Reconstruction Methods for Magnetic Resonance Imaging

Martin Uecker
Introduction

Topics:

- Image Reconstruction as Inverse Problem
- Parallel Imaging
- Non-Cartesian MRI
- Subspace Methods
- Model-based Reconstruction
- Compressed Sensing
Tentative Syllabus

- 01: Apr 12 Introduction
- 02: Apr 19 Parallel Imaging as Inverse Problem
- 03: Apr 26 Iterative Reconstruction Algorithms
- 04: May 03 Non-Cartesian MRI
- 05: May 10 Nonlinear Inverse Reconstruction
- 06: May 17 Reconstruction in k-space
- 07: May 24 Reconstruction in k-space
- 08: May 31 Subspace methods
- 09: Jun 07 Subspace methods - ESPIRiT
- 10: Jun 14 Model-based Reconstruction
- 11: Jun 21 Model-based Reconstruction
- 12: Jun 28 Compressed Sensing
- 13: Jul 05 Compressed Sensing
- 14: Jul 12 TBA
http://wwwuser.gwdg.de/~muecker1/class.html

Available:

▶ Slides
▶ References
▶ Optional homework project
Today

- Review of last lecture
- Statistical Interpretation
- Parallel Imaging as Inverse Problem
- SENSE
Direct Image Reconstruction

- **Assumption:** Signal is Fourier transform of the image:

\[
s(t) = \int d\vec{x} \rho(\vec{x}) e^{-i2\pi \vec{x} \cdot \vec{k}(t)}
\]

- Image reconstruction with inverse DFT

\[
\vec{k}(t) = \gamma \int_0^t d\tau \, \vec{G}(\tau)
\]

⇒ sampling

⇒ k-space

⇒ iDFT

⇒ image
Requirements

- Short readout (signal equation holds for short time spans only)
- Sampling on a Cartesian grid

⇒ Line-by-line scanning

\[
\text{measurement time:}
\begin{align*}
\text{TR} & \geq 2ms \\
\text{2D: } N & \approx 256 \Rightarrow \text{seconds} \\
\text{3D: } N & \approx 256 \times 256 \Rightarrow \text{minutes}
\end{align*}
\]
Sampling
Dirichlet Kernel

\[ D_n(x) = \sum_{k=-n}^{n} e^{2\pi ikx} = \frac{\sin(\pi(2n+1)x)}{\sin(\pi x)} \]

\[ (D_n * f)(x) = \int_{-\infty}^{\infty} dy \ f(y) D_n(x - y) = \sum_{k=-n}^{n} \hat{f}(k) e^{2\pi ikx} \]

FWHM \approx 1.2
Gibbs Phenomenon

- Truncation of Fourier series
  ⇒ **Ringing** at jump discontinuities

Rectangular wave and Fourier approximation
Direct Image Reconstruction

**Assumptions:**

- Signal is Fourier transform of the image
- Sampling on a Cartesian grid
- Signal from a limited (compact) field of view
- Missing high-frequency samples are small
- Noise is neglectable
Inverse Problem

Definition (Hadamard): A problem is called well-posed if

1. there exists a solution to the problem (existence),
2. there is at most one solution to the problem (uniqueness),
3. the solution depends continuously on the data (stability).

(we will later see that all three conditions can be violated)
Moore-Penrose Generalized Inverse

\((A\text{ linear and bounded})\)

\(x\) is least-squares solution, if

\[\|Ax - y\| = \inf \{ \|A\hat{x} - y\| : \hat{x} \in X \}\]

\(x\) is best approximate solution, if

\[\|x\| = \inf \{ \|\hat{x}\| : \hat{x} \text{ least-squares solution} \}\]

Generalized inverse \(A^\dagger : D(A^\dagger) := R(A) + R(A)^\perp \mapsto X\) maps data to the best approximate solution.
Tikhonov Regularization

(inverse not continuous: \( \|A^{-1}\| = \infty \))

Regularized optimization problem:

\[
x_{\alpha}^{\delta} = \arg\min_{x} \|Ax - y^{\delta}\|_2^2 + \alpha \|x\|_2^2
\]

Explicit solution:

\[
x_{\alpha}^{\delta} = \left( A^{H} A + \alpha I \right)^{-1} A^{H} y^{\delta}
\]

Generalized inverse:

\[
A^{\dagger} = \lim_{\alpha \to 0} A_{\alpha}^{\dagger}
\]
Tikhonov Regularization: Bias vs Noise

Noise contamination:

\[ y^\delta = y + \delta n \]
\[ = Ax + \delta n \]

Reconstruction error:

\[ x - x_\alpha^\delta = x - (A^H A + \alpha I)^{-1} A^H y^\delta \]
\[ = x - (A^H A + \alpha I)^{-1} A^H Ax + (A^H A + \alpha I)^{-1} A^H \delta n \]
\[ = \alpha (A^H A + \alpha I)^{-1} x + (A^H A + \alpha I)^{-1} A^H \delta n \]

approximation error \hspace{1cm} data noise error
Today

- Review of last lecture
- Statistical Interpretation
- Parallel Imaging as Inverse Problem
- SENSE
Today

- Review of last lecture
- **Statistical Interpretation**
- Parallel Imaging as Inverse Problem
- SENSE
Statistical Interpretation

**Probability** of $x$ given $y$: $p(x|y)$

**Likelihood:** Probability of a specific outcome given a parameter: $p(y|x)$

**Maximum-Likelihood-Estimator:** The estimate is the parameter $x$ which maximizes the likelihood.
**Statistical Model**

**Linear measurements** contaminated by **noise**:

\[ y = Ax + n \]

**Gaussian white noise**: \( n \sim \mathcal{N}(0, \sigma^2) \)

with \( T \sim \mathcal{N}(\mu, \sigma^2) \) means \( p(T = t) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \)

Probability distr. of an outcome (measurement) given the image \( x \):

\[ y \sim \mathcal{N}(Ax, \sigma^2) \]
Prior knowledge as a probability distribution for the image:

\[ p(x) = \cdots \]

A posterior probability distribution \( p(x|y) \) given the data \( y \) can be computed using Bayes’ theorem:

\[ p(x|y) = \frac{p(y|x)p(x)}{p(y)} \]

A point estimate for the image can then be obtained, for example by maximum a posteriori estimation (MAP), or by minimizing the posterior expected value of a loss function, e.g. a minimum mean squared error (MMSE) estimator.
Bayesian Prior

- **L₂-Regularization: Ridge Regression**

\[ p(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{Gaussian prior} \]

- **L₁-Regularization: LASSO**

\[ p(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}} \quad \text{Laplacian prior} \]

- **L₂ and L₁: Elastic net¹**

---

Today

- Review of last lecture
- Statistical Interpretation
- Parallel Imaging as Inverse Problem
- SENSE
Today

- Review of last lecture
- Statistical Interpretation
- Parallel Imaging as Inverse Problem
- SENSE
Today

- Review of last lecture
- Statistical Interpretation
- **Parallel Imaging as Inverse Problem**
- SENSE
Phased Array

Advantages:

- SNR of small surface coils
- FOV of large coils
- Parallel Imaging

Roemer, The NMR Phased Array, MRM 1990.
Phased Array

Signal is Fourier transform of magnetization image $m$ weighted by coil sensitivities $c_j$:

$$s_j(t) = \int d\vec{x} \rho(\vec{x}) c_j(\vec{x}) e^{-i2\pi\vec{k}(t)\cdot\vec{x}}$$

Images of a human brain from an eight channel array:
Image Combination

What to do with all those images?

Estimation of a single image:

- Minimum-variance unbiased estimator
- Root-of-sum-of-squares
- (Minimum mean-squared error)
Minimum-Variance Unbiased Estimator

- Noise-optimal unbiased estimation:

\[
\hat{m}(x) = \sum_i c_i^*(x) m_i(x) \frac{\sum_i |c_i(x)|^2}{\sum_i |c_i(x)|^2}
\]

- Proof: Gauss-Markov theorem.

- **Coil sensitivities needed!**

  (Assumption: uncorrelated noise)
Minimum-Variance Unbiased Estimator

- Noise-optimal unbiased estimation:

\[
\hat{m}(x) = \sum_i \frac{c_i^*(x)m_i(x)}{\sum_i |c_i(x)|^2}
\]

\[
= \left(C_x^H C_x\right)^{-1} C_x^H \begin{pmatrix} m_1(x) \\ \vdots \\ m_N(x) \end{pmatrix}
\]

\[
C_x = \begin{pmatrix} c_1(x) \\ \vdots \\ c_N(x) \end{pmatrix}
\]

- Proof: Gauss-Markov theorem.

- **Coil sensitivities needed!**

(Assumption: uncorrelated noise)
Minimum Mean Squared Error

- Biased estimation:

\[ \hat{m}_\alpha(x) = \sum_i \frac{c_i^*(x)m_i(x)}{\lambda + \sum_i |c_i(x)|^2} \]

- \( \lambda \) controls tradeoff between bias and SNR (can achieve optimal MSE)

- Differs from MVUE only by (spatially variant) scaling!

- **Coil sensitivities needed!**

- Tikhonov regularization: \((C^HC + \lambda I)^{-1}C^H\)
Nonlinear estimation of magnitude image:

\[ \hat{m}(x) = \sqrt{\sum_i |m_i(x)|^2} \]

Basic idea: Pixel values as estimates for sensitivities:

\[ c_i(x) \approx \frac{m_i(x)}{\sqrt{\sum_i |m_i(x)|^2}} \]

Final image is weighted by coil sensitivities (bias):

\[ \hat{m}(x) = \sqrt{\sum_i |m_i(x)|^2} = |m(x)| \sqrt{\sum_i |c_i(x)|^2} \]

No coil sensitivities needed!
Channel Combination

RSS

MVUE

MMSE
Parallel MRI

**Goal:** Reduction of measurement time
- Subsampling of k-space
- Simultaneous acquisition with multiple receive coils

- Coil sensitivities provide spatial information
- Compensation for missing k-space data

Parallel MRI: Undersampling

Undersampling

\[ k_{\text{phase}} \]

\[ k_{\text{read}} \]

Aliasing

\[ k_{\text{partition}} \]

\[ k_{\text{phase}} \]
Parallel Imaging

- SMASH\(^1\)
- SENSE, CG-SENSE\(^2\)
- GRAPPA\(^3\), ARC
- JSENSE, NLINV
- SPIRiT
- ESPIRiT
- SAKE, CLEAR
- ...

Image Domain and k-Space Domain

Convolution theorem: $\mathcal{F}\{m \cdot c\} = \mathcal{F}\{m\} \ast \mathcal{F}\{c\}$
Parallel Imaging as Inverse Problem

**Model:** Signal from multiple coils (image \( \rho \), sensitivities \( c_j \)):

\[
s_j(t) = \int_\Omega d\vec{x} \, \rho(\vec{x}) c_j(\vec{x}) e^{-i2\pi \vec{x} \cdot \vec{k}(t)} + n_j(t)
\]

**Assumptions:**
- Image is square-integrable function \( \rho \in L^2(\Omega, \mathbb{C}) \)
- Additive multi-variate Gaussian white noise \( n \)

**Problem:** Find best approximate/regularized solution in \( L^2(\Omega, \mathbb{C}) \).
Parallel Imaging as Inverse Problem

Signals from multiple receive coils (image $\rho$, sensitivities $c_j$):

$$s_j(t) = \int_\Omega d\vec{x} \rho(\vec{x}) c_j(\vec{x}) e^{-i2\pi\vec{x} \cdot \vec{k}(t)} + n_j(t)$$

**Assumption:** known sensitivities $c_j$

$\Rightarrow$ linear relationship between image and data

Image reconstruction is a **linear inverse problem**:

$$\arg \min_\rho \| P\mathcal{F}S\rho - y \|_2^2 + R(\rho)$$

$\rho \text{ image, } S \text{ multiplication with sensitivities, } \mathcal{F} \text{ Fourier transform, } P \text{ sampling operator, } y \text{ data}$

$\Rightarrow$ Generalized/iterative (minimum-norm) SENSE

SENSE: Basic Idea

System of **decoupled** equations:

\[
\begin{pmatrix}
m_1(x, y) \\
\vdots \\
m_n(x, y)
\end{pmatrix}
= 
\begin{pmatrix}
c_1(x, y_1) & c_1(x, y_2) & \cdots & c_n(x, y_1) & c_n(x, y_2)
\end{pmatrix}
\cdot
\begin{pmatrix}
\rho(x, y_1) \\
\vdots \\
\rho(x, y_2)
\end{pmatrix}
\]
Discretization of Linear Inverse Problems

Integral operator \( F : f \mapsto g \) on continuous domain with kernel \( K \):

\[
g(t) = \int_a^b ds \, K(t, s)f(s)
\]

Discrete system of linear equations:

\[
y = Ax
\]

Considerations:
- Discretization error
- Efficient computation
- Implicit regularization

<table>
<thead>
<tr>
<th>Consideration</th>
<th>Operator</th>
<th>Unknown</th>
<th>Continuous</th>
<th>Discrete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discretization error</td>
<td>( F )</td>
<td>( f )</td>
<td>( g )</td>
<td>( A )</td>
</tr>
<tr>
<td>Efficient computation</td>
<td></td>
<td></td>
<td></td>
<td>( x )</td>
</tr>
<tr>
<td>Implicit regularization</td>
<td></td>
<td></td>
<td></td>
<td>( y )</td>
</tr>
</tbody>
</table>
Discretization for Parallel Imaging

Forward operator $F$ with kernel:

$$K_j(\vec{k}, \vec{x}) = \underbrace{c_j(\vec{x}) e^{-i2\pi \vec{k} \vec{x}}}^{\text{enc}_{k,j}}$$

Finite number of discrete samples $k \in S$ and channels $j = 1, \cdots, N$ $\Rightarrow$ $F$ already has finite rank

Expansion of image $f$ using basis functions $b_l$:

$$f(\vec{x}) = \sum_l a_l b_l(\vec{x})$$
Discretization for Parallel Imaging

$F$ finite rank

Tikhonov regularization: $F_{\alpha}^\dagger y \in R(F^H)$

Finite dimensional space:

$$R(F_{\alpha}^\dagger) = R(F^H) = \text{span}_{k,j}\{enc_{k,j}\}$$

$\Rightarrow$ Exact solution can be expressed in a finite basis!

Problem: $enc_{k,j}$ are usually only known approximately
Discretization for Parallel Imaging

Discrete Fourier basis:

\[ f(x, y) \approx \sum_{l=-N}^{N} \sum_{k=-N}^{N} \hat{a}_{l,k} e^{i2\pi \left( \frac{lx}{\text{FOV}_x} + \frac{ky}{\text{FOV}_y} \right)} \]

- Efficient computation (FFT)
- Approximates \( R(F^H) \) extremely well
- Voxels: Dirichlet kernel \( D_N\left( \frac{x}{\text{FOV}_x} \right)D_N\left( \frac{y}{\text{FOV}_y} \right) \)

\[ \text{FOV}_x \]
\[ \Omega \]
\[ \text{FOV}_y \]
Discretization for Parallel Imaging

Slightly more elements than extend of k-space data:

\[ m \cdot c \]
\[ \mathcal{F}\{m\} \star \mathcal{F}\{c\} \]
SENSE: Classical Theory

**Encoding functions:**

\[ y_{jk} = \int d\vec{x} \ c_j(\vec{x}) e^{-i2\pi \vec{k} \cdot \vec{x}} \rho(\vec{x}) \]

\[ = \langle enc_{jk}, \rho \rangle \]

**Discretization:**

\[ \rho(\vec{x}) \approx \sum_l a_l b_l(\vec{x}) \]

(In SENSE, \( b_l \) are called ideal voxel functions)

**Encoding matrix:**

\[ E_{jk,l} = \langle enc_{jk}, b_l \rangle \]

maps coefficients to data:

\[ y_{jk} = \langle enc_{jk}, \rho \rangle \approx \sum_l \langle enc_{jk}, b_l \rangle a_l = \sum_l E_{jk,l} a_l \]
Linear reconstruction of coefficients from samples using a reconstruction matrix $R$:

$$a_l = \sum_{jk} R_{l,jk} y_{jk}$$

Encoding functions transformed to voxel functions:

$$f_l = \sum_{jk} R_{l,jk} enc_{jk}$$

Encoding functions: spatial sensitivity of samples $y_{jk} = \langle enc_{jk} , \rho \rangle$

Voxel functions: spatial sensitivity of coefficients $a_l = \langle f_l , \rho \rangle$
**SENSE: Classical Theory**

Still missing: basis $b_l$ and reconstruction $F$

Given: basis $b_l$ (ideal voxel functions)

- **Strong voxel condition**: Voxel functions should be close to basis elements $b_l$ in least-squares sense:
  $$R = \arg\min_R \sum_l \| b_l - f_l \|^2$$

- **Weak voxel condition**: Images from discretized subspace should be recovered exactly (from noiseless data):
  $$RE = I$$

$$a_l = \sum_{jk} R_{l,jk} y_{jk} = \sum_{jk} R_{l,jk} \sum_l E_{jk,l} a_l$$
**SENSE: Strong Voxel Condition**

**Strong voxel condition:** Voxel functions should be close to basis elements $b_l$ in least-squares sense:

$$R = \arg\min_R \sum_l \|b_l - f_l\|^2$$

With the correlation matrix $K_{il,jk} = \langle enc_{il}, enc_{jk} \rangle$, this leads to (see Pruessmann et al.):

$$R = E^H K^{-1}$$

With forward operator $F$ and basis $B$:

$$K = FF^H \quad E = FB$$

**Projection of best approximation to discretized space:**

$$B^H F^H \left( FF^H \right)^{-1} = B^H F^\dagger$$
Geometry

The diagram illustrates the span of vector expressions:

- \( \text{span } \text{enc}_{jk} \)
- \( F^\dagger y \)
- \( B^H F^\dagger y \)
- \( \text{span } b_l \)
Weak voxel condition: Images from discretized subspace should be recovered exactly (from noiseless data):

$$RE = I$$

Best approximate solution in discretized subspace:

$$R = E^\dagger = \left( E^H E \right)^{-1} E^H$$

$$= \left( B^H F^H FB \right)^{-1} B^H F^H$$
Continuous reconstruction is ill-posed!
Discretized problem might even be well-conditioned

Attention: Be careful when simulating data! Same discretization for simulation and reconstruction ⇒ misleading results (inverse crime)
SENSE: Discretization

Common choice:

\[ f(x, y) \approx \sum_{r, s} \delta(x - r \frac{\text{FOV}_x}{N_x})\delta(y - s \frac{\text{FOV}_y}{N_y}) \]

- Efficient computation using FFT algorithm
- Periodic sampling (⇒ decoupling)

Problem: Periodically extended k-space.
⇒ Error at the k-space boundary!
System of **decoupled** equations:

\[
\begin{pmatrix}
  m_1(x, y) \\
  \vdots \\
  m_n(x, y)
\end{pmatrix}
= \begin{pmatrix}
  c_1(x, y_1) & c_1(x, y_2) & \cdots \\
  \cdots & \cdots & \cdots \\
  c_n(x, y_1) & c_n(x, y_2)
\end{pmatrix}
\cdot
\begin{pmatrix}
  \rho(x, y_1) \\
  \rho(x, y_2)
\end{pmatrix}
\]
Project 1: Iterative SENSE

**Project:** Implement and study Cartesian iterative SENSE

- **Tools:** Matlab, reconstruction toolbox, python, ...
- **See website for data and instructions.**
Project 1: Iterative SENSE

Step 1: Implement Model

\[ A = P \mathcal{F} S \]
\[ A^H = S^H \mathcal{F}^{-1} P^H \]

Hints:
- Use unitary and centered (fftshift) FFT \[ \| \mathcal{F} x \|_2 = \| x \|_2 \]
- Implement undersampling as a mask, store data with zero
- Check \[ \langle x, Ay \rangle = \langle A^H x, y \rangle \] for random vectors \( x, y \)
Step 2: Implement Reconstruction

Landweber (gradient descent)$^1$

$$x_{n+1} = x_n + \alpha A^H(y - Ax_n)$$

Conjugate gradient algorithm$^2$

Next lesson: Iterative Algorithms

Step 3: Experiments

- Noise, errors, and convergence speed
- Different sampling
- Regularization

---

Tentative Syllabus

- **01:** Apr 12 Introduction
- **02:** Apr 19 Parallel Imaging as Inverse Problem
- **03:** Apr 26 Iterative Reconstruction Algorithms
- **04:** May 03 Non-Cartesian MRI
- **05:** May 10 Nonlinear Inverse Reconstruction
- **06:** May 17 Reconstruction in k-space
- **07:** May 24 Reconstruction in k-space
- **08:** May 31 Subspace methods
- **09:** Jun 07 Subspace methods - ESPIRiT
- **10:** Jun 14 Model-based Reconstruction
- **11:** Jun 21 Model-based Reconstruction
- **12:** Jun 28 Compressed Sensing
- **13:** Jul 05 Compressed Sensing
- **14:** Jul 12 TBA