Optimization transfer approach to joint registration / reconstruction for motion-compensated image reconstruction

Jeffrey A. Fessler

EECS Dept.
University of Michigan

ISBI
Apr. 15, 2010

Introduction

Image reconstruction of moving objects with unknown motion

Joint estimation of motion parameters and object

- Jacobson *et al.*, IEEE NSS 2003
- Taguchi *et al.* SPIE 2007
- Odille *et al.* MRM Jul. 2008
- Also super-resolution problems with unknown motion
  (cf starting with low-resolution images vs starting with sinograms or k-space)
- ...

Computational challenge: motion operator in forward model.

We use optimization transfer to put motion estimation step in image domain

Measurement model

M "frames" (defined generally)

$$\boldsymbol{y}_m = \boldsymbol{A}_m \boldsymbol{x}_m + \boldsymbol{\varepsilon}_m, \qquad m = 1, \ldots, M$$

- $\boldsymbol{y}_m$ measured data for mth frame
- $\boldsymbol{A}_m$ system matrix for mth frame
- $\boldsymbol{x}_m$ unknown image for mth frame
- $\boldsymbol{\varepsilon}_m$ measurement noise for mth frame

Nominal goal:
reconstruct image frames $\{\boldsymbol{x}_m\}$ from measured data $\{\boldsymbol{y}_m\}$.

Object model

Assume each frame is a spatial transformation of one base image:

$$\boldsymbol{x}_m = \boldsymbol{T}(\boldsymbol{\alpha}_m)\,\boldsymbol{c}$$

- $\boldsymbol{\alpha}_m$ motion parameters for mth frame
- $\boldsymbol{T}(\cdot)$ nonrigid warp operator
- $\boldsymbol{c}$ base image coefficient vector (*e.g.*, for B-splines)

$$\boldsymbol{x}_0 = \boldsymbol{T}(\mathbf{0})\,\boldsymbol{c}$$

Motion-compensated image reconstruction goal:
reconstruct base image coefficients $\boldsymbol{c}$ and motion parameters $\{\boldsymbol{\alpha}_m\}$ from measured data $\{\boldsymbol{y}_m\}$.

Joint registration/reconstruction

Combined measurement model / object model:

$$\boldsymbol{y}_m = \boldsymbol{A}_m \boldsymbol{T}(\boldsymbol{\alpha}_m)\,\boldsymbol{c} + \boldsymbol{\varepsilon}_m, \qquad m = 1, \ldots, M$$

Stacking, where $\boldsymbol{\alpha} \triangleq (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_M)$:

$$\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{T}(\boldsymbol{\alpha})\,\boldsymbol{c} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{y} \triangleq \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_M \end{bmatrix} \quad \boldsymbol{A} \triangleq \begin{bmatrix} \boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_M \end{bmatrix} \quad \boldsymbol{T}(\boldsymbol{\alpha}) \triangleq \begin{bmatrix} \boldsymbol{T}(\boldsymbol{\alpha}_1) \\ \vdots \\ \boldsymbol{T}(\boldsymbol{\alpha}_M) \end{bmatrix} \quad \boldsymbol{\varepsilon} \triangleq \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \vdots \\ \boldsymbol{\varepsilon}_M \end{bmatrix}$$

Penalized weighted least-squares (PWLS) estimation:

$$(\hat{\boldsymbol{c}}, \hat{\boldsymbol{\alpha}}) = \arg\min_{\boldsymbol{c},\boldsymbol{\alpha}} \Psi(\boldsymbol{c}, \boldsymbol{\alpha})$$

$$\Psi(\boldsymbol{c}, \boldsymbol{\alpha}) = \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{A}\,\boldsymbol{T}(\boldsymbol{\alpha})\,\boldsymbol{c}\|_{\boldsymbol{W}}^2 + R_1(\boldsymbol{c}) + R_2(\boldsymbol{\alpha})$$

Optimization by alternation

Initialize base image $\boldsymbol{c}^0$ and motion parameters $\boldsymbol{\alpha}^0$.

Alternating updates:

$$\begin{aligned}\boldsymbol{c}^{n+1} &= \arg\min_{\boldsymbol{c}} \Psi(\boldsymbol{\alpha}^n, \boldsymbol{c}) \\ \boldsymbol{\alpha}^{n+1} &= \arg\min_{\boldsymbol{\alpha}} \Psi\left(\boldsymbol{\alpha}, \boldsymbol{c}^{n+1}\right)\end{aligned}$$

Update $\boldsymbol{c}$ using standard image reconstruction methods:

$$\boldsymbol{c}^{n+1} = \arg\min_{\boldsymbol{c}} \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{A}\,\boldsymbol{T}(\boldsymbol{\alpha}^n)\,\boldsymbol{c}\|_{\boldsymbol{W}}^2 + R_1(\boldsymbol{c})$$

Updating motion parameters $\boldsymbol{\alpha}$ is challenging:

$$\boldsymbol{\alpha}^{n+1} = \arg\min_{\boldsymbol{\alpha}} \psi(\boldsymbol{\alpha}), \quad \psi(\boldsymbol{\alpha}) = \frac{1}{2} \left\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{T}(\boldsymbol{\alpha})\boldsymbol{c}^{n+1}\right\|_{\boldsymbol{W}}^2 + R_2(\boldsymbol{\alpha})$$

Optimization transfer / Majorize-minimize

Alternative to minimizing cost function $\psi(\boldsymbol{\alpha})$ directly:

- S-step: find a surrogate function (majorizer) $\phi(\boldsymbol{\alpha}; \boldsymbol{\alpha}^n)$ s.t.

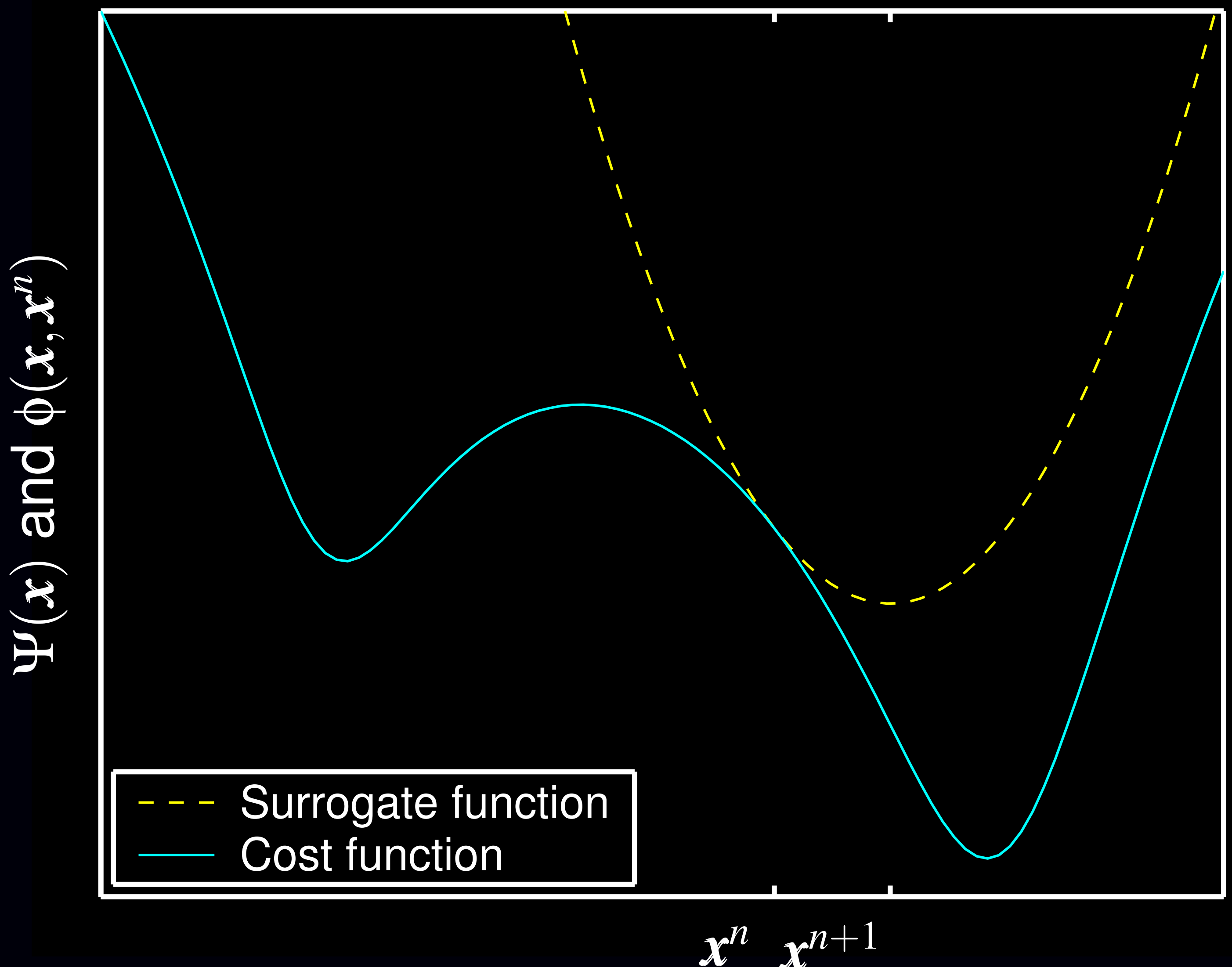
$$\begin{aligned} \phi(\boldsymbol{\alpha}; \boldsymbol{\alpha}^n) &\geq \psi(\boldsymbol{\alpha}), \qquad \forall \boldsymbol{\alpha} \\ \phi(\boldsymbol{\alpha}^n; \boldsymbol{\alpha}^n) &= \psi(\boldsymbol{\alpha}^n) \end{aligned}$$

- M-step: minimize surrogate function:

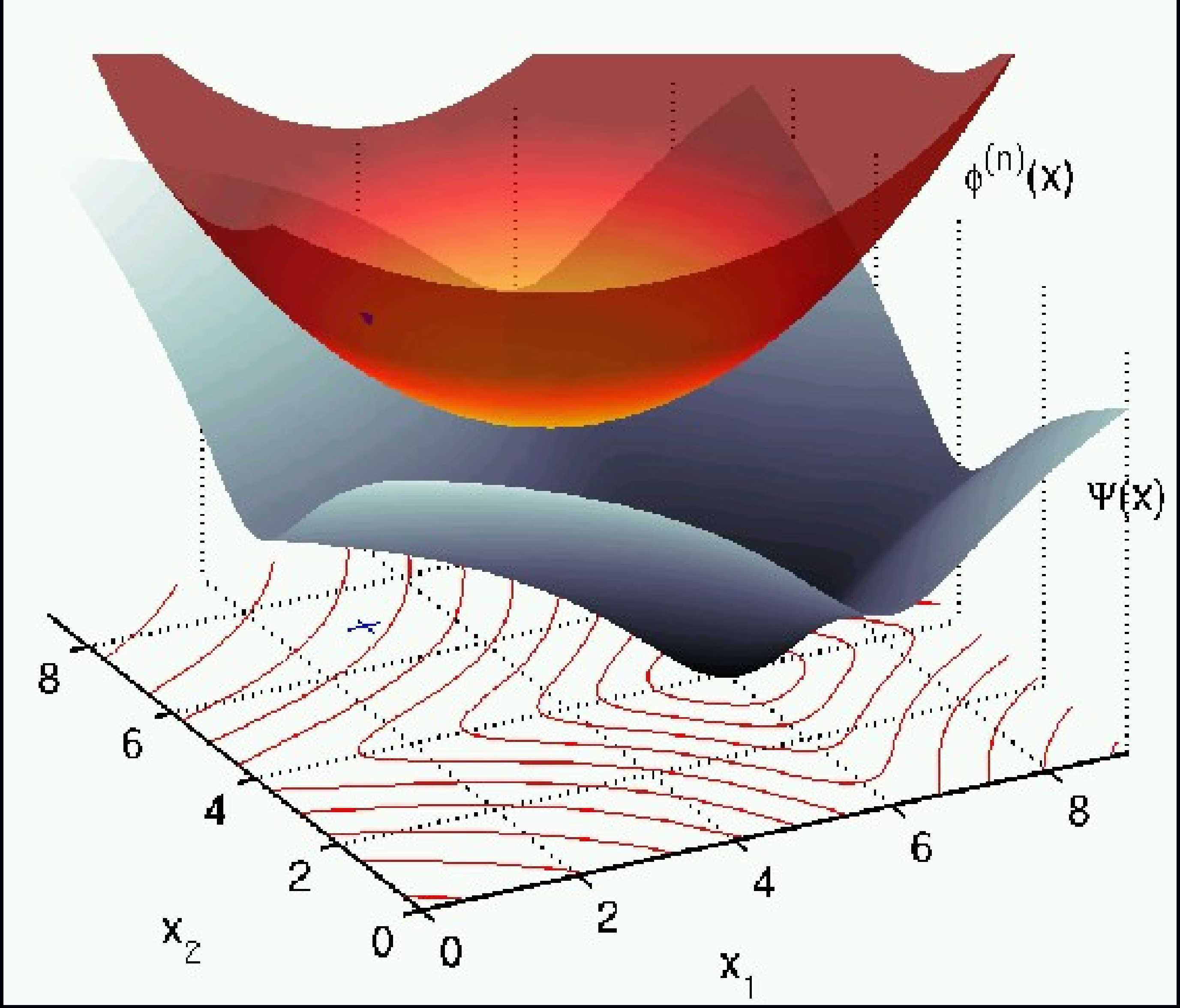
$$\boldsymbol{\alpha}^{n+1} = \arg\min_{\boldsymbol{\alpha}} \phi(\boldsymbol{\alpha}; \boldsymbol{\alpha}^n).$$

Guaranteed to decrease cost function $\psi(\boldsymbol{\alpha}^n)$ *monotonically*.

Optimization Transfer Illustrated



Optimization Transfer in 2d



Separable quadratic surrogate for WLS

$$
\begin{aligned}
L(\boldsymbol{x}) &= \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|_{\boldsymbol{W}}^2 = \sum_i \frac{w_i}{2}(y_i - [\boldsymbol{A}\boldsymbol{x}]_i)^2 \\
&= L(\boldsymbol{x}^n) + (\boldsymbol{x}-\boldsymbol{x}^n)'\boldsymbol{g}^n + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^n)'\boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{x}^n) \\
&\leq L(\boldsymbol{x}^n) + (\boldsymbol{x}-\boldsymbol{x}^n)'\boldsymbol{g}^n + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^n)'\boldsymbol{D}(\boldsymbol{x}-\boldsymbol{x}^n) \\
&\equiv \frac{1}{2}\left\|\boldsymbol{D}^{1/2}\left[\boldsymbol{x}-\left(\boldsymbol{x}^n+\boldsymbol{D}^{-1}\boldsymbol{g}^n\right)\right]\right\|^2 \triangleq Q(\boldsymbol{x};\boldsymbol{x}^n),
\end{aligned}
$$

when $\boldsymbol{D}$ is any matrix that satisfies $\boldsymbol{D} \succeq \boldsymbol{A}'\boldsymbol{W}\boldsymbol{A}$, where

$$
\boldsymbol{g}^n \triangleq -\nabla L(\boldsymbol{x}^n) = \boldsymbol{A}'\boldsymbol{W}\left(\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}^n\right).
$$

Useful choice (Erdoğan and Fessler, PMB 1999):

$$
\boldsymbol{D} = \operatorname{diag}\{d_j\}, \quad d_j \triangleq \sum_i w_i |a_{ij}| \left(\sum_k |a_{ik}|\right).
$$

Surrogate for motion parameters

Recall: $\psi(\boldsymbol{\alpha}) = \frac{1}{2}\left\| \boldsymbol{y} - \boldsymbol{A}\,\underbrace{\boldsymbol{T}(\boldsymbol{\alpha})\boldsymbol{c}^{n+1}}_{\text{“}\boldsymbol{x}\text{”}} \right\|_{\boldsymbol{W}}^2 + R_2(\boldsymbol{\alpha})$

Using result for WLS, the following majorizes ψ:

$$\begin{aligned}
\phi(\boldsymbol{\alpha}, \boldsymbol{\alpha}^n) &\triangleq Q(\boldsymbol{T}(\boldsymbol{\alpha})\,\boldsymbol{c}^n, \boldsymbol{T}(\boldsymbol{\alpha}^n)\,\boldsymbol{c}^n) + R_2(\boldsymbol{\alpha}) \\
&= \frac{1}{2}\left\| \boldsymbol{D}^{1/2}\left[\boldsymbol{T}(\boldsymbol{\alpha})\,\boldsymbol{c}^n - \left(\boldsymbol{T}(\boldsymbol{\alpha}^n)\,\boldsymbol{c}^n + \boldsymbol{D}^{-1}\boldsymbol{g}^n\right)\right] \right\|^2 + R_2(\boldsymbol{\alpha}),
\end{aligned}$$

where the gradient depends on *previous* estimates:

$$\boldsymbol{g}^n \triangleq -\nabla L(\boldsymbol{x}^n) = \boldsymbol{A}'\boldsymbol{W}\left(\boldsymbol{y} - \boldsymbol{A}\,\boldsymbol{T}(\boldsymbol{\alpha}^n)\,\boldsymbol{c}^n\right).$$

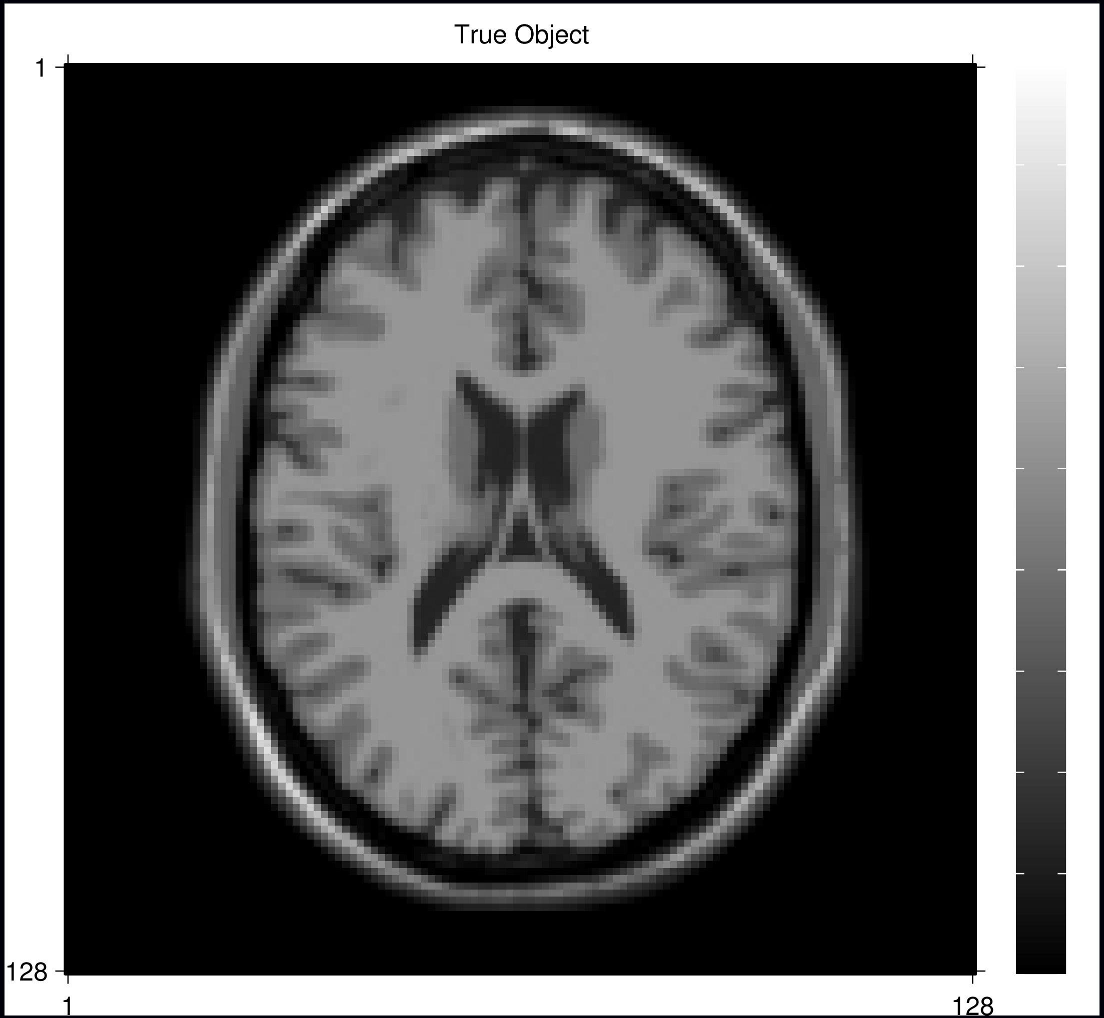
M-step is entirely image-domain operations:

$$\boldsymbol{\alpha}^{n+1} = \arg\min_{\boldsymbol{\alpha}} \phi(\boldsymbol{\alpha}, \boldsymbol{\alpha}^n).$$

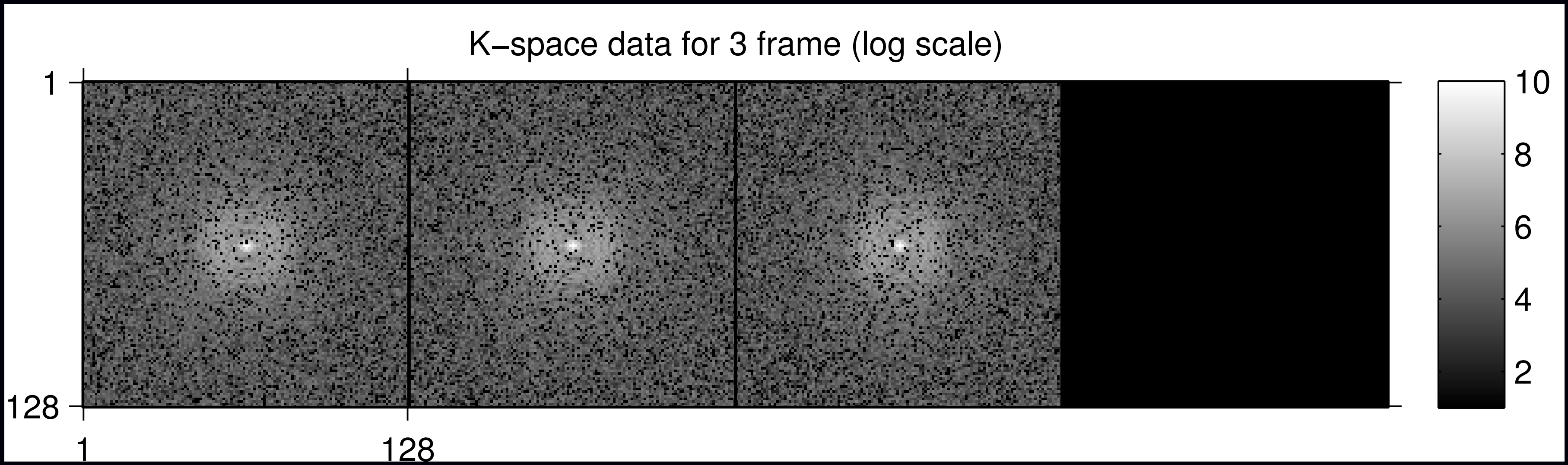
(Guobao Wang & Jinyi Qi did something similar for PET kinetics in ISBI 2008)

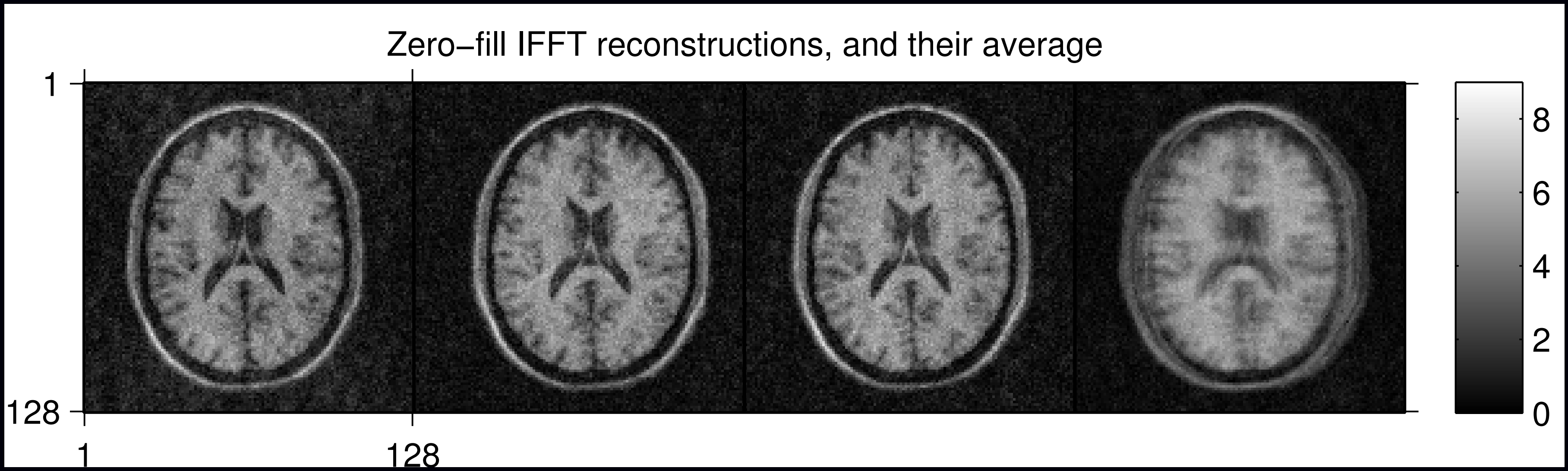
Simulation Example

MRI with randomly ordered k_y,k_z phase encodes



Simulation Example



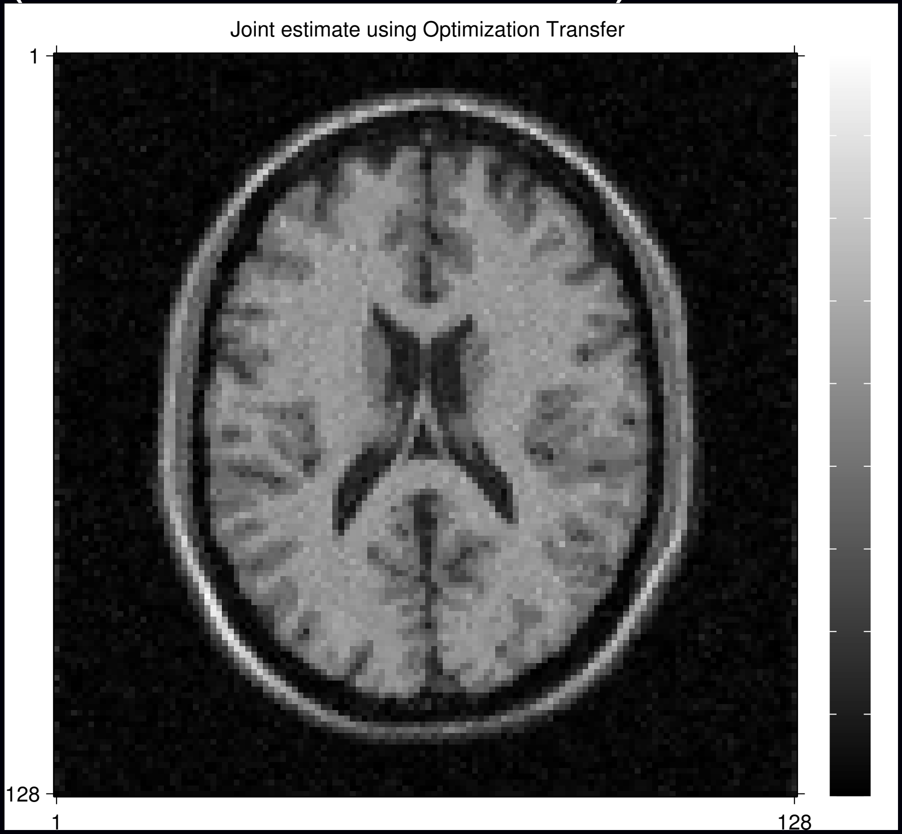


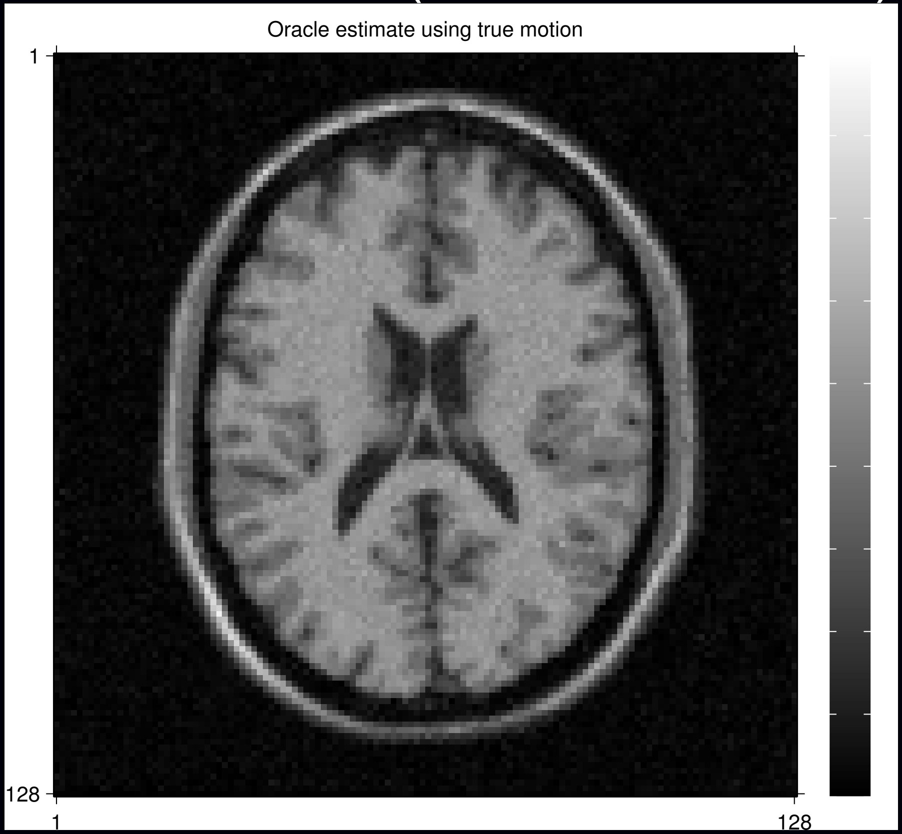
True motion: horizontal translation: -2, 4, 0

Joint motion estimation / reconstruction

Joint estimate:
(3 alternations, 9 CG)

Oracle reconstruction
("known" motion)





Estimated horizontal translation: -1.4, 4.4, 0.4
True horizontal translation: -2, 4, 0

Discussion

- Proposed optimization transfer approach to joint image reconstruction and motion estimation
- Puts motion estimation step in image domain
  - akin to image registration
  - avoids expensive forward projections in registration step
  - promotes software modularity

    (*e.g.*, importance sampling, Bhagalia *et al.*, IEEE T-MI, Aug. 2009)
- Optimization transfer approach may require “more” iterations
- Usual nonconvexity issues (multi-resolution...)
- Generalizes readily to non-quadratic log-likelihood functions (*e.g.*, PET, polyenergetic CT, ...)
- Awaits evaluation with real data...
- Awaits comparisons with nonlinear CG
- Open problem: using ordered-subsets approaches efficiently