Tomographic Image Reconstruction Using Few Projections
A Review on CS and Numerical Algorithms

Han WANG
CEA-LID/Thales
2 février 2009
Overview of Reconstruction Problem

Beer-Lambert’s law:

\[ m(X) = N_0 \exp(-AX) \]  \hspace{1cm} (1)

Physical model (Sampling under Poisson law):

\[ Y \sim \mathcal{P}(m(X)) \]  \hspace{1cm} (2)
\[ B = \log N_0 - \log Y \]  \hspace{1cm} (3)

MAP Estimation:

\[ \min_{x} - \log \mathbb{P}(y|x) - \log \mathbb{P}(x) \]  \hspace{1cm} (4)
Approximation & Choice of prior

High dose + Edge-preservation reconstruction:

- Approximation of $-\log \mathbb{P}(y|x)$ by quadratic term
- Choice of regularization prior $TV(x)$

$$\min_{x} \frac{1}{2} \| Ax - b \|_2^2 + \lambda TV(x)$$  \hspace{1cm} (5)

Constrainted optimization formulation:

$$\min_{x} TV(x) \quad \text{s.t.} \quad \| Ax - b \|_2^2 \leq \varepsilon^2$$  \hspace{1cm} (6)

Good reconstruction quality observed with small number of projections for sparse object.
Sparsity Prior

$TV$ semi-norm prompts "cartoon-like" object:

$$TV(x) := \sum_k |\nabla x_k|$$  \hspace{1cm} (7)

$$TV(x) := \sum_k \sqrt{\partial_x x_k^2 + \partial_y x_k^2}$$  \hspace{1cm} (8)

Sparsifying transform: basis, frames (redundant basis), or sub-basis.

- Natural basis: Spikes
- $TV$: Sharp edges/Time-sparsity
- DCT: Periodic smooth
- Wavelet: Texture...
Sparsifying transform : Analysis vs. Synthesis

- Analysis. For invertible transform (basis):
  \[ x = D\alpha \iff D^{-1}x = \alpha \]  
  \hspace{1cm} (9)

- Synthesis. For non invertible transform (frame, sub-basis):
  \[ x = D\alpha \]  
  \hspace{1cm} (10)

Transformed coefficients \( \alpha \) are:
- \( S \)-sparse: only \( S \) non zero terms
- approximatively sparse: rapide decreasing rate → sparse approximation error
Motivation

Medical objects are often sparse:
- Blood vessels: natural basis
- Tissues: wavelet
- Organs/Bones: TV
- ...

Medical measurements are expensive to take:
- X-ray: dose
- MRI: time

Taking advantage of sparsity to reduce measurements?
Find the sparse solution(s)

Measurement matrix $A$ (Radon transform) is underdetermined.

- MSE solution (ART, FBP)

$$\min_{x} \|x\|_2^2 \quad \text{s.t.} \quad Ax = b \iff x^* = A^\dagger b \quad (11)$$

- Sparest solution

$$\min_{x} N(x) \quad \text{s.t.} \quad Ax = b \quad (12)$$

Sparsity measure $N(x)$:

- Natural: $\|x\|_0 := \# \{ k : x_k \neq 0 \}$
- Nonconvex $l_p$ norm: $\|x\|_p^p := \sum_k |x_k|^p$, for $0 < p < 1$
- Convex $l_1$ norm: $\|x\|_1 := \sum_k |x_k|$
Why use $l_p(0 \leq p \leq 1)$ norm as sparsity measure?

Theorem (Decreasing rate of $R$-sphere vector)

Given $x \in \mathbb{R}^N$, with $\|x\|_p = R$, and $|x_1| \geq |x_2| \geq \ldots \geq |x_N|$, then

$$|x_k| \leq Rk^{-1/p}$$

- High dimension ball is **almost empty**.
- $l_p(p \leq 1)$ ball is **near the axis**.
Sparse reconstruction with $l_p$ norm

The true object $x$ is (app.) sparse under sparsifying transform:

$$\min_{\alpha} \|\alpha\|_p^p \quad \text{s.t.} \quad D\alpha = x \quad (13)$$

Applied to sensing matrix $A$:

$$\min_{\alpha} \|\alpha\|_p^p \quad \text{s.t.} \quad AD\alpha = b \quad (14)$$

Noisy case:

$$\min_{\alpha} \|\alpha\|_p^p \quad \text{s.t.} \quad \|AD\alpha - b\|_2^2 \leq \epsilon^2 \quad (15)$$

$A, D$ are generally underdetermined.
Reconstruction: $x^* = D\alpha^*$
Reconstruction error & choice of $l_p$ norm

Reconstruction error $\|x - x^*\|_2$ depends on:

- Uniqueness of sparse representation: Sparsifying transform $D$
- Dimension of solution space: Measurement matrix $A$
- Observation noise $\varepsilon = b - Ax$

Choice of $l_p$-norm ($0 \leq p \leq 1$):

- Sparsity or approximatively sparsity
- Solution uniqueness condition: local optimum for $p < 1$
- Efficient numerical algorithm
Idea of CS

To recovery a $S$-sparse signal $x$ with underdetermined sensing matrix $A$:

- $P_0$ problem (NP-hard): $\min_x \|x\|_0$ s.t. $Ax = b$
- $P_1$ problem (convex): $\min_x \|x\|_1$ s.t. $Ax = b$

Equivalence of $P_0$ and $P_1$?

Definition (Restricted Isometry Property)

RIP of $A$ is the smallest $\delta_S > 0$ s.t. for all $S$-sparse signal $x$:

$$(1 - \delta_S)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_S)\|x\|_2^2$$

(16)

i.e., all $S$-submatrix of $A$ is a ”restricted” isometry.
**$P_0$ and $P_1$ equivalence**

**Theorem (Perfect reconstruction)**

Given that $2S$-RIP of $A$ is $\delta_{2S} \leq \sqrt{2} - 1$, for a $S$-sparse true signal $x$, $P_1$ solution $x^*$ is exactly $x$.

**Theorem (Almost Perfect reconstruction)**

Given that $2S$-RIP of $\Phi$ is $\delta_{2S} \leq \sqrt{2} - 1$, for a general true signal $x$, $P_1$ solution $x^*$ obeys:

\[
\|x^* - x\|_1 \leq C\|x - x_S\|_1 \quad \text{and} \quad \|x^* - x\|_2 \leq C\frac{\|x - x_S\|_1}{\sqrt{S}}
\]

where $x_S = (x_1, x_2, ..x_S, 0..)$ is the $S$ biggest term approximation of $x$. 
Robustness

For noisy observation $b = Ax + \varepsilon$:

$$\min_x \|x\|_1 \text{ s.t. } \|Ax - b\|_2 \leq \varepsilon \quad (NP_1)$$

Theorem (Robust reconstruction)

*Under the same hypothesis, the solution $x^*$ to $(NP_1)$ obeys:*

$$\|x^* - x\|_2 \leq C_1 \varepsilon + C_2 \frac{\|x - x_S\|_1}{\sqrt{S}} \quad (17)$$
Random matrices

Sensing matrices $A$ satisfy RIP with probability $1 - O(e^{-N})$:

- Gaussian matrix: $S \leq C.K/\log(N/K)$
- Binary matrix: $\mathbb{P}(A_{i,j} = \pm 1/\sqrt{K}) = 0.5$
- Fourier matrix: $S \leq C.K/\log N$ (conjectured)

$K$ randomly chosen observations can measure the essential part of $x$!
$P_0$ and $P_p$ equivalence

$P_p$ problem:

$$\min_{x} \left\| x \right\|_p^p \quad \text{s.t.} \quad Ax = b$$  \hspace{1cm} (18)

Theorem ($P_p$ robust reconstruction)

Given that following holds for some $k > 1$, and $kS \in \mathbb{Z}_+$:

$$\delta_{kS} + k^{2/p-1} \delta_{(k+1)S} < k^{2/p-1} - 1$$  \hspace{1cm} (19)

Then for arbitrary true signal $x$, the solution $x^*$ to $P_p$ obeys:

$$\|x^* - x\|_2^p \leq C_1 \epsilon^p + C_2 \frac{\|x - xS\|_p^p}{S^{1-p/2}}$$  \hspace{1cm} (20)

Particularly, one gets $P_1$ results for $k = 3$ and $p = 1$. 
Sparse reconstruction with $TV$

For partial Fourier measurement $A$, $TV$ regularization equals to:

$$\min_{\alpha} \|\alpha\|_1 \quad \text{s.t.} \quad A\alpha = b', \text{ with } \alpha = \nabla x$$

(21)

**Fig.** $P_1$ reconstruction with 18 projections, 15 projections and 13 projections. For the left one, $P_p$ reconstruction with $p = 0.5$ achieves the same result using only 10 projections.
Problems & variant models

For general model:

$$\min_{\alpha} \|\alpha\|_p^p \quad \text{s.t.} \quad AD\alpha = b \quad (22)$$

RIP property of sensing matrix $A$ could be violated since:

- $D$ is deterministic sparsifying transform
- Unique sparse representation doesn’t imply unique sparse reconstruction

Analysis model (example):

$$\min_x TV(x) + \lambda \|F x\|_1 \quad \text{s.t.} \quad Ax = b \quad (23)$$

No theoretical estimation available for the reconstruction error?
Sparsifying transform

A dictionary $D = [d_1...d_M]$ of $\mathbb{C}^N$ satisfies:

$$\|d_m\|_2 = 1, \quad \text{and} \quad \text{Span}(D) = \mathbb{C}^N \quad (24)$$

Find the sparsest representation(s) when dictionary is redundant ($M > N$):

- $P_0 : \min_{\alpha} \|\alpha\|_0$ s.t. $D\alpha = x$
- $P_1 : \min_{\alpha} \|\alpha\|_1$ s.t. $D\alpha = x$

Equivalence: the \textbf{same unique} solution $\alpha^*$. 
$P_0$ and $P_1$ equivalence

Mutual Coherence of a dictionary:

$$\mu(D) \triangleq \max_{i \neq j} |\langle d_i, d_j \rangle|$$ (25)

$P_0 \ P_1$ equivalence conditions:

- General dictionary: $\|x\|_0 < \left( 1 + \mu(D)^{-1} \right)/2$
- Union of 2 orthonormal bases: $\|x\|_0 < (\sqrt{2} - 0.5)\mu(D)^{-1}$
- Union of $L$ orthonormal bases:

$$\|x\|_0 < \left( \sqrt{2} - 1 + \frac{1}{2(L - 1)} \right) \mu(D)^{-1}$$
Incoherence between measurement and representation

RIP property remains true for general model:

$$\min_{\alpha} \|\alpha\|_1 \quad \text{s.t.} \quad AD\alpha = b$$

with overwhelming probability if:

$$K \geq \mu(A, D)^2 CS(\log N)^4$$

where $$\mu(A, D) := \max_{i,j} |\langle a_i, d_j \rangle|.$$ 

Design the dictionary $$D$$ to be incoherent with $$A.$$
Robust $P_1$ problem

Robust $P_1$ problem:

$$\min \|x\|_1 \quad \text{s.t.} \quad \|Ax - b\|_2^2 \leq \varepsilon^2$$

(NP$_1$)

Equivalent formulation (Lagrangian Relaxation):

$$\min_x \lambda \|x\|_1 + \frac{1}{2}\|Ax - b\|_2^2$$

$(Q_\lambda)$

LR is easier to solve than $P_1$.
Gradient Projection Sparse Reconstruction (GPSR)

LR can be reformulated to

$$\min_{z \geq 0} \frac{1}{2} z^* B z + c^* z$$

(BCQP)

with

$$z = \begin{bmatrix} x^+ \\ x^- \end{bmatrix}, \quad c = \begin{bmatrix} \lambda + A^* b \\ \lambda - A^* b \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} A^* A & -A^* A \\ -A^* A & A^* A \end{bmatrix}$$

(28)

Efficiently resolved (BCQP) by the general Gradient Projection method.
Direct solution: Linear Programming

\( P_1 \) can be transformed into Linear Programming:

\[
\min \sum_{k}(x^+_k + x^-_k) \quad \text{s.t.} \quad A(x^+ - x^-) = b, \quad \text{and} \quad x^+, x^- \geq 0 \quad (29)
\]

then solved by:
- Simplex method
- Interior Point method

Generally slow when \( A \) is
- large scale
- dense
- submatrix of fast orthonormal transform (DFT, Wavelet).
Indirect solution

General convex optimization problem:

$$\min_x J(x) + H(x) \quad (B_0)$$

- data fitting term: $H(x) = \|Ax - b\|_2^2/2$
- regularization term: $J(x) = \lambda \|x\|_1$

Bregman relaxation:

$$x^* = \arg \min_x B^p_J(x, x') + H(x) \quad (B_1)$$

with Bregman distance:

$$B^p_J(u, v) = J(u) - J(v) - \langle p, u - v \rangle, \quad p \in \partial J(v) \quad (30)$$
Bregman Iterative Regularization

Idea of BIR:

\[ x^* \text{ solves } (B_1) \iff 0 \in \partial \{ B_P^p(x, x') + H(x) \} \]

which gives:

\[ p - \nabla H(x^*) \in \partial J(x^*) \]

**Bregman Iterative Regularization**:

1. Initialization: \( k = 0, p_0 = 0, x_0 = 0 \)
2. while Not converged do
3. \( x_{k+1} \leftarrow \arg \min_x B_J^{p_k}(x, x_k) + H(x) \)
4. \( p_{k+1} \leftarrow p_k - \nabla H(x_{k+1}) \in \partial J(x_{k+1}) \)
5. \( k \leftarrow k + 1 \)
6. end while

Convergence guaranteed in Bregman distance.
Bregman Iterative Regularization

At step $k$, $\min_x B^{p_k}_J(x, x_k) + \frac{1}{2} \|Ax - b\|_2^2$ is equivalent to:

$$\min_x J(x) + \frac{1}{2} \|Ax - b_{k+1}\|_2^2,$$

with $b_{k+1} = b + b_k - Ax_k$ (31)

which can be efficiently solved by GPSR.

1: Initialization : $k = 0, b_0 = 0, x_0 = 0$
2: while Not converged do
3: $b_{k+1} \leftarrow b + b_k - Ax_k$
4: $x_{k+1} \leftarrow \arg \min_x J(x) + \|Ax - b_{k+1}\|_2^2 / 2$
5: $k \leftarrow k + 1$
6: end while

In finite steps, $x_k$ converges to the true solution of $P_1$. 
Conclusion

Sparse reconstruction works thanks:
- Sparse representation of object in (redundant)dictionary
- Efficient numerical algorithms
- Sensing matrix $A$ well-behaved

Challenges:
- Tomography projector $A$ : Polar Fourier Transform
- Numerical methods non predicted by theoretical analysis
- Bridging the representation and reconstruction problems