Biomedical image reconstruction: From the foundations to deep neural nets

Michael Unser
Biomedical Imaging Group
EPFL, Lausanne, Switzerland

Pol del Aguila Pla
Center for Biomedical Imaging
Mathematical Imaging Section
EPFL, Lausanne, Switzerland

Tutorial 10, IEEE Int. Conf. Audio, Speech & Signal Processing (ICASSP 2020), 4-8 May 2020, Barcelona

OUTLINE

1. Imaging as an inverse problem
   - Basic imaging operators
   - Discretization of the inverse problem

2. Classical image reconstruction (1st gen.)
   - Backprojection
   - Tikhonov regularization; Wiener / LMSE solution

3. Sparsity-based image reconstruction (2nd gen.)

   Specific examples:
   - Magnetic resonance imaging
   - Computed tomography
   - Differential phase-contrast tomography

4. The learning (R)evolution (3rd gen.)
Inverse problems in bio-imaging

- Linear forward model
  \[ y = Hs + n \]

- The easy scenario
  Inverse problem is well posed
  \[ s \approx H^{-1}y \]

- Backprojection (poor man’s solution)

**Basic limitations**

1. Inherent noise amplification
2. Difficulty to invert H (too large or non-square)
3. All interesting inverse problems are ill-posed

Part 1:
Setting up the problem
Forward imaging model (noise-free)

Unknown molecular/anatomical map: $s(r), r = (x, y, z, t) \in \mathbb{R}^d$

\textit{defined over a continuum in space-time}

$$s \in L_2(\mathbb{R}^d) \quad \text{(space of finite-energy functions)}$$

Imaging operator $H: s \mapsto y = (y_1, \cdots, y_M) = H\{s\}$

\textit{from continuum to discrete (finite dimensional)}

$$H: L_2(\mathbb{R}^d) \to \mathbb{R}^M$$

Linearity assumption: for all $s_1, s_2 \in L_2(\mathbb{R}^d), \alpha_1, \alpha_2 \in \mathbb{R}$

$$H\{\alpha_1 s_1 + \alpha_2 s_2\} = \alpha_1 H\{s_1\} + \alpha_2 H\{s_2\}$$

$$\Rightarrow \quad [y]_m = y_m = \langle \eta_m, s \rangle = \int_{\mathbb{R}^d} \eta_m(r)s(r)\,dr$$

(by the Riesz representation theorem)

Images are obviously made of sine waves ...
Basic operator: Fourier transform

\[ \mathcal{F} : L_2(\mathbb{R}^d) \rightarrow L_2(\mathbb{R}^d) \]

\[ \hat{f}(\omega) = \mathcal{F}\{f\}(\omega) = \int_{\mathbb{R}^d} f(x)e^{-j\langle \omega, x \rangle} \, dx \]

Reconstruction formula (inverse Fourier transform)

\[ f(x) = \mathcal{F}^{-1}\{f\}(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\omega)e^{j\langle \omega, x \rangle} \, d\omega \quad \text{(a.e.)} \]

Equivalent analysis functions: \( \eta_m(x) = e^{j\langle \omega_m, x \rangle} \) (complex sinusoids)

2D Fourier reconstruction

Original image:

\( f(x) \)

Reconstruction using \( N \) largest coefficients:

\[ \tilde{f}(x) = \frac{1}{(2\pi)^2} \sum_{\text{subset}} \hat{f}(\omega) e^{j\langle x, \omega \rangle} \]
Magnetic resonance imaging

- Magnetic resonance: $\omega_0 = \gamma B_0$

Frequency encode:

- Linear forward model for MRI

$$\hat{s}(\omega_m) = \int_{\mathbb{R}^3} s(r)e^{-j(\omega_m, r)} dr$$

- Extended forward model with coil sensitivity

$$\hat{s}_w(\omega_m) = \int_{\mathbb{R}^3} w(r)s(r)e^{-j(\omega_m, r)} dr$$

Basic operator: Windowing

$$W : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$$

$$W\{f\}(x) = w(x)f(x)$$

Positive window function (continuous and bounded): $w \in C_b(\mathbb{R}^d), w(x) \geq 0$

- Special case: modulation

$$w(r) = e^{j(\omega_0, r)}$$

$$e^{j(\omega_0, r)}f(r) \iff \hat{f}(\omega - \omega_0)$$

Application: Structured illumination microscopy (SIM)
Basic operator: Convolution

$$H : L_2(\mathbb{R}^d) \rightarrow L_2(\mathbb{R}^d)$$

$$H\{f\}(x) = (h \ast f)(x) = \int_{\mathbb{R}^d} h(x-y)f(y)\,dy$$

Impulse response: $$h(x) = H\{\delta\}$$

Equivalent analysis functions: $$\eta_m(x) = h(x_m - \cdot)$$

Frequency response: $$\hat{h}(\omega) = \mathcal{F}\{h\}(\omega)$$

Convolution as a frequency-domain product

$$\mathcal{F}\{(h \ast f)(x)\} \leftrightarrow \hat{h}(\omega) \hat{f}(\omega)$$

Modeling of optical systems

Diffraction-limited optics = LSI system

$$f(x, y) \rightarrow g(x, y) = (h \ast f)(x, y)$$

$$h(x, y):$$ Point Spread Function (PSF)

Aberation-free point spread function (in focal plane)

$$h(x, y) = h(r) = C \cdot \left[\frac{2J_1(\pi r)}{\pi r}\right]^2$$

where $$r = \sqrt{x^2 + y^2}$$ (radial distance)

Effect of misfocus

Point source (in focus) output (defocus)
Basic operator: X-ray transform

Projection geometry: \( x = t\theta + r\theta^\perp \) with \( \theta = (\cos \theta, \sin \theta) \)

- **Radon transform (line integrals)**

\[
R_\theta \{s(x)\}(t) = \int_{\mathbb{R}} s(t\theta + r\theta^\perp) dr \\
= \int_{\mathbb{R}^2} s(x) \delta(t - \langle x, \theta \rangle) dx
\]

Equivalent analysis functions: \( \eta_m(x) = \delta(t_m - \langle x, \theta_m \rangle) \)

Central slice theorem

- Measurements of line integrals (Radon transform)

\( p_\theta(t) = R_\theta \{f\}(t, \theta) \)

- 1D and 2D Fourier transforms

\[
\hat{p}_\theta(\omega) = \mathcal{F}_{1D}\{p_\theta\}(\omega) \\
\hat{f}(\omega) = \mathcal{F}_{2D}\{f\}(\omega) = \hat{f}_{\text{pol}}(\omega, \theta)
\]

- Central-slice theorem

\[
\hat{p}_\theta(\omega) = \hat{f}(\omega \cos \theta, \omega \sin \theta) = \hat{f}_{\text{pol}}(\omega, \theta)
\]

Proof: for \( \theta = 0 \)

\[
\hat{f}(\omega, 0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-j\omega x} dx dy = \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{+\infty} f(x, y) dy \right) e^{-j\omega x} dx = \hat{p}_0(\omega)
\]

then use rotation property of Fourier transform...
<table>
<thead>
<tr>
<th>Modality</th>
<th>Radiation</th>
<th>Forward model</th>
<th>Variations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D or 3D tomography</td>
<td>coherent x-ray</td>
<td>$y_i = R_{\theta_i}x$</td>
<td>parallel, cone beam, spiral sampling</td>
</tr>
<tr>
<td>3D deconvolution microscopy</td>
<td>fluorescence</td>
<td>$y = Hx$</td>
<td>brightfield, confocal, light sheet</td>
</tr>
<tr>
<td>structured illumination microscopy (SIM)</td>
<td>fluorescence</td>
<td>$y_i = HW_i x$</td>
<td>full 3D reconstruction, non-sinusoidal patterns</td>
</tr>
<tr>
<td>Positron Emission Tomography (PET)</td>
<td>gamma rays</td>
<td>$y_i = H_{\theta_i} x$</td>
<td>list mode with time-of-flight</td>
</tr>
<tr>
<td>Magnetic resonance imaging (MRI)</td>
<td>radio frequency</td>
<td>$y = Fx$</td>
<td>uniform or non-uniform sampling in k space</td>
</tr>
<tr>
<td>Cardiac MRI (parallel, non-uniform)</td>
<td>radio frequency</td>
<td>$y_{t,i} = F_t W_i x$</td>
<td>gated or not, retrospective registration</td>
</tr>
<tr>
<td>Optical diffraction tomography</td>
<td>coherent light</td>
<td>$y_i = W_i F_i x$</td>
<td>with holography or grating interferometry</td>
</tr>
</tbody>
</table>

**Discretization: Finite dimensional formalism**

$$s(r) = \sum_{k \in \Omega} s[k] \beta_k(r)$$

Signal vector: $s = (s[k])_{k \in \Omega}$ of dimension $K$

- Measurement model (image formation)

  $$y_m = \int_{\mathbb{R}^d} s(r) \eta_m(r) dr + n[m] = \langle s, \eta_m \rangle + n[m], \quad (m = 1, \ldots, M)$$

  $\eta_m$: sampling/imaging function ($m$th detector)

  $n[\cdot]$: additive noise

  $$y = y_0 + n = Hs + n$$

  $$(M \times K)$$ system matrix: $[H]_{m,k} = \langle \eta_m, \beta_k \rangle = \int_{\mathbb{R}^d} \eta_m(r) \beta_k(r) dr$$
Example of basis functions

Shift-invariant representation: \( \beta_k(x) = \beta(x - k) \)

Separable generator: \( \beta(x) = \prod_{n=1}^{d} \beta(x_n) \)

- Pixelated model
  \( \beta(x) = \text{rect}(x) \)

- Bilinear model
  \( \beta(x) = (\text{rect} \ast \text{rect})(x) = \text{tri}(x) \)

- Bandlimited representation
  \( \beta(x) = \text{sinc}(x) \)

Part 2:

Classical image reconstruction

Discretized forward model: \( y = Hs + n \)

Inverse problem: How to efficiently recover \( s \) from \( y \)?
Vector calculus

- Scalar cost function $J(v) : \mathbb{R}^N \rightarrow \mathbb{R}$
- Vector differentiation: $
\begin{bmatrix}
\frac{\partial J(v)}{\partial v_1} \\
\vdots \\
\frac{\partial J(v)}{\partial v_N}
\end{bmatrix} = \nabla J(v)$ (gradient)

- Necessary condition for an unconstrained optimum (minimum or maximum)
  $$\frac{\partial J(v)}{\partial v} = 0$$ (also sufficient if $J(v)$ is convex in $v$)

- Useful identities
  $$\begin{align*}
  \frac{\partial}{\partial v} (a^T v) &= \frac{\partial}{\partial v} (v^T a) = a \\
  \frac{\partial}{\partial v} (v^T A v) &= (A + A^T) \cdot v \\
  \frac{\partial}{\partial v} (v^T A v) &= 2A \cdot v \quad \text{if } A \text{ is symmetric}
  \end{align*}$$

Basic reconstruction: least-squares solution

- Least-squares fitting criterion:
  $$J_{LS}(\hat{s}, y) = \|y - H\hat{s}\|^2$$
  $$\min_{\hat{s}} \|y - \hat{y}\|^2 = \min_{s} J_{LS}(s, y) \quad \text{(maximum consistency with the data)}$$

- Formal least-squares solution
  $$J_{LS}(s, y) = \|y - Hs\|^2 = \|s\|^2 + s^T H^T H s - 2y^T H s$$
  $$\frac{\partial J_{LS}(s, y)}{\partial s} = 2H^T H s - 2H^T y$$

- Backprojection (poor man’s solution):
  $$s \approx H^T y$$

  OK if $H$ is unitary $\iff$

Basic limitations

1) Inherent noise amplification
2) Difficulty to invert $H$ (too large or non-square)
3) All interesting inverse problems are ill-posed
Linear inverse problems (20th century theory)

- Dealing with **ill-posed problems**: Tikhonov regularization
  \[ \mathcal{R}(s) = \|Ls\|_2^2: \text{regularization (or smoothness) functional} \]
  \[ L: \text{regularization operator (i.e., Gradient)} \]
  \[
  \min_s \mathcal{R}(s) \text{ subject to } \|y - Hs\|_2^2 \leq \sigma^2
  \]
- Equivalent variational problem
  \[
  s^* = \arg \min_s \underbrace{\|y - Hs\|_2^2}_{\text{data consistency}} + \underbrace{\lambda\|Ls\|_2^2}_{\text{regularization}}
  \]
  \[
  \text{Formal linear solution: } s = (H^TH + \lambda L^TL)^{-1}H^Ty = R_\lambda \cdot y
  \]
  \[
  \text{Interpretation: “filtered” backprojection}
  \]

Statistical formulation (20th century)

- Linear measurement model: \( y = Hs + n \)
  \[ n: \text{additive white Gaussian noise (i. i. d.)} \]
  \[ s: \text{realization of Gaussian process with zero-mean and covariance matrix } \mathbb{E}\{s \cdot s^T\} = C_s \]

  \[\begin{aligned}
  \text{Wiener (LMMSE) solution} &= \text{Gauss MMSE} = \text{Gauss MAP} \\
  s_{\text{MAP}} &= \arg \min_s \underbrace{\frac{1}{\sigma^2}\|y - Hs\|_2^2}_{\text{Data Log likelihood}} + \underbrace{\|C_s^{-1/2}s\|_2^2}_{\text{Gaussian prior likelihood}}
  \end{aligned}\]

  \[\uparrow \quad L = C_s^{-1/2}: \text{Whitening filter} \]

- Quadratic regularization (Tikhonov)
  \[
  s_{\text{Tik}} = \arg \min_s \left(\|y - Hs\|_2^2 + \lambda\mathcal{R}(s)\right) \quad \text{with} \quad \mathcal{R}(s) = \|Ls\|_2^2
  \]
  \[
  \text{Linear solution: } s = (H^TH + \lambda L^TL)^{-1}H^Ty = R_\lambda \cdot y
  \]

Andrey N. Tikhonov (1906-1993)

Norbert Wiener (1894-1964)
Iterative reconstruction algorithm

- Generic minimization problem: \( s_{\text{opt}} = \arg \min_s J(s, y) \)

- Steepest-descent solution
  \[
  s^{(k+1)} = s^{(k)} - \gamma \nabla J(s^{(k)}, y)
  \]

- Iterative constrained least-squares reconstruction
  \[
  J_{\text{TK}}(s, y) = \frac{1}{2} \| y - Hs \|^2 + \frac{\lambda}{2} \| Ls \|^2
  \]
  Gradient: \( \frac{\partial J_{\text{TK}}(s, y)}{\partial s} = -s_0 + (H^T H + \lambda L^T L)s \) with \( s_0 = H^T y \)
  Steepest-descent algorithm
  \[
  s^{(k+1)} = s^{(k)} + \gamma \left( s_0 - (H^T H + \lambda L^T L)\tilde{s}^{(k)} \right)
  \]
  Positivity constraint (IC): \( [\tilde{s}^{(k+1)}]_i = \begin{cases} 0, & [s^{(k+1)}]_i < 0 \\ [s^{(k+1)}]_i, & \text{otherwise.} \end{cases} \) (projection on convex set)

Iterative deconvolution: unregularized case

Degraded image: Gaussian blur + additive noise

van Cittert animation

Ground truth
Effect of regularization parameter

Degraded image: Gaussian blur + additive noise

Optimal regularization: $\lambda = 2$

Selecting the regularization operator

- Translation, rotation and scale-invariant operators
  - Laplacian: $\Delta s = (\nabla^T \nabla)s \iff -\|\omega\|^2 \hat{s}(\omega)$
  - Modulus of gradient: $|\nabla s|
  - Fractional Laplacian: $(-\Delta)^{\gamma/2} \iff \|\omega\|^\gamma \hat{s}(\omega)$

- TRS-invariant regularization functional
  $$\|\nabla s\|_{L_2(\mathbb{R}^d)}^2 = \|(-\Delta)^{1/2} s\|_{L_2(\mathbb{R}^d)}^2 \implies L: \text{discrete version of gradient}$$

- Fractional Brownian motion field
  - Statistical decoupling/whitening: $(-\Delta)^{\gamma/2} s = w \iff \frac{1}{|\omega|^\gamma}$ spectral decay
Relevance of self-similarity for bio-imaging

- Fractals and physiology

Designing fast reconstruction algorithms

Normal matrix: \( \mathbf{A} = \mathbf{H}^T \mathbf{H} \) (symmetric)

Formal linear solution: \( \mathbf{s} = (\mathbf{A} + \lambda \mathbf{L}^T \mathbf{L})^{-1} \mathbf{H}^T \mathbf{y} = \mathbf{R}_\lambda \cdot \mathbf{y} \)

Generic form of the iterator: \( \mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \gamma (\mathbf{s}_0 - (\mathbf{A} + \lambda \mathbf{L}^T \mathbf{L}) \mathbf{s}^{(k)}) \)

- Recognizing structured matrices
  - \( \mathbf{L} \): convolution matrix \( \Rightarrow \mathbf{L}^T \mathbf{L} \): symmetric convolution matrix
  - \( \mathbf{L}, \mathbf{A} \): convolution matrices \( \Rightarrow (\mathbf{A} + \lambda \mathbf{L}^T \mathbf{L}) \): symmetric convolution matrix

- Fast implementation
  - Diagonalization of convolution matrices \( \Rightarrow \) FFT-based implementation
  - Applicable to: - deconvolution microscopy (Wiener filter)
  - parallel rays computer tomography (FBP)
  - MRI, including non-uniform sampling of \( k \)-space
Part 3:

Sparsity-based image reconstruction (2nd generation)

**Linear inverse problems: Sparsity**

(20th Century) \( p = 2 \rightarrow 1 \) (21st Century)

\[
\hat{s}_{\text{rec}} = \arg \min_s (\|y - Hs\|_2^2 + \lambda \mathcal{R}(s))
\]

- **Non-quadratic regularization regularization**

\[
\mathcal{R}(s) = \|Ls\|_2^2 \rightarrow \|Ls\|_p^p \rightarrow \|Ls\|_1
\]

- **Total variation** (Rudin-Osher, 1992)

\[
\mathcal{R}(s) = \|Ls\|_1 \text{ with } L: \text{gradient}
\]

- **Wavelet-domain regularization** (Figuereido et al., Daubechies et al. 2004)

\[
v = W^{-1}s: \text{wavelet expansion of } s \text{ (typically, sparse)}
\]

\[
\mathcal{R}(s) = \|v\|_1
\]

- **Compressed sensing/sampling** (Candes-Romberg-Tao; Donoho, 2006)
Sparsifying transforms

Biomedical images are well described by few basis coefficients

Prior = sparse representation

\[ R(s) = \lambda \| W^T s \|_1 \]

Advantages:
- convex
- favors sparse solutions
- Fast: WFISTA

(Guerquin-Kern *IEEE TMI* 2011)

Theory of compressive sensing

- Generalized sampling setting (after discretization)
  - Linear inverse problem: \( y = Hs + n \)
  - Sparse representation of signal: \( s = Wx \) with \( \| x \|_0 = K \ll N_x \)
  - \( N_y \times N_x \) system matrix: \( A = HW \)

- Formulation of ill-posed recovery problem when \( 2K < N_y \ll N_x \)
  \[
  \begin{align*}
  \text{(P0)} & \quad \min_x \| y - Ax \|_2^2 \quad \text{subject to} \quad \| x \|_0 \leq K \\
  \text{(P1)} & \quad \min_x \| y - Ax \|_2^2 \quad \text{subject to} \quad \| x \|_1 \leq C_1
  \end{align*}
  \]

[Donoho et al., 2005
Candès-Tao, 2006, ...]
Compressive sensing (CS) and $l_1$ minimization

\[
\begin{align*}
\begin{array}{ccc}
\text{y} & \overset{\text{A}}{=} & \text{x} \\
\end{array}
\end{align*}
\]

+ “noise”

Sparse representation of signal: $s = Wx$ with $\|x\|_0 = K \ll N_x$

Equivalent $N_y \times N_x$ **sensing matrix**: $A = HW$

- Constrained (synthesis) formulation of recovery problem

\[
\begin{align*}
\min_x & \|x\|_1 \quad \text{subject to} \quad \|y - Ax\|_2^2 \leq \sigma^2 \\
\end{align*}
\]

**Classical regularized least-squares estimator**

- Linear measurement model:

\[
y_m = \langle h_m, x \rangle + n[m], \quad m = 1, \ldots, M
\]

- System matrix: $H = [h_1 \cdots h_M]^T \in \mathbb{R}^{N \times N}$

\[
x_{LS} = \arg \min_{x \in \mathbb{R}^N} \|y - Hx\|_2^2 + \lambda \|x\|_2^2
\]

\[
\Rightarrow \quad x_{LS} = (H^T H + \lambda I_N)^{-1} H^T y
\]

\[
= H^T a = \sum_{m=1}^M a_m h_m \quad \text{where} \quad a = (HH^T + \lambda I_M)^{-1} y
\]

Interpretation: $x_{LS} \in \text{span}\{h_m\}_{m=1}^M$

**Lemma**

\[
(H^T H + \lambda I_N)^{-1} H^T = H^T (HH^T + \lambda I_M)^{-1}
\]
Generalization: constrained $l_2$ minimization

- Discrete signal to reconstruct: $x = (x[n])_{n \in \mathbb{Z}}$
- Sensing operator $H : \ell_2(\mathbb{Z}) \rightarrow \mathbb{R}^M$
  \[ x \mapsto z = H\{x\} = (\langle x, h_1 \rangle, \ldots, \langle x, h_M \rangle) \text{ with } h_m \in \ell_2(\mathbb{Z}) \]
- Closed convex set in measurement space: $C \subset \mathbb{R}^M$

Example: \[ C_y = \{ z \in \mathbb{R}^M : \|y - z\|_2^2 \leq \sigma^2 \} \]

Representer theorem for constrained $\ell_2$ minimization

\[ (P2) \quad \min_{x \in \ell_2(\mathbb{Z})} \|x\|_{\ell_2}^2 \quad \text{s.t.} \quad H\{x\} \in C \]

The problem (P2) has a unique solution of the form
\[ x_{LS} = \sum_{m=1}^{M} a_m h_m = H^*\{a\} \]
with expansion coefficients \( a = (a_1, \ldots, a_M) \in \mathbb{R}^M \).


Constrained $l_1$ minimization $\Rightarrow$ sparsifying effect

- Discrete signal to reconstruct: $x = (x[n])_{n \in \mathbb{Z}}$
- Sensing operator $H : \ell_1(\mathbb{Z}) \rightarrow \mathbb{R}^M$
  \[ x \mapsto z = H\{x\} = (\langle x, h_1 \rangle, \ldots, \langle x, h_M \rangle) \text{ with } h_m \in \ell_\infty(\mathbb{Z}) \]
- Closed convex set in measurement space: $C \subset \mathbb{R}^M$

Representer theorem for constrained $\ell_1$ minimization

\[ (P1) \quad V = \arg \min_{x \in \ell_1(\mathbb{Z})} \|x\|_{\ell_1} \quad \text{s.t.} \quad H\{x\} \in C \]

is convex, weak*-compact with extreme points of the form
\[ x_{\text{sparse}}[\cdot] = \sum_{k=1}^{K} a_k \delta[\cdot - n_k] \quad \text{with} \quad K = \|x_{\text{sparse}}\|_0 \leq M. \]

If CS condition is satisfied, then solution is unique

Controlling sparsity

Measurement model: \( y_m = \langle h_m, x \rangle + n[m], \quad m = 1, \ldots, M \)

\[
x_{\text{sparse}} = \arg \min_{x \in \ell_1(\mathbb{Z})} \left( \sum_{m=1}^{M} \left| y_m - \langle h_m, x \rangle \right|^2 + \lambda \| x \|_{\ell_1} \right)
\]

Geometry of \( l_2 \) vs. \( l_1 \) minimization

- Prototypical inverse problem

\[
\min_x \{ \| y - Hx \|_2^2 + \lambda \| x \|_2^2 \} \Leftrightarrow \min_x \| x \|_2 \text{ subject to } \| y - Hx \|_2^2 \leq \sigma^2
\]

\[
\min_x \{ \| y - Hx \|_2^2 + \lambda \| x \|_1 \} \Leftrightarrow \min_x \| x \|_1 \text{ subject to } \| y - Hx \|_2^2 \leq \sigma^2
\]

\( C \)

\[ y_1 = h_1^T x \]

\( \ell_2\)-ball: \( |x_1|^2 + |x_2|^2 = C_2 \)

\( \ell_1\)-ball: \( |x_1| + |x_2| = C_1 \)
Geometry of $l_2$ vs. $l_1$ minimization

- Prototypical inverse problem

\[
\min_x \{ \|y - Hx\|_2^2 + \lambda \|x\|_2^2 \} \ \Leftrightarrow \ \min_x \|x\|_2 \ \text{subject to} \ \|y - Hx\|_2^2 \leq \sigma^2
\]

\[
\min_x \{ \|y - Hx\|_2^2 + \lambda \|x\|_1 \} \ \Leftrightarrow \ \min_x \|x\|_1 \ \text{subject to} \ \|y - Hx\|_2^2 \leq \sigma^2
\]

Configuration for non-unique $l_1$ solution

Variational-MAP formulation of inverse problem

- Linear forward model

\[ y = Hs + n \]

- Reconstruction as an optimization problem

\[
s_{\text{rec}} = \arg \min \left\{ \|y - Hs\|_2^2 + \lambda \|Ls\|_p^p \right\}, \quad p = 1, 2
\]

\[- \log \text{Prob}(s) : \text{prior likelihood} \]
Discretization of reconstruction problem

Spline-like reconstruction model: \( s(r) = \sum_{k \in \Omega} s[k] \beta_k(r) \) \( \leftrightarrow \) \( s = (s[k])_{k \in \Omega} \)

Statistical innovation model

\[
\begin{align*}
\text{Discretization} & \quad u = Ls \quad \text{(matrix notation)} \\
Ls &= w \\
s &= L^{-1}w
\end{align*}
\]

\( p_U \) is part of \textit{infinitely divisible} family

Physical model: image formation and acquisition

\( y_m = \int_{\mathbb{R}^d} s(x) \eta_m(x) \text{d}x + n[m] = \langle s, \eta_m \rangle + n[m], \quad (m = 1, \ldots, M) \)

\[
y = y_0 + n = Hs + n
\]

\( n: \) i.i.d. noise with pdf \( p_N \)

Posterior probability distribution

\[
p_{S|Y}(s|y) = \frac{p_{Y|S}(y|s)p_{S}(s)}{p_{Y}(y)} = \frac{p_N(y - Hs)p_S(s)}{p_Y(y)}
\]

\[
= \frac{1}{Z}p_N(y - Hs)p_S(s)
\]

\( \text{Bayes’ rule} \)

Statistical decoupling

\[
u = Ls \quad \Rightarrow \quad p_S(s) \propto p_U(Ls) \approx \prod_{k \in \Omega} p_U([Ls]_k)
\]

Additive white Gaussian noise scenario (AWGN)

\[
p_{S|Y}(s|y) \propto \exp \left( -\frac{\|y - Hs\|^2}{2\sigma^2} \right) \prod_{k \in \Omega} p_U([Ls]_k)
\]

... and then take the log and maximize ...

... and then take the log and maximize ...

Unser and Tafti An Introduction to Sparse Stochastic Processes
General form of MAP estimator

\[ s_{\text{MAP}} = \arg\min \left( \frac{1}{2} \| y - Hs \|^2_2 + \sigma^2 \sum_n \Phi_U ([Ls]_n) \right) \]

- Gaussian: \( p_U(x) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-x^2/(2\sigma_0^2)} \) \( \Rightarrow \Phi_U(x) = \frac{1}{2\sigma_0^2} x^2 + C_1 \)
- Laplace: \( p_U(x) = \frac{1}{2} e^{-\lambda|x|} \) \( \Rightarrow \Phi_U(x) = \lambda|x| + C_2 \)
- Student: \( p_U(x) = \frac{1}{B(r, \frac{1}{2})} \left( \frac{1}{x^2 + 1} \right)^{r+\frac{1}{2}} \) \( \Rightarrow \Phi_U(x) = (r + \frac{1}{2}) \log(1 + x^2) + C_3 \)

Potential: \( \Phi_U(x) = -\log p_U(x) \)

Proximal operator: pointwise denoiser

\[ \text{prox}_{\Phi_U} (y; \sigma^2) = \arg\min_{u \in \mathbb{R}} \frac{1}{2} \| y - u \|^2 + \sigma^2 \Phi_U(u) \]

\( \tilde{u} = \text{prox}_{\Phi_U} (y; 1) \)

- linear attenuation
- soft-threshold
- shrinkage function

\( \ell_2 \) minimization
\( \ell_1 \) minimization
\( \approx \ell_p \) relaxation for \( p \rightarrow 0 \)
Maximum a posteriori (MAP) estimation

- Constrained optimization formulation

Auxiliary innovation variable: $u = Ls$

$$s_{\text{MAP}} = \arg \min_{s \in \mathbb{R}^K} \left( \frac{1}{2} \| y - Hs \|_2^2 + \sigma^2 \sum_n \Phi_U([u]_n) \right) \quad \text{subject to} \quad u = Ls$$

- Augmented Lagrangian method

Quadratic penalty term: $\frac{\mu}{2} \| Ls - u \|_2^2$

Lagrange multiplier vector: $\alpha$

$$\mathcal{L}_A(s, u, \alpha) = \frac{1}{2} \| y - Hs \|_2^2 + \sigma^2 \sum_n \Phi_U([u]_n) + \alpha^T(Ls - u) + \frac{\mu}{2} \| Ls - u \|_2^2$$

(Bostan et al. IEEE TIP 2013)

Alternating direction method of multipliers (ADMM)

$$\mathcal{L}_A(s, u, \alpha) = \frac{1}{2} \| y - Hs \|_2^2 + \sigma^2 \sum_n \Phi_U([u]_n) + \alpha^T(Ls - u) + \frac{\mu}{2} \| Ls - u \|_2^2$$

Sequential minimization

$$s^{k+1} \leftarrow \arg \min_{s \in \mathbb{R}^N} \mathcal{L}_A(s, u^k, \alpha^k)$$

$$\alpha^{k+1} = \alpha^k + \mu(Ls^{k+1} - u^k)$$

$$u^{k+1} \leftarrow \arg \min_{u \in \mathbb{R}^N} \mathcal{L}_A(s^{k+1}, u, \alpha^{k+1})$$

Linear inverse problem: $s^{k+1} = (H^T H + \mu L^T L)^{-1} (H^T y + z^{k+1})$

with $z^{k+1} = L^T (\mu u^k - \alpha^k)$

Nonlinear denoising: $u^{k+1} = \text{prox}_{\Phi_U}(Ls^{k+1} + \frac{1}{\mu} \alpha^{k+1}, \frac{\sigma^2}{\mu})$

- Proximal operator tailored to stochastic model

$$\text{prox}_{\Phi_U}(y; \lambda) = \arg \min_u \frac{1}{2} \| y - u \|^2 + \lambda \Phi_U(u)$$

Cauchy prior with increasing $\alpha$
Deconvolution in widefield microscopy

Physical model of a diffraction-limited microscope

\[ g(x, y, z) = (h_{3D} * s)(x, y, z) \]

3-D point spread function (PSF)

\[ h_{3D}(x, y, z) = I_0 \left| p_\lambda \left( \frac{x}{M}, \frac{y}{M}, \frac{z}{M^2} \right) \right|^2 \]

\[ p_\lambda(x, y, z) = \int_{\mathbb{R}^2} P(\omega_1, \omega_2) \exp \left( j2\pi z \frac{\omega_1^2 + \omega_2^2}{2\lambda f_0^2} \right) \exp \left( -j2\pi \frac{x\omega_1 + y\omega_2}{\lambda f_0} \right) d\omega_1 d\omega_2 \]

Optical parameters

- \( \lambda \): wavelength (emission)
- \( M \): magnification factor
- \( f_0 \): focal length
- \( P(\omega_1, \omega_2) = 1_{||\omega|| < R_0} \): pupil function
- \( NA = n \sin \theta = R_0 / f_0 \): numerical aperture

2-D (in focus) convolution model

Thin specimen

\[ h_{2D}(x, y) \]

- Airy disk: \( h_{2D}(x, y) = I_0 \left| 2 \frac{J_1(r/r_0)}{r/r_0} \right|^2 \), with \( r = \sqrt{x^2 + y^2} \)

\( J_1(r) \): first-order Bessel function, and \( r_0 = \frac{\lambda f_0}{2\pi R_0} \)

Optical parameters

- \( \lambda \): wavelength (emission)
- \( f_0 \): focal length
- \( R_0 \): radius of aperture

Modulation transfer function

\[ |\hat{h}_{2D}(\omega)| = \begin{cases} \frac{2}{\pi} \left( \arccos \left( \frac{||\omega||}{\omega_0} \right) - \frac{||\omega||}{\omega_0} \sqrt{1 - \left( \frac{||\omega||}{\omega_0} \right)^2} \right), & \text{for } 0 \leq ||\omega|| < \omega_0 \\ 0, & \text{otherwise} \end{cases} \]

Cut-off frequency (Rayleigh): \( \omega_0 = \frac{2R_0}{\lambda f_0} = \frac{\pi}{r_0} \approx \frac{2NA}{\lambda} \)
2-D deconvolution: numerical set-up

- Discretization

  \[ \omega_0 \leq \pi \] and representation in (separable) sinc basis

  \[ \beta_k(x) = \text{sinc}(x - k) \text{ with } k \in \mathbb{Z}^2 \]

  Analysis functions (impulse response): \( \eta_m(x, y) = h_{2D}(x - m_1, y - m_2) \)

  \[
  [H]_{m,k} = \langle \eta_m, \beta_k \rangle = \langle \eta_m, \text{sinc}(\cdot - k) \rangle
  = \langle h_{2D}(\cdot - m), \text{sinc}(\cdot - k) \rangle
  = (\text{sinc} * h_{2D})(m - k) = h_{2D}(m - k).
  \]

- Linear step of ADMM algorithm implemented using the FFT

  \[
  s^{k+1} = (H^T H + \mu L^T L)^{-1} (H^T y + z^{k+1})
  \]

  with \( z^{k+1} = L^T (\mu u^k - \lambda^k) \)

2D deconvolution experiment

- Disk-shaped PSF \((7 \times 7)\), \(L\): gradient (TV-like), optimized parameters

<table>
<thead>
<tr>
<th>Deconvolution results (SNR in dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astrocytes cells</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>12.18</td>
</tr>
<tr>
<td>Pulmonary cells</td>
</tr>
<tr>
<td>Stem cells</td>
</tr>
</tbody>
</table>
3D deconvolution of a widefield stack

C. Elegans embryo. 3 stacks obtained by a Olympus CellR. Pixel size: 64.5 nm, Z-step: 200 nm (3.1 ratio)

PSF from an analytical model (see PSF Generator). Deconvolution with GlobalBioIm.

\[
\mathbf{s} = \arg\min_{\mathbf{s} \in \mathbb{R}^K} \left( \frac{1}{2} \| \mathbf{y} - \mathbf{SHs} \|_2^2 + \lambda \| \mathbf{Ls} \|_{2,1} + \delta_{\mathbb{R}_+^K}(\mathbf{s}) \right)
\]

- Practical considerations
  - \( \mathbf{H} \) (convolution) and \( \mathbf{L} \) (gradient) as explained
  - \( \mathbf{S} \): patch extraction / masking (remove padding of the FFT implementation)
  - \( \| \cdot \|_{2,1} \): group-sparse norm for isotropic TV
  - \( \delta_{\mathbb{R}_+^K} : \mathbb{R} \to \{0, \infty\} \): flurophore concentrations are not negative

- and more...
  - implementing proximal optimization is hard
  - memory management, convergence criteria, GPU?
  - efficient implementations of linear operators
  - beyond ADMM...? Trying different splittings?

GlobalBioIm
A unifying Matlab library for imaging inverse problems

Download/Clone the latest version
GlobalBioIm

- Three main abstract classes:
  - Linear operators (LinOp)
  - Cost functions (Cost)
  - Optimization algorithms (Opti)

- Packaged with everything needed
  - Operators: efficient implementations of $Hx$, $H^*y$, $H^*Hx$, norm, ...
  - Cost functions: gradient, prox, Lipschitz constant, ...
  - Optimization algorithms: automagically use all of the above for pain-free prototyping.

3D deconvolution of a widefield stack

$$s = \arg \min_{s \in \mathbb{R}^K} \left( \frac{1}{2} \| y - SHs \|_2^2 + \lambda \| Ls \|_{2,1} + \delta_{\mathbb{R}_+^N} (s) \right)$$

- ADMM with 3-way splitting
  - $u_1 = Hs$, $u_2 = Ls$ and $u_3 = s$
  - $\min_{s \in \mathbb{R}^N} \mathcal{L}_A \left( s, \left\{ u_n^k \right\}_{n=1}^3, \left\{ \alpha_n^k \right\}_{n=1}^3 \right)$ in Fourier.
  - $\mathcal{L}_A \left( s, \left\{ u_n^k \right\}_{n=1}^3, \left\{ \alpha_n^k \right\}_{n=1}^3 \right) = \frac{1}{2} \| y - S u_1 \|_2^2 + \lambda \| u_2 \|_{2,1} + \delta_{\mathbb{R}_+^N} (u_3)$

- Configure convergence criteria
- 300 iterations or relative cost under 1e-4 or relative step under 1e-4
- Run ADMM
- With initialization at zero
- ADMM.run( zeros_( var_size ) ); step_decay( 1e-4 / 30 );
- Configure algorithm output while running
- Report costs (1 and 2 in cost_functions, corresponding to least squares and TV regularizer), but don't store them, 30 times in the number of maximum iterations.
- ADMM.OutputOpt = OutputOpt( true, [], round( ADMM.maxiter / 30 ), [1, 2] );
- ADMM.InitOutPut = ADMM.maxiter / 30;
- least_squares_cost = l2_cost * S;
Welcome to the GlobalBioIm Library Webpage

This is a free Matlab library. It contains generic modules that facilitate the implementation of forward models and optimization algorithms. It also capitalizes on the strong commonalities between the various image-formation models that can be exploited to build a fast, streamlined code.

This page contains the detailed documentation of each function/class of the library. The documentation is generated automatically from comments within M-files.

Releases

- v1.1.2 (April 2019).
- v1.1.1 (September 2018).
- v1.1 (July 2018). Speed up your codes using the library with GPU (read more).
- v1.0.1 (May 2018).
- v1.0 milestone (March 2018).
- v0.2 (November 2017). New tools, more flexibility, improved composition.
- v0.1 (June 2017), First public release of the library.

Reference

Pocket Guide to Solve Inverse Problems with GlobalBioIm,
Differential phase-contrast tomography

Paul Scherrer Institute (PSI), Villigen

Mathematical model

\[ y(t, \theta) = \frac{\partial}{\partial t} R_\theta \{ s \}(t) \]

\[ y = H s \]

\[ [H]_{(i,j),k} = \frac{\partial}{\partial t} P_{\theta, j} \beta_k(t_j) \]

Reducing the numbers of views

Rat brain reconstruction with 181 projections

ADMM-PCG

g-FBP

Collaboration: Prof. Marco Stampanoni, TOMCAT PSI / ETHZ

(Nichian et al. Optics Express 2013)
Performance evaluation

Goldstandard: high-quality iterative reconstruction with 721 views

⇒ Reduction of acquisition time by a factor 10 (or more)?

Compressed sensing: Applications in imaging

- Digital holography (Brady, *Opt. Express* 2009; Marim 2010)
- Spectral-domain OCT (Liu, *Opt. Express* 2010)
- Ultrafast photography (Gao, *Nature* 2014)
Conceptual summary of 2nd generation methods

\[
J(x, u) = \frac{1}{2} \|y - Hx\|_2^2 + \lambda R(u) + \mu \|Lx - u\|_2^2
\]

Physical model \quad \text{Statistical model of signal}

\[J(x, u) = \frac{1}{2} \|y - Hx\|_2^2\]

consistency

\[+ \lambda R(u)\]

prior constraints

\[+ \mu \|Lx - u\|_2^2\]

algorithmic coupling

Schematic structure of reconstruction algorithm:

\[
\begin{align*}
& \text{Repeat} \\
& \quad n = 1, \ldots, N_{\text{iter}} \\
& \quad x^{(n)} = \arg \min_x J(x, u^{(n-1)}) : \text{Linear step (problem specific)} \\
& \quad u^{(n)} = \arg \min_u J(x^{(n)}, u) : \text{Statistical or “denoising” step} \\
& \text{until stop criterion}
\end{align*}
\]

Inverse problems in imaging: Current status

- **Higher reconstruction quality**: Sparsity-promoting schemes almost systematically outperform the classical linear reconstruction methods in MRI, x-ray tomography, deconvolution microscopy, etc... (Lustig et al. 2007)

- **Faster imaging, reduced radiation exposure**: Reconstruction from a lesser number of measurements supported by **compressed sensing**. (Candes-Romberg-Tao; Donoho, 2006)

- **Increased complexity**: Resolution of linear inverse problems using $\ell_1$ regularization requires more sophisticated algorithms (iterative and non-linear); efficient solutions (FISTA, ADMM) have emerged during the past decade. (Chambolle 2004; Figueiredo 2004; Beck-Teboule 2009; Boyd 2011)

- **Outstanding research issues**
  - Beyond $\ell_1$ and TV: Connection with **statistical modeling & learning**
  - Beyond matrix algebra: **Continuous-domain** formulation (Unser, *SIAM Rev* 2017)
Part 4:

The (deep) learning (r)evolution

⇒ Emergence of 3rd generation methods

Learning within the current paradigm

- Data-driven tuning of parameters: $\lambda$, calibration of forward model
  Semi-blind methods, sequential optimization

- Improved decoupling/representation of the signal
  Data-driven **dictionary learning** (based of sparsity or statistics/ICA)  
  $\Rightarrow$ “optimal” $L$

  (Elad 2006, Ravishankar 2011, Mairal 2012)

- Learning of non-linearities / Proximal operators
  CNN-type parametrization, backpropagation  
  $\Rightarrow$ “optimal” potential $\Phi$

  (Chen-Pock 2015-2016, Kamilov 2016)
Structure of iterative reconstruction algorithm

\[ s_{\text{sparse}} = \arg \min_{s \in \mathbb{R}^K} \left( \frac{1}{2} \| y - Hs \|^2_2 + \lambda \| u \|_1 \right) \text{ subject to } u = Ls \]

ADMM

\[ \mathcal{L}_A(s, u, \alpha) = \frac{1}{2} \| y - Hs \|^2_2 + \lambda \sum_n |u_n| + \alpha^T(Ls - u) + \frac{\mu}{2} \| Ls - u \|^2_2 \]

For \( k = 0, \ldots, K \)

\[ s^{k+1} = (H^TH + \mu L^TL)^{-1} (z_0 + z^{k+1}) \]

\[ \alpha^{k+1} = \alpha^k + \mu (Ls^{k+1} - u^k) \]

Linear step

Nonlinear step \approx \text{“denoising” of } u

\[ u^{k+1} = \text{prox}_{\lambda} (Ls^{k+1} + \frac{1}{\mu} \alpha^{k+1}; \frac{\lambda}{\mu}) \]

Connection with deep neural networks

(Gregor-LeCun 2010)

Unrolled Iterative Shrinkage Thresholding Algorithm (ISTA)

**LISTA**: learning-based ISTA

**ISTA with sparsifying transformation**

**FBPConvNet structures**

**Unrolled** Iterative Shrinkage Thresholding Algorithm (ISTA)
Recent appearance of Deep ConvNets

(Jin et al. 2016; Adler-Öktem 2017; Chen et al. 2017; ...)

- CT reconstruction based on Deep ConvNets
  - Input: Sparse view FBP reconstruction
  - Training: Set of 500 high-quality full-view CT reconstructions
  - Architecture: U-Net with skip connection

(Jin et al., IEEE TIP 2017)

CT data

Dose reduction by 7: 143 views

<table>
<thead>
<tr>
<th>Ground truth</th>
<th>FBP SNR 24.06</th>
<th>TV SNR 29.64</th>
</tr>
</thead>
</table>

Reconstructed from from 1000 views
Dose reduction by 7: 143 views

Reconstructed from 1000 views

(Jin et al., *IEEE Trans. Im Proc.*, 2017)

Dose reduction by 20: 50 views

Reconstructed from 1000 views

Dose reduction by 14: 51 views

\[ \mu \text{CT data} \]

Ground truth

FBP
SNR 3.265

TV
SNR 7.481

FBPConvNet
SNR 9.003

Reconstructed from
from 721 views

COMPARISON OF SNR BETWEEN DIFFERENT RECONSTRUCTION ALGORITHMS FOR EXPERIMENTAL DATASET.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Methods</th>
<th>FBP</th>
<th>TV [13]</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg. SNR (dB)</td>
<td>145 views (x5)</td>
<td>5.38</td>
<td>8.25</td>
<td>11.34</td>
</tr>
<tr>
<td></td>
<td>51 views (x14)</td>
<td>3.29</td>
<td>7.25</td>
<td>8.85</td>
</tr>
</tbody>
</table>

**CNN algorithms: Conditions of utilization**

- Standard “regression” setting
  - Mapping of an image into an image
    \[ f_\theta : \mathbb{R}^N \to \mathbb{R}^N : y \mapsto s = f_\theta(y) \]

- Fundamental change of paradigm
  - Requires **extensive sets of representative training data**
  - together with **gold-standards** = desired high-quality reconstruction

- Application niches
  - Denoising
  - Super-resolution (data extrapolation)
  - Reconstruction from **fewer measurements**
    (trained on high-quality full-view data sets)
  - Use of CNN to **emulate/speedup** some well-performing, but “slow”, reference reconstruction methods
Design of CNN algorithms: General principles

■ Data preparation
  ■ Backprojection or classical linear reconstruction
    ⇒ Use of feedforward CNN to correct artifacts of first-generation methods

■ Connection with second-generation methods
  ■ Conceptual: **unrolling** to justify deep architecture

■ **Hybrid** methods (“plug & play”):
  Enforce consistency, while using CNN as “regularizer” or projector
  (Tezcan--Konukoglu, *IEEE TMI* 2018)
  (Gupta--Unser, *IEEE TMI* 2018)

■ Training
  ■ Choice of suitable cost: SNR or perceptual loss
  ■ Availability of extensive data set: \((s_k, y_k), k = 1, \ldots, K\)
  ■ Use of data augmentation: translations, rotations, deformations

Deep CNNs for bioimage reconstruction images

- X-ray tomography \((\text{Jin--Unser, } IEEE TIP 2017)\)
  \((\text{Chen--Wang, } Biomed Opt. Exp 2017)\)

- Magnetic resonance imaging (MRI) \((\text{Hammernik--Pock, } Mag Res Med 2018)\)
  \((\text{Tezcan--Konukoglu, } IEEE TMI 2018)\)

- Dynamic MRI (cardial imaging) \((\text{Schlemper--Rueckert, } IEEE TMI 2018)\)
  \((\text{Hauptmann--Arridge, } Mag Res Med 2019)\)

- 2D microscopy \((\text{Rivenson--Ozcan, } Optica 2017)\)

- 3D fluorescence microscopy \((\text{Weigert--Jug, Myers, } Nature Meth. 2018)\)

- Super-resolution microscopy \((\text{Nehme--Shechtman, } Optica 2018)\)

- Diffraction tomography \((\text{Sun--Kamilov, } Optics Express 2018)\)

- Ultrasound \((\text{Yoon--Ye, } IEEE TMI 2019)\)
Example: MRI reconstruction

Group of Thomas Pock, Univ. Graz


Example: Dynamic MRI reconstruction

Group of Simon Arridge, UCL

(Hauptmann et al., Mag Res Med 2019)
Example: Axial super-resolution in 3D fluorescence microscopy

Group of Florian Jug, Max Planck, Desden


Learning the complete sensor-to-image map, including the physics!

**LETTER**

*Image reconstruction by domain-transform manifold learning*

Bo Zhu1,2,3, Avinash T. Lila4, Stephen F. Cauley1,2, Bruce R. Rosen1,2 & Matthew S. Rosen1,2,3

Image reconstruction is essential for imaging applications across the physical and life sciences, including optical and radar systems, magnetic resonance imaging, X-ray computed tomography, positron emission tomography, ultrasonic imaging and radio astronomy. During image acquisition, the sensor encodes an intermediate representation of an object in the sensor domain, which is subsequently reconstructed into an image by an inversion function. Image reconstruction is challenging because exact knowledge of the exact inverse transformation may not exist a priori, especially in the presence of sensor non-idealities and noise. Thus, the standard reconstruction approach involves approximating the inverse function with multiple ad hoc stages in a signal processing chain (Fig. 1a). The composition of which depends on the details of each acquisition strategy, and often requires expert parameter tuning to optimize reconstruction performance. Here we present a unified framework for image reconstruction—automated transform by manifold approximation (AUTOMAP)—which recasts image reconstruction as a data-driven supervised learning task that allows a mapping between the sensor and the image domain to emerge from an appropriate corpus of training data. We implemented AUTOMAP with a deep neural network and exhibit its flexibility in learning reconstruction transforms for various magnetic resonance imaging acquisition strategies, using the same network architecture and hyperparameters. We further demonstrate that manifold learning during training results in sparse representations of domain transforms along low-dimensional data manifolds, and observe superior immunity to noise and a reduction in reconstruction artefacts compared with conventional handcrafted reconstruction methods. In addition to improving the reconstruction performance of existing acquisition methodologies, we anticipate that AUTOMAP and other learned reconstruction approaches will accelerate the development of new acquisition strategies across imaging modalities.

```
Fundamental limitation: O(n^{2d}) memory requirement ⇒ Does not scale well!
```

---

*doi:10.1038/nature25988*
Deep networks can behave erratically (instability)

Tiny adversarial perturbations of increasing strength

Ground-Truth

State-of-the-art

Compressed Sensing


Conclusion: Frontiers in bioimage reconstruction

- Opportunities for learning-based techniques
  - Faster, higher-resolution, lower-dose imaging
- How the newer methods profit from the older ones

Important open issues

- How does one assess reconstruction quality?
  - Should be “task oriented”!!!
- Improving the stability of CNNs
- Theory to guide the design: What is the optimal architecture?
- Theory to explain the regularization effect of CNNs, and their ability to generalize

Can we trust the results?

- Infrastructure requirements
  - Extensive database of high-quality data (including goldstandard)
  - Development of more realistic simulators
    - both “ground truth” images + physical forward model
  - True 3D CNN toolbox (still missing)
References

■ Foundations


■ Algorithms and imaging applications


References (Cont’d)

■ Deep neural networks

Acknowledgments

Many thanks to (former) members of EPFL’s Biomedical Imaging Group

- Dr. Emmanuel Soubies
- Dr. Kyong Jin
- Dr. McCann
- Dr. Pouya Tafti
- Dr. Julien Fageot
- Prof. Emrah Bostan
- Dr. Masih Nilchian
- Prof. Ulugbek Kamilov
- Dr. Cédric Vonesch

and collaborators ...

- Prof. Demetri Psaltis
- Prof. Marco Stampanoni
- Prof. Carlos-Oscar Sorzano
- Dr. Arne Seitz
- ....

Preprints and demos: [http://bigwww.epfl.ch/](http://bigwww.epfl.ch/)