Assignment 3: CS 754, Advanced Image Processing

Group Details:

Vinit Awale (18D070067)
Piyush Bharambe (18D070019)

Question 1

- Your task here is to implement the ISTA algorithm for the following three cases:

  1. Consider the image from the homework folder. Add iid Gaussian noise of mean 0 and variance 3 (on a [0,255] scale) to it, using the 'randn' function in MATLAB. Thus $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\eta}$ where $\boldsymbol{\eta} \sim \mathcal{N}(0, 3)$ blue(earlier the variance was mistakenly marked as 4). You should obtain $\boldsymbol{x}$ from $\boldsymbol{y}$ using the fact that patches from $\boldsymbol{x}$ have a sparse or near-sparse representation in the 2D-DCT basis.

  2. Divide the image shared in the homework folder into patches of size 8×8. Let $\boldsymbol{x_i}$ be the vectorized version of the i^{th} patch. Consider the measurement $\boldsymbol{y_i} = \boldsymbol{\Phi x_i}$ where $\boldsymbol{\Phi}$ is a 32×64 matrix with entries drawn iid from $\mathcal{N}(0, 1)$. Note that $\boldsymbol{x_i}$ has a near-sparse representation in the 2D-DCT basis $\boldsymbol{U}$ which is computed in MATLAB as 'kron(dctmtx(8)',dctmtx(8)')'. In other words, $\boldsymbol{x_i} = \boldsymbol{U\theta_i}$ where $\boldsymbol{\theta_i}$ is a near-sparse vector. Your job is to reconstruct each $\boldsymbol{x_i}$ given $\boldsymbol{y_i}$ and $\boldsymbol{\Phi}$ using ISTA. Then you should reconstruct the image by averaging the overlapping patches. You should choose the α parameter in the ISTA algorithm judiciously. Choose $\lambda = 1$ (for a [0,255] image). Display the reconstructed image in your report. State the RMSE given as $\|X(:) - \hat{X}(:)\|_2 / \|X(:)\|_2$ where $\hat{X}$ is the reconstructed image and X is the true image. [15 points]

Answer:

1. The original and the noisy (Gaussian Noise of standard deviation 3) Barbara images are shown below.

Original Barbara Image

Noisy Barbara Image

The RMSE between the original and the noisy image is 0.012663

Now, we will use the noisy image as the input image and the original image as the ground truth. Using the input image and using the prior information that the patches in the ground truth image have a sparse representation in the 2D-DCT basis, we will reconstruct the image using ISTA algorithm.

Note: In the ISTA algorithm, we have used $\lambda = 1$ and $\alpha = 1 + max(eigenvalue of A^T A)$ where A is the measurement matrix.

The reconstructed image using ISTA algorithm is as shown below.

Reconstructed Image using ISTA

The RMSE between the original and the reconstructed image using ISTA algorithm is 0.012366

2. In this part, we consider 32 compressive measurements for each 8×8 patch. The entries of the measurement matrix are drawn iid from $\mathcal{N}(0, 1)$. The ground truth image is the original Barbara image. We used ISTA algorithm on each overlapping patch and averaged the results. The reconstructed image is shown below.
   Note: This program takes roughly 124 seconds to run.

Reconstructed Image using ISTA on compressive measurements

The RMSE between the original and the reconstructed image using ISTA algorithm on compressive measurements is 0.42538

Q 2)

(a) Def^n: $X \in \mathbb{R}^{N \times p}$, $\beta \in \mathbb{R}^p$, $y \in \mathbb{R}^N$

We say the model matrix X satisfies the restricted eigenvalue condition with respect to subset $C \in \mathbb{R}^p$ if there is a constant $\gamma > 0$ such that

$$\frac{v^T X^T X v}{N \|v\|_2^2} \geq \gamma \text{ for all non zero } v \in C.$$

Explanation according to text:

We want our loss function $f_N(\beta)$ to be strictly convex where $f_N(\beta) = \frac{\|y - X\beta\|_2^2}{2N}$

Then the hessian $\nabla^2 f_N(\beta) = \left(\frac{X^T X}{N}\right)$ should have its eigenvalues uniformly bounded away from 0. But rank of $X^T X$ is $\min(N, p)$ which makes it rank deficient and hence not strongly convex, whenever $N < p$.

So, we relax the condition by allowing the function to be strictly convex in subset $C \in \mathbb{R}^p$. By restrictly strong convexity at β^* w.r.t. C if there is constant $\gamma > 0$ s.t.

$$\frac{v^T \nabla^2 f(\beta) v}{\|v\|_2^2} \geq \gamma \text{ for all non zero } v \in C$$

& for all $\beta \in \mathbb{R}^p$ in neighbourhood of β.

In specific case of linear regression, where $\nabla^2 f(\beta) = X^T X / N$, we get the restricted eigenvalue condition.

(b) $$G(v) = \frac{1}{2N}\|y - X(\beta^* + v)\|_2^2 + \lambda_N\|\beta^* + v\|_1$$

$$G(0) = \frac{1}{2N}\|y - X\beta^*\|_2^2 + \lambda_N\|\beta^*\|_1 \quad \text{— (1)}$$

Now, $\hat{v} = \hat{\beta} - \beta^*$.

$$G(\hat{v}) = \frac{1}{2N}\|y - X(\beta^* + \hat{\beta} - \beta^*)\|_2^2 + \lambda_N\|\beta^* + \hat{\beta} - \beta^*\|_1$$

$$= \frac{1}{2N}\|y - X\hat{\beta}\|_2^2 + \lambda_N\|\hat{\beta}\|_1 \quad \text{— (2)}.$$

We know $\hat{\beta}$ is minimizer of $J(\beta)$

where $J(\beta) = \frac{1}{2N}\|y - X\beta\|_2^2 + \lambda_N\|\beta\|_1$

$\therefore\ J(\hat{\beta}) \leq J(\beta^*)$

From (1) & (2),

$\therefore\ \underline{G(\hat{v}) \leq G(0)}$.

(c) From previous part $G(\hat{v}) \leq G(0)$

$$\therefore \frac{1}{2N}\|y - X(\beta^* + \hat{v})\|_2^2 + \lambda_N\|\beta^* + \hat{v}\|_1 \leq \frac{1}{2N}\|y - X\beta^*\|_2^2 + \lambda_N\|\beta^*\|_1$$

Since $y = X\beta^* + w$,

$$\frac{1}{2N}\|w - X\hat{v}\|_2^2 + \lambda_N\|\beta^* + \hat{v}\|_1 \leq \frac{1}{2N}\|w\|_2^2 + \lambda_N\|\beta^*\|_1$$

$$\frac{1}{2N}\left(\|w - X\hat{v}\|_2^2 - \|w\|_2^2\right) \leq \lambda_N\left(\|\beta^*\|_1 - \|\beta^* + \hat{v}\|_1\right)$$

$$\frac{1}{2N}\left((w - X\hat{v})^T(w - X\hat{v}) - w^Tw\right) \leq \lambda_N\left(\|\beta^*\|_1 - \|\beta^* + \hat{v}\|_1\right)$$

$$\frac{1}{2N}\left(-w^TX\hat{v} - \hat{v}^TX^Tw + (X\hat{v})^TX\hat{v}\right) \leq \lambda_N\left(\|\beta^*\|_1 - \|\beta^* + \hat{v}\|_1\right)$$

$$\frac{1}{2N}\left(-2w^TX\hat{v} + (X\hat{v})^TX\hat{v}\right) \leq \lambda_N\left(\|\beta^*\|_1 - \|\beta^* + \hat{v}\|_1\right)$$

(since $w^TX\hat{v}$ is scalar & $\text{scalar}^T = \text{scalar}$)

Rearranging we get,

$$\frac{\hat{v}^T X^T X \hat{v}}{2N} \le \frac{w^T X \hat{v}}{N} + \lambda_N \left(\|\beta^*\|_1 - \|\beta^* + \hat{v}\|_1 \right)$$

$$\frac{\|X\hat{v}\|_2^2}{2N} \le \frac{w^T X \hat{v}}{N} + \lambda_N \left(\|\beta^*\|_1 - \|\beta^* + \hat{v}\|_1 \right)$$

(d) β^* is a s-sparse vector.
$S = S(\beta^*)$ is the set of indices which represent the non-zero support of β^*. $|S(\beta^*)| = s$
$x_S \in \mathbb{R}^{|S|}$ is the subvector indexed by the elements of S and x_{S^c} is defined in analogous manner.
Now, since $\beta^*_{S^c} = 0$, we have $\|\beta^*\|_1 = \|\beta^*_S\|_1$
Using this, we have

$$\|\beta^* + \hat{v}\|_1 = \|\beta^*_S + \hat{v}_S\|_1 + \|\hat{v}_{S^c}\|_1$$
$$\ge \|\beta^*_S\|_1 - \|\hat{v}_S\|_1 + \|\hat{v}_{S^c}\|_1$$

{by triangle inequalities}

Thus,

$$\|\beta^*\|_1 - \|\beta^* + \hat{v}\|_1 \le \|\hat{v}_S\|_1 - \|\hat{v}_{S^c}\|_1$$

~~Putting~~ Using above inequality on the previous part we get,

$$\frac{\|X\hat{v}\|_2^2}{2N} \le \frac{w^T X \hat{v}}{N} + \lambda_N \left(\|\hat{v}_S\|_1 - \|\hat{v}_{S^c}\|_1 \right) \quad (1)$$

Holder's inequality: (from wikipedia).

Let (S, Σ, μ) be measure space and $p, q \in [1, \infty)$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$. Then,

for all measurable real or complex valued functions f and g,

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

Using holder's inequality for $p = \infty, q = 1$, & $f = w^T X$, $g = \hat{v}$,

$$\|w^T X \hat{v}\|_1 \leq \|w^T X\|_\infty \|\hat{v}\|_1$$

Since $w^T X \hat{v}$ is a scalar,

$$|w^T X \hat{v}| \leq \|w^T X\|_\infty \|\hat{v}\|_1$$

Also $\|w^T X\|_\infty = \|X w^T\|_\infty$

Using this in eqn ①,

$$\frac{\|X\hat{v}\|_2^2}{2N} \leq \frac{\|X^T w\|_\infty}{N} \|\hat{v}\|_1 + \lambda_n (\|\hat{v}_S\|_1 - \|\hat{v}_{S^c}\|_1)$$

(e) By assumption, $\frac{\|X^T w\|_\infty}{N} \leq \frac{\lambda_N}{2}$

$$\therefore \frac{\|X\hat{v}\|_2^2}{2N} \leq \frac{\lambda_N}{2}\|\hat{v}\|_1 + \lambda_N\left(\|\hat{v}_S\|_1 - \|\hat{v}_{S^c}\|_1\right)$$

$$\leq \frac{\lambda_N}{2}\left(\|\hat{v}_S\|_1 + \|\hat{v}_{S^c}\|_1\right) + \lambda_N\left(\|\hat{v}_S\|_1 - \|\hat{v}_{S^c}\|_1\right)$$

$$\leq \frac{3\lambda_N}{2}\|\hat{v}_S\|_1 - \frac{\lambda_N}{2}\|\hat{v}_{S^c}\|_1$$

$$\leq \frac{3}{2}\lambda_N\|\hat{v}_S\|_1 \qquad \left\{\frac{\lambda_N}{2}\|\hat{v}_{S^c}\|_1 \text{ is a positive term}\right\}$$

$$\therefore \frac{\|X\hat{v}\|_2^2}{2N} \leq \frac{3}{2}\lambda_N\|\hat{v}_S\|_1.$$

~~Now, as $\hat{v}$ we know $\|\hat{v}_S\|_1 \leq \|\hat{v}\|$~~

Now, we use the fact that if v is a k sparse vector then by a Cauchy Schwartz inequality $\|v\|_1 = \sum_{i=1}^{k} |v_i| \cdot 1 \leq \left(\sum |v_i|^2\right)^{1/2}\left(\sum 1^2\right)^{1/2}$

$$\leq \sqrt{k}\,\|v\|_2$$

Therefore,

$$\frac{\|X\hat{v}\|_2^2}{2N} \leq \frac{3}{2}\lambda_N\|\hat{v}_S\|_1 \leq \frac{3}{2}\lambda_N\sqrt{k}\,\|\hat{v}_S\|_2$$

$$\leq \frac{3}{2}\lambda_N\sqrt{k}\,\|\hat{v}\|_2 \qquad \left\{\|\hat{v}_S\|_2 < \|\hat{v}\|_2\right\}$$

$$\therefore \frac{\|X\hat{v}\|_2^2}{2N} \leq \frac{3}{2}\lambda_N\sqrt{k}\,\|\hat{v}\|_2$$

(f) From previous part,

$$\frac{\|X\hat{v}\|_2^2}{2N} \leq \frac{3}{2}\lambda_N \sqrt{k}\,\|\hat{v}\|_2 \quad \text{--- (1)}$$

Lemma 11.1 states that if $\lambda_N \geq 2\left\|\frac{X^T w}{N}\right\|_\infty > 0$,

then the error $\hat{v} = \hat{\beta} - \beta^*$ associated with any lasso solution $\hat{\beta}$ belongs to cone set $C(S;3)$

Now, we ~~a~~ apply the restricted eigenvalue condition to $\hat{v}$ belonging to the cone set

$$\therefore \frac{1}{N}\|X\hat{v}\|_2^2 \geq \gamma \|\hat{v}\|_2^2$$

$$\frac{1}{2N}\|X\hat{v}\|_2^2 \geq \frac{\gamma}{2}\|\hat{v}\|_2^2 . \quad \text{--- (2)}$$

From (1) & (2),

$$\frac{\gamma}{2}\|\hat{v}\|_2^2 \leq \frac{3}{2}\lambda_N \sqrt{k}\,\|\hat{v}\|_2$$

Since $\hat{v} = \hat{\beta} - \beta^*$,

$$\frac{\gamma}{2}\|\hat{v}\|_2 \leq \frac{3}{2}\lambda_n \sqrt{k}$$

$$\|\hat{\beta} - \beta^*\|_2 \leq \frac{3}{\gamma}\lambda_N \sqrt{k}$$

$$\|\hat{\beta} - \beta^*\|_2 \leq \frac{3}{\gamma}\sqrt{\frac{k}{N}}\,\sqrt{N}\,\lambda_N$$

Hence proved eqn 11.14 b [i.e. Theorem 11.1 (b)]

(g) The inequality ~~sho~~ $\lambda_N \geq \frac{2\|X^T w\|_\infty}{N}$

shows up in part (e); which gives us the cone set $C(S;3)$ for $\hat{v}$.
It is explained as follows,
from part (e), we have,

$$0 \leq \frac{\|X\hat{v}\|_2^2}{2N} \leq \frac{3\lambda_N}{2}\|\hat{v}_S\|_1 - \frac{\lambda}{2}\|\hat{v}_{S^c}\|_1$$

$$\therefore \frac{\lambda}{2}\|\hat{v}_{S^c}\|_1 \leq \frac{3\lambda_N}{2}\|\hat{v}_S\|_1$$

$$\therefore \|\hat{v}_{S^c}\|_1 \leq 3\|\hat{v}_S\|_1$$

This is the cone constraint for $C(S;3)$

(h) As explained in part (a), whenever $N<p$, the matrix X^TX is rank deficient and hence not strongly convex.
We relax the condition, and impose the strong convexity condition on a subset $C \in \mathbb{R}^p$ for all possible vectors $\hat{v}$.
For the function $y = X\beta^* + w$, we get restricted strong convexity ~~for~~ with respect to the ~~cone~~ subset C if there is a constant $\gamma > 0$ s.t.

$$\frac{\hat{v}^T X^T X v}{N\|\hat{v}\|_2^2} \geq \gamma \quad \text{for all } v \in C;\ v \neq 0$$

Now, it turns out that if we solve the regularised version of lasso with a "suitable" choice of λ_N, the lasso error satisfies the constraint $\|\hat{v}_{S^c}\|_1 \leq 3\|\hat{v}_S\|_1$ which is the cone constraint.

(i) Example 11.1 for classical linear Gaussian model says,

$$\|\hat{\theta} - \theta^*\|_2 \le \frac{c\sigma}{\gamma} \sqrt{\frac{\tau k \log n}{m}}$$

with probability $1 - 2e^{-\frac{1}{2}(\tau-2)\log p}$ and $w \in R^m$ is Gaussian with iid $N(0,\sigma^2)$ distribution.

Theorem 3 $A = \Phi\Psi$ of ~~some~~ size $m \times n$ has RIP of order $2s$ where $\delta_{2s} < 0.41$. Let solution of following be θ^* (measurement vector $y = \Phi\Psi\theta + \eta$):

$$\min \|\theta\|_1 \text{ s.t. } \|y - \Phi\Psi\theta\|_2^2 \le \epsilon$$

Then we have,

$$\|\theta^* - \theta\|_2 \le \frac{C_0}{\sqrt{s}} \|\theta - \theta_s\|_1 + C_1 \epsilon.$$

Advantages of theorem 3 over ex 11.1 bound are:

(i) The ~~sparsest~~ theorem 3 works for any sparsity level. ~~given it follows RIP with $\delta_{2s} < 0.41$~~, But ex 11.1 requires the sparsity level to be known.

ii) Theorem 3 also give bounds for compressible signals (which have small values in some ~~all~~ orthonormal basis) 11.1 considers only sparse signals.

iii) Theorem 3 is more general than 11.1. 11.1 requires proper choice of λ_N ~~[illegible]~~ ~~[illegible]~~ according to theorem 11.1 b & restricted eigenvalue condition.

(iv) 11.1 is only for gaussian noise which theorem 3 is for any general noise.

Advantages of 11.1 over ~~th~~ theorem 3:

(i) 11.1 requires restricted eigenvalue condition, required by theorem 11.1 (b), instead of RIP condition. RIP can be costly.

(ii) The ε value needs to be fixed properly according to noise level for theorem 3. This can be avoided by 11.1.

(iii) Even with knowledge of support set, since model has k free parameters no method can achieve squared l_2-error that decays more quickly than $\frac{k}{N}$.

Apart from logarithmic factor, the lasso rate matches the best possible one could achieve even if subset $S(\beta^*)$ were known a-priori. The rate given by 11.1 ea. is known to be minimax optimal meaning it cannot be improved upon by any estimator.

So the bound by 11.1 gives good gaurantees.

(j) ~~Da~~ Lasso: Constraint: $\|X^T w\|_\infty \le \frac{N\lambda_N}{2}$

Bound: $\|\hat{\beta} - \beta^*\|_2 \le \frac{3}{\gamma}\sqrt{K}\lambda_N$

Dantzig: Constraint: $\|A^T e\|_\infty \le \lambda$.

Bound: $\cancel{\|\hat{\beta}-\beta^*\|_2}$

$$\|\hat{x} - x\|_2 \le \frac{C_0 \sigma_K(x)_1}{\sqrt{K}} + C_3\sqrt{K}\lambda.$$

Both λ in Dantzig & λ_N in Lasso are similar. Both give upper bound on the largest value of $\|X^T w\|_\infty$ i.e. interaction between measurement matrix & noise vector.

The term $\sqrt{K}\lambda$ appear in both.

In fact for case when $\sigma_K(x)_1 = 0$ for Dantzig, we get $\|\hat{x} - x\|_2 \le C_3\sqrt{K}\lambda$ ~~&~~ similar to lasso $\|\hat{\beta} - \beta^*\|_2 \le \frac{3}{\gamma}\sqrt{K}\lambda_N$.

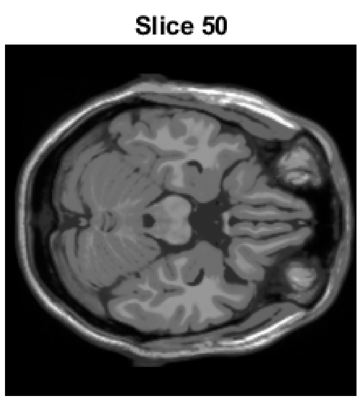
Question 3

- In this task, you will you use the well-known package L1_LS from `https://stanford.edu/~boyd/l1_ls/`. This package is often used for compressed sensing solution, but here you will use it for the purpose of tomographic reconstruction. The homework folder contains images of two slices taken from an MR volume of the brain. Create measurements by parallel beam tomographic projections at any 18 randomly angles chosen from a uniform distribution on $[0, \pi)$. Use the MATLAB function 'radon' for this purpose. Now perform tomographic reconstruction using the following method: (a) filtered back-projection using the Ram-Lak filter, as implemented in the 'iradon' function in MATLAB, (b) independent CS-based reconstruction for each slice by solving an optimization problem of the form $J(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \lambda\|\boldsymbol{x}\|_1$, (c) a coupled CS-based reconstruction that takes into account the similarity of the two slices using the model given in the lectures notes on tomography. For parts (b) and (c), use the aforementioned package from Stanford. For part (c), make sure you use a different random set of 18 angles for each of the two slices. The tricky part is careful creation of the forward model matrix $\boldsymbol{A}$ or a function handle representing that matrix, as well as the corresponding adjoint operator $\boldsymbol{A}^T$. Use the 2D-DCT basis for the image representation. Modify the objective function from the lecture notes for the case of three similar slices. Carefully define all terms in the equation but do not re-implement it. [3+7+8+7 = 25 points]

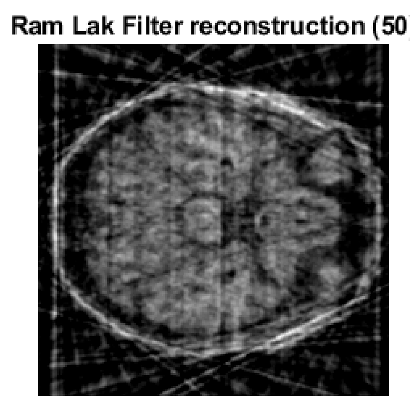
Answer:

Using 18 different angles of projection

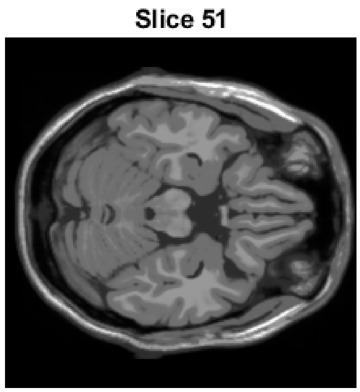
(a) In this part we perform the reconstruction using the filtered back-projection method using Ram-Lak filter. The results are shown in the figure below.
Note: For this part we have used 18 uniformly spaced angles of projection.



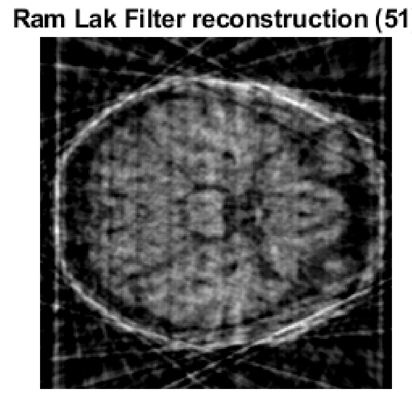
Original Slice 50



Reconstructed Slice 50 using Ram-Lak filter



Original Slice 51

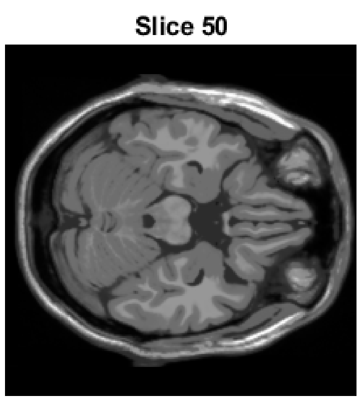


Reconstructed Slice 51 using Ram-Lak filter

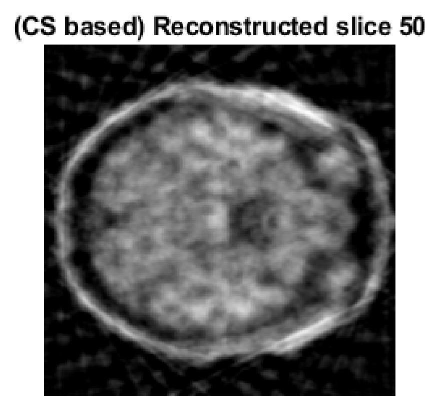
(b) In this part we have performed independent Compressed Sensing reconstruction for each slice. The optimization problem solved here is

$$J(x) = ||y - Ax||_2^2 + \lambda ||x||_1$$

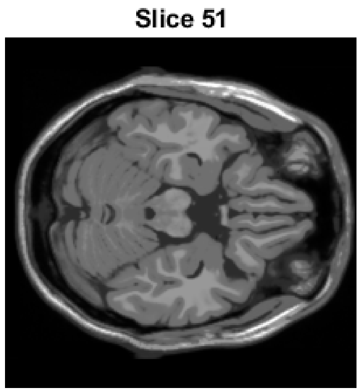
The results are shown in the figure below.
Note: For this part we have used 18 uniformly spaced angles of projection.



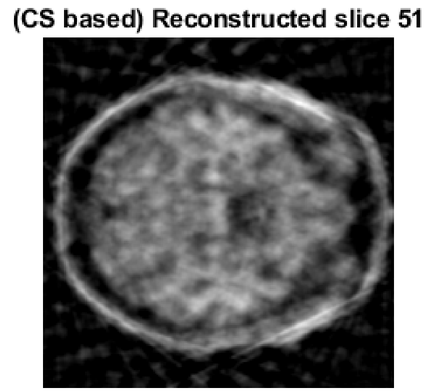
Original Slice 50



CS based reconstruction (Slice 50)



Original Slice 51



CS based reconstruction (Slice 51)

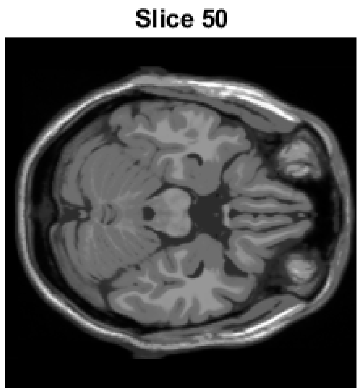
(c) In this part we perform a coupled CS-based reconstruction. The optimization problem solved here is

$$E(\beta_1, \beta_2) = \left\| \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} R_1 U & 0 \\ R_2 U & R_2 U \end{pmatrix} \begin{pmatrix} \beta_1 \\ \Delta\beta \end{pmatrix} \right\|_2^2 + \lambda \left\| \begin{pmatrix} \beta_1 \\ \Delta\beta \end{pmatrix} \right\|_1$$

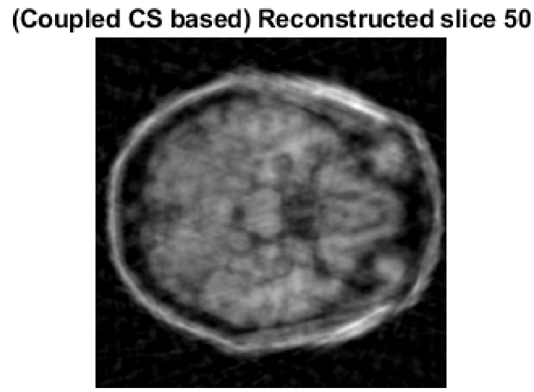
where,

- $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ is the measurement vector,
- $\begin{pmatrix} R_1 U & 0 \\ R_2 U & R_2 U \end{pmatrix}$ is the forward model matrix where U represents the inverse 2D DCT matrix and R_1 and R_2 are the radon projection matrices,
- $\begin{pmatrix} \beta_1 \\ \Delta\beta \end{pmatrix}$ is the vector of unknown parameters
  $\Delta\beta = \beta_2 - \beta_1$ is the difference between the two unknown parameters.

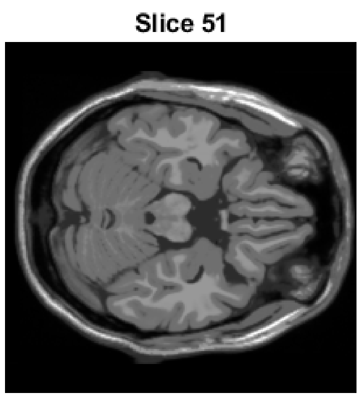
The result of the reconstruction is shown in the figure below.
Note: For this part we have used 18 randomly generated angles of projection.



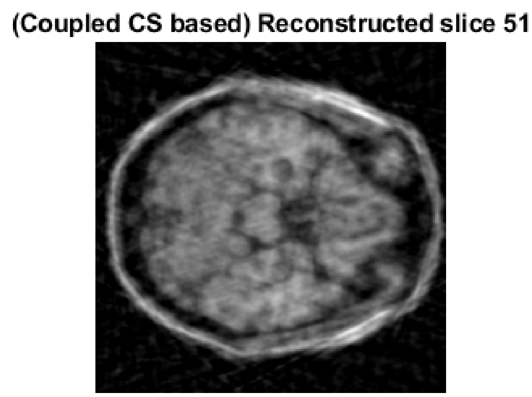
Original Slice 50



Coupled CS based reconstruction (Slice 50)



Original Slice 51



Coupled CS based reconstruction (Slice 51)

(d) In this part we derive the optimization problem for the coupled CS-based reconstruction using 3 consecutive slices. The optimization problem we solve is given as

$$E(\beta_1,\beta_2,\beta_3) = \|y_1 - R_1U\beta_1\|_2^2 + \|y_2 - R_2U\beta_2\|_2^2 + \|y_3 - R_3U\beta_3\|_2^2 + \lambda\left\|\beta_1\right\|_1 + \lambda\left\|\beta_2-\beta_1\right\|_1 + \lambda\left\|\beta_3-\beta_1\right\|_1$$
$$= \left\| \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} - \begin{pmatrix} R_1U & 0 & 0 \\ R_2U & R_2U & 0 \\ R_3U & 0 & R_3U \end{pmatrix} \begin{pmatrix} \beta_1 \\ \Delta\beta_1 \\ \Delta\beta_2 \end{pmatrix} \right\|_2^2 + \lambda \left\| \begin{pmatrix} \beta_1 \\ \Delta\beta_1 \\ \Delta\beta_2 \end{pmatrix} \right\|_1$$

where,

- $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ is the measurement vector,

- $\begin{pmatrix} R_1U & 0 & 0 \\ R_2U & R_2U & 0 \\ R_3U & 0 & R_3U \end{pmatrix}$ is the forward model matrix where U represents the inverse 2D DCT matrix and R_1 and R_2 and R_3 are the radon projection matrices,

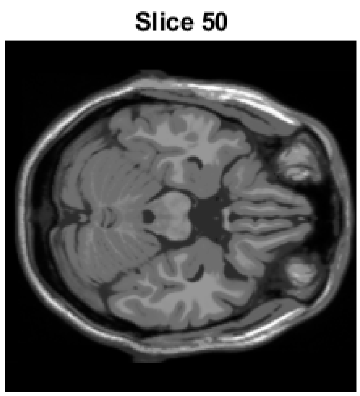
- $\begin{pmatrix} \beta_1 \\ \Delta\beta_1 \\ \Delta\beta_2 \end{pmatrix}$ is the vector of unknown parameters
  $\Delta\beta_1 = \beta_2 - \beta_1$ and $\Delta\beta_2 = \beta_3 - \beta_1$ are the difference between the two unknown parameters.

EXTRA PART

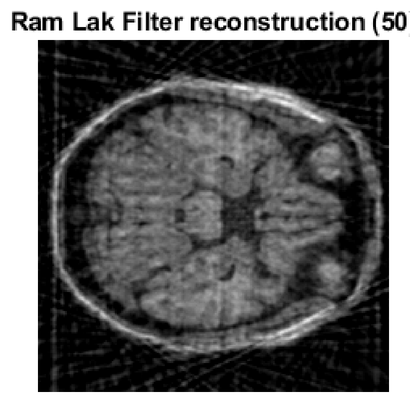
To check the results that we get by using more angles of projection, we used 30 angles of projection. Let us see the results.

Using 30 different angles of projection

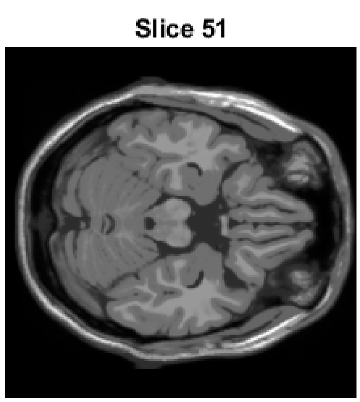
(a) FBP with Ram-Lak filter



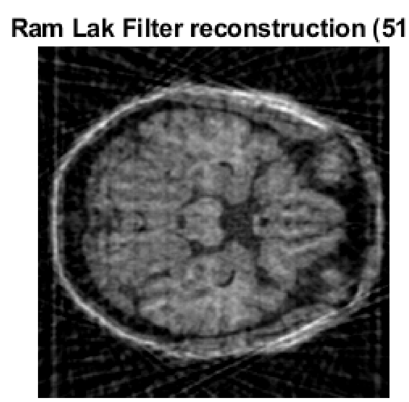
Original Slice 50



Reconstructed Slice 50 using Ram-Lak filter

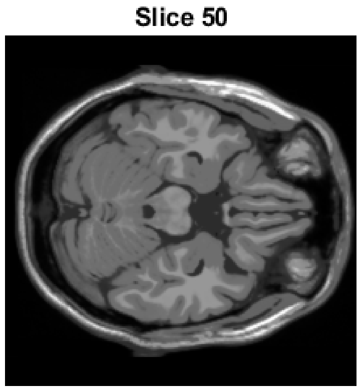


Original Slice 51

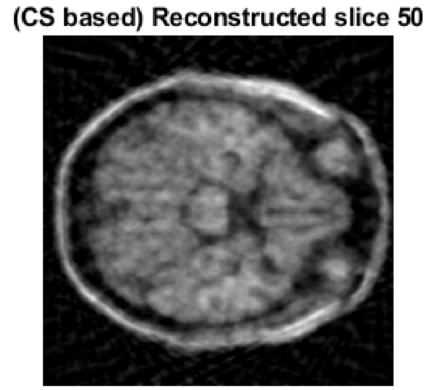


Reconstructed Slice 51 using Ram-Lak filter

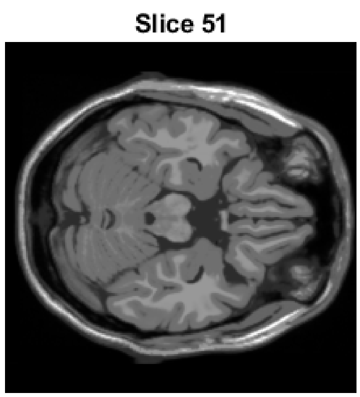
(b) CS based independent reconstruction



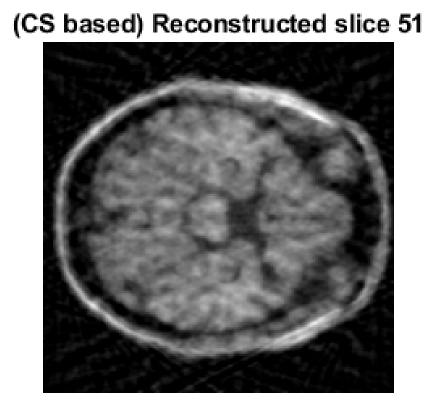
Original Slice 50



CS based independent reconstruction

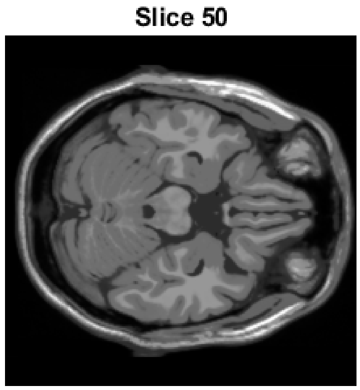


Original Slice 51

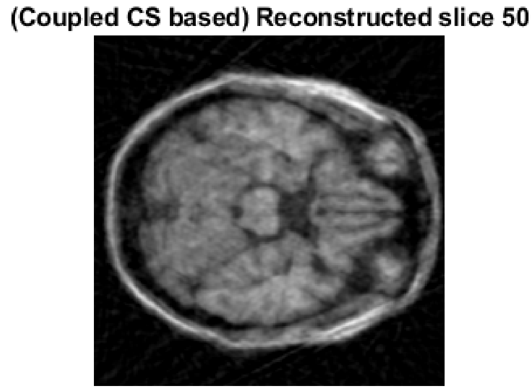


CS based independent reconstruction

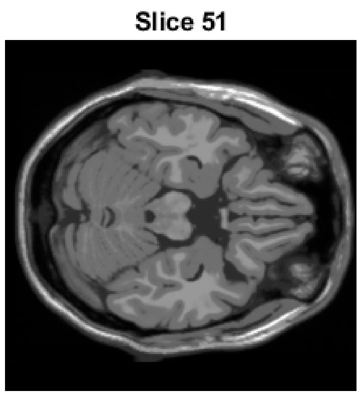
(c) CS based coupled reconstruction



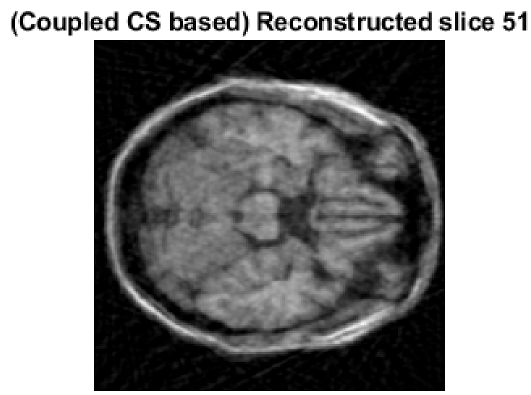
Original Slice 50



CS based coupled reconstruction



Original Slice 51



CS based coupled reconstruction

Here we can easily that the reconstruction using the coupled scheme is better than the independent scheme. Also, as the number of projection angles is increased the reconstruction quality is also improved.

Question 4

- Here is our Google search question again. You know of the applications of tomography in medicine (CT scanning) and virology/structural biology. Your job is to search for a journal paper from any other field which requires the use of tomographic reconstruction (examples: seismology, agriculture, gemology). State the title, venue and year of publication of the paper. State the mathematical problem defined in the paper. Take care to explain the meaning of all key terms clearly. State the method of optimization that the paper uses to solve the problem. [16 points]

Answer:

- Title: Ultrasound Transmission Tomography for Detecting and Measuring Cylindrical Objects Embedded in Concrete
- Journal name: Sensors (open access journal on the science and technology of sensors)
- Year of publication: 2017
- Authors: Lluveras Núñez, D.; Molero-Armenta, M.Á.; Izquierdo, M.Á.G.; Hernández, M.G.; Anaya Velayos
- Field: Civil structures and materials, Structural engineering

Brief Summary of the paper:
The paper explores the reconstruction of concrete using ultrasound attenuation tomography. The reconstruction is based on the attenuation of ultrasound signal. In the field of seismic tomography, travel time tomography is a common method. In travel time tomography, the travel time of a seismic signal is measured. The travel time of a seismic signal is the time it takes for the signal to travel from the source to the receiver. However, in attenuation tomography, along with the travel time, attenuation of elastic waves is also utilized for tomographic reconstruction. It is also stated that attenuation is vital for characterizing properties of rock and fluid because attenuation is more sensitive than velocity to some of the material condition properties, such as saturation, porosity, permeability and viscosity. The paper also evaluates the performace of ultrasound transmission based on attenuation to locate and estimate the most common materials that are embedded in concrete, reinforcements and natural and artificial voids. The paper uses fan tomographic inspection that considers the discretization of the attenuation spatial-distribution as shown below

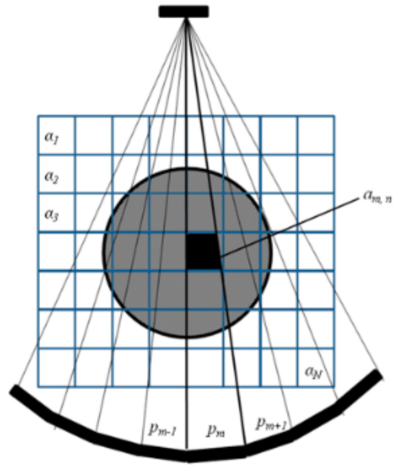


Figure 1. Fan tomographic inspection that considers the discretization of the attenuation spatial-distribution.

Mathematical Problem:

Typically, the inverse problem solved in attenuation tomography is the following:

$$p = A\alpha$$

where, p is the measurement vector, α is the vector that contains the discretized values of the attenuation spatial-distribution, and A is the matrix describing the linear relation. Since, the problem is ill-posed the inverse problem is solved using the iterative method. The iterative method is based on algebraic reconstruction methods update the solution by successively processing each equation (ray) separately. The process can be mathematically described by

$$\alpha^{k+1} = \alpha^k + \lambda \frac{p_{(k+1)} - a_{(k+1)}^T \alpha^k}{||a_{(k+1)}||^2} a_{(k+1)}$$

where $a_{(k+1)}$ is the (k + 1)th row of the matrix A. The relexation parameter is denoted by λ which varies between 0 and 1. The above mentioned iterative algorithm is also known as SIRT (Simultaneous Iterative Reconstruction Technique). In SIRT based ray-trace methodology, propagation of wave is based on Hyugen's principle

The experimental results in the paper showed that using a high-performance automatic ultrasonic inspection system the tomographic images are very useful for evaluating the internal inclusions of concrete structures. It was shown that the attenuation tomographic images allow to locate and to estimate the size of the inclusion; however, it was not possible to determine the type of embedded material.

Q5)

$\longrightarrow$ The Radon transform of $f(x,y)$ is,

$$R_\theta(f) = g(\rho,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\delta(x\cos\theta + y\sin\theta - \rho)\,dx\,dy \quad \text{—— (1)}$$

Now, $f'(x,y) = f(ax, ay)$; $a \neq 0$.

Then,

$$R_\theta(f') = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(ax, ay)\,\delta(x\cos\theta + y\sin\theta - \rho)\,dx\,dy$$

Let $v = ax$, $w = ay$

$$R_\theta(f') = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(v,w)\,\delta\left(\frac{v\cos\theta}{a} + \frac{w\sin\theta}{a} - \rho\right)\frac{1}{a^2}\,dv\,dw$$

$$= \frac{1}{a^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(v,w)\,\delta(v\cos\theta + w\sin\theta - a\rho)\,dv\,dw$$

$$= \frac{1}{a^2}\,g(a\rho,\theta) \quad \text{— from (1)}$$

$$\therefore R_\theta(f') = \frac{1}{a^2}\,g(a\rho,\theta) \,, \quad a \neq 0$$

Q6)

→ Let $f(x,y) = \delta(x,y)$

$$R_\theta(f) = g(\rho,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x,y)\,\delta(x\cos\theta + y\sin\theta - \rho)\,dx\,dy.$$

Now, we know, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x,y)\, h(x,y)\, dx\, dy$

$= h(0,0)$

Here we take $\delta(x\cos\theta + y\sin\theta - \rho)$ as $h(x,y)$

$$\therefore \boxed{R_\theta(f) = \delta(-\rho) = \delta(\rho)}$$

Now, by shifting property of Radon transform,

$$R(f(x-x_0, y-y_0))(\rho,\theta) = R(\cancel{g}f(x,y))(\rho - x_0\cos\theta - y_0\sin\theta)$$

To prove this, see ~~that~~ following,

LHS =

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x-x_0, y-y_0)\,\delta(x\cos\theta + y\sin\theta - \rho)\, dx\, dy$$

Put $u = x - x_0$, $v = y - y_0$,

$$= \iint f(u,v)\,\delta(u\cos\theta + v\sin\theta - (\rho - x_0\cos\theta - y_0\sin\theta))\,dx\,dy$$

$$= R(f(x,y))(\rho - x_0\cos\theta - y_0\sin\theta) = RHS.$$

Using this,

$$R(\delta(x-x_0, y-y_0))(\rho,\theta) = R(\delta(x,y))(\rho - x_0\cos\theta - y_0\sin\theta, \theta)$$

$$= \delta(\rho - x_0\cos\theta - y_0\sin\theta)$$

Radon transform of unit impulse is $\delta(\rho)$

& of shifted unit impulse is $\delta(\rho - x_0\cos\theta - y_0\sin\theta)$