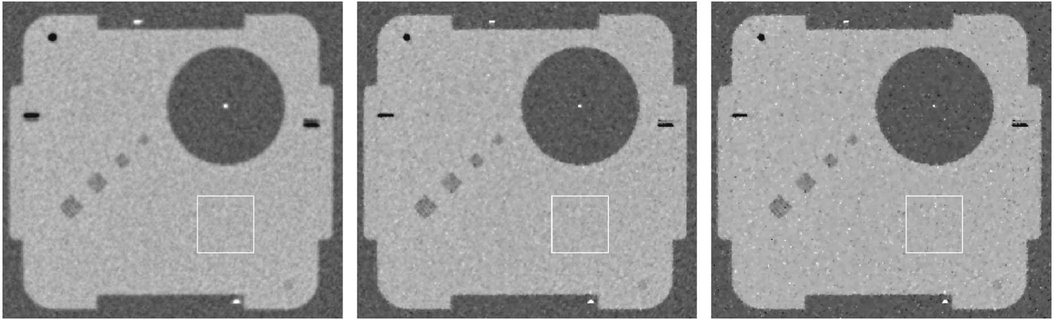
Other topics:

- system design
- motion
- spectral
- dose ...

"q generalized gaussian" potential function with tuning parameters: β, δ, p, q:

$$\beta\,\psi(t) = \beta \frac{\frac{1}{2}\,|t|^p}{1 + |t/\delta|^{p-q}}.$$

(Thibault et al., Med. Phys., Nov. 2007) [44]



	$p = q = 2$	$p = 2,\ q = 1.2,\ \delta = 10$ HU	$p = q = 1.1$
noise (HU):	11.1	10.9	10.8
(#lp/cm):	4.2	7.2	8.2

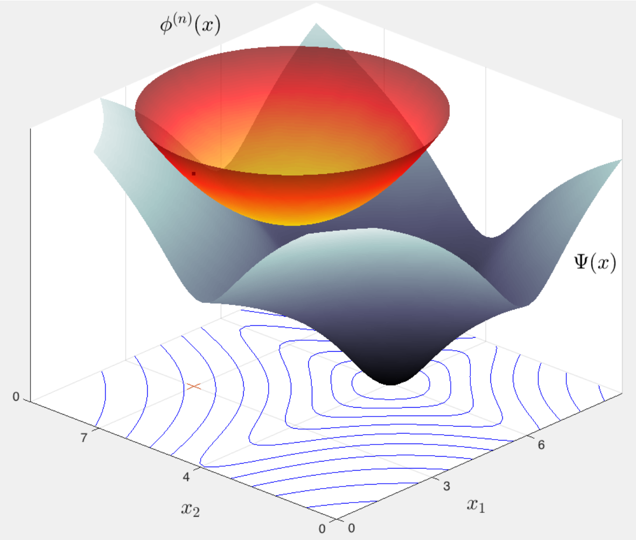
SIR for CT: Optimization challenges

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x} \geq \boldsymbol{0}}{\arg\min}\, \Psi(\boldsymbol{x}), \quad \Psi(\boldsymbol{x}) \triangleq \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_{\boldsymbol{W}}^2 + \sum_{j=1}^{N} \sum_{k} \beta_{j,k}\, \psi(x_j - x_k)$$

Optimization challenges:

- large problem size: $\boldsymbol{x} \in \mathbb{R}^{512 \times 512 \times 600}$, $\boldsymbol{y} \in \mathbb{R}^{888 \times 64 \times 7000}$
- $\boldsymbol{A}$ is sparse but still too large to store; compute $\boldsymbol{A}\boldsymbol{x}$ on-the-fly
- $\boldsymbol{W}$ has enormous dynamic range (1 to $\exp(-9) \approx 1.2 \cdot 10^{-4}$)
- Gram matrix $\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}$ highly shift variant
- Ψ is non-quadratic but convex (and often smooth)
- nonnegativity constraint
- data size grows: dual-source CT, spectral CT, wide-cone CT, ...
- Moore's law insufficient
  more cores/threads, not faster clock speeds

Optimization transfer (Majorize-Minimize) methods: 2D



$$\phi^{(n)}(\boldsymbol{x}^{(n)}) = \Psi(\boldsymbol{x}^{(n)})$$
$$\phi^{(n)}(\boldsymbol{x}) \geq \Psi(\boldsymbol{x})$$

cf. ML-EM

$$\boldsymbol{x}^{(n+1)} = \arg\min_{\boldsymbol{x}} \phi^{(n)}(\boldsymbol{x})$$

New approach by optimizing step-sizes $\{h_{n,k}\}$ analytically

Initialize: $t_0 = 1,\ \boldsymbol{z}^{(0)} = \boldsymbol{x}^{(0)}$ (Donghwan Kim and JF; 2014-2016)

$$\boldsymbol{z}^{(n+1)} = \boldsymbol{x}^{(n)} - \frac{1}{L} \nabla \Psi\left(\boldsymbol{x}^{(n)}\right) \qquad \text{(usual GD update)}$$

$$t_{n+1} = \frac{1}{2}\left(1 + \sqrt{1 + 4t_n^2}\right) \qquad \text{(momentum factors)}$$

$$\boldsymbol{x}^{(n+1)} = \boldsymbol{z}^{(n+1)} + \underbrace{\frac{t_n - 1}{t_{n+1}}\left(\boldsymbol{z}^{(n+1)} - \boldsymbol{z}^{(n)}\right)}_{\text{Nesterov}} \underbrace{+ \frac{t_n}{t_{n+1}}\left(\boldsymbol{z}^{(n+1)} - \boldsymbol{x}^{(n)}\right)}_{\text{new momentum}}$$

Smaller (worst-case) convergence bound than Nesterov by $2\times$:

$$\Psi\left(\boldsymbol{z}^{(n)}\right) - \Psi\left(\boldsymbol{x}^{(\star)}\right) \leq \frac{1L\left\|\boldsymbol{x}^{(0)} - \boldsymbol{x}^{(\star)}\right\|_2^2}{(n+1)^2}.$$

Recently Y. Drori [48] found a matching lower bound for any first-order method in high dimensions.

Ordered subsets approximation

- Data decomposition (aka incremental gradients, *cf.* stochastic GD, mini-batch):

$$\Psi(\boldsymbol{x}) = \sum_{m=1}^{M} \Psi_m(\boldsymbol{x}), \quad \Psi_m(\boldsymbol{x}) \triangleq \underbrace{\frac{1}{2} \|\boldsymbol{y}_m - \boldsymbol{A}_m \boldsymbol{x}\|^2_{\boldsymbol{W}_m}}_{1/M\text{th of measurements}} + \frac{1}{M} \mathsf{R}(\boldsymbol{x})$$

- Key idea. For $\boldsymbol{x}$ far from minimizer: $\nabla \Psi(\boldsymbol{x}) \approx M \nabla \Psi_m(\boldsymbol{x})$ [13]
- SQS (MM): [42]

$$\boldsymbol{x}^{(n+1)} = \boldsymbol{x}^{(n)} - \boldsymbol{D}^{-1} \nabla \Psi\left(\boldsymbol{x}^{(n)}\right)$$

- OS-SQS:

  for $n = 0, 1, \ldots$ (iteration)

  for $m = 1, \ldots, M$ (subset)

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^{k} - \boldsymbol{D}^{-1} M \underbrace{\nabla \Psi_m\left(\boldsymbol{x}^{k}\right)}_{\text{less work}}, \qquad k = nM + m \text{ (subiteration)}$$

- Applied coil-wise in parallel MRI (Muckley, Noll, JF, ISMRM 2014) [50]

Ordered subsets version of OGM1

For more acceleration, combine OGM1 with ordered subsets (OS). [51]

OS-OGM1:
Initialize: $t_0 = 1$, $\boldsymbol{z}^{(0)} = \boldsymbol{x}^{(0)}$
for $n = 0, 1, \ldots$ (iteration)
for $m = 1, \ldots, M$ (subset)

$$k = nM + m \text{ (subiteration)}$$

$$\boldsymbol{z}^{k+1} = \left[\boldsymbol{x}^k - \boldsymbol{D}^{-1} M \nabla \Psi_m\left(\boldsymbol{x}^k\right)\right]_+ \qquad \text{(typical OS-SQS)}$$

$$t_{k+1} = \frac{1}{2}\left(1 + \sqrt{1 + 4t_k^2}\right)$$

$$\boldsymbol{x}^{k+1} = \boldsymbol{z}^{k+1} + \frac{t_k - 1}{t_{k+1}}\left(\boldsymbol{z}^{k+1} - \boldsymbol{z}^k\right) + \frac{t_k}{t_{k+1}}\left(\boldsymbol{z}^{k+1} - \boldsymbol{x}^k\right)$$

OS-OGM1 properties

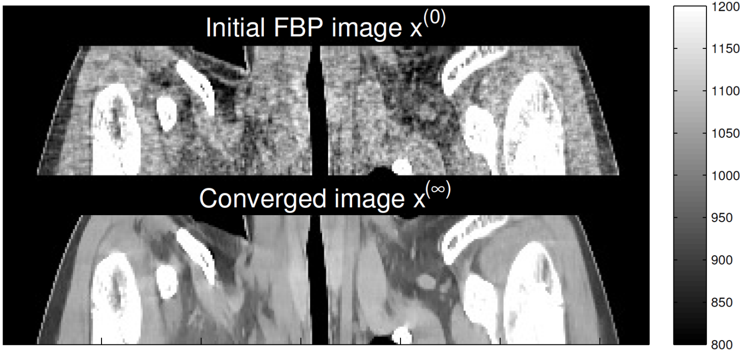
- Approximate convergence rate for Ψ: $O\left(\frac{1}{n^2 M^2}\right)$

  (Donghwan Kim and JF; IEEE T-MI 2015 [51])
- Same compute per iteration as other OS methods

  (One forward / backward projection and M regularizer gradients per iteration)
- Same memory as OGM1 (two more images than OS-SQS)
- Guaranteed convergence for $M = 1$
- No convergence theory for $M > 1$
  - unstable for large M
  - small M preferable for parallelization
- Now fast enough to show X-ray CT examples...

OS-OGM1 results: data

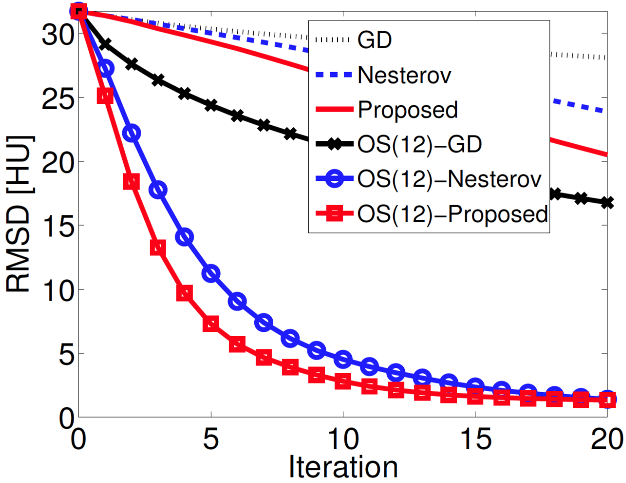
- 3D cone-beam helical X-ray CT scan
- pitch 0.5
- image $\boldsymbol{x}$: $512 \times 512 \times 109$ with 70 cm FOV and 0.625 mm slices
- sinogram : $\boldsymbol{y}$ 888 detectors $\times$ 32 rows $\times$ 7146 views



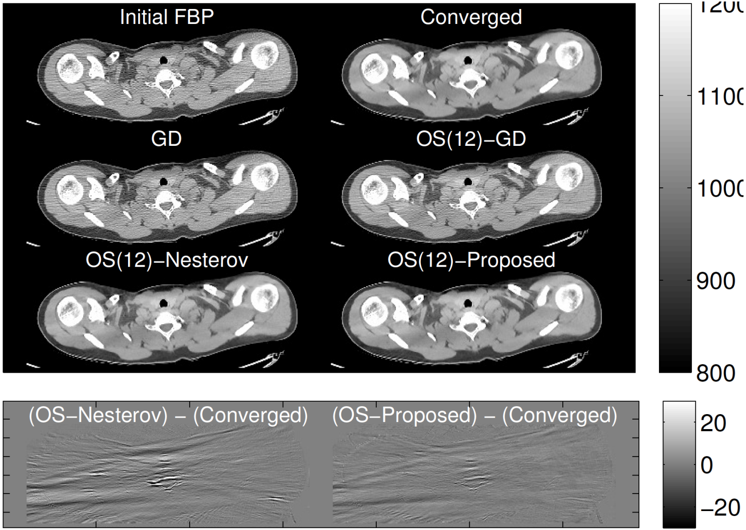
OS-OGM1 results: convergence rate

RMSD between $\boldsymbol{x}^{(n)}$ and $\boldsymbol{x}^{(\infty)}$ over ROI (in HU), versus iteration. ("Proposed" = OGM1.)

(Compute times per iteration are very similar.)



OS-OGM1 results: images



At iteration $n = 10$ with $M = 12$ subsets.

Outline

More realistic measurement model in CT with current-integrating detectors:

$$Y_i \sim \underbrace{\text{Poisson}\left\{ I_i \, \mathrm{e}^{-[\boldsymbol{A}\boldsymbol{x}]_i} \right\}}_{\text{X-ray photons}} + \underbrace{\mathcal{N}(\mu, \sigma^2)}_{\text{readout}}$$

Important for very low-dose CT scans where logarithm is problematic
Corresponding log-likelihood is complicated. Approximations:

- Shifted Poisson: [53, 54, 55]

$$Y_i - \mu + \sigma^2 \sim \text{Poisson}\left\{ I_i \, \mathrm{e}^{-[\boldsymbol{A}\boldsymbol{x}]_i} + \sigma^2 \right\}$$

- Model-dependent normal (leads to nonlinear LS): [56, 57]

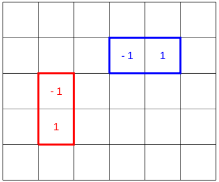
$$Y_i \sim \mathcal{N}\left(I_i \, \mathrm{e}^{-[\boldsymbol{A}\boldsymbol{x}]_i} + \mu, \, I_i \, \mathrm{e}^{-[\boldsymbol{A}\boldsymbol{x}]_i} + \mu + \sigma^2 \right)$$

- Compound Poisson and other complicated models and approximations [58, 59]

J. Fessler

- Needed for very low-dose scans and sparse-view scans

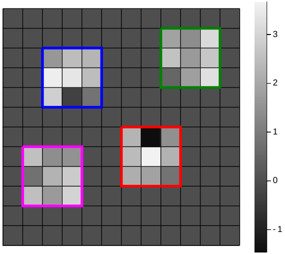
Using TV regularizer $R(\boldsymbol{x}) = \|\boldsymbol{T}\boldsymbol{x}\|_1$
where $\boldsymbol{T}$ is finite-differences
$\equiv$ patches of size 2×1.



Larger patches provide more context
for distinguishing signal from noise.

cf. CNN approaches

Patch-based regularizers:
- synthesis models
- analysis methods



Advanced regularizers II

- Patch-based dictionary synthesis models [60]

$$R(\boldsymbol{x}) = \min_{\boldsymbol{Z} \in \mathbb{R}^{K \times M}} \sum_{m=1}^{M} \frac{1}{2} \|\boldsymbol{R}_m \boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}_m\|_2^2 + \alpha \|\boldsymbol{z}_m\|_1$$

- Patch-based analysis / transform sparsity

$$R(\boldsymbol{x}) = \sum_{m=1}^{M} \|\boldsymbol{T}\boldsymbol{R}_m \boldsymbol{x}\|_1$$

- Dictionary $\boldsymbol{D}$ or transform $\boldsymbol{T}$ can be
  - learned from population training
  - adapted to each patient

Advanced regularizers III

- Convolutional dictionary sparsity [61]

$$R(\boldsymbol{x}) = \min_{\boldsymbol{z}} \frac{1}{2} \left\| \boldsymbol{x} - \sum_{k=1}^{K} \boldsymbol{h}_k * \boldsymbol{z}_k \right\|_2^2 + \alpha \sum_{k=1}^{K} \|\boldsymbol{z}_k\|_1$$

- Convolutional analysis sparsity (cf CNN) [62]

$$R(\boldsymbol{x}) = \sum_{k=1}^{K} \|\boldsymbol{h}_k * \boldsymbol{x}\|_1$$

- Filters $\{\boldsymbol{h}_k\}$ learned from population training data
- Block-matching / non-local means ... [63]
- Joint sparsity for spectral CT: mixed ℓ_2, ℓ_1 norms, or nuclear norms [64, 65]

X-ray CT with learned sparsifying transforms

- Data
  - Population adaptive methods
  - Patient adaptive methods
- Spatial structure
  - Patch-based models
  - Convolutional models
- Regularizer formulation
  - Synthesis (dictionary) approach
  - Analysis (sparsifying transform) approach

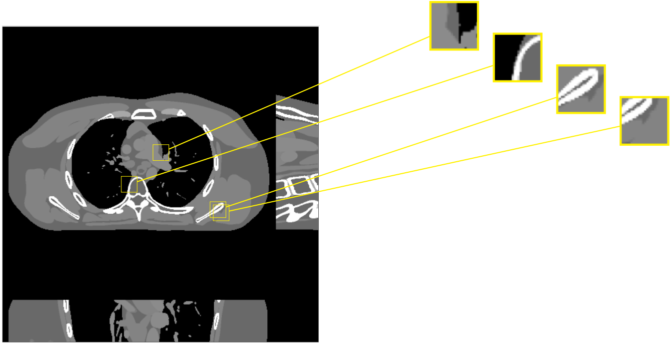
Patch-wise transform sparsity model

J. Fessler

Assumption: if $\boldsymbol{x}$ is a plausible image, then each patch transform $\boldsymbol{T}\mathbf{P}_m\boldsymbol{x}$ is sparse.

- $\mathbf{P}_m\boldsymbol{x}$ extracts the mth of M patches from $\boldsymbol{x}$
- $\boldsymbol{T}$ is a (often square) sparsifying transform matrix.

What $\boldsymbol{T}$?



Sparsifying transform learning (population adaptive)

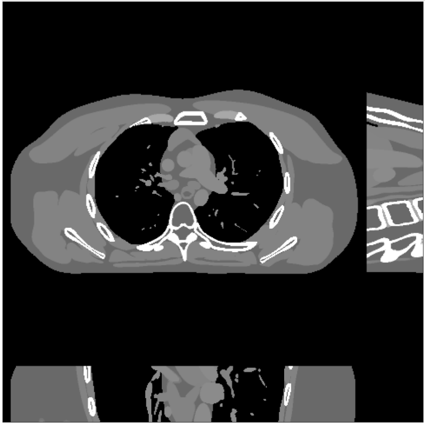
Given training images $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L$ from a representative population, find transform $\boldsymbol{T}_*$ that best sparsifies their patches:

$$\boldsymbol{T}_* = \mathop{\arg\min}_{\boldsymbol{T} \text{ unitary}} \min_{\{\boldsymbol{z}_{l,m}\}} \sum_{l=1}^{L} \sum_{m=1}^{M} \|\boldsymbol{T}\mathbf{P}_m \boldsymbol{x}_l - \boldsymbol{z}_{l,m}\|_2^2 + \alpha \|\boldsymbol{z}_{l,m}\|_0$$

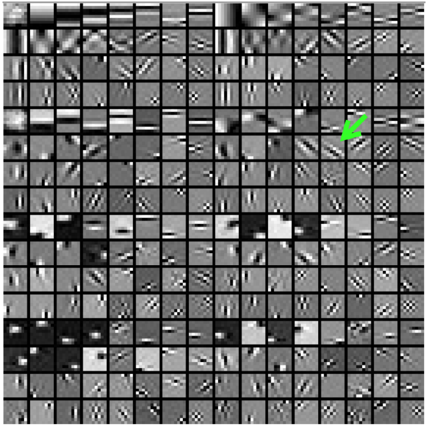
- Encourage aggregate sparsity, not patch-wise sparsity (cf K-SVD [66])
- Non-convex due to unitary constraint and $\|\cdot\|_0$
- Efficient alternating minimization algorithm [67]
  - $\boldsymbol{z}$ update : simple hard thresholding
  - $\boldsymbol{T}$ update : orthogonal Procrustes problem (SVD)
  - Subsequence convergence guarantees [67]

Example of learned sparsifying transform

3D X-ray training data



Parts of learned sparsifier $\boldsymbol{T}_*$



(2D slices in x-y, x-z, y-z, from 3D image volume)

$8 \times 8 \times 8$ patches $\Longrightarrow$ $\boldsymbol{T}_*$ is $8^3 \times 8^3 = 512 \times 512$

top 8×8 slice of 256 of the 512 rows of $\boldsymbol{T}_*$ $\uparrow$

Regularizer based on learned sparsifying transform

Regularized inverse problem [68]:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{W}}^2 + \beta \, \mathsf{R}(\boldsymbol{x})$$

$$\mathsf{R}(\boldsymbol{x}) = \min_{\{\boldsymbol{z}_m\}} \sum_{m=1}^{M} \|\boldsymbol{T}_* \mathbf{P}_m \boldsymbol{x} - \boldsymbol{z}_m\|_2^2 + \alpha \, \|\boldsymbol{z}_m\|_0 \, .$$

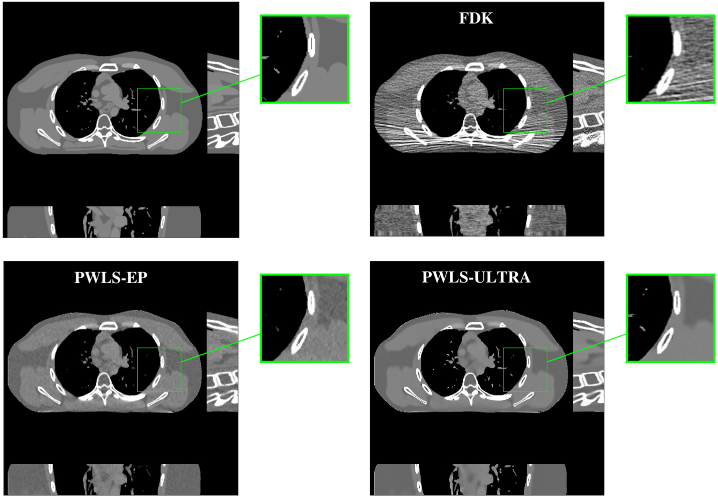
$\boldsymbol{T}_*$ adapted to population training data

Alternating minimization optimizer:

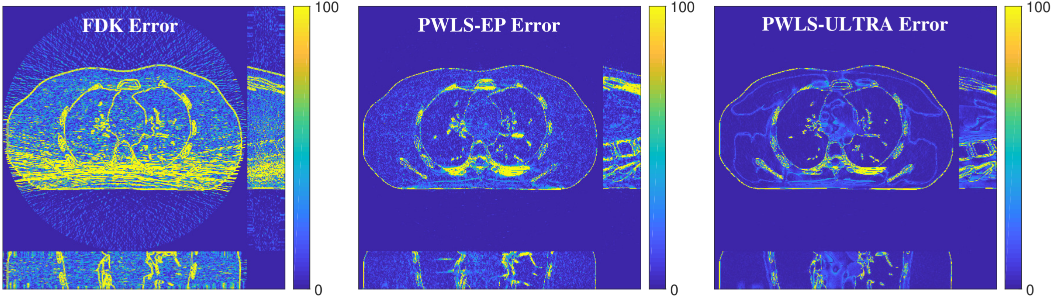
- $\boldsymbol{z}_m$ update : simple hard thresholding
- $\boldsymbol{x}$ update : quadratic problem (many options)
  Linearized augmented Lagrangian method (LALM) [69]

Example: low-dose 3D X-ray CT simulation

X. Zheng, S. Ravishankar, Y. Long, JF: IEEE T-MI, June 2018 [68].



3D X-ray CT simulation Error maps



RMSE in HU	X-ray Intensity	FDK	EP	ST T_*	ULTRA	ULTRA-$\{\tau_j\}$
	1×10^4	67.8	34.6	32.1	30.7	**29.2**
	5×10^3	89.0	41.1	37.3	35.7	**34.2**

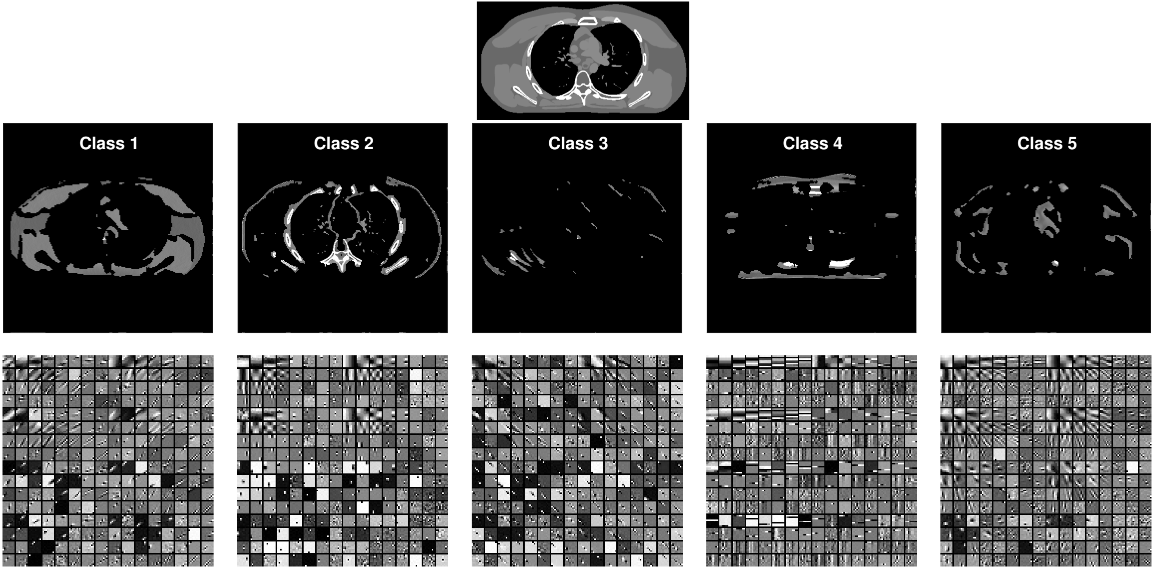
- Physics / statistics provides dramatic improvement
- Data adaptive regularization further reduces RMSE

Given training images $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L$ from a representative population, find a set of transforms $\left\{ \hat{\boldsymbol{T}}_k \right\}_{k=1}^{K}$ that best sparsify image patches:

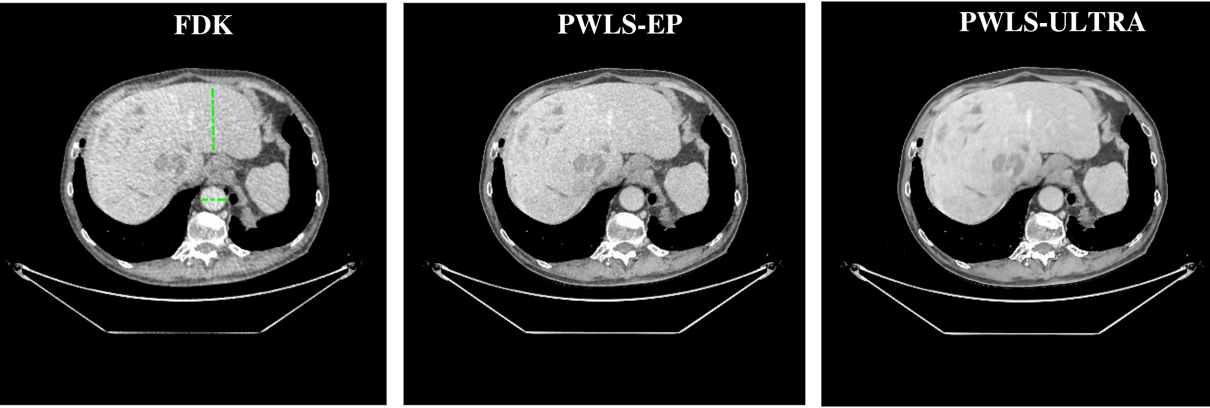
$$\left\{ \hat{\boldsymbol{T}}_k \right\} = \mathop{\arg\min}_{\{\boldsymbol{T}_k \text{ unitary}\}} \min_{\{\boldsymbol{z}_{l,m}\}} \sum_{l=1}^{L} \sum_{m=1}^{M} \left(\min_{k \in \{1,\ldots,K\}} \|\boldsymbol{T}_k \mathbf{P}_m \boldsymbol{x}_l - \boldsymbol{z}_{l,m}\|_2^2 + \alpha \|\boldsymbol{z}_{l,m}\|_0 \right)$$

- Joint unsupervised clustering / sparsification
- Further nonconvexity due to clustering
- Efficient alternating minimization algorithm [70]

Example: 3D X-ray CT learned set of transforms



Example: 3D X-ray CT ULTRA for chest scan



Zheng et al., IEEE T-MI, June 2018 [68] (Special issue on machine learning for image reconstruction)

Matlab code: `http://web.eecs.umich.edu/~fessler/irt/reproduce/`

`https://github.com/xuehangzheng/PWLS-ULTRA-for-Low-Dose-3D-CT-Image-Reconstruction`

Deep-learning approaches to CT image reconstruction

Overview:

- image-domain learning
  - arXiv papers starting in 2016 [71, 72]
  - Journal papers starting in 2017 [73, 74, 75]
  - Explosion of methods, *e.g.*, GANs [76, 77], Wasserstein loss [78]
  - beyond denoising: metal artifact reduction [79], dual energy, spectral CT...
- sinogram or data-domain learning
  denoising, "in-painting" for metal-artifact reduction [80]
- transform learning (direct from sinogram to image) ?
  in 2012 for 32×32 images [81]
  extremely difficult for 3D helical CT
  direct from sinogram to stenosis size [82, 83]
- hybrid-domain learning (unrolled loop, *e.g.*, variational network)
  alternate between denoising/destreaking and reconstruction from sinogram
  e.g., [84, 85, 86, 87, 88, 89]

Image-domain learning I

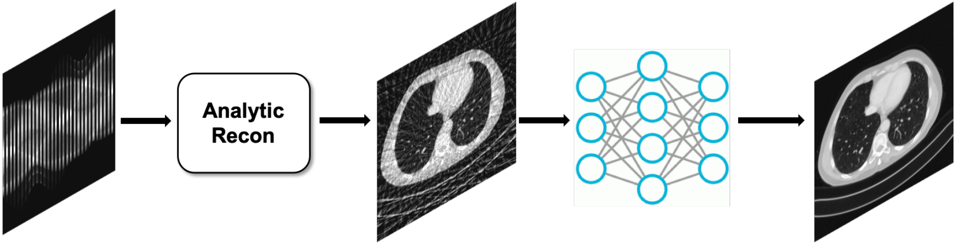
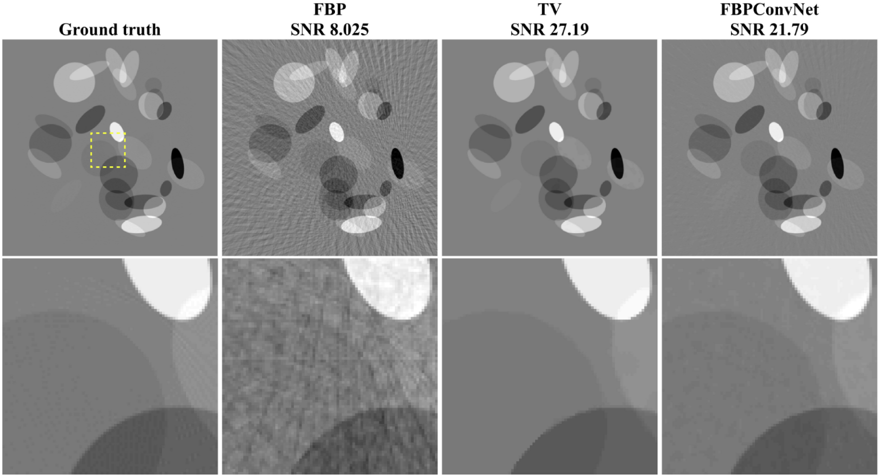
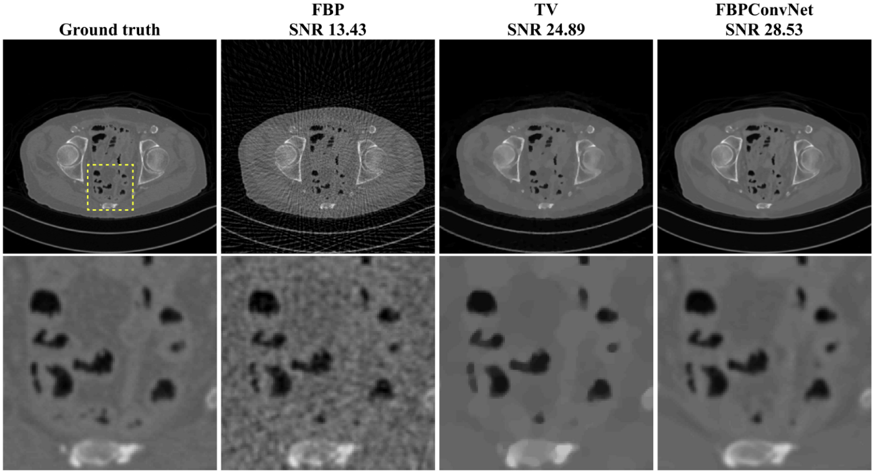


Image-domain learning II



[75]

Image-domain learning III



[75]

Convolutional sparsity revisted

Cost function for convolutional sparsity regularization:

$$\arg\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{W}}^2 + \beta \left(\min_{\boldsymbol{z}} \sum_{k=1}^K \frac{1}{2} \|\boldsymbol{h}_k * \boldsymbol{x} - \boldsymbol{z}_k\|_2^2 + \alpha \|\boldsymbol{z}_k\|_1 \right)$$

Alternating minimization updates:

$$\text{Sparse code:} \qquad \boldsymbol{z}_k^{(n+1)} = \operatorname{soft}\{\boldsymbol{h}_k * \boldsymbol{x}^{(n)}, \alpha\}$$

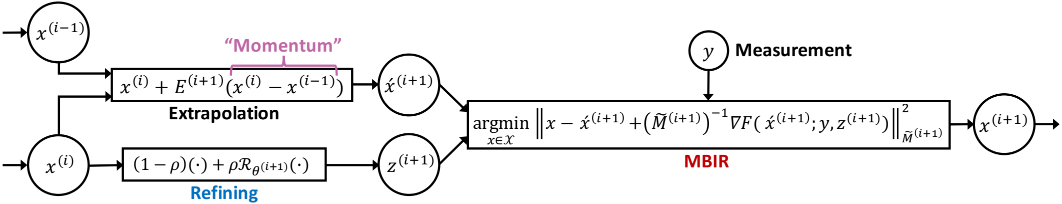
$$\text{Image:} \quad \boldsymbol{x}^{(n+1)} = \arg\min_{\boldsymbol{x}} F(\boldsymbol{x}; \boldsymbol{y}, \boldsymbol{z}^{(n)})$$

$$F(\boldsymbol{x}; \boldsymbol{y}, \boldsymbol{z}^{(n)}) \triangleq \frac{1}{2} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{W}}^2 + \beta \left(\sum_{k=1}^K \frac{1}{2} \left\|\boldsymbol{h}_k * \boldsymbol{x} - \boldsymbol{z}_k^{(n+1)}\right\|_2^2 + \alpha \left\|\boldsymbol{z}_k^{(n+1)}\right\|_1 \right)$$

$$= \frac{1}{2} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{W}}^2 + \beta \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{z}^{(n)}\|_2^2 \quad \text{(quadratic but } \textit{large} \implies \text{majorize)}$$

$$\boldsymbol{z}^{(n)} = \mathcal{R}(\boldsymbol{z}^{(n)}) = \sum_{k=1}^K \operatorname{flip}(\boldsymbol{h}_k) * \operatorname{soft}\{\boldsymbol{h}_k * \boldsymbol{x}^{(n)}\} \quad \text{(denoise} \implies \text{learn)}$$

Unrolled loop network with momentum and quadratic majorizer:



- Diagonal majorizer: $\boldsymbol{M} = \operatorname{diag}\{\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}\mathbf{1}\} + \beta \boldsymbol{I} \succeq \boldsymbol{A}'\boldsymbol{W}\boldsymbol{A} + \beta \boldsymbol{I}$
- Learn image mapper ("refiner") $\mathcal{R}$ from training data (supervised).
  cf CNN: filter $\rightarrow$ threshold $\rightarrow$ filter

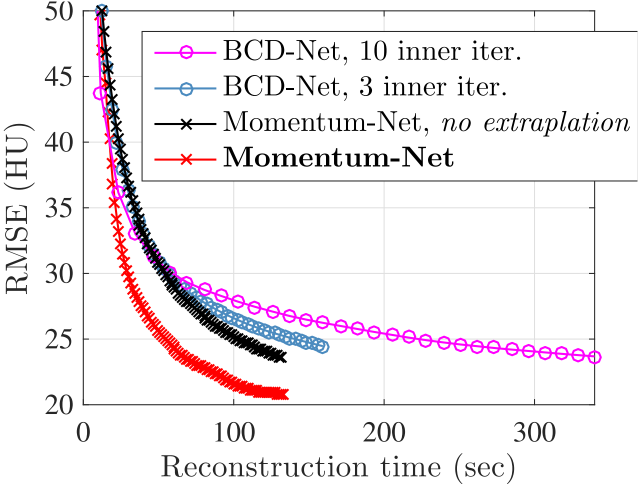
Momentum-Net benefits

- Image mapper $\mathcal{R}$ is shallow
  $\Longrightarrow$ less risk of over-fitting / hallucination
- Momentum accelerates convergence (fewer layers)
- First unrolled loop approach to have convergence theory
  (under suitable assumptions on $\mathcal{R}$)
- Image update uses original CT sinogram $\boldsymbol{y}$ and imaging physics $\boldsymbol{A}$

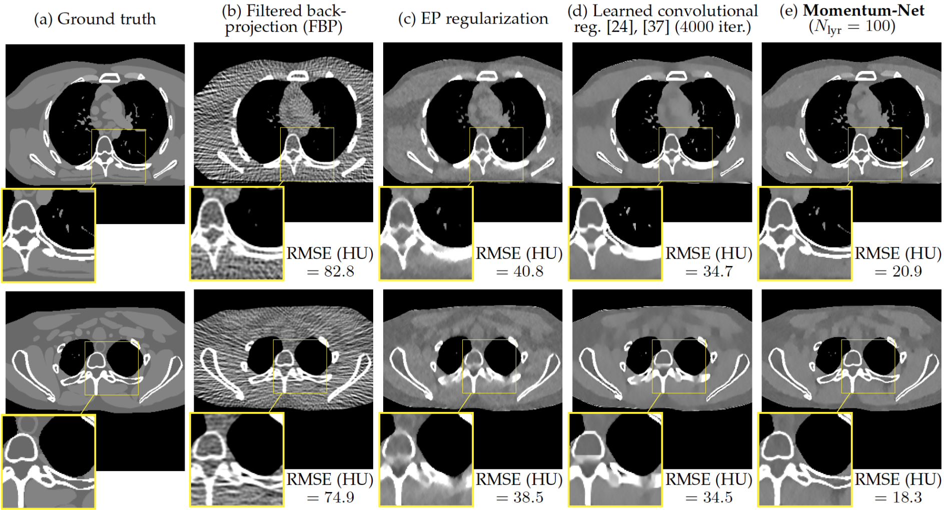
[90]
Il Yong Chun, Zhengyu Huang, Hongki Lim, J A Fessler
Momentum-Net: Fast and convergent iterative neural network for inverse problems
http://arxiv.org/abs/1907.11818

Momentum-Net preliminary results

Illustration of benefits of momentum:



Momentum-Net preliminary image results

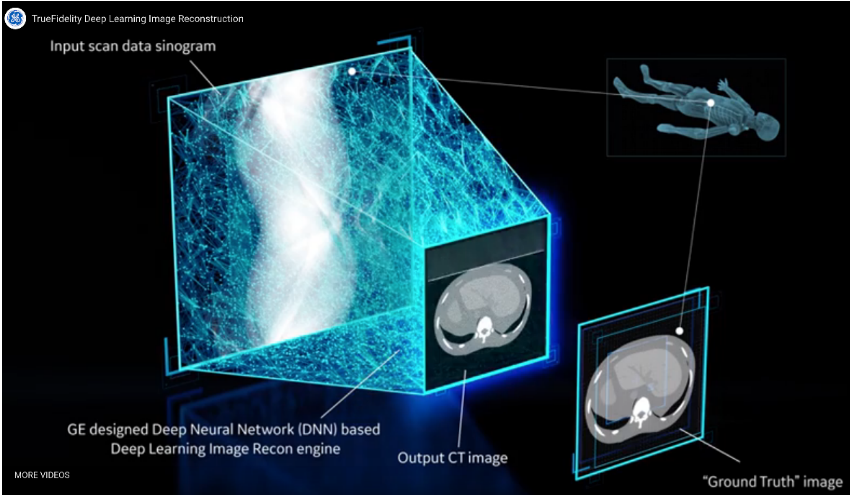


Sparse-view CT with 123/984 views, $I_0 = 10^5$, 800-1200 mod. HU display.

DL for CT now FDA approved

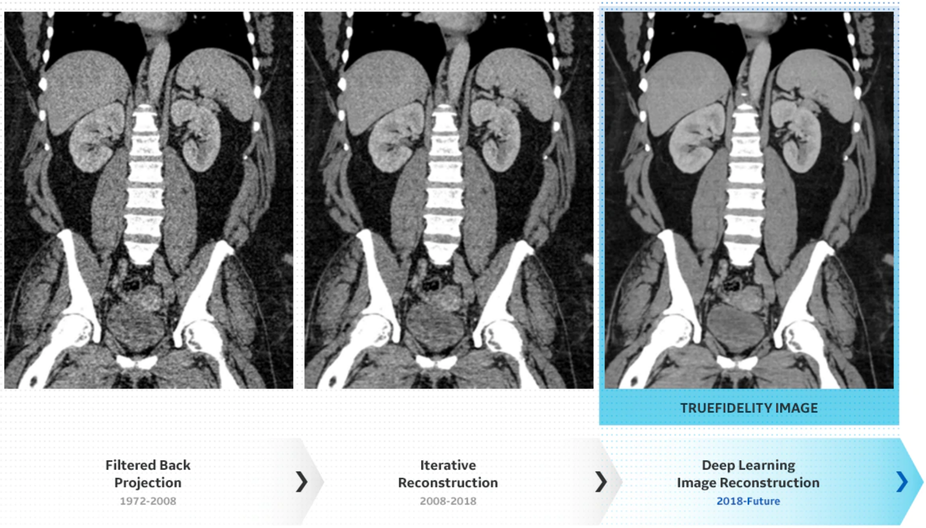
- In 2019, both Canon and GE got FDA approval for DL methods for CT [91, 92]
- Canon: "AiCE Deep Learning Reconstruction"
  Canon press release: "Advanced Intelligent Clear-IQ Engine (AiCE) uses a deep learning algorithm to differentiate signal from noise so that it can suppress noise while enhancing signal."
- GE "Deep-learning image reconstruction"
  Possibly related papers [93, 94]
  - Plug-and-play ADMM (unrolled loop) [95, 96]
  - Denoiser is 17-layer residual learning CNN, trained to map 2D noisy FBP patches to clean MBIR with squared error loss
  - Report faster "convergence" than standard MBIR
  - Sliding window of 3 slices in and 1 slice out

Glimmering neural networks



https://www.gehealthcare.com/products/truefidelity

DL for CT Example



180 kg patient

https://www.gehealthcare.com/products/truefidelity

Summary

Iterative methods for CT image reconstruction:

- have had important impact on clinical CT
- remain an active research topic
- are more painful to study realistically (than MRI) due to proprietary sinogram data
- use similar regularization methods as MRI in research
- use simpler regularization methods than MRI clinically

The future?

-

Summary

Iterative methods for CT image reconstruction:

- have had important impact on clinical CT
- remain an active research topic
- are more painful to study realistically (than MRI) due to proprietary sinogram data
- use similar regularization methods as MRI in research
- use simpler regularization methods than MRI clinically

The future?

- Apparently iterative recon for CT perished in 2018?
-

Summary

Iterative methods for CT image reconstruction:

- have had important impact on clinical CT
- remain an active research topic
- are more painful to study realistically (than MRI) due to proprietary sinogram data
- use similar regularization methods as MRI in research
- use simpler regularization methods than MRI clinically

The future?

- Apparently iterative recon for CT perished in 2018?
- Apparently CT beat MRI to FDA-approved DL recon methods?

Resources

Slides: `http://web.eecs.umich.edu/~fessler/papers/files/talk/20/sedona.pdf`

Code: Julia version of MIRT `https://github.com/JeffFessler/MIRT.jl`

Bibliography I

[1] M. Lustig, D. Donoho, and J. M. Pauly.
Sparse MRI: The application of compressed sensing for rapid MR imaging.
Mag. Res. Med., 58(6):1182–95, December 2007.

[2] FDA.
510k premarket notification of HyperSense (GE Medical Systems), 2017.

[3] FDA.
510k premarket notification of Compressed Sensing Cardiac Cine (Siemens), 2017.

[4] FDA.
510k premarket notification of Compressed SENSE, 2018.

[5] D. E. Kuhl and R. Q. Edwards.
Image separation radioisotope scanning.
Radiology, 80(4):653–62, April 1963.

[6] D. A. Chesler.
Three-dimensional activity distribution from multiple positron scintgraphs.
J. Nuc. Med., 12(6):347–8, June 1971.

[7] M. Goitein.
Three-dimensional density reconstruction from a series of two-dimensional projections.
Nucl. Instr. Meth., 101(3):509–18, June 1972.

[8] W. H. Richardson.
Bayesian-based iterative method of image restoration.
J. Opt. Soc. Am., 62(1):55–9, January 1972.

Bibliography II

[9] L. Lucy.
An iterative technique for the rectification of observed distributions.
The Astronomical Journal, 79(6):745–54, June 1974.

[10] A. J. Rockmore and A. Macovski.
A maximum likelihood approach to emission image reconstruction from projections.
IEEE Trans. Nuc. Sci., 23(4):1428–32, August 1976.

[11] L. A. Shepp and Y. Vardi.
Maximum likelihood reconstruction for emission tomography.
IEEE Trans. Med. Imag., 1(2):113–22, October 1982.

[12] S. Geman and D. E. McClure.
Bayesian image analysis: an application to single photon emission tomography.
In *Proc. of Stat. Comp. Sect. of Amer. Stat. Assoc.*, pages 12–8, 1985.

[13] H. M. Hudson and R. S. Larkin.
Accelerated image reconstruction using ordered subsets of projection data.
IEEE Trans. Med. Imag., 13(4):601–9, December 1994.

[14] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar.
High resolution 3D Bayesian image reconstruction using the microPET small-animal scanner.
Phys. Med. Biol., 43(4):1001–14, April 1998.

[15] S. Ahn, S. G. Ross, E. Asma, J. Miao, X. Jin, L. Cheng, S. D. Wollenweber, and R. M. Manjeshwar.
Quantitative comparison of OSEM and penalized likelihood image reconstruction using relative difference penalties for clinical PET.
Phys. Med. Biol., 60(15):5733–52, August 2015.

[16] J. Nuyts, D. Beque, P. Dupont, and L. Mortelmans.
A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography.
IEEE Trans. Nuc. Sci., 49(1-1):56–60, February 2002.

[17] J. Llacer, E. Veklerov, L. R. Baxter, S. T. Grafton, L. K. Griffeth, R. A. Hawkins, C. K. Hoh, J. C. Mazziotta, E. J. Hoffman, and C. E. Metz.
Results of a clinical receiver operating characteristic study comparing filtered backprojection and maximum likelihood estimator images in FDG PET studies.
J. Nuc. Med., 34(7):1198–203, July 1993.

[18] S. R. Meikle, B. F. Hutton, D. L. Bailey, P. K. Hooper, and M. J. Fulham.
Accelerated EM reconstruction in total-body PET: potential for improving tumour detectability.
Phys. Med. Biol., 39(10):1689–794, October 1994.

[19] B. Yang, L. Ying, and J. Tang.
Artificial neural network enhanced Bayesian PET image reconstruction.
IEEE Trans. Med. Imag., 37(6):1297–309, June 2018.

[20] X. Hong, Y. Zan, F. Weng, W. Tao, Q. Peng, and Q. Huang.
Enhancing the image quality via transferred deep residual learning of coarse PET sinograms.
IEEE Trans. Med. Imag., 2018.

[21] K. Kim, D. Wu, K. Gong, J. Dutta, J. H. Kim, Y. D. Son, H. K. Kim, G. E. Fakhri, and Q. Li.
Penalized PET reconstruction using deep learning prior and local linear fitting.
IEEE Trans. Med. Imag., 37(6):1478–87, June 2018.

[22] K. Gong, J. Guan, K. Kim, X. Zhang, J. Yang, Y. Seo, G. E. Fakhri, J. Qi, and Q. Li.
Iterative PET image reconstruction using convolutional neural network representation.
IEEE Trans. Med. Imag., 38(3):675–85, March 2019.

[23] H. Lim, I. Y. Chun, Y. K. Dewaraja, and J. A. Fessler.
Improved low-count quantitative PET reconstruction with a variational neural network, 2019.

[24] I. Haggstrom, C. R. Schmidtlein, G. Campanella, and T. J. Fuchs.
DeepPET: A deep encoder-decoder network for directly solving the PET image reconstruction inverse problem.
Med. Im. Anal., 54:253–62, May 2019.

[25] C. E. Floyd.
An artificial neural network for SPECT image reconstruction.
IEEE Trans. Med. Imag., 10(3):485–7, September 1991.

[26] S. Lam, R. Gupta, H. Kelly, H. D. Curtin, and R. Forghani.
Multiparametric evaluation of head and neck squamous cell carcinoma using a single-source dual-energy CT with fast kVp switching: state of the art.
Cancers, 7(4):2201–16, 2015.

[27] A. J. Mathews, G. Gang, R. Levinson, W. Zbijewski, S. Kawamoto, J. H. Siewerdsen, and J. W. Stayman.
Experimental evaluation of dual multiple aperture devices for fluence field modulated x-ray computed tomography.
In *Proc. SPIE 10132 Medical Imaging: Phys. Med. Im.*, page 101322O, 2017.

[28] D. P. Cormode, S. Si-Mohamed, D. Bar-Ness, M. Sigovan, P. C. Naha, J. Balegamire, F. Lavenne, P. Coulon, E. Roessl, M. Bartels, M. Rokni, I. Blevis, L. Boussel, and P. Douek.
Multicolor spectral photon-counting computed tomography: in vivo dual contrast imaging with a high count rate scanner.
Nature, 7:4784, 2017.

[29] M. J. Muckley, B. Chen, T. Vahle, T. O'Donnell, F. Knoll, A. Sodickson, D. Sodickson, and R. Otazo.
Image reconstruction for interrupted-beam X-ray CT on diagnostic clinical scanners.
Phys. Med. Biol., 64(15):155007, 2019.

[30] G. Hounsfield.
A method of apparatus for examination of a body by radiation such as x-ray or gamma radiation, 1972.
US Patent 1283915. British patent 1283915, London.

[31] S. Kaczmarz.
Angenaherte auflosung von systemen linearer gleichungen.
Bull. Acad. Polon. Sci. Lett. A, 35:355–7, 1937.
Approximate solution to systems of linear equations.

[32] R. Gordon, R. Bender, and G. T. Herman.
Algebraic reconstruction techniques (ART) for the three-dimensional electron microscopy and X-ray photography.
J. Theor. Biol., 29(3):471–81, December 1970.

[33] R. Gordon and G. T. Herman.
Reconstruction of pictures from their projections.
Comm. ACM, 14(12):759–68, December 1971.

[34] G. T. Herman, A. Lent, and S. W. Rowland.
ART: mathematics and applications (a report on the mathematical foundations and on the applicability to real data of the algebraic reconstruction techniques).
J. Theor. Biol., 42(1):1–32, November 1973.

[35] R. Gordon.
A tutorial on ART (algebraic reconstruction techniques).
IEEE Trans. Nuc. Sci., 21(3):78–93, June 1974.

[36] R. L. Kashyap and M. C. Mittal.
Picture reconstruction from projections.
IEEE Trans. Comp., 24(9):915–23, September 1975.

Bibliography VI

[37] A. J. Rockmore and A. Macovski.
A maximum likelihood approach to transmission image reconstruction from projections.
IEEE Trans. Nuc. Sci., 24(3):1929–35, June 1977.

[38] K. Lange and R. Carson.
EM reconstruction algorithms for emission and transmission tomography.
J. Comp. Assisted Tomo., 8(2):306–16, April 1984.

[39] K. Sauer and C. Bouman.
A local update strategy for iterative reconstruction from projections.
IEEE Trans. Sig. Proc., 41(2):534–48, February 1993.

[40] S. H. Manglos, G. M. Gagne, A. Krol, F. D. Thomas, and R. Narayanaswamy.
Transmission maximum-likelihood reconstruction with ordered subsets for cone beam CT.
Phys. Med. Biol., 40(7):1225–41, July 1995.

[41] C. Kamphuis and F. J. Beekman.
Accelerated iterative transmission CT reconstruction using an ordered subsets convex algorithm.
IEEE Trans. Med. Imag., 17(6):1001–5, December 1998.

[42] H. Erdogan and J. A. Fessler.
Ordered subsets algorithms for transmission tomography.
Phys. Med. Biol., 44(11):2835–51, November 1999.

[43] E. Hansis, J. Bredno, D. Sowards-Emmerd, and L. Shao.
Iterative reconstruction for circular cone-beam CT with an offset flat-panel detector.
In *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, pages 2228–31, 2010.

[44] J-B. Thibault, K. Sauer, C. Bouman, and J. Hsieh.
A three-dimensional statistical approach to improved image quality for multi-slice helical CT.
Med. Phys., 34(11):4526–44, November 2007.

[45] D. Kim and J. A. Fessler.
Optimized first-order methods for smooth convex minimization.
Mathematical Programming, 159(1):81–107, September 2016.

[46] D. Kim and J. A. Fessler.
Optimized first-order methods for smooth convex minimization, 2014.

[47] D. Kim and J. A. Fessler.
An optimized first-order method for image restoration.
In *Proc. IEEE Intl. Conf. on Image Processing*, pages 3675–9, 2015.

[48] Y. Drori.
The exact information-based complexity of smooth convex minimization.
J. Complexity, 39:1–16, April 2017.

[49] A. B. Taylor, J. M. Hendrickx, and Francois Glineur.
Smooth strongly convex interpolation and exact worst-case performance of first- order methods.
Mathematical Programming, 161(1):307–45, January 2017.

[50] M. Muckley, D. C. Noll, and J. A. Fessler.
Accelerating SENSE-type MR image reconstruction algorithms with incremental gradients.
In *Proc. Intl. Soc. Mag. Res. Med.*, page 4400, 2014.

Bibliography VIII

[51] D. Kim, S. Ramani, and J. A. Fessler.
Combining ordered subsets and momentum for accelerated X-ray CT image reconstruction.
IEEE Trans. Med. Imag., 34(1):167–78, January 2015.

[52] D. Kim and J. A. Fessler.
Optimized momentum steps for accelerating X-ray CT ordered subsets image reconstruction.
In *Proc. 3rd Intl. Mtg. on Image Formation in X-ray CT*, pages 103–6, 2014.

[53] M. Yavuz and J. A. Fessler.
New statistical models for randoms-precorrected PET scans.
In J Duncan and G Gindi, editors, *Information Processing in Medical Im.*, volume 1230 of *Lecture Notes in Computer Science*, pages 190–203. Springer-Verlag, Berlin, 1997.

[54] S. Ye, S. Ravishankar, Y. Long, and J. A. Fessler.
Adaptive sparse modeling and shifted-Poisson likelihood based approach for low-dose CT image reconstruction.
In *Proc. IEEE Wkshp. Machine Learning for Signal Proc.*, pages 1–6, 2017.

[55] S. Ye, S. Ravishankar, Y. Long, and J. A. Fessler.
SPULTRA: low-dose CT image reconstruction with joint statistical and learned image models.
IEEE Trans. Med. Imag., 2019.
To appear.

[56] Q. Ding, Y. Long, X. Zhang, and J. A. Fessler.
Modeling mixed Poisson-Gaussian noise in statistical image reconstruction for X-ray CT.
In *Proc. 4th Intl. Mtg. on Image Formation in X-ray CT*, pages 399–402, 2016.

Bibliography IX

[57] Q. Ding, Y. Long, X. Zhang, and J. A. Fessler.
Statistical image reconstruction using mixed Poisson-Gaussian noise model for X-ray CT.
Inverse Prob. and Imaging, 2019.
Submitted.

[58] B. R. Whiting.
Signal statistics in x-ray computed tomography.
In *Proc. SPIE 4682 Medical Imaging: Med. Phys.*, pages 53–60, 2002.

[59] J. Xu and B. M. W. Tsui.
Electronic noise modeling in statistical iterative reconstruction.
IEEE Trans. Im. Proc., 18(6):1228–38, June 2009.

[60] S. Ravishankar, J. C. Ye, and J. A. Fessler.
Image reconstruction: from sparsity to data-adaptive methods and machine learning.
Proc. IEEE, 108(1):86–109, January 2020.

[61] I. Y. Chun and J. A. Fessler.
Convolutional dictionary learning: acceleration and convergence.
IEEE Trans. Im. Proc., 27(4):1697–712, April 2018.

[62] I. Y. Chun and J. A. Fessler.
Convolutional analysis operator learning: acceleration and convergence.
IEEE Trans. Im. Proc., 29(1):2108–22, January 2020.

[63] C. Ailong, L. Lei, Z. Zhizhong, W. Linyuan, and Y. Bin.
Block-matching sparsity regularization-based image reconstruction for low-dose computed tomography.
Med. Phys., 45(6):2439–52, June 2018.

[64] B. Zhao, H. Ding, Y. Lu, G. Wang, J. Zhao, and S. Molloi.
Dual-dictionary learning-based iterative image reconstruction for spectral computed tomography application.
Phys. Med. Biol., 57(24):8217–30, December 2012.

[65] Y. Zhang, X. Mou, G. Wang, and H. Yu.
Tensor-based dictionary learning for spectral CT reconstruction.
IEEE Trans. Med. Imag., 36(1):142–54, January 2017.

[66] M. Aharon, M. Elad, and A. Bruckstein.
K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation.
IEEE Trans. Sig. Proc., 54(11):4311–22, November 2006.

[67] S. Ravishankar and Y. Bresler.
l_0 sparsifying transform learning with efficient optimal updates and convergence guarantees.
IEEE Trans. Sig. Proc., 63(9):2389–404, May 2015.

[68] X. Zheng, S. Ravishankar, Y. Long, and J. A. Fessler.
PWLS-ULTRA: An efficient clustering and learning-based approach for low-dose 3D CT image reconstruction.
IEEE Trans. Med. Imag., 37(6):1498–510, June 2018.

[69] H. Nien and J. A. Fessler.
Relaxed linearized algorithms for faster X-ray CT image reconstruction.
IEEE Trans. Med. Imag., 35(4):1090–8, April 2016.

[70] S. Ravishankar and Y. Bresler.
Data-driven learning of a union of sparsifying transforms model for blind compressed sensing.
IEEE Trans. Computational Imaging, 2(3):294–309, September 2016.

[71] H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, and G. Wang.
Low-dose CT denoising with convolutional neural network, 2016.

[72] Y. S. Han, J. Yoo, and J. C. Ye.
Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis, 2016.

[73] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang.
Low-dose CT with a residual encoder-decoder convolutional neural network (RED-CNN).
IEEE Trans. Med. Imag., 36(12):2524–35, December 2017.

[74] E. Kang, J. Min, and J. C. Ye.
A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction.
Med. Phys., 44(10):e360–75, October 2017.

[75] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser.
Deep convolutional neural network for inverse problems in imaging.
IEEE Trans. Im. Proc., 26(9):4509–22, September 2017.

[76] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Isgum.
Generative adversarial networks for noise reduction in low-dose CT.
IEEE Trans. Med. Imag., 36(12):2536–45, December 2017.

[77] E. Kang, H. J. Koo, D. H. Yang, J. B. Seo, and J. C. Ye.
Cycle-consistent adversarial denoising network for multiphase coronary CT angiography.
Med. Phys., 46(2):550–62, February 2019.

[78] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang.
Low dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss.
IEEE Trans. Med. Imag., 37(6):1348–57, June 2018.

[79] Y. Zhang and H. Yu.
Convolutional neural network based metal artifact reduction in X-ray computed tomography.
IEEE Trans. Med. Imag., 37(6):1370–81, June 2018.

[80] M. U. Ghani and W. C. Karl.
Fast accurate CT metal artifact reduction using data domain deep learning, 2019.

[81] M. Argyrou, D. Maintas, C. Tsoumpas, and E. Stiliaris.
Tomographic image reconstruction based on artificial neural network (ANN) techniques.
In *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, pages 3324–7, 2012.

[82] E. Haneda, B. Claus, P. FitzGerald, G. Wang, and B. De Man.
CT sinogram analysis using deep learning.
In *Proc. 5th Intl. Mtg. on Image Formation in X-ray CT*, pages 419–22, 2018.

[83] Q. De Man, E. Haneda, B. Claus, P. Fitzgerald, B. De Man, G. Qian, H. Shan, J. Min, M. Sabuncu, and G. Wang.
A two-dimensional feasibility study of deep learning-based feature detection and characterization directly from CT sinograms.
Med. Phys., 46(12):e790–800, December 2019.

[84] D. Wu, K. Kim, G. E. Fakhri, and Q. Li.
Iterative low-dose CT reconstruction with priors trained by artificial neural network.
IEEE Trans. Med. Imag., 36(12):2479–86, December 2017.

[85] J. Adler and O. Oktem.
Learned primal-dual reconstruction.
IEEE Trans. Med. Imag., 37(6):1322–32, June 2018.

[86] H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang.
LEARN: Learned experts: assessment-based reconstruction network for sparse-data CT.
IEEE Trans. Med. Imag., 37(6):1333–47, June 2018.

[87] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser.
CNN-based projected gradient descent for consistent image reconstruction.
IEEE Trans. Med. Imag., 37(6):1440–53, June 2018.

[88] C. Shen, Y. Gonzalez, L. Chen, S. B. Jiang, and X. Jia.
Intelligent parameter tuning in optimization-based iterative CT reconstruction via deep reinforcement learning.
IEEE Trans. Med. Imag., 37(6):1430–9, June 2018.

[89] H. Shan, A. Padole, F. Homayounieh, U. Kruger, R. D. Khera, C. Nitiwarangkul, M. K. Kalra, and G. Wang.
Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction.
Nature Mach. Intel., 1(6):269–76, 2019.

[90] I. Y. Chun, Z. Huang, H. Lim, and J. A. Fessler.
Momentum-Net: Fast and convergent iterative neural network for inverse problems, 2019.

[91] FDA.
510k premarket notification of AiCE Deep Learning Reconstruction (Canon), 2019.

[92] FDA.
510k premarket notification of Deep Learning Image Reconstruction (GE Medical Systems), 2019.

[93] D. H. Ye, S. Srivastava, J. Thibault, K. Sauer, and C. Bouman.
Deep residual learning for model-based iterative CT reconstruction using plug-and-play framework.
In *Proc. IEEE Conf. Acoust. Speech Sig. Proc.*, pages 6668–72, 2018.

[94] A. Ziabari, D. H. Ye, S. Srivastava, K. D. Sauer, J. Thibault, and C. A. Bouman.
2.5D deep learning for CT image reconstruction using A multi-GPU implementation.
In *asccs*, pages 2044–9, 2018.

[95] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg.
Plug-and-play priors for model based reconstruction.
In *IEEE GlobalSIP*, pages 945–8, 2013.

[96] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman.
Plug-and-play priors for bright field electron tomography and sparse interpolation.
IEEE Trans. Computational Imaging, 2(4):408–23, December 2016.