Adaptive Reconstruction Methods for Low-Dose Computed Tomography

Joseph Shtok

Ph.D. supervisors: Prof. Michael Elad, Dr. Michael Zibulevsky.

Technion IIT, Computer Science dept.

Israel, 2011

Contents of this talk

Intro to Computed Tomography	• Scan model, Noise, Local reconstruction
Framework of adaptive reconstruction	• General scheme of supervised learning
Adaptive FBP	• Learned FBP filter for local reconstruction
Sparsity-based sinogram restoration	• Adaptation of K-SVD to low-dose CT reconstruction
Learned shrinkage in a transform domain	• Adaptation of the method to low-dose CT reconstruction
Performance boosting of existing algorithms	• Local fusion of multiple versions of the algorithm output
Then you get tired of me.	

Short Intro to Computed Tomography

And as usual you have scheduled a CT scan for today.

Short Intro to Computed Tomography

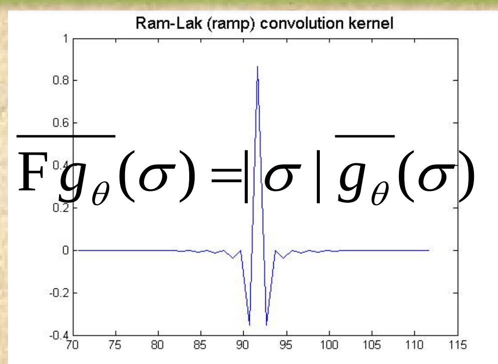
Filtered Back-Projection (FBP)

Back-Projection

Ram-Lak filter

Sum contributions to x from all incident rays

$$\left(\mathbf{R}^* g\right)x = \int_\theta g_\theta(x \cdot \theta)d\theta$$



$$\overline{\mathrm{F} g_\theta(\sigma)} = |\sigma| \, \overline{g_\theta(\sigma)}$$

X-ray source

λ_0 photons

$f(x)$

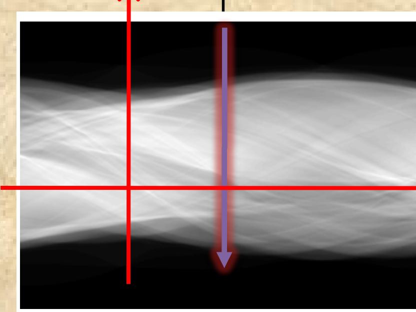
Detectors

Line l

photon counts y_l

Sinogram values

$$g_l = -\log\left(\frac{y_l}{\lambda_0}\right)$$



$$g_l = [\mathbf{R}\, f]_l = \int_l f(x)dl$$

Radon (X-ray) transform

Noise in Low-Dose Reconstruction

Accepted model for detector measurements (similar to one in CCD sensors):

$y_l \xleftarrow{\text{instance}} Y_l \sim Poisson(\lambda_l) + \mathcal{N}(0, \sigma_n)$

$\lambda_l = \lambda_0 e^{-[\mathrm{R}\, f]_l}$ - ideal count

Poor photon statistics due to low counts

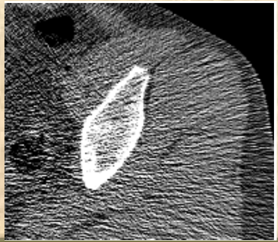
Electronic noise in the hardware

$\overline{Y_l} = Y_l + \sigma_n^2 \approx Poisson(\lambda_l + \sigma_n^2) \xrightarrow{\text{instance}} \overline{y}_l \quad \mathbf{var}(\overline{y}_l) = \lambda_l + \sigma_n^2$

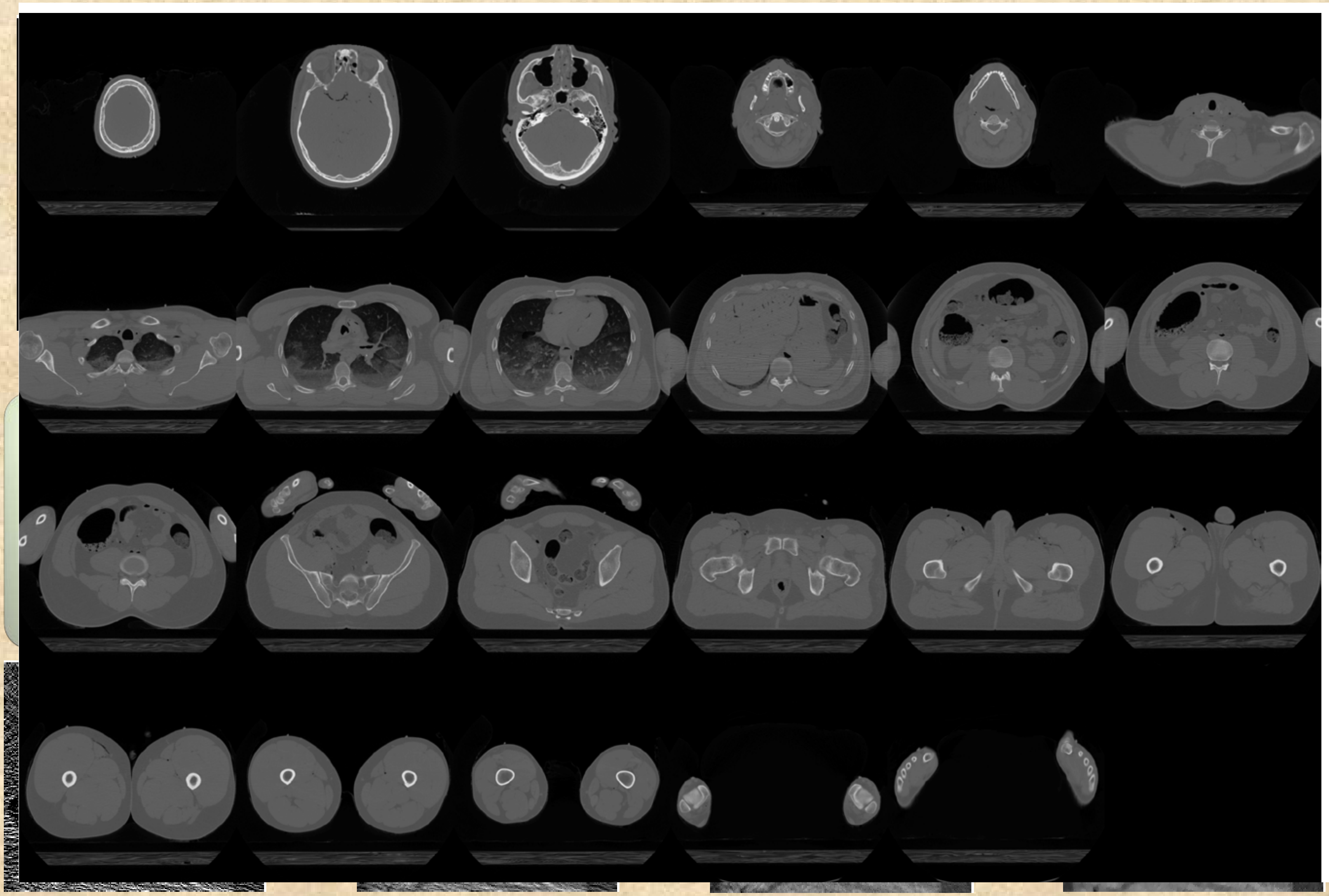
$Z_l = Anscombe(\overline{Y_l}) = \sqrt{\overline{Y_l} + 3/8} \xrightarrow{\text{instance}} z_l \quad \mathbf{var}(z_l) = 1$

Large attenuation → High integral value → Low count → High noise variance → Streak artifacts

$g_l = \int_l f(x)dl \quad y_l = I_0 e^{-g_l} \quad \mathbf{var}(g_l) = y_l^{-1}$

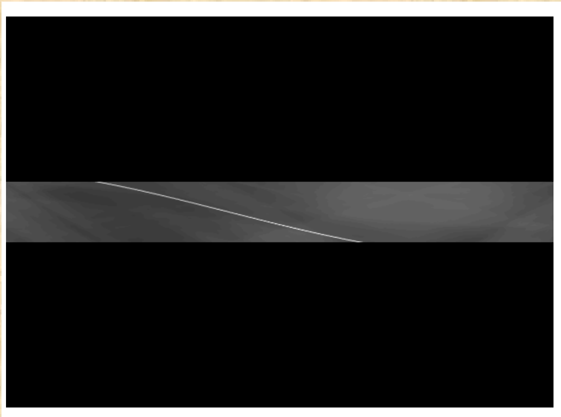


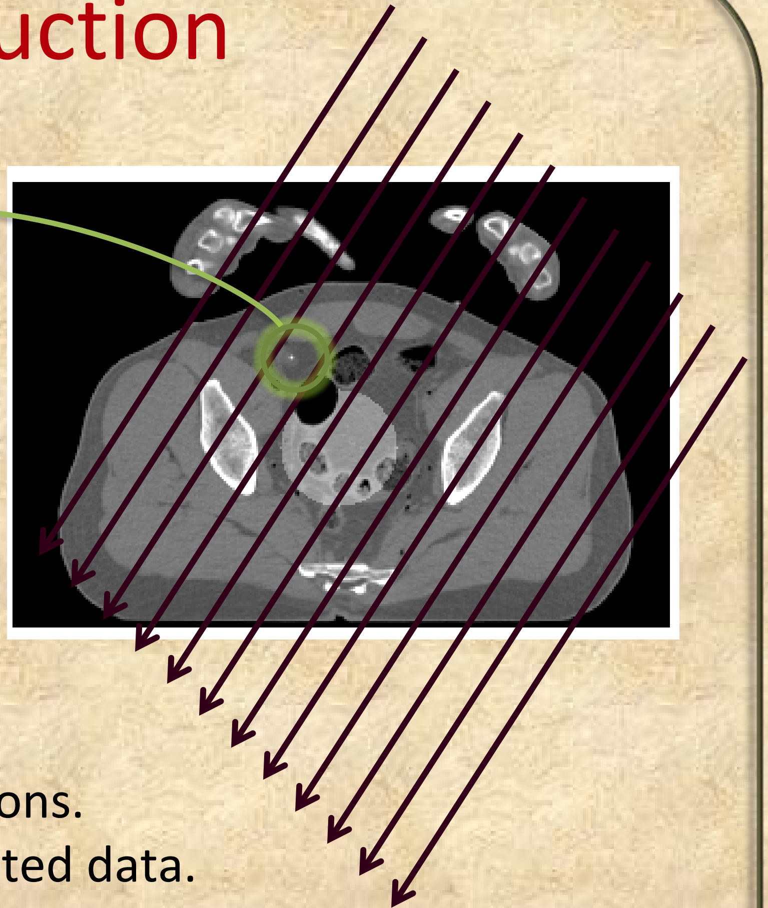
Noise in Low-Dose Reconstruction



Problem of local reconstruction

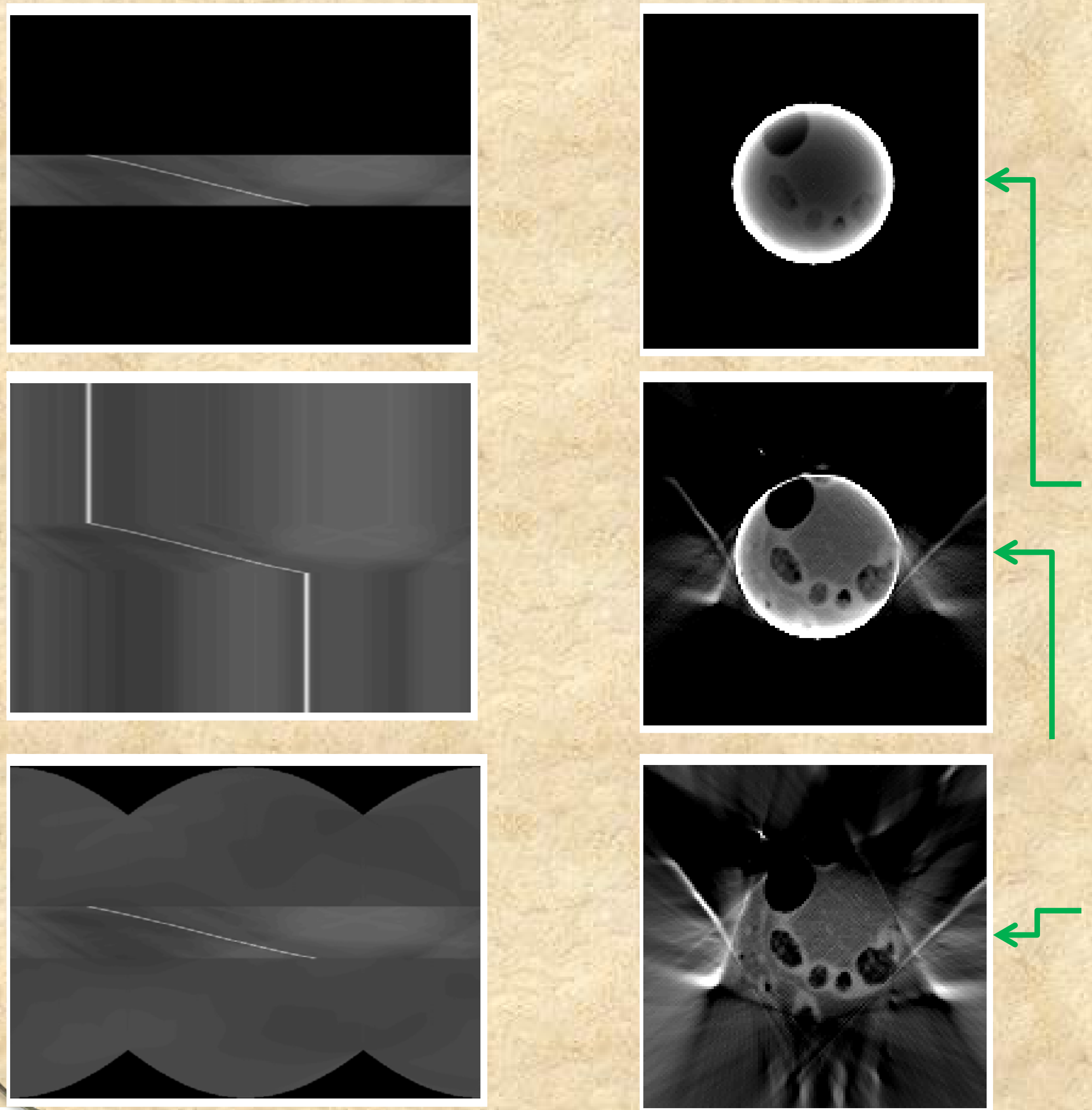
A point in the image draws a sine.





Points outside the ROI contribute to its projections.
ROI is not uniquely determined from the truncated data.

Problem of local reconstruction



FBP reconstruction from zero-padded truncated projections

Basic sinogram completion: duplicate the margins.

Non-linear sine-based sinogram completion

Intro to Computed Tomography	• Scan model, Noise, Local reconstruction
Framework of adaptive reconstruction	• General scheme of supervised learning
Adaptive FBP	• Learned FBP filter for local reconstruction
Sparsity-based sinogram restoration	• Adaptation of K-SVD to low-dose CT reconstruction
Learned shrinkage in a transform domain	• Adaptation of the method to low-dose CT reconstruction
Performance boosting of existing algorithms	• Local fusion of multiple versions of the algorithm output

Error measure for CT reconstruction

$f -$ reference image $\quad \tilde{f} -$ reconstructed image

Basic error measure: Mean Square Error (MSE)

$$\varphi_1(\tilde{f}) = \sum_x \left(f(x) - \tilde{f}(x)\right)^2 = \left\| f - \tilde{f} \right\|_2^2$$

Problem: MSE can be reduced by blurring the image.

Sharpness-promoting penalty: the gradient norm in $\tilde{f}$ should not fall below the gradient norm in f.

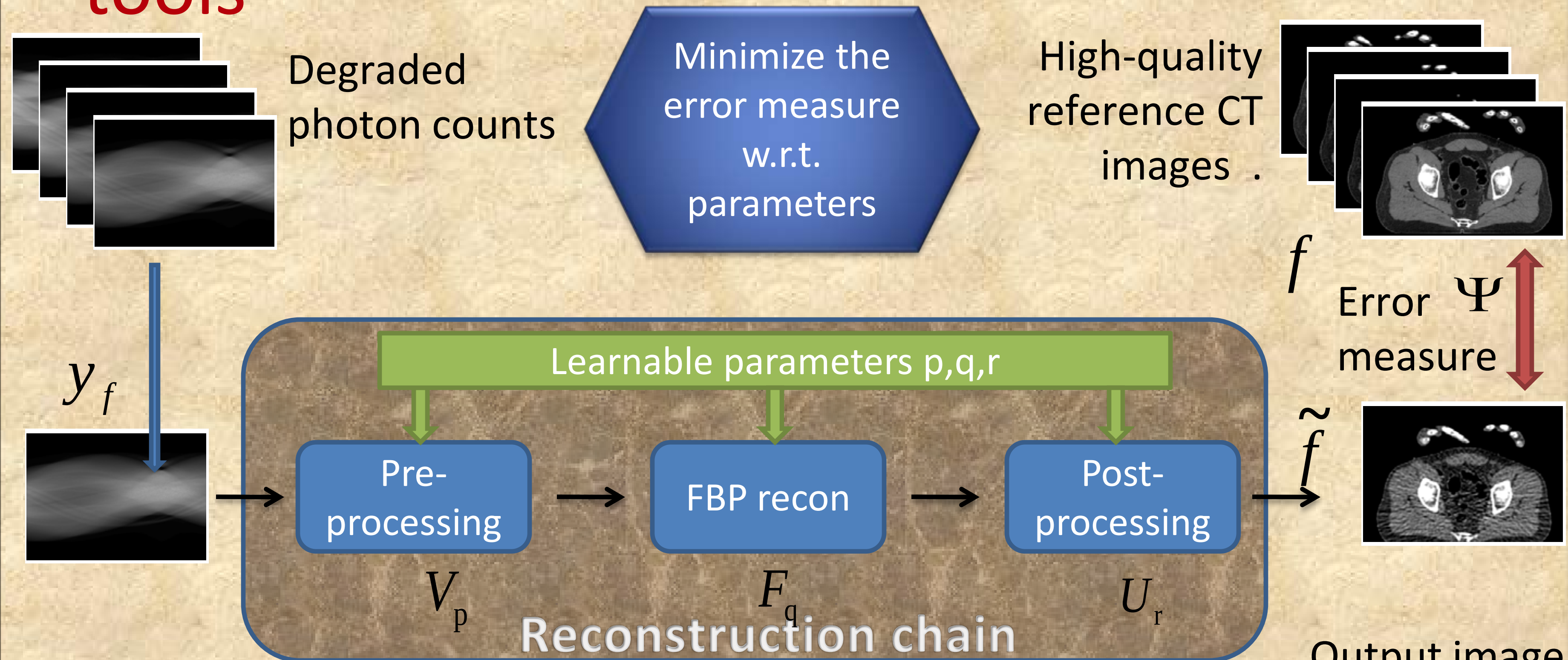
$$\varphi_2(\tilde{f}) = \left\| f - \tilde{f} \right\|_2^2 + \mu\left(J - \tilde{J}\right)_+$$

$$J = \left\| \nabla_x f \right\|_2^2, \quad \tilde{J} = \left\| \nabla_x \tilde{f} \right\|_2^2$$

Nuances:

- The MSE component is restricted to regions of interest
- The gradient-based component is restricted to fine edges.
- The non-negativity function $(\ )_+$ is smoothed for better optimization.

Supervised learning of adaptive processing tools



$$\Psi(\mathrm{p},\mathrm{q},\mathrm{r}) = \sum_{image\ f} \left\| f - U_{\mathrm{r}} F_{\mathrm{q}} V_{\mathrm{p}}(y_f) \right\|_2^2 + \mu\left(J - \tilde{J}\right)_+$$

Output image - a function of p, q, r.

$$J = \left\|\nabla_x f\right\|_2^2, \quad \tilde{J} = \left\|\nabla_x \tilde{f}\right\|_2^2$$

Intro to Computed Tomography	• Scan model, Noise, Local reconstruction
Framework of adaptive reconstruction	• General scheme of supervised learning
Adaptive FBP	• Learned FBP filter for local reconstruction
Sparsity-based sinogram restoration	• Adaptation of K-SVD to low-dose CT reconstruction
Learned shrinkage in a transform domain	• Adaptation of the method to low-dose CT reconstruction
Performance boosting of existing algorithms	• Local fusion of multiple versions of the algorithm output

Learned FBP filter for ROI reconstruction

FBP operator: $\mathbf{T}_\kappa(g) = \mathbf{R}^*(\kappa * g).$

Train the convolution kernel κ to pursuit reconstruction goals.

Training objective for ROI reconstruction:

$$\Psi(\kappa) = \left\| \mathbf{T}_\kappa(g_f) - f \right\|_{2,Q}^2$$

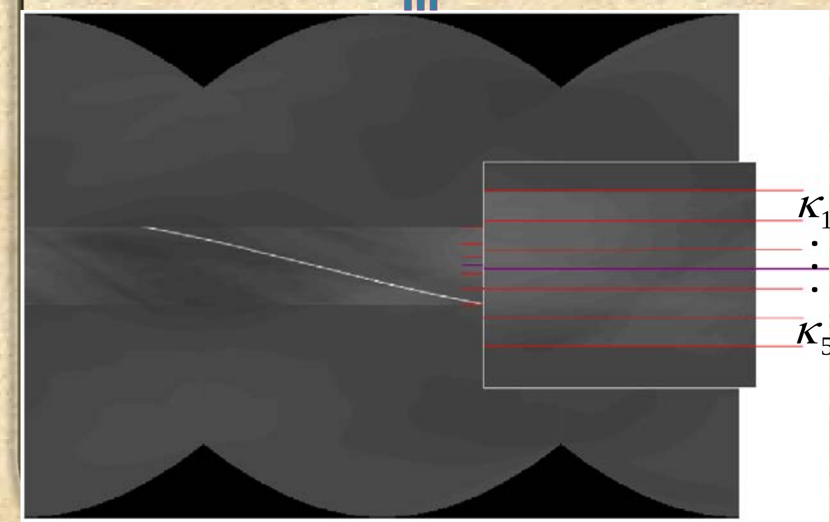
$\{\kappa_1, \dots, \kappa_5\}$

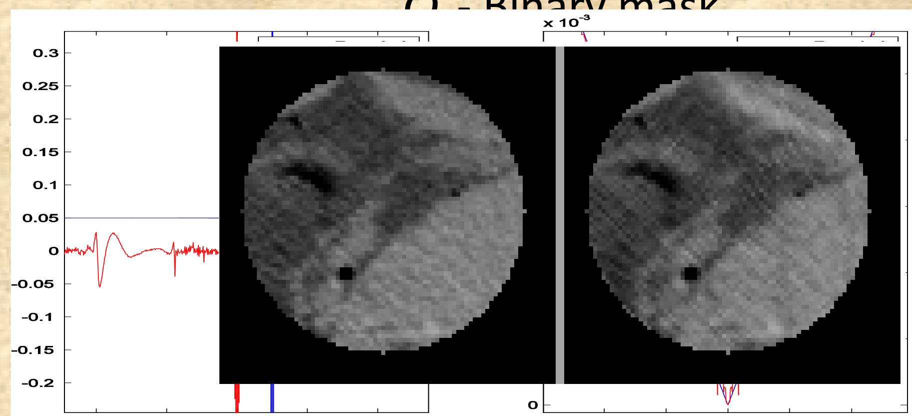
g_f -Truncated sinogram after completion.

Q - Binary mask

Truncated sinogram with completion

1. Use 2-D kernel.





Reference AFBP reconstruction

2. Filter the sinogram with radially-variant convolution kernel.

ROI reconstruction

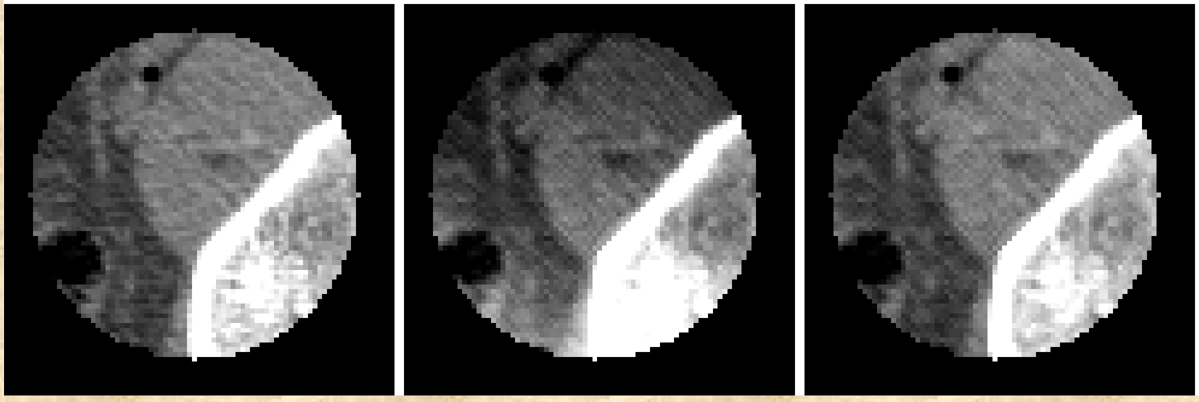
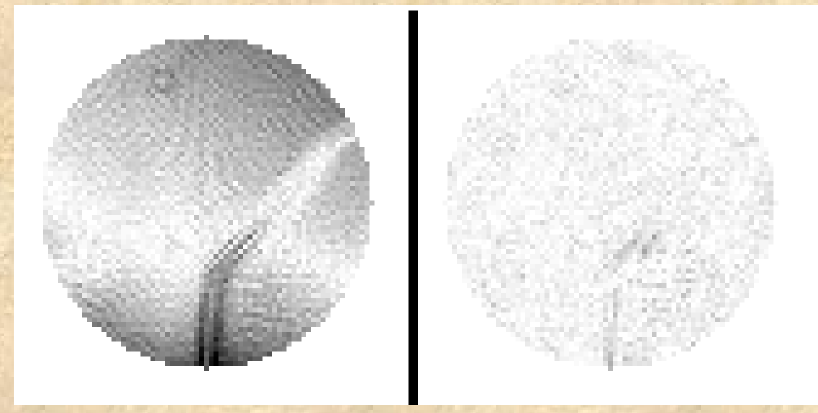


Image size = 461 pixels.
ROI radius = 34 pixels,
Margin = 3 pixels.



True ROI image	FBP reconstruction 22.9 dB	AFBP reconstruction 34.68 dB

ROI reconstruction

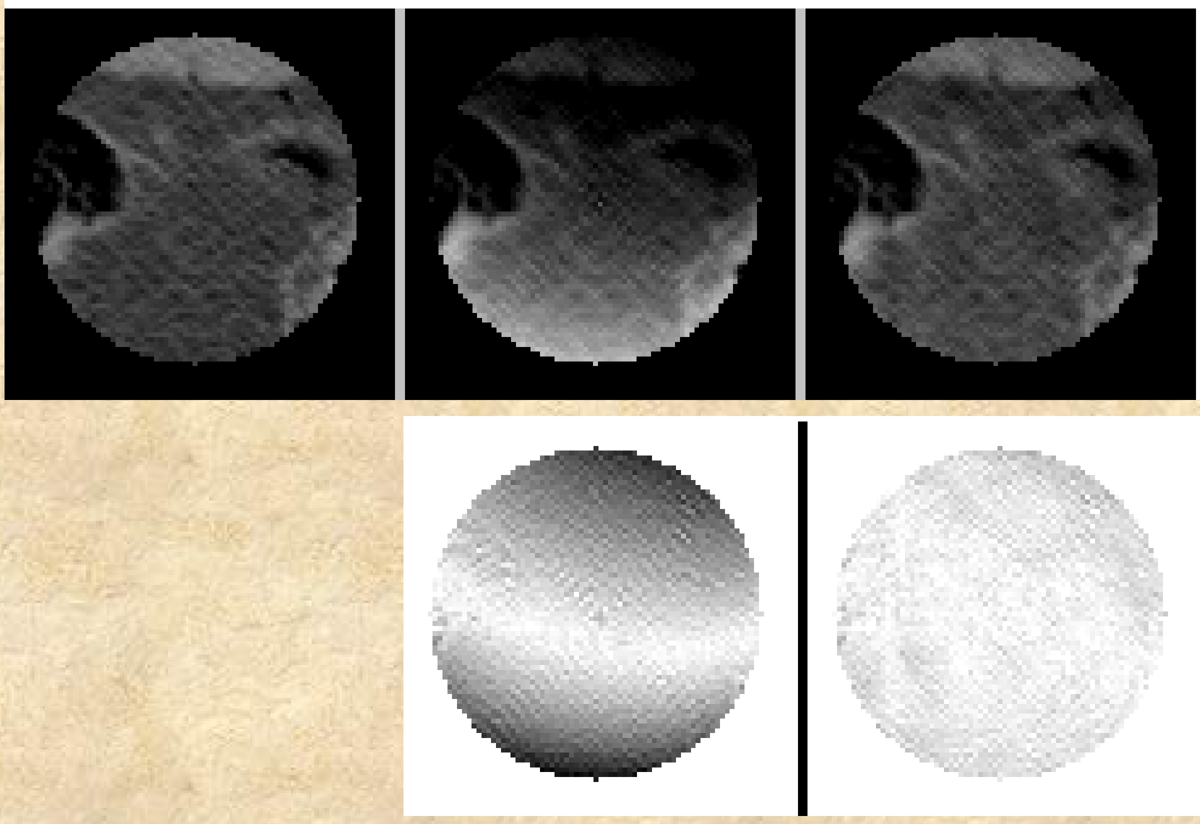


Image size = 461 pixels.
ROI radius = 34 pixels,
Margin = 3 pixels.

True ROI image	FBP reconstruction 18.04 dB	AFBP reconstruction 29.63 dB

ROI reconstruction

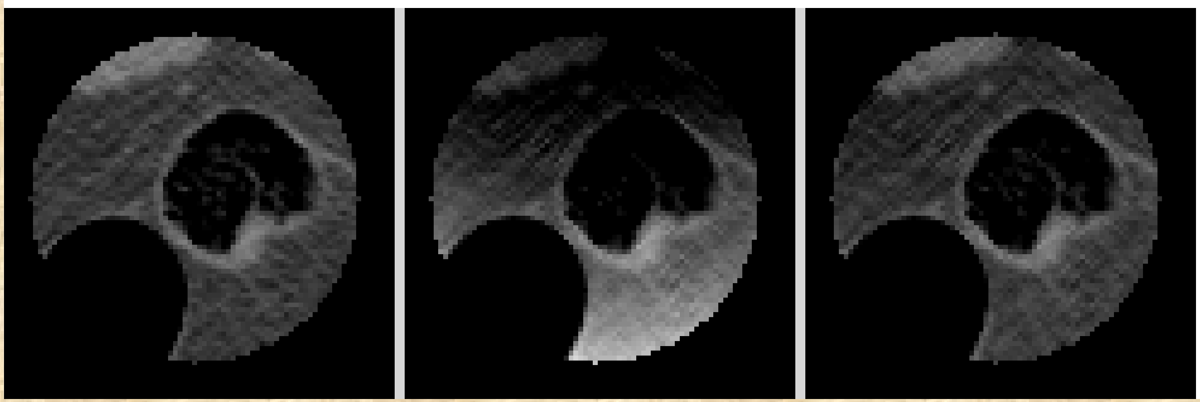
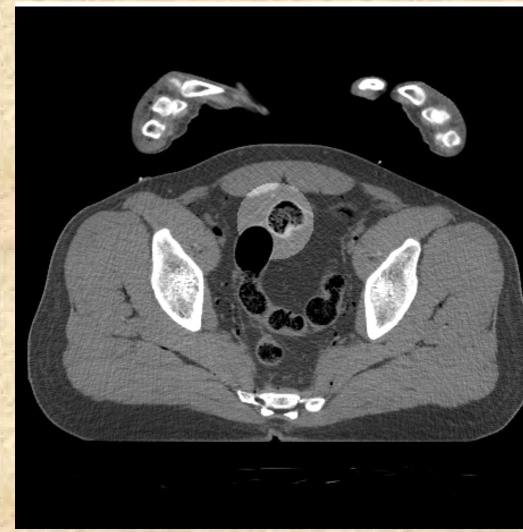
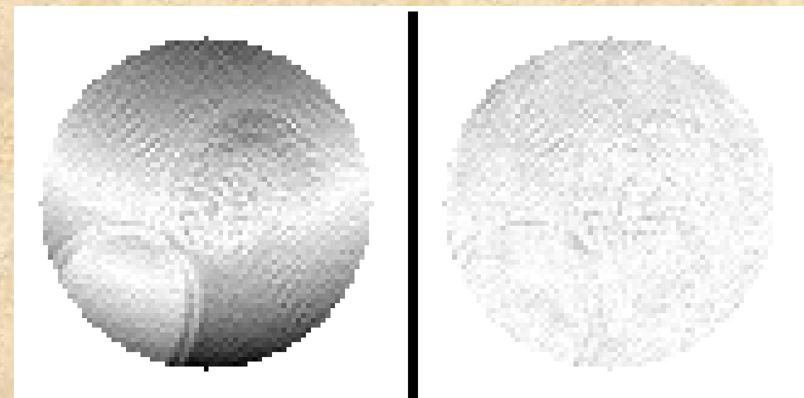


Image size = 461 pixels.
ROI radius = 34 pixels,
Margin = 3 pixels.





True ROI image	FBP reconstruction 19.48 dB	AFBP reconstruction 31.44 dB

Intro to Computed Tomography	• Scan model, Noise, Local reconstruction
Framework of adaptive reconstruction	• General scheme of supervised learning
Adaptive FBP	• Learned FBP filter for local reconstruction
Sparsity-based sinogram restoration	• Adaptation of K-SVD to low-dose CT reconstruction
Learned shrinkage in a transform domain	• Adaptation of the method to low-dose CT reconstruction
Performance boosting of existing algorithms	• Local fusion of multiple versions of the algorithm output

Sparse-Land model for signals

The concept: natural signals admit a faithful representation using only few columns (atoms) from a dedicated overcomplete dictionary.

Natural dictionaries: Wavelets, Haar functions, Discrete Cosines, Fourier.

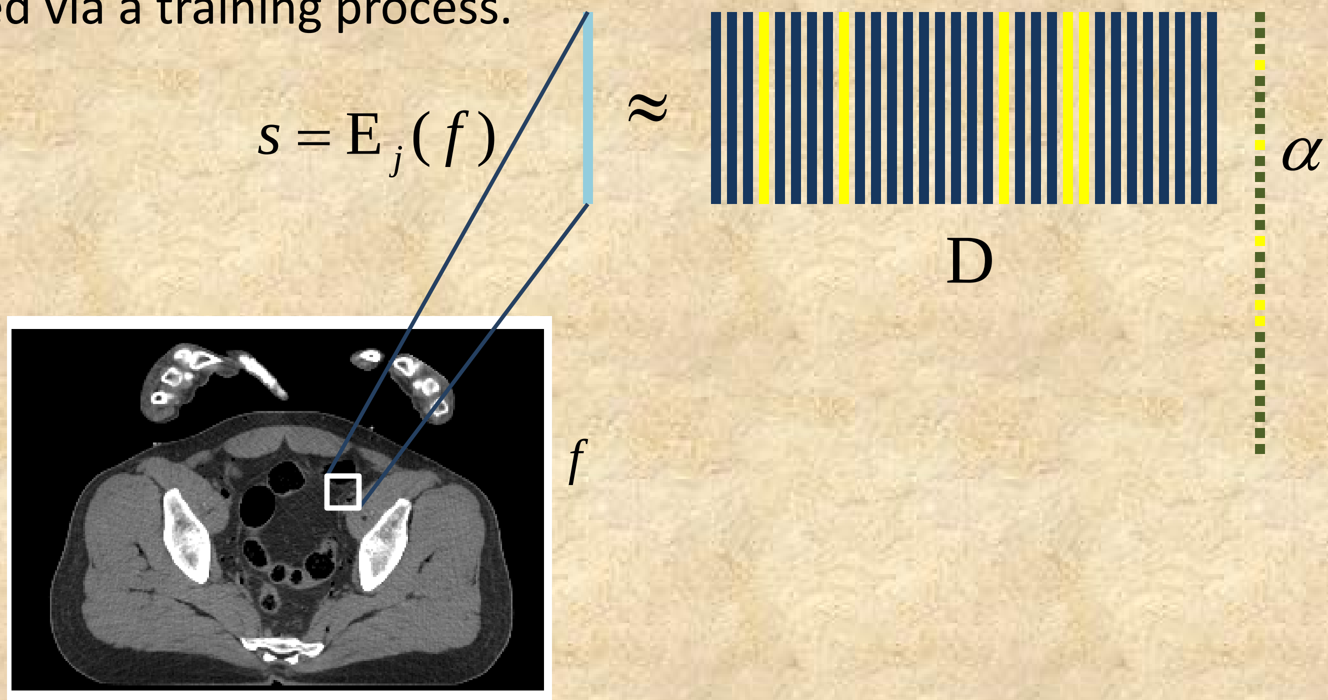
Dictionaries tailored to the specific family of signals: obtained via a training process.

$\|\alpha\|_0 \le k$ Number of non-zeros is small

$\|v\|_2 \le \varepsilon$ Residual is small

$$s + v = \mathrm{D}\alpha$$

$$s = \mathrm{E}_j(f)$$



Sparse-Land model for signals

Denoising technique (Elad, Aharon, 2006):

$$\Phi(\mathrm{D}, f, \alpha) = \delta \left\| f - \tilde{f} \right\|_2^2 + \sum_{patch\ j} \mu_j \left\| \alpha_j \right\|_0 + \sum_{patch\ j} \left\| \mathrm{D}\alpha_j - \mathrm{E}_j f \right\|_2^2$$

$\tilde{f}$ - Noisy image

Minimize w.r.t f

Minimize w.r.t $\mathbf{D}, \{\alpha_j\}$

1. (**K-SVD**) Train a dictionary D along with sparse representations $\{\alpha\}$.
2. Compute the image estimate (closed-form solution).

Sparse coding: $\alpha_j = \mathbf{arg}\min_{\alpha} \left\| \alpha \right\|_0 \quad s.t. \left\| \mathbf{D}\alpha_j - \mathbf{E}_j f \right\|_2^2 \le \varepsilon_j$

Dictionary update: $d_i = \arg\min_{d} \sum_{patch\ j} \left\| \mathrm{D}\alpha_j - \mathrm{E}_j f \right\|_2^2$ Variable d is the i-th column in $\mathbf{D}$.

- State-of-the-art noise reduction.
- Adaptive to current image or training set.
- Uniform noise assumption.

Application to CT reconstruction

Previous work (Liao, Sapiro, 2007):

$$\left\{\mathrm{D}^*, f^*, \alpha^*\right\} = \arg\min_{\mathrm{D}, f, \alpha} \left\{ \delta \left\| \mathrm{R} f - \tilde{g} \right\|_2^2 + \sum_{patch\ j} \mu_j \left\| \alpha_j \right\|_0 + \sum_{patch\ j} \left\| \mathrm{D} \alpha_j - \mathrm{E}_j f \right\|_2^2 \right\}$$

- Patch-wise sparse coding of CT image f.
- Online learning from noisy data.
- Very nice results on geometric images under severe angular subsampling.

Drawbacks:

- Data fidelity term in the sinogram domain.
- No reference to statistical model of the noise.
- Sparse coding thresholds not treated.

Application to CT reconstruction

Our approach:

1. check data fidelity and perform sparse coding in the domain of noise-normalized raw data : $\tilde{z} = \sqrt{(\tilde{y} + \sigma_n^2) + 3/8}$

$$\left\{ \mathrm{D}_1^*, \alpha^*, z \right\} = \arg\min_{\mathrm{D}_1, \alpha, z} \left\{ \lambda \left\| z - \tilde{z} \right\|_2^2 + \mu \sum_{patch\ j} \left\| \alpha_j \right\|_0 + \sum_{patch\ j} \left\| \mathrm{D}_1 \alpha_j - \mathrm{E}_j \tilde{z} \right\|_2^2 \right\}$$

Solve for D_1, α using K-SVD, but allow to use a different dictionary D_2 at restoration stage:

$$z^* = \arg\min_z \left\{ \lambda \left\| z - \tilde{z} \right\|_2^2 + \sum_{patch\ j} \left\| \mathrm{D}_2 \alpha_j - \mathrm{E}_j \tilde{z} \right\|_2^2 \right\} =$$

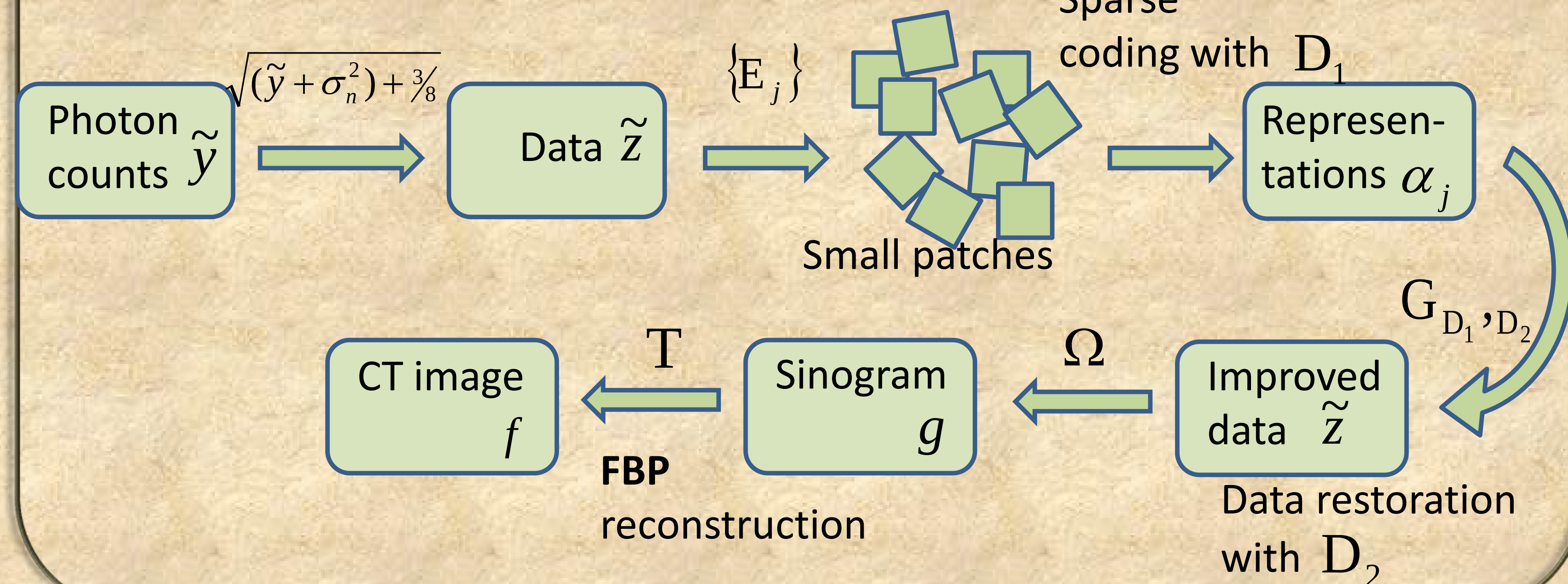
$$= G_{\mathrm{D}_1, \mathrm{D}_2}(\alpha) \equiv \left(\sum_j \mathrm{E}_j^T \mathrm{E}_j \right)^{-1} \left(\sum_{patch\ j} \mathrm{E}_j^T \mathrm{D}_2 \alpha_j + \lambda \tilde{z} \right)$$

Application to CT reconstruction

2. Train a second dictionary D_2 optimized for image reconstruction using a designed error measure and pre-computed repersentations α :

$$D_2^* = \arg\min_{D_2} \left\| f - T\,\Omega\,G_{D_1,D_2}\,\alpha \right\|_2^2 + \mu (J - \tilde{J})_+ \qquad \Omega : z \to y \to g$$

Reconstruction chain:



Compared algorithms

Adaptive Trimmed Mean (ATM) Filter *Hsieh, '98.*

- Extract M values from the neighborhood of a photon count y_l
- Remove $2\alpha \mathrm{M}$ extreme values and compute the average of the rest.
- M, α are data-dependent; computed through

$$\mathrm{M}(y_l) = \frac{2\beta\lambda}{2\lambda + [y_l - \delta]_+}, \qquad \alpha(y_l) = \frac{\alpha_m y_l}{\lambda}.$$

In our experience: depends highly on the parameters.

Penalized Weighted Least Squares (PWLS) *Elbakri, Fessler, '02.*

2-nd order aproximation of a penalized log-likelihood expression for photon counts data:

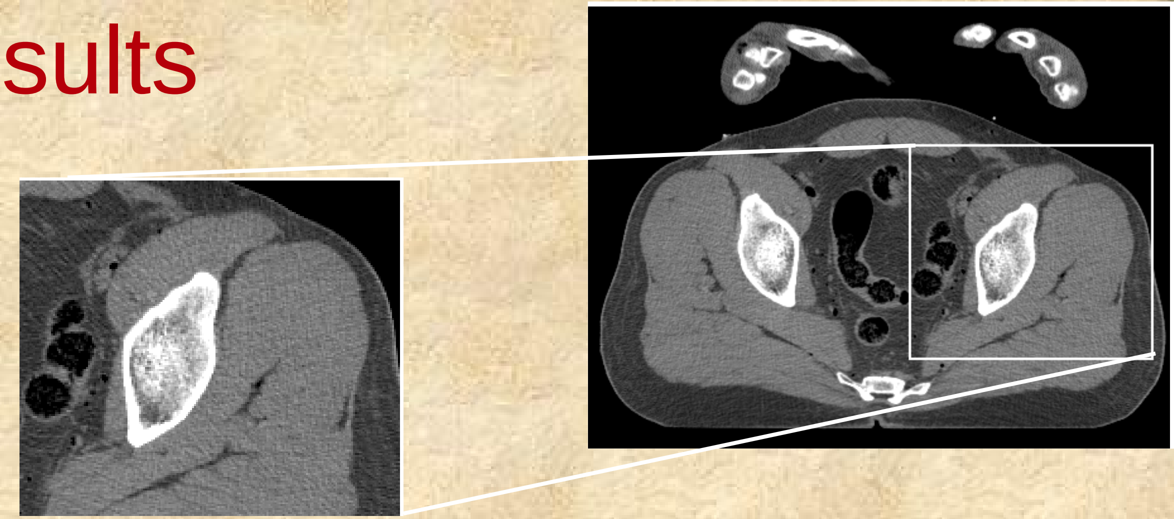
$$PWLS(y \mid f) = \tfrac{1}{2}\sum_l W_l([\mathrm{R}\, f]_l - g_l)^2 + \lambda \sum_p \sum_{k \in N(p)} \psi(f_p - f_k)$$

Works quite well.

Penalty weight. Controls variance-resolution tradeoff. (λ)

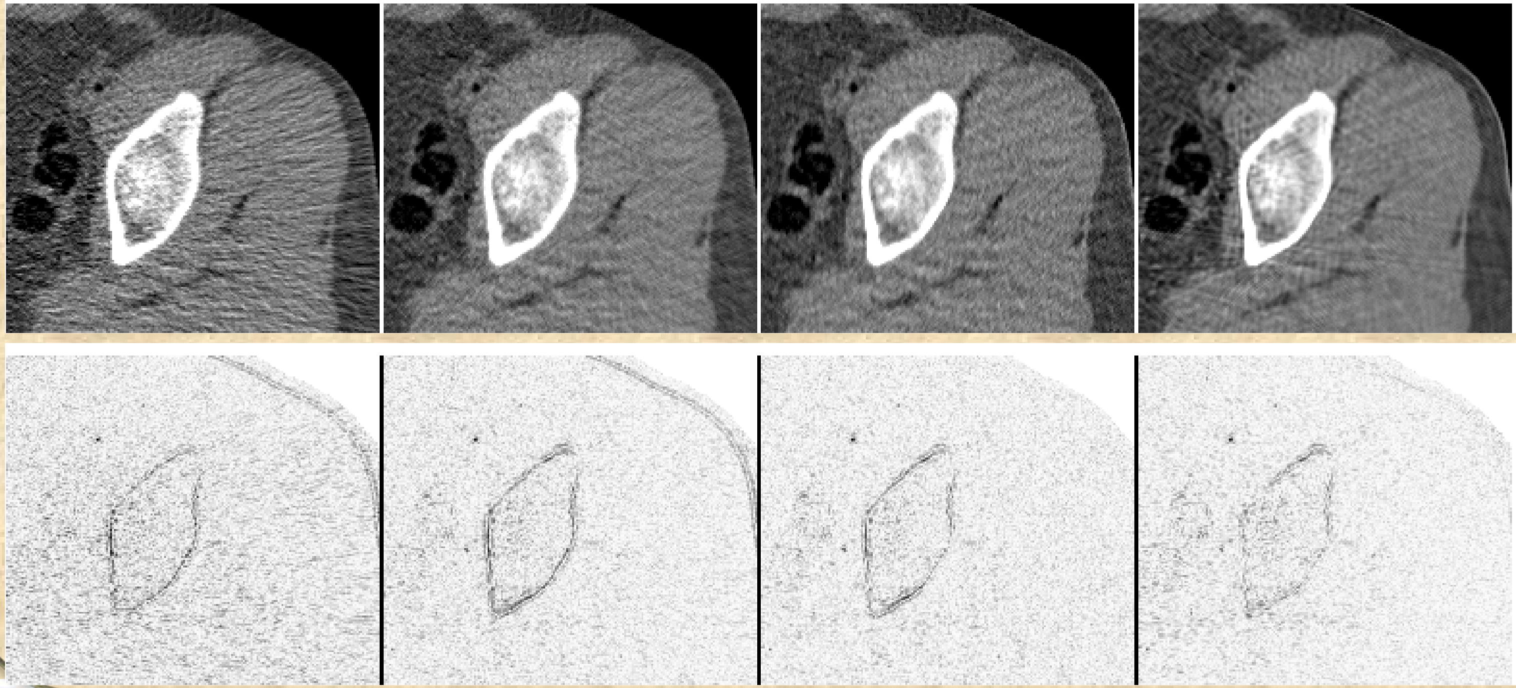
Huber penalty (smoothed L_1 norm). (ψ)

Empirical results



Thighs section

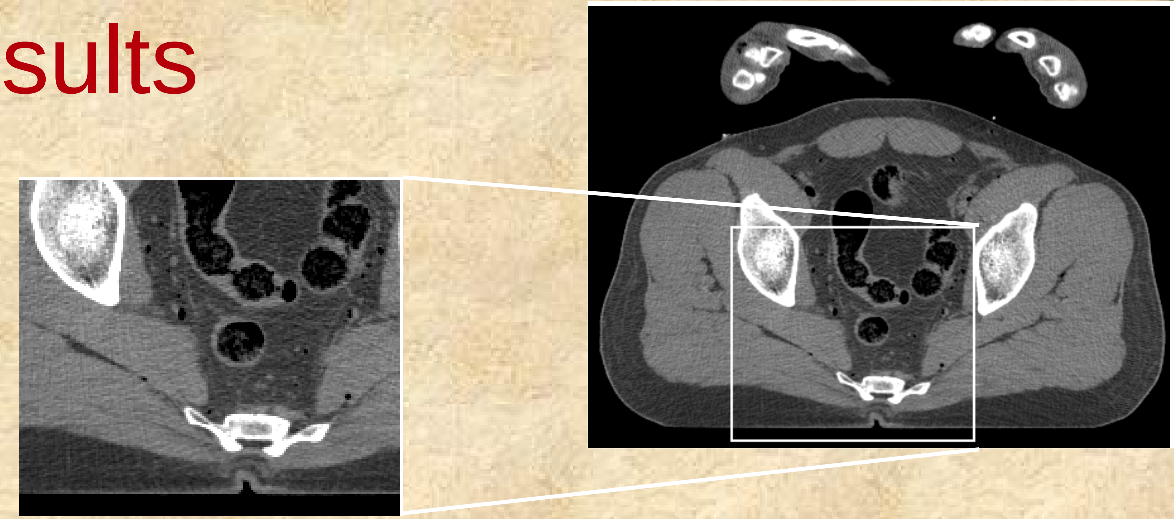
FBP, 25.76 dB	ATM, 28.32 dB	PWLS, 28.90 dB	Sparse, 29.62 dB



Recon. in [-220,350] HU

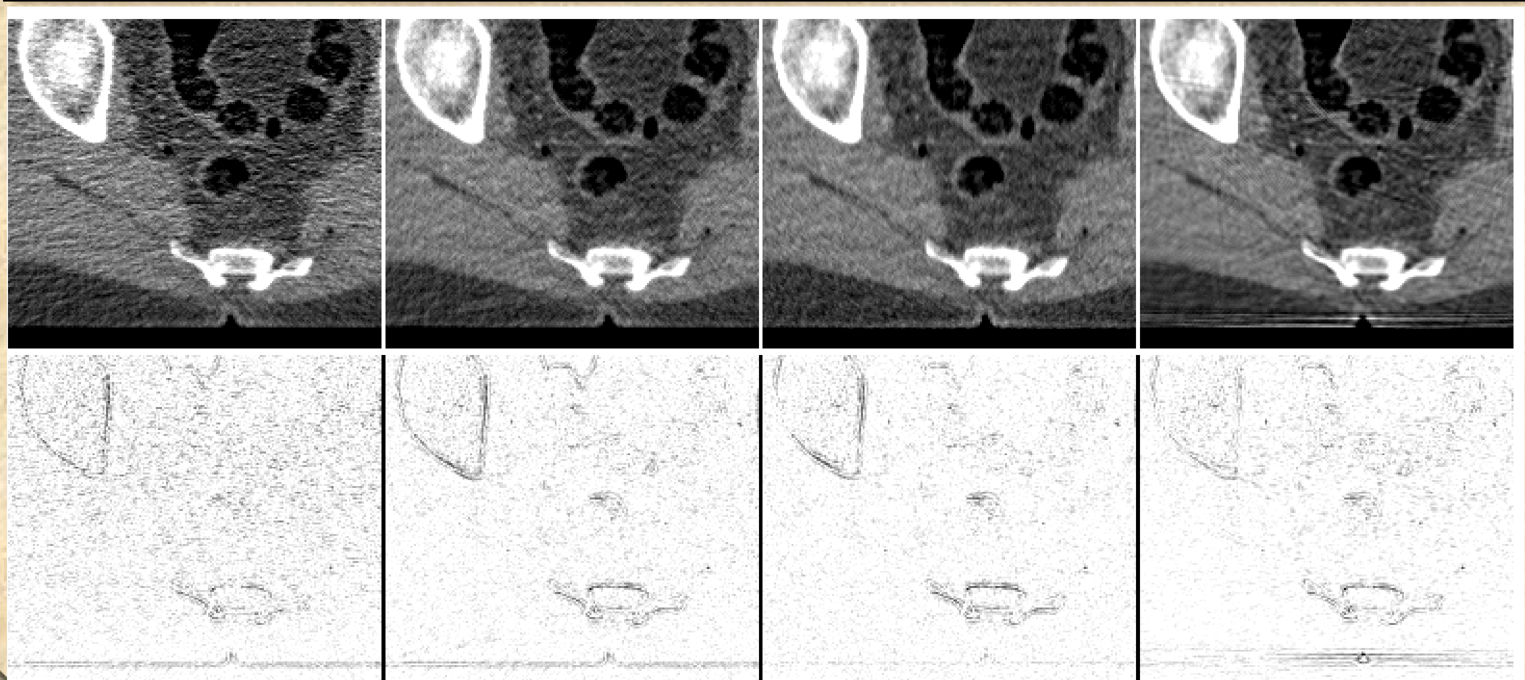
Error images

Empirical results



Thighs section

FBP, 25.26 dB	ATM, 27.46 dB	PWLS, 28.26 dB	Sparse, 27.94 dB

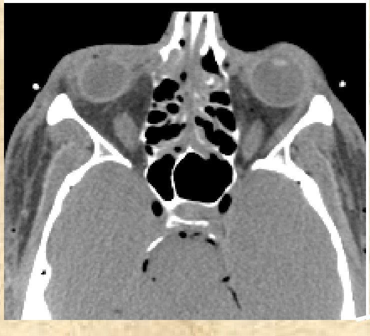


Recon. in [-220,350] HU

Error images

Empirical results

Same parameters, new anatomical region.



Head section

FBP, 29.84 dB	ATM, 29.83 dB	PWLS, 31.02 dB	Sparse, 32.36 dB

Recon. in [-170,250] HU

Error images

Intro to Computed Tomography	• Scan model, Noise, Local reconstruction
Framework of adaptive reconstruction	• General scheme of supervised learning
Adaptive FBP	• Learned FBP filter for local reconstruction
Sparsity-based sinogram restoration	• Adaptation of K-SVD to low-dose CT reconstruction
Learned shrinkage in a transform domain	• Adaptation of the method to low-dose CT reconstruction
Performance boosting of existing algorithms	• Local fusion of multiple versions of the algorithm output

Learned shrinkage in a transform domain

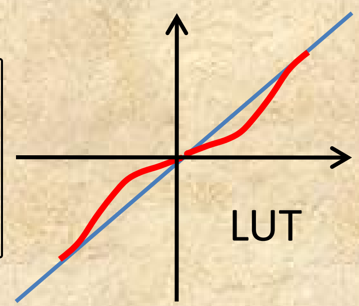
Convert to z-domain → Analysis → Shrinkage → Synhtesis → Recon-struction

Analysis: D^{+} (pseudo-inverse)

Shrinkage: Scalar shrinkage functions

Synhtesis: Dictionary D

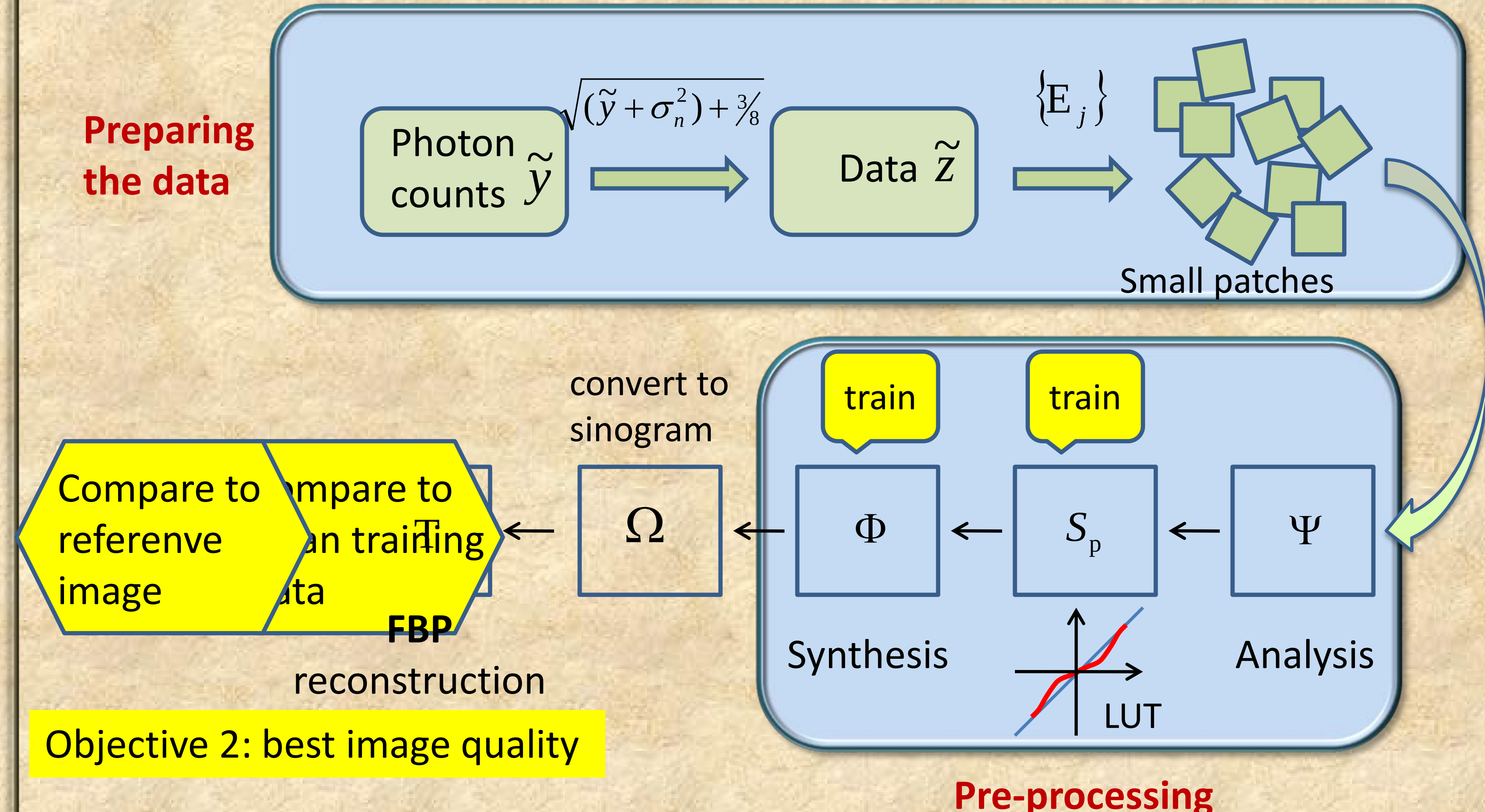
Denoising by supression of small coefficients, which usually contain the noise.



Examples of D: Discrete Cosines, Wavelets, etc.

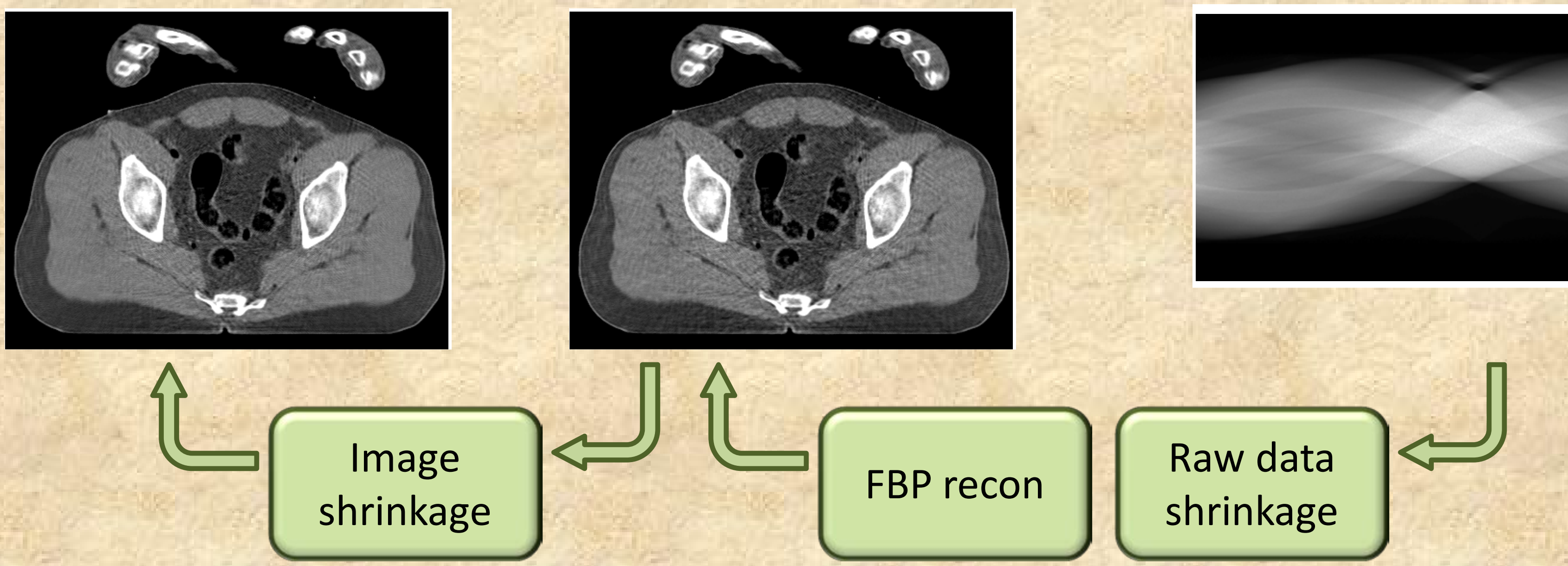
- Denosing by shrinkage of wavelet coeffs: Donoho & Johnston, 1994.
  The tool: Descriptive functions for descriptive dictionary.
- Denoising with learned shrinkage functions: Hel-Or and Shaked, 2002.
  The tool: Learned functions for descriptive dictionary.
- **Our goal: Solving non-linear inverse problems.**
  The tool: Learned functions for learned dictionary in a look-ahead training.

Learned shrinkage in a transform domain



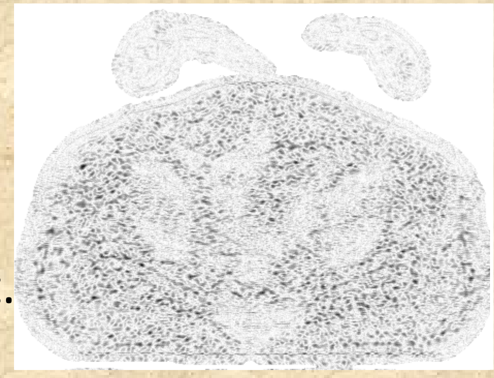
$$p,\Phi = \arg\min_{p,\Phi} \left\| f - T\Omega\Phi S_p \Psi E \tilde{z} \right\|_2^2 + \mu (J - \tilde{J})_+$$

Why not repeat the trick?



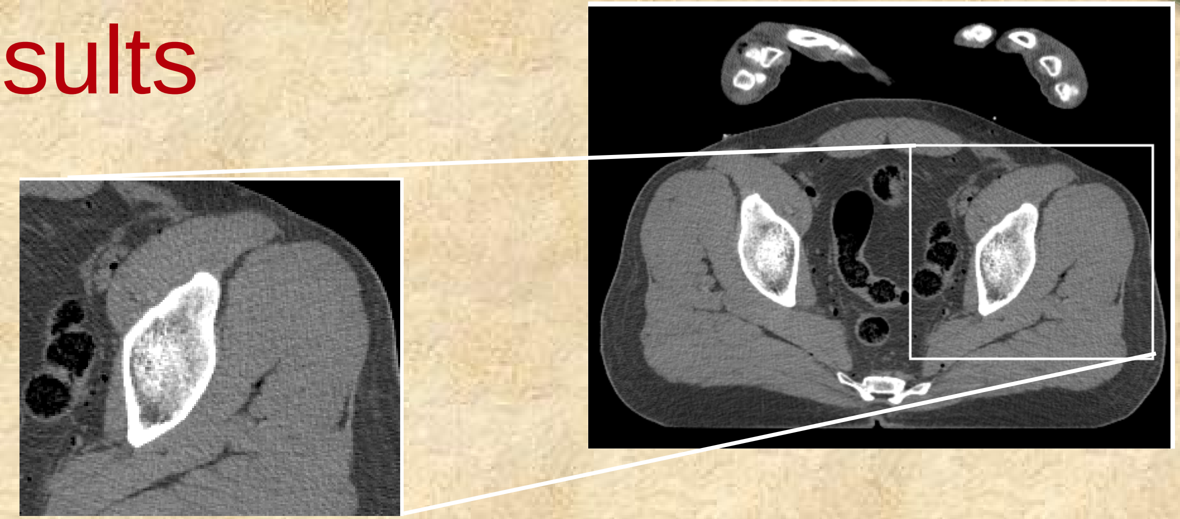
Post-processing with shrinkage functions, also trained by comparing to reference images.

Difference made by the post-processing : no image structure lost.



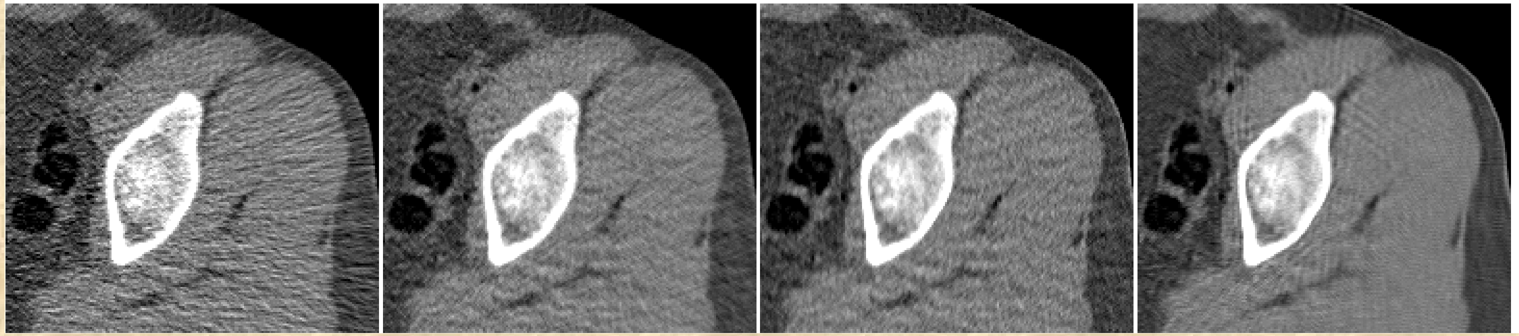
$$p,\Phi = \arg\min_{p,\Phi} \left\| f - \Phi S_P \Psi \mathrm{E} \hat{f} \right\|_2^2 + \mu (J - \tilde{J})_+$$

Empirical results

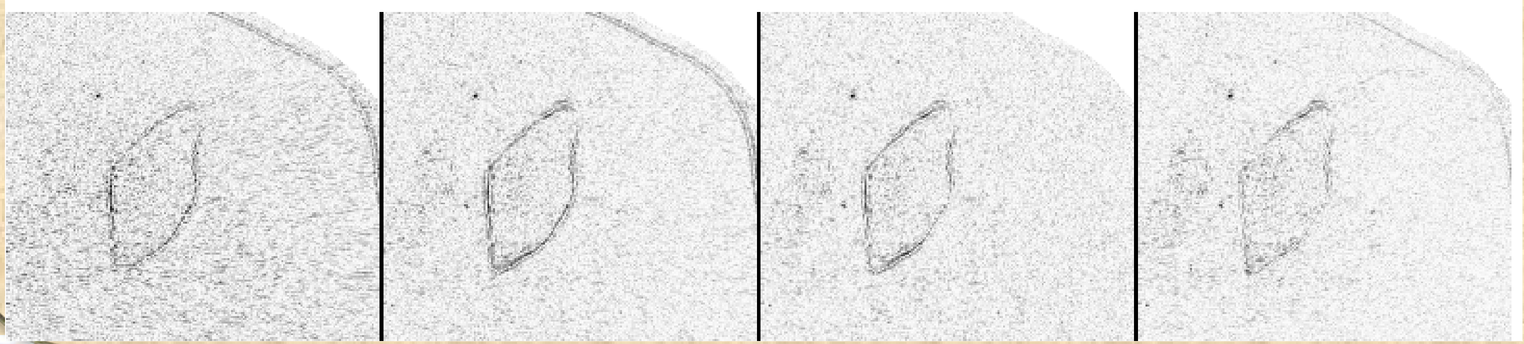


Thighs section (male)

FBP, 25.76 dB	ATM, 28.32 dB	PWLS, 28.90 dB	Shrinkage 30.05 dB



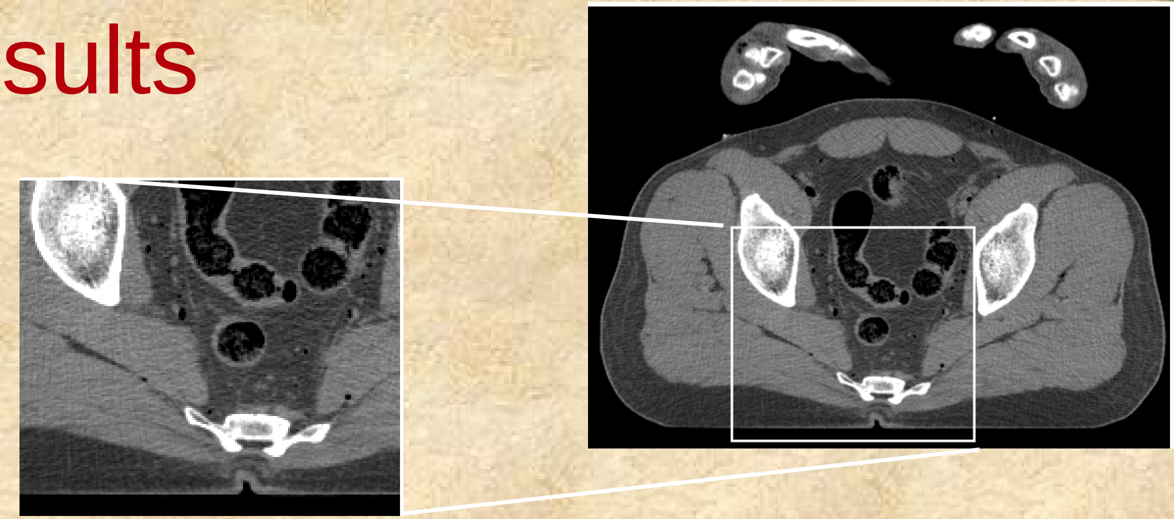
Recon. in [-220,350] HU



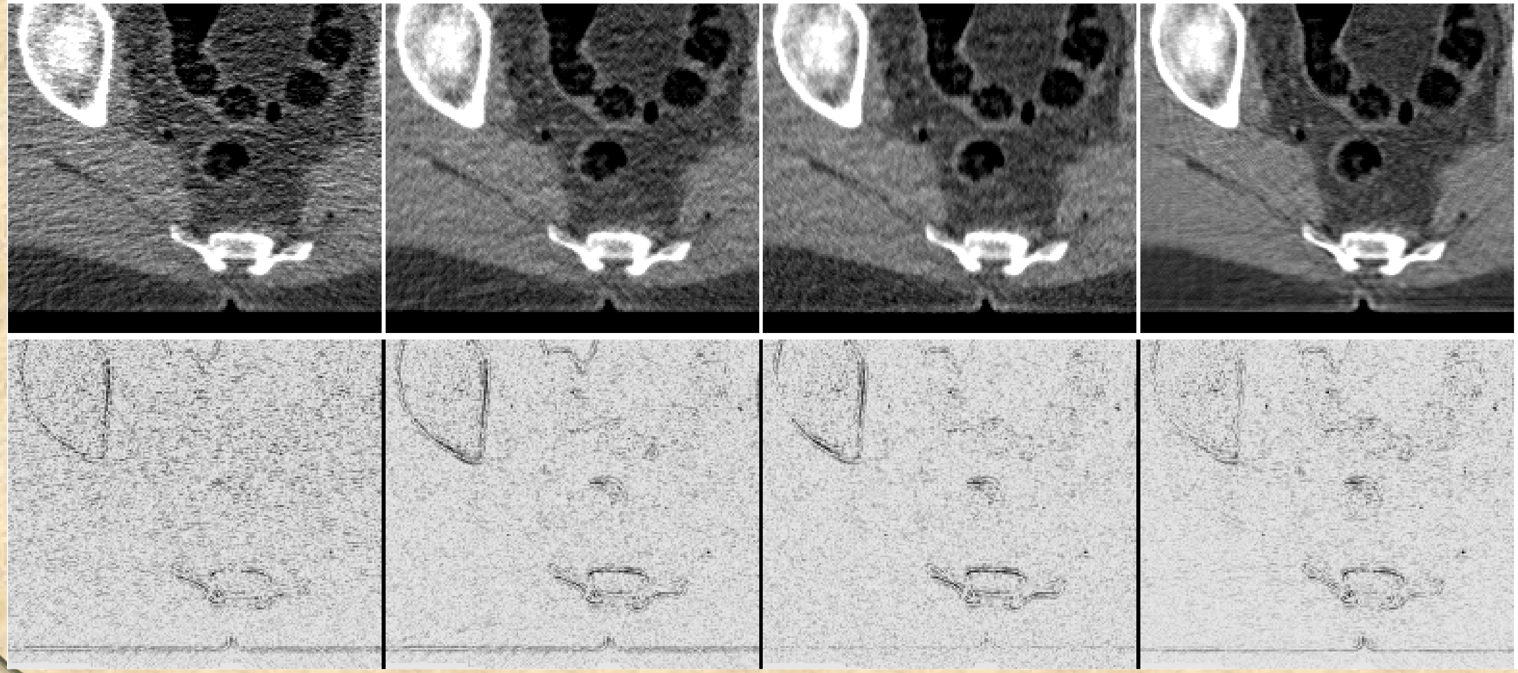
Error images

Empirical results

Thighs section



FBP, 25.26 dB	ATM, 27.46 dB	PWLS, 28.26 dB	Shrinkage 28.94 dB



Recon. in [-220,350] HU

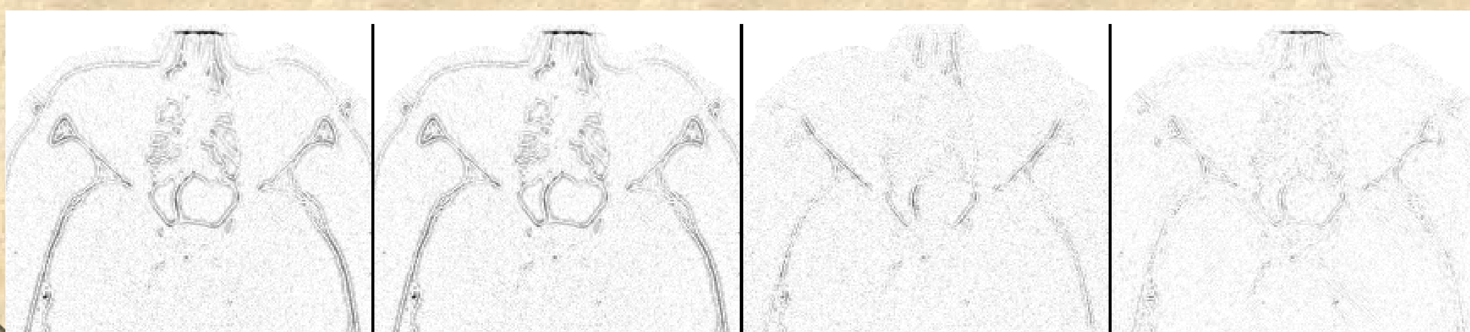
Error images

Empirical results

Head section

FBP, 29.84 dB	ATM, 29.83 dB	PWLS, 31.02 dB	Shrinkage 32.81dB

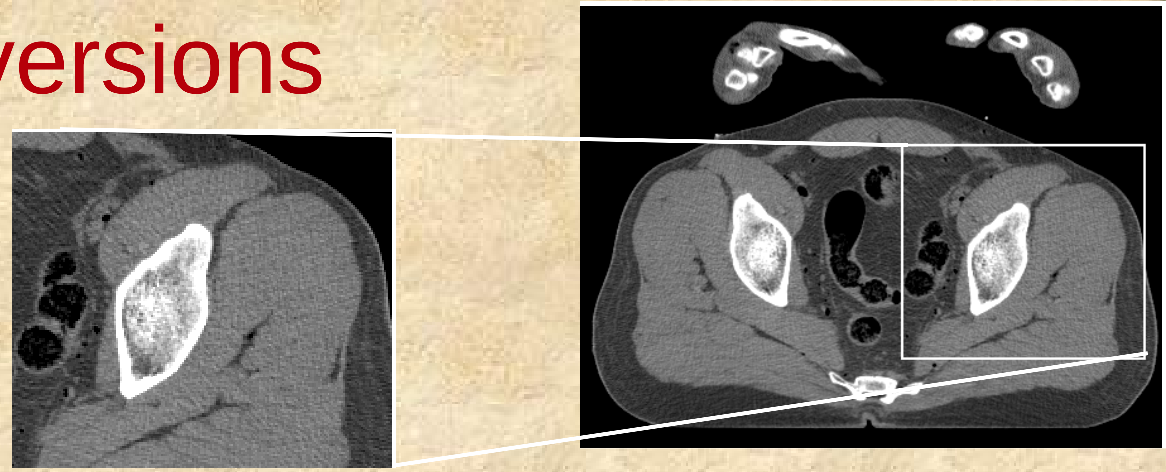
Recon. in [-170,250] HU



Error images

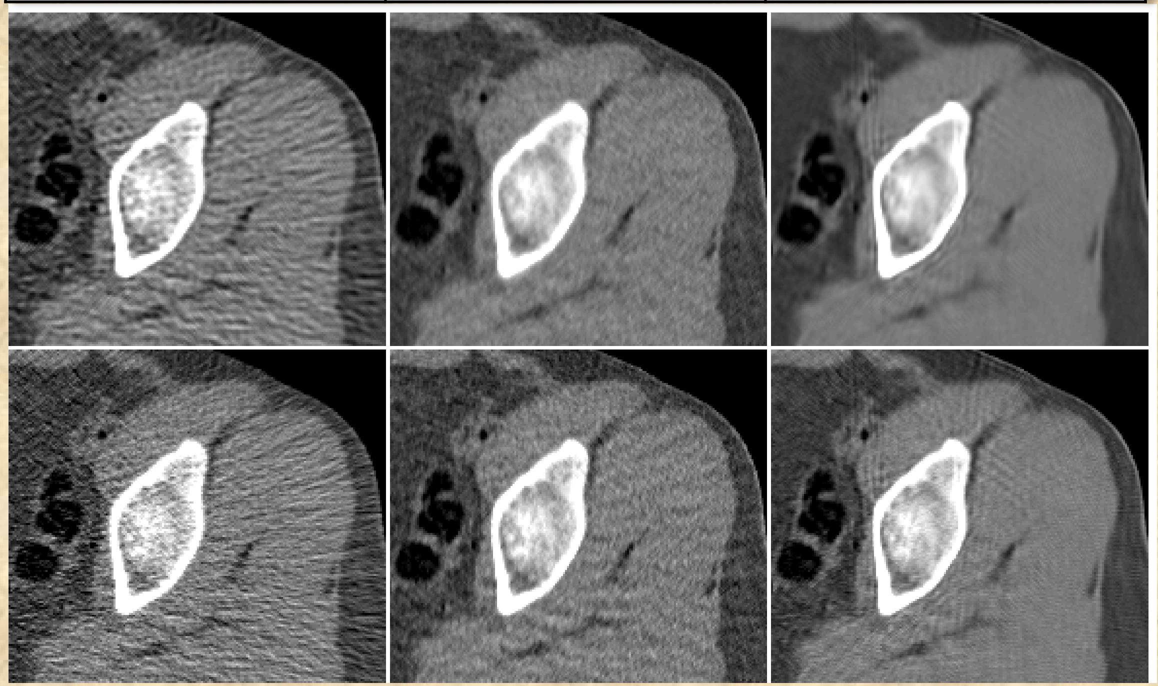
Alternative versions

Optimizing for MSE introduces a blur into the image.



Thighs section

FBP, 27.70 dB 25.76 dB	PWLS, 30.45 dB 28.90 dB	Shrinkage 30.84 dB 30.05 dB



MSE-optimized versions

Tuned by visual appearance

Effective dose reduction

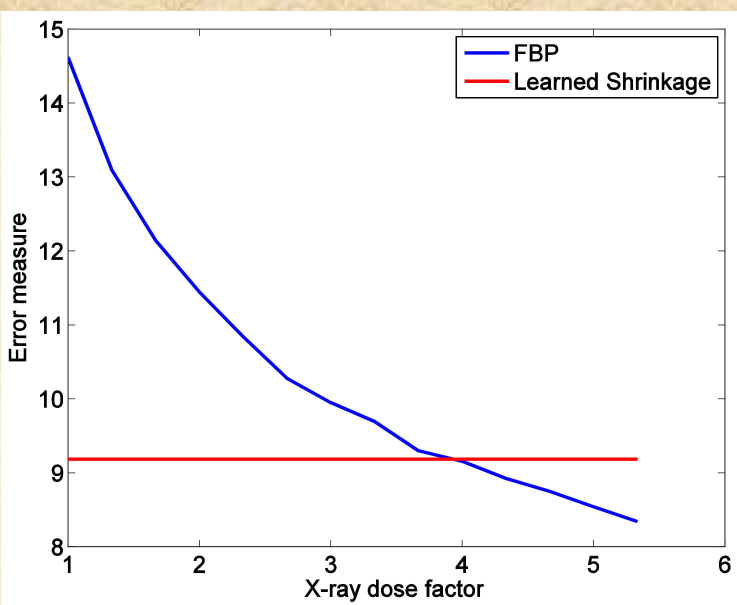
Estimating dose reduction factor:

•For each noise level, sweep over a range of FBP parameter and chose a reconstruction with minimal error measure.

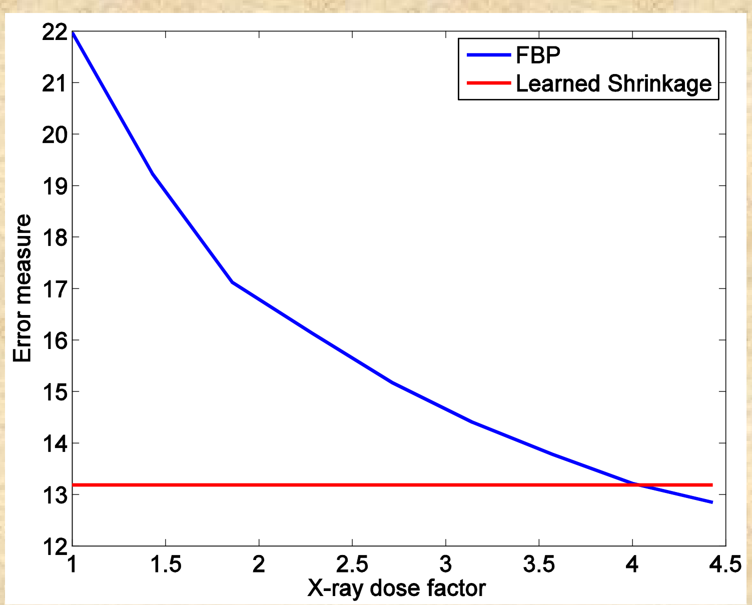
•Sweep over a range of the noise level and compare to learned shrinkage.

$$Error(f,\tilde{f}) = \left\| f - \tilde{f} \right\|_2^2 + \mu (J - \tilde{J})_+$$

Normal X-ray dose range



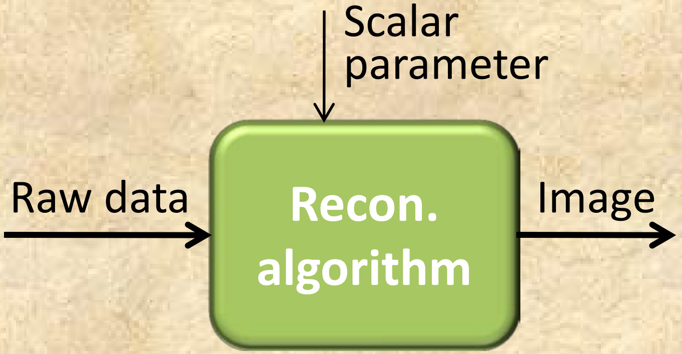
Low X-ray dose range

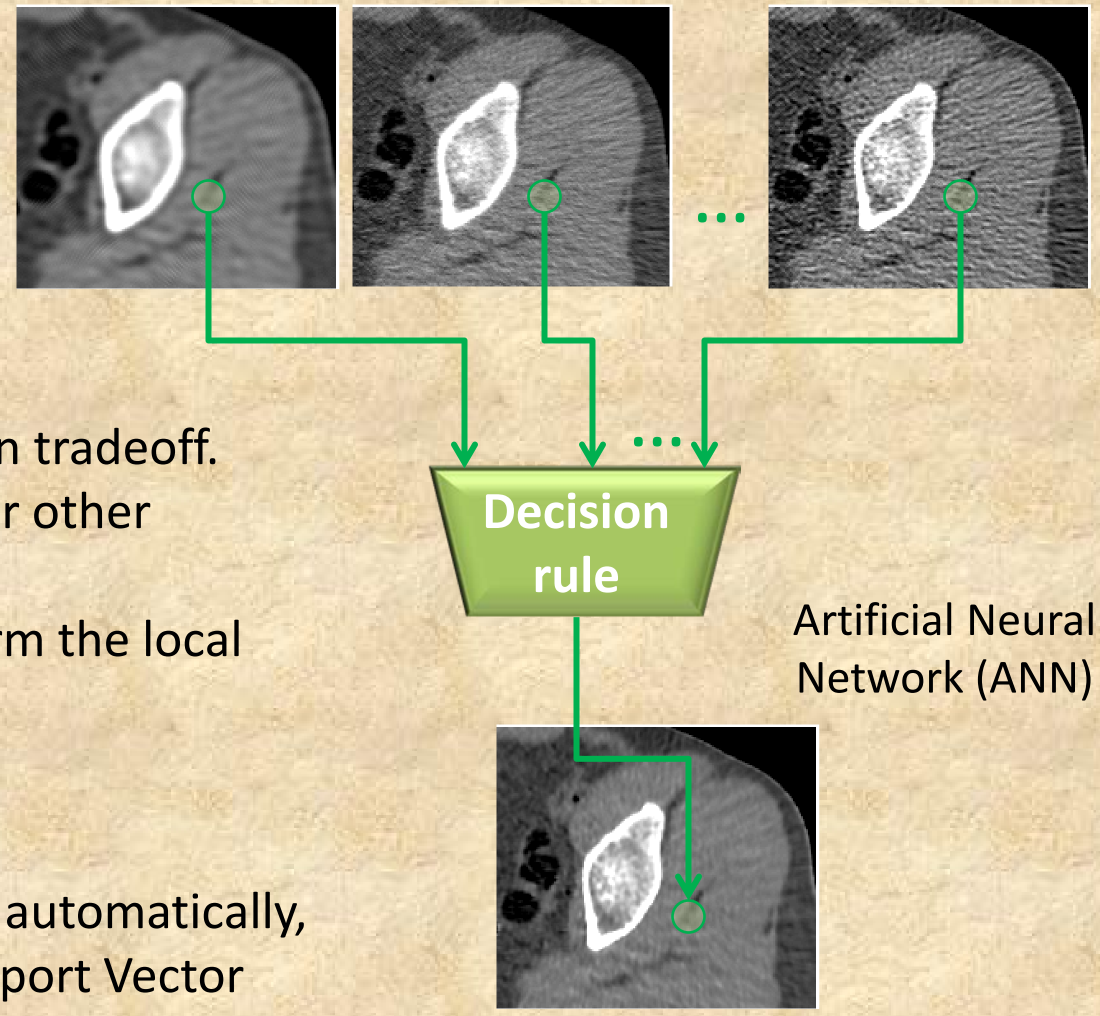


Contents of this talk

Intro to Computed Tomography	• Scan model, Noise, Local reconstruction
Framework of adaptive reconstruction	• General scheme of supervised learning
Adaptive FBP	• Learned FBP filter for local reconstruction
Sparsity-based sinogram restoration	• Adaptation of K-SVD to low-dose CT reconstruction
Learned shrinkage in a transform domain	• Adaptation of the method to low-dose CT reconstruction
Performance boosting of existing algorithms	• Local fusion of multiple versions of the algorithm output

Fusion over a smoothing parameter





The algorithm:

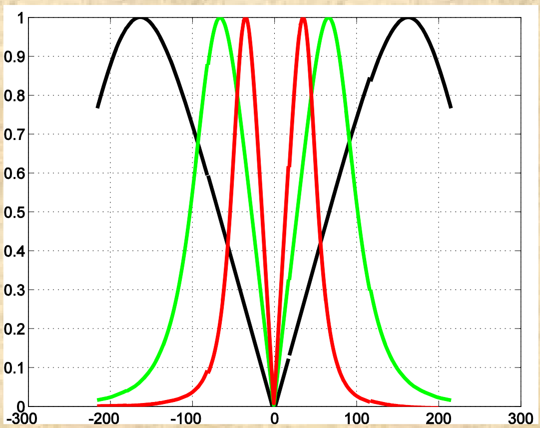
- Sweep the variance-resolution tradeoff.
- Extract pixel neighborhood (or other features) from each version.
- Build a decision rule to perform the local fusion.

Decision rule:

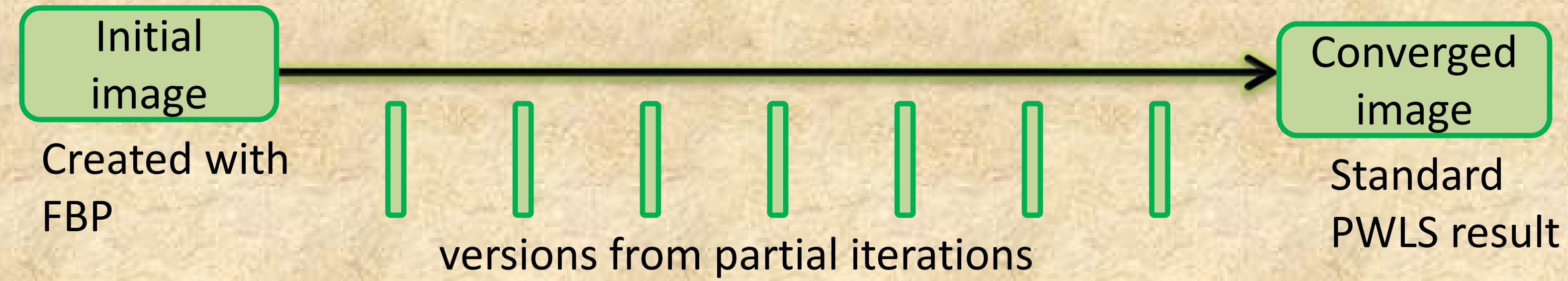
Use a regression to build one automatically, with a Neural network or Support Vector Regression.

Fusion over a smoothing parameter

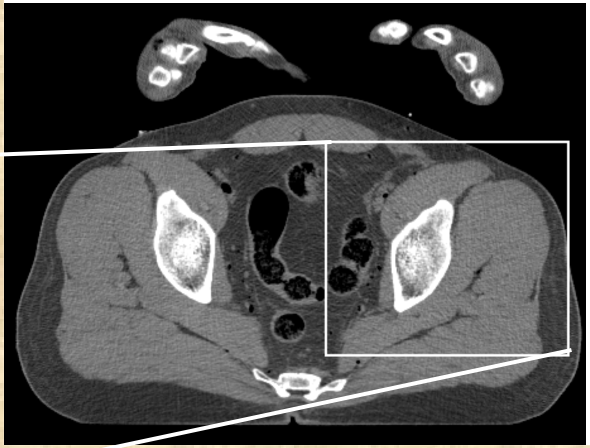
FBP algorithm: sweep the cut-off frequency of the low-pass sinogram filter.
Collect few images with different resolution-variance trade-off.



PWLS algorithm: perform the regular reconstruction while collecting versions along the iterations.

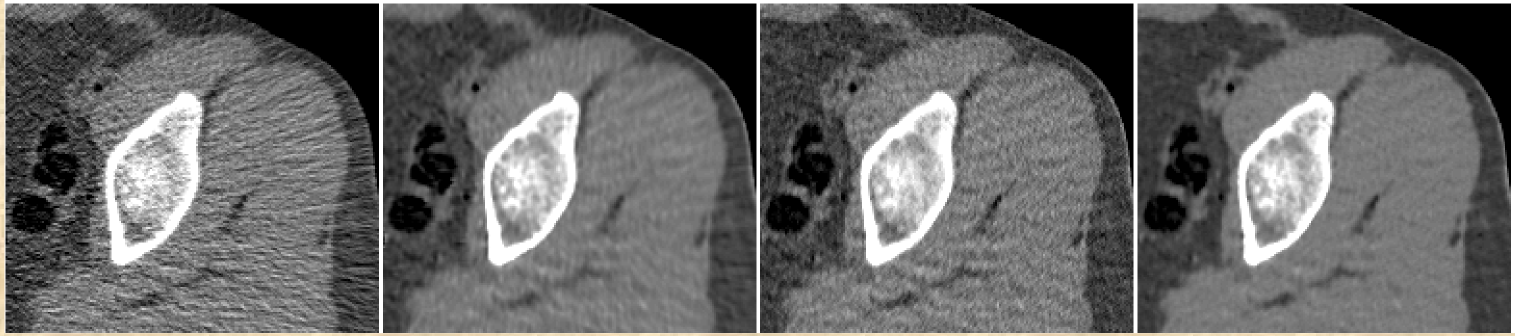


Empirical results

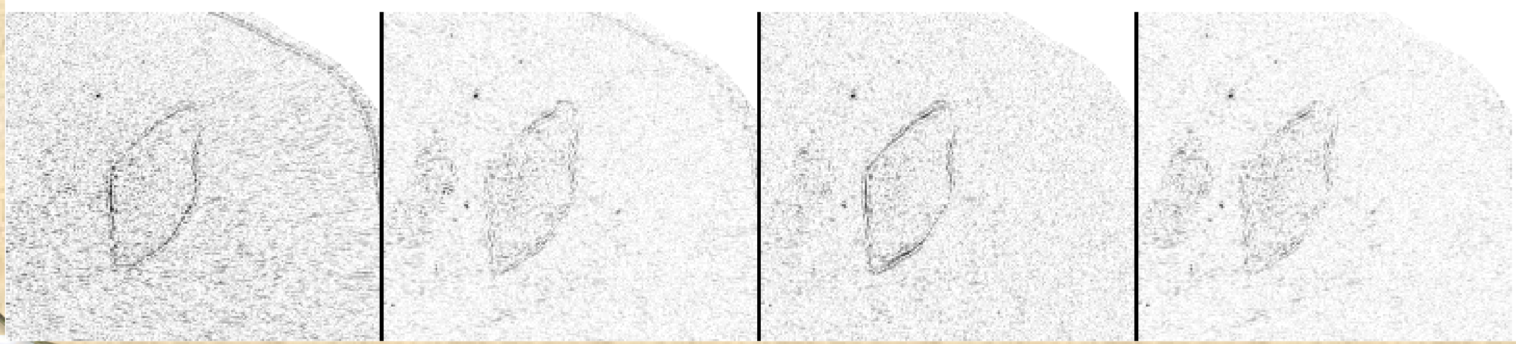


Thighs section

FBP 25.76 dB	FBP-ANN 30.62 dB	PWLS 28.90 dB	PWLS-ANN 31.11 dB

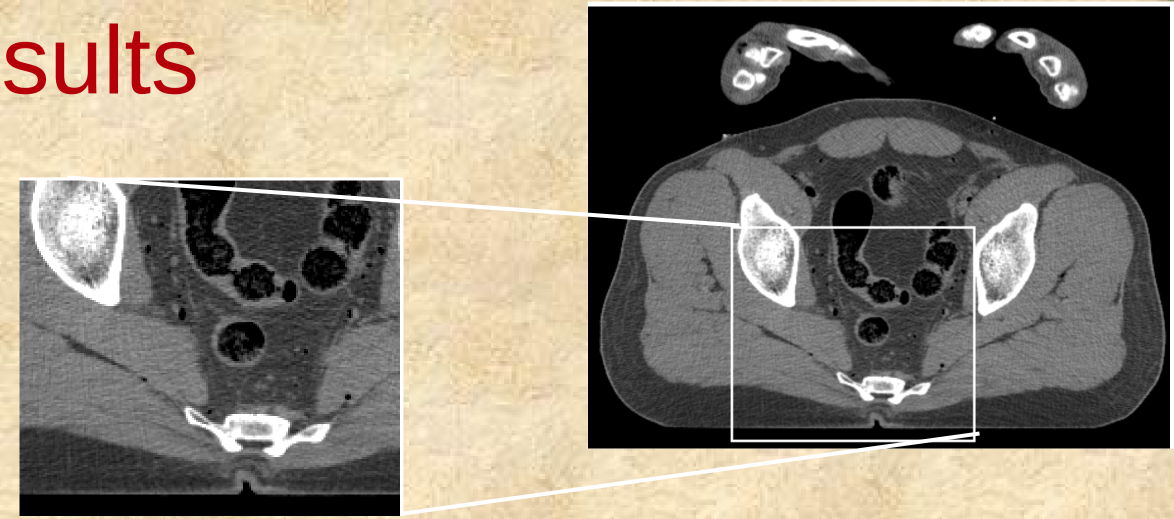


Recon. in [-220,350] HU



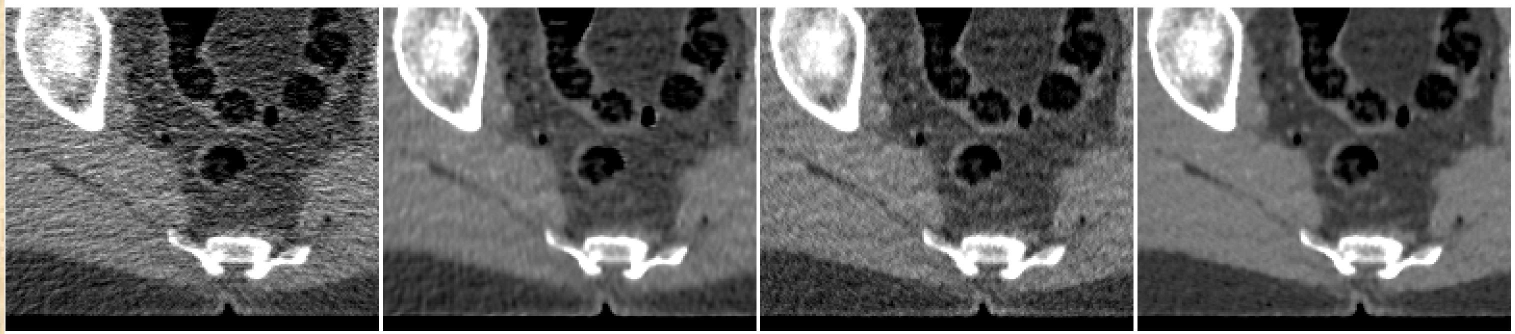
Error images

Empirical results

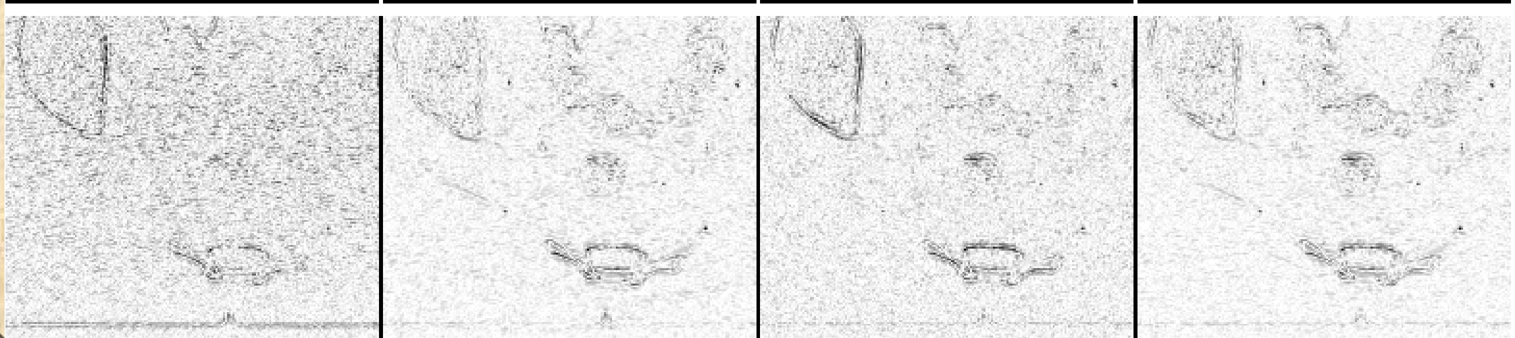


Thighs section

FBP 25.26 dB	FBP-ANN 29.67 dB	PWLS 28.26 dB	PWLS-ANN 30.03 dB



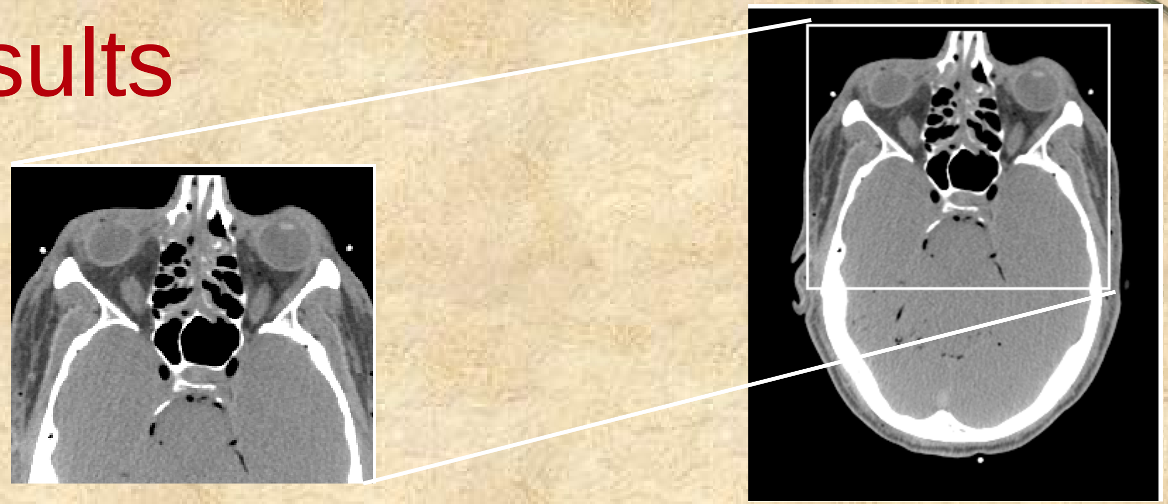
Recon. in [-220,350] HU



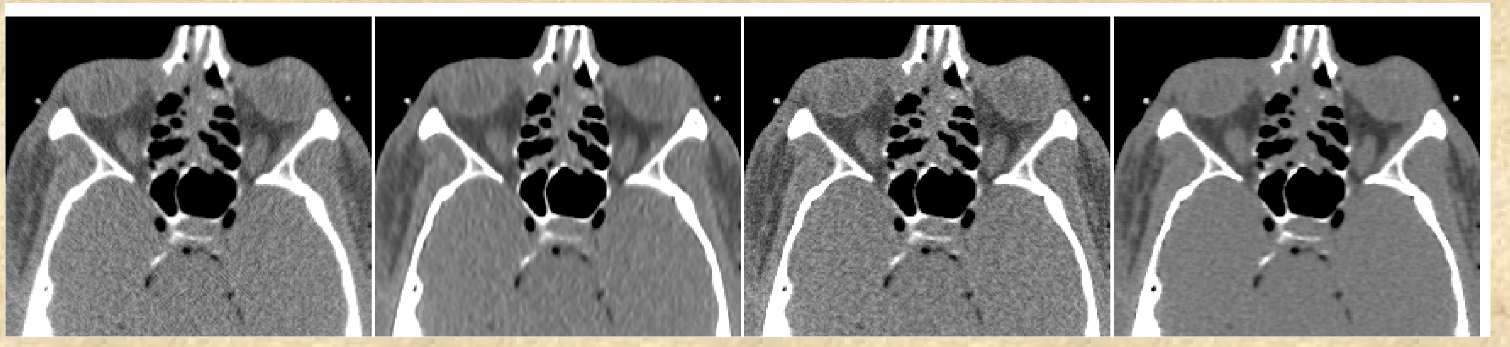
Error images

Empirical results

Head section



FBP 29.84 dB	FBP-ANN 33.41dB	PWLS 31.02 dB	PWLS-ANN 33.66 dB



Recon. in [-170,250] HU

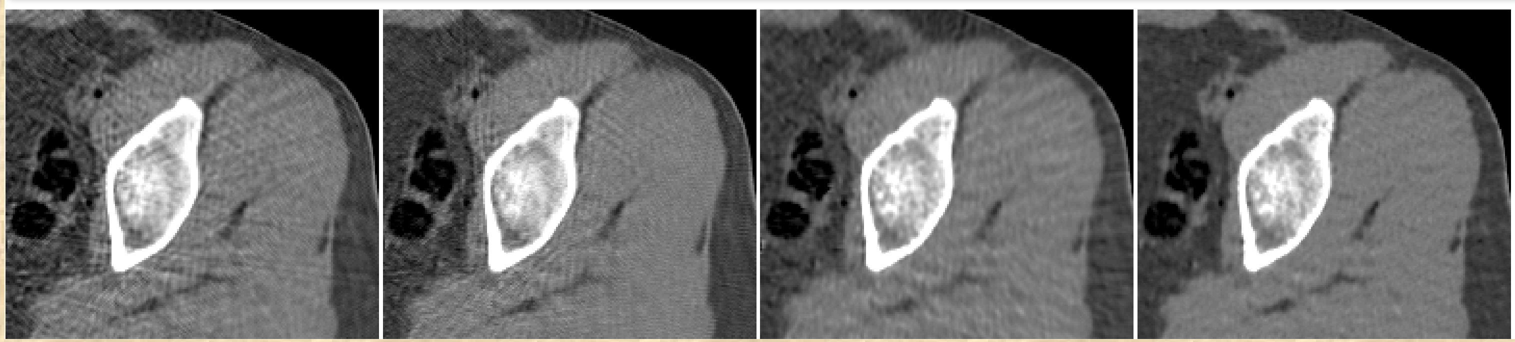
Error images

Intro to Computed Tomography	• Scan model, Noise, Local reconstruction
Framework of adaptive reconstruction	• General scheme of supervised learning
Adaptive FBP	• Learned FBP filter for local reconstruction
Sparsity-based sinogram restoration	• Adaptation of K-SVD to low-dose CT reconstruction
Learned shrinkage in a transform domain	• Adaptation of the method to low-dose CT reconstruction
Performance boosting of existing algorithms	• Local fusion of multiple versions of the algorithm output
Conclusions	

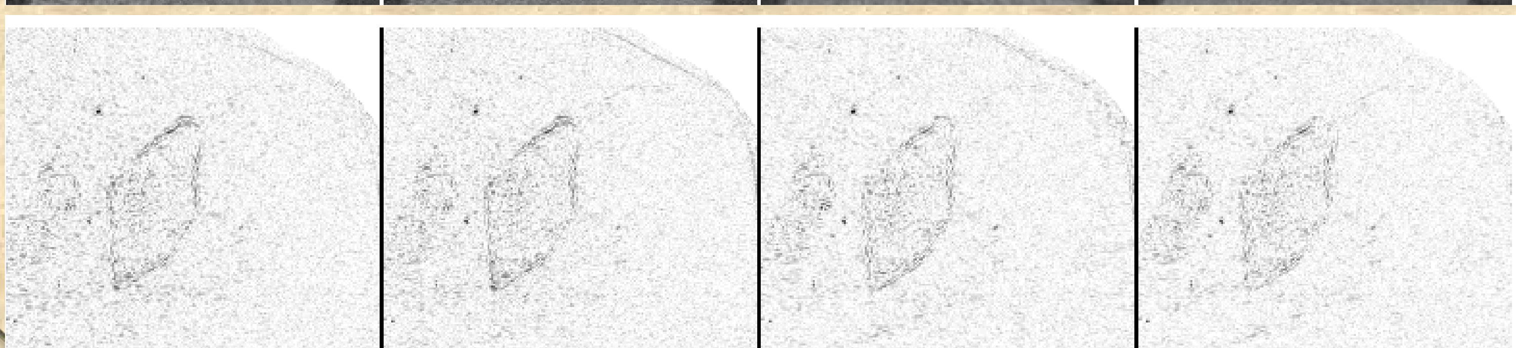
Parade of proposed methods

Thighs section

Sparse 29.62 dB	Shrinkage 30.05 dB	FBP-ANN 30.62 dB	PWLS-ANN 31.11 dB



Recon. in [-220,350] HU



Error images

Summary

Adaptive methods can help improving CT reconstruction.

FBP needs only a little help to allow truly local reconstruction.

Once the raw data is variance-normalized, the sparsity-based denoising mends most of the damage done by the low-dose scan.

When the smootheness parameter is swept, reconstruction algorithms supply more information about the image; it is easily extracted by a regression function using only the intensity values.

Example-based training does not jeopardize the image content (in the presented algorithms) and can be allowed for clinical use.

Thank you.


©BillyBear4Kids.com


CT SLICKER


© BillyBear4Kids.com


ARKive
www.arkive.org
© Anthony Bannister / www.nhpa.co.uk