Iterative Reconstruction Methods in Computed Tomography

J. Webster Stayman
Dept. of Biomedical Engineering, Johns Hopkins University

Power of Iterative Reconstruction

FBP reconstruction
Iterative Reconstruction

0.8 mAs/frame
0.1 mAs/frame

(80 kVp, 360 projections)

Learning Objectives

- Fundamentals of iterative methods
  - Approach to forming an iterative algorithm
  - Identify particular classes of methods
- Advantages of iterative approaches
  - Intuition behind what these algorithms do
  - Flexibility of these techniques
- Images produced by iterative methods
  - Differences from traditional reconstruction
  - Image properties associated with iterative reconstruction
**Iterative Reconstruction**

- What is iterative reconstruction?
  - Requires a data model
  - Enforce desirable image properties
  - Encourage smoothness, edges, etc.
  - Need a measure of "better"

- How can we make the image better?
  - Get a better match to the data
  - Enforce desirable image properties
  - Encourage smoothness, edges, etc.
  - Need a measure of "better"

**Building an Iterative Technique**

- Define the objective
  - Find the volume that best fits the data and desired image quality
    
    \[
    \text{volume} = \arg \max \{ \text{data}, \text{model} \} \& \text{image properties} \}
    
    \[
    \hat{\mu} = \arg \max \| \mathbf{y} - \mathbf{y}(\mu) \| \}
    
- Devise an algorithm that solves the objective
  - Iteratively solve for \( \hat{\mu} \)
  - Decide when to stop iterating (and how to start)

**Reconstruction Choices**

- Forward Model Driven
- Regularized ART
- PICCS
- Image Properties
- Image Restoration
- Maximum Likelihood
- Penalized Likelihood
- MAP
- PWLS
- Veo SAFIRE
- ASIR AIDR
- Image Denoising

*Disclaimer: The exact details of commercially available reconstruction methods are not known by the author.*
Model-based Approaches

- Transmission Tomography Forward Model
  - Projection Physics, Beer’s Law
  \[ P(\text{photon survives}) = \exp\left(-\int \bar{\mu}(x,y) \, dl \right) \]
  \[ \bar{y}_j = \lambda \{\text{number of photons} \} = l_0 \exp\left(-\int \bar{\mu}(x,y) \, dl \right) \]
- Need a Parameterization of \( \bar{\mu} \)

Parameterization of the Object

- Continuous-domain object
- Want finite number of parameters
- Choices of basis functions:
  - Point Samples - Bandlimited
  - Contours – Piecewise constant
  - Blobs, Wavelets, “Natural Pixels,” ...
- Voxel Basis

\[ \bar{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \end{bmatrix} \]

\[ \bar{\mu}(x,y) \equiv \mathbf{B}\bar{\mu} \]

Projection

- Linear operation
  \[ \bar{\gamma}_j = l_0 \exp\left(-\int \bar{\mu}(x,y) \, dl \right) \]
- Discrete-Discrete for parameterized problem
  \[ \bar{\gamma}_j = l_0 \exp\left(-\sum_{i=1}^{p} a_{ij}\bar{\mu}_i \right) \]
- System matrix

\[ \mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pp} \end{bmatrix} \]
**Forward Model**
- Mean measurements as a function of parameters
  \[ y(\mu) = I_0 \exp(-A\mu) \]

\[ \begin{array}{ccc}
\mu & A\mu & I_0 \exp(-A\mu) \\
\end{array} \]

**Aside: Backprojection**
- Projection
- Backprojection

**A Simple Model-Driven Approach**
- Forward Model
  \[ \tilde{y}(\mu) = I_0 \exp(-A\mu) \]
- Objective and Estimator
  \[ \hat{\mu} = \arg\min \| y - \tilde{y}(\mu) \| \quad \hat{\mu} = \arg\min -\log \left( \frac{y}{I_0} \right) + A\mu \]
  \[ \hat{\mu} = \arg\min \| A\mu - \vec{l} \| = \left[ A^T A \right]^{-1} A^T \vec{l} \]
  \[ \begin{array}{ccc}
\text{Projection-Backprojection} & \text{Backprojection} \\
\end{array} \]
- For a squared-distance metric:
  - Many possible algorithms
  - Can be equivalent to ART, which seeks \( A\mu = \vec{l} \)
Flexibility of Iterative Methods

\[ \hat{\mu} = \arg \min |A\mu - l| = [A^T A]^{-1} A^T l \]

- For a well-sampled, parallel beam case:
  \[ A^T A \approx \frac{1}{r} \rightarrow [A^T A]^{-1} x = F^{-1} [\phi] + x \]

- For other cases:
  \[ [A^T A]^{-1} A^T l \]

- Iterative methods implicitly handle the geometry
  - Find the correct inversion for the specific geometry
  - Cannot overcome data nullspaces (needs complete data)

More Complete Forward Models

- More physics
  - Polyenergetic beam, energy-dependent attenuation
  - Detector effects, finite size elements, blur
  - Source effects, finite size element
  - Scattered radiation
- Noise
  - Data statistics
  - Quantum noise = x-ray photons
  - Detector Noise

Toy Problem

3 Random Variables
Different std dev (\( \sigma_1, \sigma_2, \sigma_3 \))

Best way to estimate \( \mu \)!?
**Maximum Likelihood Estimation**

- Find the parameter values most likely to be responsible for the observed measurements.

- Properties
  - Asymptotically unbiased and efficient under general conditions

- Likelihood Function
  \[ L(y; \mu) = p(y_1, y_2, \ldots, y_N | \mu_1, \mu_2, \ldots, \mu_N) \]

- Maximum Likelihood Objective Function
  \[ \hat{\mu} = \arg \max L(y; \mu) \]

**ML Estimate for the Toy Problem**

Likelihood function:

\[ L(y, \mu) = p(y | \mu) = \prod_{i=1}^{N} p(y_i | \mu) \]

\[ p(y_i | \mu) = \frac{1}{\sqrt{2\pi} \sigma_i} \exp \left\{ -\frac{1}{2} \left[ \frac{y_i - \mu}{\sigma_i} \right]^2 \right\} \]

Log-Likelihood function:

\[ \log L(y, \mu) = -\frac{1}{2} \sum_{i=1}^{N} \log(2\pi\sigma_i^2) - \sum_{i=1}^{N} \frac{1}{2} \left( \frac{y_i - \mu}{\sigma_i} \right)^2 \]

Maximize over \( \mu \):

\[ \hat{\mu} = \arg \max \log L(y, \mu) \]

\[ \frac{\partial}{\partial \mu} \log L(y, \mu) = \sum_{i=1}^{N} \frac{y_i - \mu}{\sigma_i^2} \]

**ML for Tomography**

- Need a noise model
- Depends on the statistics of the measurements
- Depends on the detection process

- Common choices:
  - Poisson - x-ray photon statistics
    \[ y_i \sim \text{Poisson}(\mu_i) \]
    \[ p(y_i | \mu_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!} \]
  - Poisson-Gaussian mixtures - photons + readout noise
    \[ y_i \sim \text{Poisson}(\mu_i) + \mathcal{N}(0, \sigma_i^2) \]
  - Gaussian (variable variances) - approx. many effects
    \[ y_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \]
    \[ p(y_i | \mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left\{ -\frac{1}{2} \frac{(y_i - \mu_i)^2}{\sigma_i^2} \right\} \]
**Poisson Likelihood Function**

- Marginal Likelihoods
  \[ p(y_i | \mu) = \exp(-y_i \mu) \frac{[\lambda(\mu)]^{y_i}}{y_i!} \]
- Likelihood
  \[ L(y; \mu) = \prod_{i=1}^{n} p(y_i | \mu) = \prod_{i=1}^{n} \exp(-y_i \mu) \frac{[\lambda(\mu)]^{y_i}}{y_i!} \]
- Log-Likelihood
  \[ \log L(y; \mu) = \sum_{i=1}^{n} y_i \log \lambda(\mu) - y_i \mu - \log y_i! \]
- Objective Function
  \[ \hat{\mu} = \arg \max \log L(y | \mu) \]

**Gaussian Likelihood Function**

- Marginal Likelihoods
  \[ l_i = -\log \left( \frac{y_i}{\mu} \right) \quad p(l_i | \mu) = \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2\mu} (l_i - \mu)^2 \right] \]
- Likelihood
  \[ L(l; \mu) = p(l | \mu) = \prod_{i=1}^{n} p(l_i | \mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2\mu} (l_i - \mu)^2 \right] \]
- Log-Likelihood
  \[ \log L(l; \mu) = \sum_{i=1}^{n} \left( -\frac{1}{2\mu} (l_i - \mu)^2 \right) \]
- Objective Function and Estimator
  \[ \hat{\mu} = \arg \max \log L(l | \mu) = \left[ (X^T X)^{-1} X^T \right] l \]

**Iterative Algorithms**

- Plethora of iterative approaches
  - Expectation-Maximization – General Purpose Methodology
  - Gradient-based Methods – Coordinate Ascent/Descent
  - Optimization Transfer – Paraboloidal Surrogates
  - Ordered-Subsets methods
- Properties of iterative algorithms
  - Monotonicity
  - True convergence
  - Speed
  - Complexity
Statistical Reconstruction Example

- **Test Case**
  - Single slice x-ray transmission problem
  - 512 x 512 0.3 mm volume
  - 400 detector bins over 180 angles (360 degrees)
  - Poisson noise: 1e5 counts per 0.5 mm detector element
  - SO: 380 mm, DO: 220 mm

- **Reconstruction**
  - Voxel basis: 512 x 512 0.3 mm voxels
  - Maximum-likelihood objective
  - EM-type algorithm
  - Initial image – constant value
  - Lots of iterations

ML-EM Iterations

FBP vs ML-EM Comparison
Enforcing Desirable Properties

- FBP
  - Filter designs – cutoff frequencies
- Iterative methods
  - Modify the objective to penalize “bad” images
  - Discourage noise
  - Preserve desirable image features
  - Other prior knowledge

Image-domain Denoising:
\[ \hat{\mu} = \arg \max \mathcal{F}(\mu) - \beta R(\mu) \]

Modal-based Reconstruction:
\[ \hat{\mu} = \arg \max \mathcal{F}(y, \mu) - \beta R(\mu) \]

Local Control of Image Properties

- Pairwise Penalty
  - Penalize the difference between neighboring voxels
  \[ R(\mu) = \sum_{j \in \mathcal{N}} w_j \psi(\mu_j - \mu_k) \]
  \[ k_{4x4} = \begin{bmatrix} -1 & -1 & -1 & -1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \]

- Quadratic Penalty
  \[ R(\mu) = \sum_{j \in \mathcal{N}} w_j (\mu_j - \mu_k)^2 \]

Penalized-Likelihood Example

- Forward Model
  - Fan-beam Transmission Tomography
  - 512 x 512 0.3 mm volume
  - 400 detector bins over 180 angles (360 degrees)
  - Poisson: \{1e4, 1e3\} counts per 0.5 mm detector element
  - SO: 380 mm, DO: 220 mm

- Objective Function
  - Poisson Likelihood
  - Quadratic, first-order Penalty

- Algorithm
  - Separable Paraboloidal Surrogates
  - 400 iterations – well-converged
Other Penalties/Energy Functions

\[ R(\mu) = \sum \sum w_{jk} \phi(\mu_j - \mu_k) \]

- Quadratic penalty
  - Tends to enforce smoothness throughout image
  - Increasingly penalizes larger pixel differences

- Non-quadratic penalties
  - Attempt to preserve edges in the image
  - Once pixel differences become large allow for decreased penalty (perhaps a relative decrease)

- Flexibility: Even more penalties
  - Wavelets and other bases, non-local means, etc.
Nonquadratic Penalties

- **Choices**
  - Truncated Quadratic
    \[ \psi(t; \delta) = |t|_{\delta} \]
  - Lange Penalty
    \[ \psi(t; \delta) = \delta \left[ \frac{1}{\delta} \log(1 + |t|) \right] \]
  - P-Norm
    \[ \psi(t; p) = |t|^p \]

**Truncated Quadratic**

**Lange**

**P-Norm**

---

**Test Case**
- Single slice x-ray transmission problem
- 480 x 480 0.8 mm volume
- 1000 detector bins over 360 angles (360 degrees)
- Poisson noise: 1e5 counts / 0.76 mm detector element
- SO: 600 mm, DO: 600 mm

**Reconstruction**
- Penalized-likelihood objective
- Shift-Invariant Quadratic Penalty
- Separable paraboloidal surrogates
- 200 iterations
- Voxel basis: 480 x 480 0.8 mm voxels

---

Closer Look at Image Properties

- **Test Case**
  - Single slice x-ray transmission problem
  - 480 x 480 0.8 mm volume
  - 1000 detector bins over 360 angles (360 degrees)
  - Poisson noise: 1e5 counts / 0.76 mm detector element
  - SO: 600 mm, DO: 600 mm

- **Reconstruction**
  - Penalized-likelihood objective
  - Shift-Invariant Quadratic Penalty
  - Separable paraboloidal surrogates
  - 200 iterations
  - Voxel basis: 480 x 480 0.8 mm voxels
FBP vs PL, Noise Properties

- Noise in FBP
  - Shift-variant variance
  - Shift-variant covariance

- Noise in Quadratic PL
  - Relatively shift-invariant variance (in object)
  - Shift-variant covariance

FBP vs PL, Resolution Properties

- Filtered-Backprojection
  - Largely shift-invariant spatial resolution
  - Shift-variant, object-dependent noise

- Uniform Quadratic Penalized Likelihood
  - Shift-variant, object-dependent spatial resolution
  - Shift-variant, object-dependent noise

- Edge-preserving Penalty Methods
  - Shift-variant, object-dependent spatial resolution
  - Shift-variant, object-dependent noise
  - Noise-resolution properties may not even be locally smooth

Image Properties
Learning Objectives I

- Fundamentals of iterative methods
  - Approach to forming an iterative algorithm
    - Forward Model
    - Objective Function
    - Optimization Algorithm
  - Identify particular classes of methods
    - Model-based vs. Image Denoising approaches
    - Statistical vs. Nonstatistical approaches
    - Kinds of regularization

Learning Objectives II

- Advantages of iterative approaches
  - Intuition behind what these algorithms do
    - Fitting reconstruction to observations
    - Data weighting by information content
    - Importance of regularization
  - Flexibility of these techniques
    - Arbitrary geometries
    - Sophisticated modeling of physics
    - General incorporation of desired image properties through regularization

Learning Objectives III

- Images produced by iterative methods
  - Differences from traditional reconstruction
    - Regularization is key
    - Image properties are tied to statistical weighting
    - Can depend on algorithm when iterative solution has not yet converged
  - Image properties associated with iterative reconstruction
    - Highly dependent on regularization
    - Can be more shift-variant (edge-preservation)
    - Different noise, artifacts
Thank You