Sparse sensing: Fundamentals and applications in X-ray CT

Hyoung Suk Park

National Institute for Mathematical Sciences

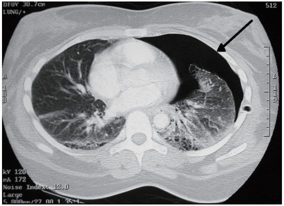
copyright © 2017 by MEDIVIEWSOFT

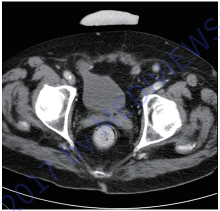
January 9, 2017

2017 Winter School- CT Session

X-ray Computed Tomography

- X-ray CT: widely used imaging modality for screening and diagnosis, emergency medicine, image-guided interventions, and monitoring of therapeutic responses.



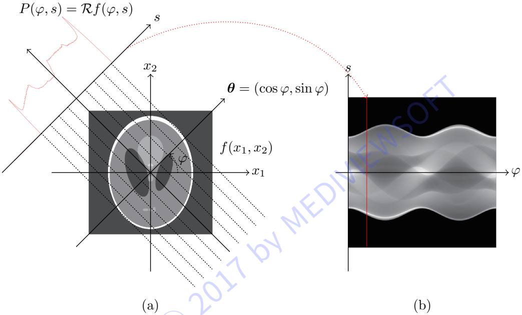



copyright© 2017 by MEDIVIEWSOFT

- Disadvantage: excessive radiation dose, risk of cancer from medical radiation exposures
  - X-ray:0.1mSV, head CT: 2, chest CT: 8, abdomen CT: 10, nature: $2.5 \sim 10$
  - Dose:1000mSV, cancer rates are 2 to 3 times higher
  - GE, Chest low dose CT: < 1 mSV

Y. Sagara, *et al* , Abdominal CT..., Medical Physics, 2010.

Basic CT Reconstruction Algorithm



copyright© 2017 by MEDIMEWSOFT

$$\mathcal{R}f(\varphi, s) = \int_{\mathbb{R}^2} f(\mathbf{x})\delta(\mathbf{x}\cdot\boldsymbol{\theta} - s)d\mathbf{x}$$: Radon transform

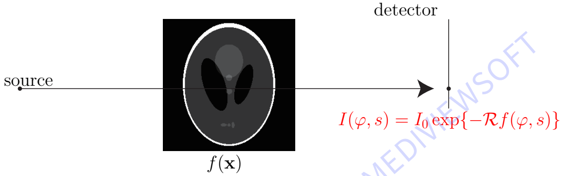
$$f(\mathbf{x}) = \frac{1}{4\pi}\mathcal{R}^*\left(\mathcal{F}^{-1}[|\omega|] * \mathcal{R}f\right)(\mathbf{x})$$: Filtered backprojection

-$f(\mathbf{x})$: attenuation coefficient of material at energy E and position $\mathbf{x}$

-$\mathcal{R}^* g(\mathbf{x}) = \int_0^{2\pi} g(\varphi, \mathbf{x}\cdot\boldsymbol{\theta})d\varphi$: Backprojection operator

Basic CT reconstruction algorithm

- Lambert-Beer's law:



I_0 : X-ray intensity generated at X-ray source.

$f(\mathbf{x})$: attenuation coefficient at position $\mathbf{x}$.

- Measured X-ray data:

$$P(\varphi, s) = -\ln \frac{I(\varphi, s)}{I_0} + \text{Poisson noise} \quad (\neq \mathcal{R}f(\varphi, s)).$$

- The FBP yields

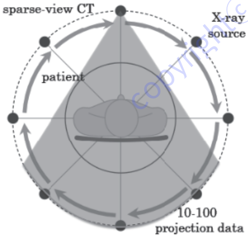
$$\text{(CT image)} \quad f_{\mathrm{CT}}(\mathbf{x}) := \frac{1}{4\pi}\mathcal{R}^* \left(\mathcal{F}^{-1}[|\omega|] * P\right)(\mathbf{x}).$$

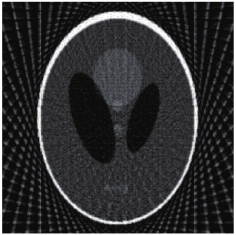
copyright © 2017 by MEDIVIEWSOFT

Low dose CT

- How to reduce X-ray dose
  1. reduce X-ray current (or I_0) $\rightarrow$ increase noise, photon starvation

  2. sparse view (under sampling) $\rightarrow$ causes streaking artifacts (studied as a potential strategy [1])





30 view

copyright © 2017 by MEDIVIEWSOFT

[1] H. Qi, *et al* , Sparse-view computed tomography image reconstruction..., Bio-Medical Materials and Engineering, 2015.

Noise Correction Methods

copyright© 2017 by MEDIVIEWSOFT

Maximum-likelihood for X-ray CT

- Given measured $\mathbf{y} = \{y_j\}_{j=1}^{M}$ with $y_j = \text{poisson}\{I_j\}$, find $\mathbf{f} = \{f_j\}_{j=1}^{N}$ that maximize the probability $\Pr(\mathbf{f}|\mathbf{y})$

$$\hat{\mathbf{f}} = \arg\max_{\mathbf{f}} \Pr(\mathbf{f}|\mathbf{y})$$

- From Bayesian rule, it is equivalent to

$$\hat{\mathbf{f}} = \arg\max_{\mathbf{f}} \{\log\Pr(\mathbf{y}|\mathbf{f}) + \log\Pr(\mathbf{f}) - \log\Pr(\mathbf{y})\}$$

- The problem can be reduced to

$$\hat{\mathbf{f}} = \arg\max_{\mathbf{f}} \ \log\Pr(\mathbf{y}|\mathbf{f}) \ \ : \ \text{Maximum likelihood approach}$$

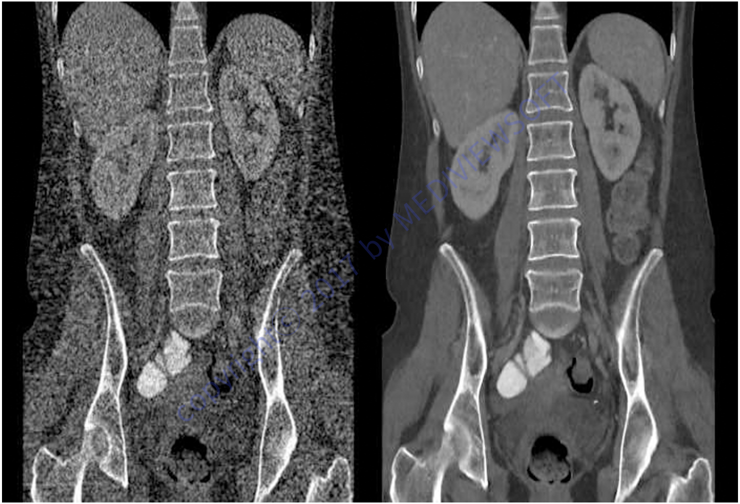
- With $\Pr(y_j|\mathbf{f}) = \frac{I_j^{y_j} e^{-I_j}}{y_j!}$, second order Taylor expansion gives

$$\log\Pr(\mathbf{y}|\mathbf{f}) \approx -\frac{1}{2}(\mathbf{P} - \mathbf{L}\mathbf{f})^T \mathbf{W}(\mathbf{P} - \mathbf{L}\mathbf{f}),$$

  - $\mathbf{W}$: diagonal matrix whose values $w_j = y_j$.
  - $\mathbf{L}$: projection operator.

copyright © 2017 by MEDIVIEWSOFT

Maximum-likelihood for X-ray tomography



FBP MBIR

copyright(c) 2017 by MEDVIEWSOFT

One-step iterative method

- Noise reduction model:

$$\hat{\mathbf{f}} = \arg\min_{\mathbf{f}} (\mathbf{P} - \mathbf{L}\mathbf{f})^T \mathbf{W} (\mathbf{P} - \mathbf{L}\mathbf{f})$$

- Gradient-decent method leads to

$$\mathbf{f}^{(k)} = \Delta t \left[\sum_{n=0}^{k-1} (\mathbf{I}_d - \Delta t\, \mathbf{L}^T \mathbf{W} \mathbf{L})^n \right] \mathbf{L}^T \mathbf{W} \mathbf{P}$$
$$= \left[\mathbf{I}_d - (\mathbf{I}_d - \Delta t\, \mathbf{L}^T \mathbf{W} \mathbf{L})^k \right] \mathbf{L}^{-1} \mathbf{P}.$$

- Using some properties of Radon transform, we obtain

$$f^{(k)}(\mathbf{x}) = \frac{1}{4\pi} \mathcal{R}^* \left(\mathcal{F}^{-1}[H_k] * P \right)(\mathbf{x}),$$

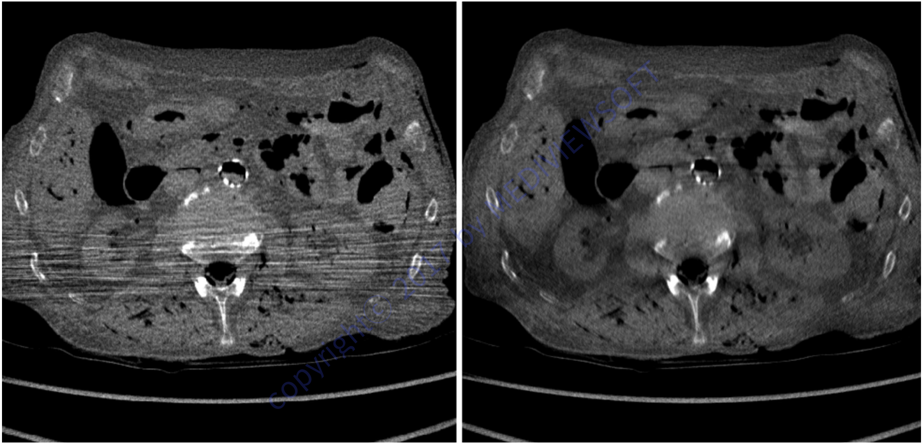
where H_k is the adaptive filter given by

$$H_k(\omega) = (1 - (1 - \Delta t \frac{\sigma_{ray}}{|\omega|}))^k)|\omega|, \quad H_k(0) = 0.$$

G. L. Zeng, A filtered backprojection MAP algorithm..., Medical Physics, 2012.

Copyright © 2017 by MEDINEWSOFT

One-step iterative method



copyright© 2017 by MEDIVIEWSOFT

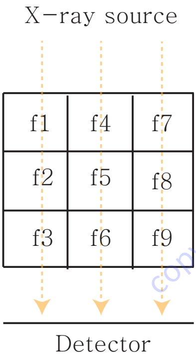
G. L. Zeng, A filtered backprojection MAP algorithm..., Medical Physics, 2012.

Sparse-view CT reconstruction methods

copyright© 2017 by MEDIVIEWSOFT

Spare view X-ray CT

- Can we reconstruct the unique solution $\mathbf{f}$ from the under-determined system $\mathbf{L}_s\mathbf{f} = \mathbf{P}_s$?

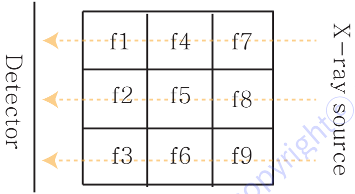


$$\underset{\text{Ls}}{\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}} \underset{\text{f}}{\begin{bmatrix} f1 \\ f2 \\ f3 \\ f4 \\ f5 \\ f6 \\ f7 \\ f8 \\ f9 \end{bmatrix}} = \underset{\text{Ps}}{\begin{bmatrix} p1 \\ p2 \\ p3 \end{bmatrix}}$$

copyright© 2017 by MEDIVIEWSOFT

Spare view X-ray CT

- Can we reconstruct the unique solution $\mathbf{f}$ from the under-determined system $\mathbf{L}_s\mathbf{f} = \mathbf{P}_s$?



$$
\begin{array}{cccc}
\text{Ls} & \text{f} & & \text{Ps} \\
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1
\end{bmatrix}
&
\begin{bmatrix} f1 \\ f2 \\ f3 \\ f4 \\ f5 \\ f6 \\ f7 \\ f8 \\ f9 \end{bmatrix}
& = &
\begin{bmatrix} p1 \\ p2 \\ p3 \\ p4 \\ p5 \\ p6 \end{bmatrix}
\end{array}
$$

copyright 2017 by MEDIVIEWSOFT

Sparse driven reconstruction

- Let $\|\mathbf{x}\|_0$= number of non-zeros of $\mathbf{x}$. Let $\mathrm{spark}(\mathbf{A}) = \min\{\|\mathbf{x}\|_0 : \mathbf{x} \in N(\mathbf{A}) \setminus \{\mathbf{0}\}\}$.

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots x_n\boldsymbol{a}_n = 0$$

  $\mathrm{spark}(\mathbf{A})$: smallest number of columns that are linearly dependent.

- If the under-determined system $\mathbf{A}\mathbf{x} = \mathbf{b}$ has the solution $\mathbf{x}$ obeying $\|\mathbf{x}\|_0 < 1/2\mathrm{spark}(\mathbf{A})$, then $\mathbf{x}$ is the unique solution to the following minimization problem.

$$\min \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \boldsymbol{b}.$$

copyright© 2017 by MEDIVIEWSOFT

Relaxation of l_0-minimization problem

- Let $\mu(\mathbf{A})$ be a mutual coherence of an m by n matrix $\mathbf{A}$ given by

$$\mu(\mathbf{A}) = \max_{j \neq k} \frac{\mathbf{a}_j \cdot \mathbf{a}_k}{\|\mathbf{a}_j\|_2 \|\mathbf{a}_k\|_2}.$$

- If $\mathbf{x} \in \mathbb{R}^n$ is a s-sparse solution ($\|\mathbf{x}\|_0 = s$) to $\mathbf{A}\mathbf{x} = \mathbf{b}$ and

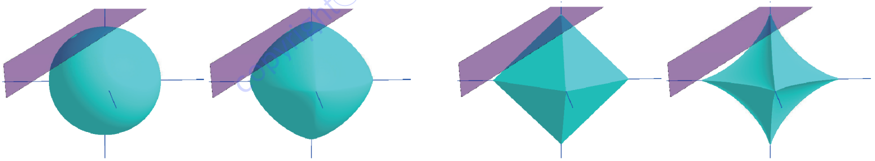
$$s < \frac{1}{2}(1 + \mu(\mathbf{A})),$$

  then $\mathbf{x}$ is the unique solution to l_1-minimization problem

$$\min \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \boldsymbol{b},$$

  and this is also solution to l_0 problem.

Copyright© 2017 by MEDIVIEWSOFT



p=2 p=1.5 p=1 p=0.7

M. Elad, sparse and redundant representation, 2010

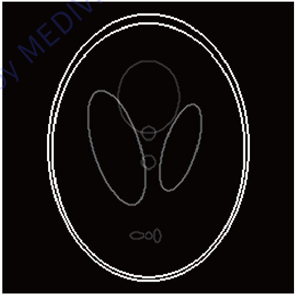
Sparse driven reconstruction

- With the aid of sparsity of $|\nabla f|$, sparse image can be reconstructed by solving the following problem.

$$\min \|\nabla f\|_1 \quad \text{subject to} \quad \|\mathcal{R}f - P\|_2^2 \leq \varepsilon.$$

copyright© 2017 by MEDIVIEWSOFT

Original image



Absolute value of gradient

- One can use Ψf (wavelet, DCT, rank,...)

Gradient-descent method

- With L1 regularization term, data-fitting model is given by

$$\min_f \|\nabla f\|_1 + \frac{\lambda}{2}\|\mathcal{R}f - P_s\|_2^2$$

- Euler-Lagrange equation provides

$$-\nabla\cdot\left(\frac{\nabla f}{|\nabla f|}\right) + \lambda\mathcal{R}^*(\mathcal{R}f - P_s) = 0$$

- Using the gradient descent method, we have

$$\frac{\partial}{\partial t}f(\mathbf{x},t) = \nabla\cdot\left(\frac{\nabla f}{\sqrt{|\nabla f|^2+\beta}}\right) - \lambda\mathcal{R}^*(\mathcal{R}f - P_s), \quad \mathbf{x} = (x,y)$$

- Using the finite difference scheme, we have

$$f(\mathbf{x},t^{n+1}) = f(\mathbf{x},t^n) + \Delta t\left[\frac{f_{xx}(f_y^2+\beta) - 2f_xf_yf_{xy} + f_{yy}(f_x^2+\beta)}{(f_x^2+f_y^2+\beta)^{3/2}} - \lambda\mathcal{R}^*(\mathcal{R}f(\mathbf{x},t^n) - P_s)\right]$$

Copyright © 2017 by MDIVIEWSOFT

Gradient-descent method

- Two disadvantages: not exact solution, convergence speed.

  $\rightarrow \frac{\partial}{\partial t} f(\mathbf{x}, t) = \nabla \cdot \left(\frac{\nabla f}{\sqrt{|\nabla f|^2 + \beta}} \right) - \lambda \mathcal{R}^* (\mathcal{R} f - P_s)$

  $\rightarrow$ step size: $\Delta t \leq c \Delta x^2 \sqrt{|\nabla f|^2 + \beta}$

copyright© 2017 by MELVIEWSOFT

Augmented Lagrangian method

$\min_f \|\nabla f\|_1 + \frac{\lambda}{2}\|\mathcal{R}f - P\|_2^2$

- Augmented Lagrangian method overcomes these defects. By introducing auxiliary variable $\boldsymbol{q} = \nabla f$, we solve

$$\min_{f,\boldsymbol{q}} \max_{\boldsymbol{\tau}} \int_\Omega |\boldsymbol{q}| d\mathbf{x} + \frac{\gamma}{2}\int_0^{2\pi}\int_{\mathbb{R}} |P - \mathcal{R}f|^2 d\varphi ds + \int_\Omega \boldsymbol{\tau}\cdot(\boldsymbol{q} - \nabla f) d\mathbf{x} + \frac{r}{2}\int_\Omega |\boldsymbol{q} - \nabla f|^2 d\mathbf{x}.$$

- AL algorithm
  1. Initialization: $\boldsymbol{q}^0 = \boldsymbol{\tau}^0 = \mathbf{0}$.
  2. For fixed multiplier $\boldsymbol{\tau}^k$, solve

$$f^{k+1} = \arg\min_f \frac{\gamma}{2}\int_0^{2\pi}\int_{\mathbb{R}} |P - \mathcal{R}f|^2 d\varphi ds - \int_\Omega \boldsymbol{\tau}^k \cdot \nabla f d\mathbf{x} + \frac{r}{2}\int_\Omega |\boldsymbol{q} - \nabla f|^2 d\mathbf{x} \tag{2}$$

  and solve

$$\boldsymbol{q}^{k+1} = \arg\min_{\boldsymbol{q}} \int_\Omega |\boldsymbol{q}| d\mathbf{x} + \int_\Omega \boldsymbol{\tau}^k \cdot \boldsymbol{q} d\mathbf{x} + \frac{r}{2}\int_\Omega |\boldsymbol{q} - \nabla f|^2 d\mathbf{x} \tag{3}$$

  3. Update $\boldsymbol{\tau}^{k+1}$ by

$$\boldsymbol{\tau}^{k+1} = \boldsymbol{\tau}^k + r(\boldsymbol{q}^{k+1} - \nabla f^{k+1})$$

copyright© 2017 by MEDVIEWSOFT

Augmented Lagrangian method

- The optimality condition of the quadratic problem (2) gives

$$(-r\nabla^2 + \gamma \mathcal{R}^*\mathcal{R})f = \gamma \mathcal{R}^* P - \nabla \cdot (r\boldsymbol{q} + \boldsymbol{\tau}^k),$$

- Since (3) can be rewritten as

$$\arg\min_{\boldsymbol{q}} \ \frac{1}{r}\int_\Omega |\boldsymbol{q}| d\mathbf{x} + \frac{1}{2}\int_\Omega |\boldsymbol{q} - (\nabla f - \frac{1}{r}\boldsymbol{\tau}^k)|^2 d\mathbf{x},$$

optimal value in (2) can be explicitly computed using Shrinkage operator:

$$\boldsymbol{q} = \frac{1}{r} shrink(r\nabla f - \boldsymbol{\tau}^k, 1),$$

where Shrinkage operator *shrink* defined by

$$shrink(\mathbf{x}, \gamma) = \frac{\mathbf{x}}{|\mathbf{x}|}\max(|\mathbf{x}| - \gamma, 0).$$

Copyright© 2017 by MEDIVIEWSOFT

copyright© 2017 by MEDIVIEWSOFT