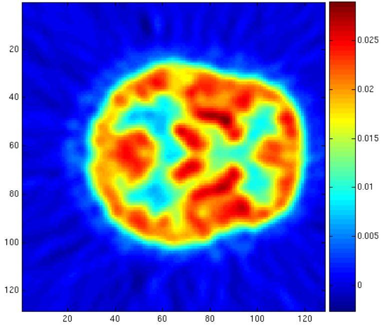
Information Extraction from PET Images

Rosemary A Renaut and Wolfgang Stefan

Arizona State University

May 16, 2006, SIAM Imaging Conference 2006

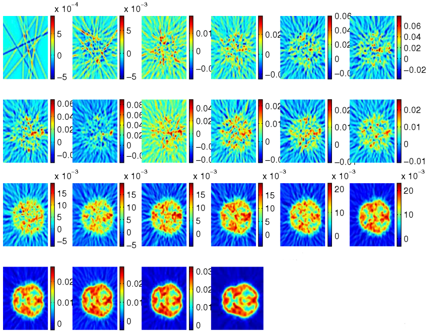
Example of typical PET scan



Typical PET Images show

- High noise content (non Gaussian)
- High blurring
- Reconstruction artifacts
- Reconstruction using filtered backprojection

Dynamic PET scan



Dynamic data

- Very poor initial scans
- Noise levels change across scans
- Solve inverse problem to estimate kinetic parameters.
- What does it mean to improve these images?

Goals of the Study

- Improved Sensitivity for identifying features in images:

Goals of the Study

- Improved Sensitivity for identifying features in images:
  - Longitudinal Studies of AD progression show changes
  - Assessing disease state- AD or MCI?
  - Assess impact of drug treatment

Goals of the Study

- Improved Sensitivity for identifying features in images:
  - Longitudinal Studies of AD progression show changes
  - Assessing disease state- AD or MCI?
  - Assess impact of drug treatment
- Equivalently, noninvasive assistance in AD studies.

Goals of the Study

- Improved Sensitivity for identifying features in images:
  - Longitudinal Studies of AD progression show changes
  - Assessing disease state- AD or MCI?
  - Assess impact of drug treatment
- Equivalently, noninvasive assistance in AD studies.
- $\Rightarrow$ **use imaging**

Goals of the image processing

- **Improve** images to

Goals of the image processing

- **Improve** images to
  - **Quantify** data from dynamic scans.

Goals of the image processing

- **Improve** images to
  - **Quantify** data from dynamic scans.
  - **Register and compare** baseline and followup parametric scans

Goals of the image processing

- **Improve** images to
  - **Quantify** data from dynamic scans.
  - **Register and compare** baseline and followup parametric scans
- **remove noise and artifacts from difference scan**

Goals of the image processing

- **Improve** images to
  - **Quantify** data from dynamic scans.
  - **Register and compare** baseline and followup parametric scans
- **remove noise and artifacts from difference scan**
- **Deblur images by deconvolution**

Inverse Problem

- Find f from $\boxed{g = f * h + n}$ given g and h with unknown n.
- g is the recorded image, f the unknown real image, h the point spread function (PSF) and n unknown noise.

Inverse Problem

- Find f from $\boxed{g = f * h + n}$ given g and h with unknown n.
- g is the recorded image, f the unknown real image, h the point spread function (PSF) and n unknown noise.
- Assuming normal distributed n yields including regularization

$$\hat{f} = \arg\min_{f}\{\|g - f * h\|_2^2 + \lambda R(f)\}$$

Regularization Methods (review)

- Common methods are Tikhonov (TK).

$$R(f) = \mathrm{TK}(f) = \int_{\Omega} |\nabla f(x)|^2 dx.$$

Regularization Methods (review)

- Common methods are Tikhonov (TK).

$$R(f) = \mathrm{TK}(f) = \int_\Omega |\nabla f(x)|^2 dx.$$

- Total Variation (TV)

$$R(f) = \mathrm{TV}(f) = \int_\Omega |\nabla f(x)| dx.$$

Regularization Methods (review)

- Common methods are Tikhonov (TK).

$$R(f) = \mathrm{TK}(f) = \int_\Omega |\nabla f(x)|^2 dx.$$

- Total Variation (TV)

$$R(f) = \mathrm{TV}(f) = \int_\Omega |\nabla f(x)| dx.$$

- Sparse deconvolution (L^1) (not relevant for PET images)

$$R(f) = \|f\|_1 = \int_\Omega |f(x)| dx.$$

Regularization Review

$$\hat{f} = \arg\min_{f}\{\|g - f * h\|_2^2 + \lambda R(f)\}$$

- λ **Governs the trade off** between the fit to the data and the smoothness of the reconstruction and can be picked by the L-curve approach, (see Hansen, Inverse Problems)

Regularization Review

$$\hat{f} = \arg\min_{f} \{ \|g - f * h\|_2^2 + \lambda R(f) \}$$

- λ **Governs the trade off** between the fit to the data and the smoothness of the reconstruction and can be picked by the L-curve approach, (see Hansen, Inverse Problems)
- **TV yields a piece wise constant** reconstruction and preserves the edges of the image.

Regularization Review

$$\hat{f} = \arg\min_{f}\{\|g - f * h\|_2^2 + \lambda R(f)\}$$

- λ **Governs the trade off** between the fit to the data and the smoothness of the reconstruction and can be picked by the L-curve approach, (see Hansen, Inverse Problems)
- **TV yields a piece wise constant** reconstruction and preserves the edges of the image.
- **TK yields a smooth** reconstruction.

Regularization Review

$$\hat{f} = \arg\min_{f}\{\|g - f * h\|_2^2 + \lambda R(f)\}$$

- λ **Governs the trade off** between the fit to the data and the smoothness of the reconstruction and can be picked by the L-curve approach, (see Hansen, Inverse Problems)
- **TV yields a piece wise constant** reconstruction and preserves the edges of the image.
- **TK yields a smooth** reconstruction.
- **L1 yields spike trains**

Optimization

- Objective function of the data fit term is **convex**

Optimization

- Objective function of the data fit term is **convex**
- TK is a linear least squares (LS) problem

$$\hat{f} = \arg\min_{f}\{\|g - Hf\|_2^2 + \lambda\|\nabla f\|_2^2)\}$$

Optimization

- Objective function of the data fit term is **convex**
- TK is a linear least squares (LS) problem

$$\hat{f} = \arg\min_{f}\{\|g - Hf\|_2^2 + \lambda\|\nabla f\|_2^2)\}$$

- The TV objective function is **non differentiable**

$$J(f) = \|g - Hf\|_2^2 + \lambda\|\nabla f\|_1$$

Differentiability of TV - $1D$ (tensor product in $2D$)

$$R(f) = \sum_i \|f_{i+1} - f_i\|$$

- for a small β define

$$R_\beta = \sum_i \sqrt{(f_{i+1} - f_i)^2 + \beta}$$

Differentiability of TV - $1D$ (tensor product in $2D$)

$$R(f) = \sum_i \|f_{i+1} - f_i\|$$

- for a small β define

$$R_\beta = \sum_i \sqrt{(f_{i+1} - f_i)^2 + \beta}$$

- choose β in 10^{-5} to 10^{-9}

Numerical optimization scheme

- To find the minimum we use a limited memory BFGS (see Vogel, Computational Methods for Inverse Problems and Nocedal)

Numerical optimization scheme

- To find the minimum we use a limited memory BFGS (see Vogel, Computational Methods for Inverse Problems and Nocedal)
- A quasi Newton Method where the estimated Hessian in each step is updated by a rank 2 update.

Numerical optimization scheme

- To find the minimum we use a limited memory BFGS (see Vogel, Computational Methods for Inverse Problems and Nocedal)
- A quasi Newton Method where the estimated Hessian in each step is updated by a rank 2 update.
- Only a limited number of update vectors are kept, e.g. 10.

Numerical optimization scheme

- To find the minimum we use a limited memory BFGS (see Vogel, Computational Methods for Inverse Problems and Nocedal)
- A quasi Newton Method where the estimated Hessian in each step is updated by a rank 2 update.
- Only a limited number of update vectors are kept, e.g. 10.
- **Evaluation** of the OF and its gradient **is cheap** (some FFTs and sparse matrix-vector multiplications)

ARIZONA STATE UNIVERSITY

Numerical optimization scheme

- To find the minimum we use a limited memory BFGS (see Vogel, Computational Methods for Inverse Problems and Nocedal)
- A quasi Newton Method where the estimated Hessian in each step is updated by a rank 2 update.
- Only a limited number of update vectors are kept, e.g. 10.
- **Evaluation** of the OF and its gradient **is cheap** (some FFTs and sparse matrix-vector multiplications)
- Problems are usually large and many iterations are needed.

Point Spread Function

- The PSF is usually unknown or only estimated
- Estimates exist for PET scanners from phantom scans
- PSF is spatially variant and also depends on the scanned object

Point Spread Function

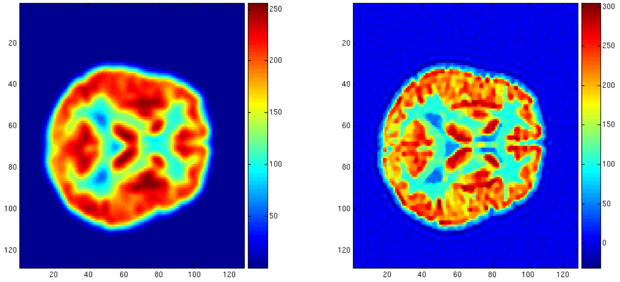
- The PSF is usually unknown or only estimated
- Estimates exist for PET scanners from phantom scans
- PSF is spatially variant and also depends on the scanned object
- Hence even if provided PSF is always **only an estimate**
- for the PET scans presented here, a $6mm$ half width Gaussian was assumed.

Simulated PET

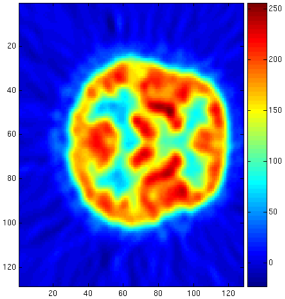
- On the left simulated PET from blurred segmented MRI scan using **Gaussian PSF** and noise added.

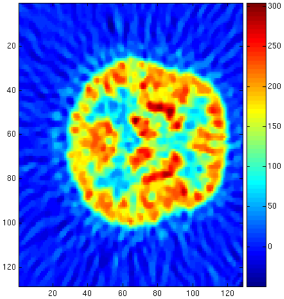
Simulated PET

- On the left simulated PET from blurred segmented MRI scan using **Gaussian PSF** and noise added.
- On the right deblurred PET with TV and known PSF.



Recover real PET image





- Reconstruction done using Filtered Back Projection
- PSF estimated by a Gaussian
- TV regularization

Observations

- Image improvement is possible even with a rough estimation of the PSF (non-blind decon)

Observations

- Image improvement is possible even with a rough estimation of the PSF (non-blind decon)
- Total Variation regularization (piecewise constant solution) is appropriate: intensity levels depend on the tissue type.

Observations

- Image improvement is possible even with a rough estimation of the PSF (non-blind decon)
- Total Variation regularization (piecewise constant solution) is appropriate: intensity levels depend on the tissue type.
- Improvement requires better approximation of the PSF

Observations

- Image improvement is possible even with a rough estimation of the PSF (non-blind decon)
- Total Variation regularization (piecewise constant solution) is appropriate: intensity levels depend on the tissue type.
- Improvement requires better approximation of the PSF
- Total Least Squares

Total least squares (TLS)

- Rewrite convolution as matrix vector product:

$$g = Hf + n$$

- H is a Toeplitz matrix of Point Spread Function

Total least squares (TLS)

- Rewrite convolution as matrix vector product:

$$g = Hf + n$$

- H is a Toeplitz matrix of Point Spread Function
- TLS assumes error in H and g i.e.

$$g = (H + E)f + n$$

Total least squares (TLS)

- Rewrite convolution as matrix vector product:

$$g = Hf + n$$

- H is a Toeplitz matrix of Point Spread Function
- TLS assumes error in H and g i.e.

$$g = (H + E)f + n$$

- Total least squares solution f_{TLS} solves

$$min\|E|n\|_F^2 \quad \text{subject to} \quad g = (H + E)f + n$$

Total least squares (TLS)

- Rewrite convolution as matrix vector product:

$$g = Hf + n$$

- H is a Toeplitz matrix of Point Spread Function
- TLS assumes error in H and g i.e.

$$g = (H + E)f + n$$

- Total least squares solution f_{TLS} solves

$$min \| E|n \|_F^2 \quad \text{subject to} \quad g = (H + E)f + n$$

- f_{TLS} can be found from SVD of $[H, g]$ (Golub et al)

Rayleigh Quotient Formulation

- The TLS solution minimizes Rayleigh Quotient:

$$min_f \frac{\|Hf - g\|_2^2}{1 + \|f\|_2^2}$$

Rayleigh Quotient Formulation

- The TLS solution minimizes Rayleigh Quotient:

$$min_f \frac{\|Hf - g\|_2^2}{1 + \|f\|_2^2}$$

- include regularization:

$$min_f \frac{\|Hf - g\|_2^2}{1 + \|f\|_2^2} + \lambda R(f)$$

Rayleigh Quotient Formulation

- The TLS solution minimizes Rayleigh Quotient:

$$min_f \frac{\|Hf - g\|_2^2}{1 + \|f\|_2^2}$$

- include regularization:

$$min_f \frac{\|Hf - g\|_2^2}{1 + \|f\|_2^2} + \lambda R(f)$$

- p=2 e.g. Golub et al (1999), Renaut et al (2005)

Scaled TLS: different noise levels

- Theory (Paige and Strakos, Numerische Mathematik)

$$min\|E|n\|_F^2 \quad \text{subject to} \quad g = (H+E)f + \frac{n}{\gamma}$$

Scaled TLS: different noise levels

- Theory (Paige and Strakos, Numerische Mathematik)

$$min\|E|n\|_F^2 \quad \text{subject to} \quad g = (H+E)f + \frac{n}{\gamma}$$

- Minimum is obtained as the minimum singular value of $[H, \gamma g]$

Scaled TLS: different noise levels

- Theory (Paige and Strakos, Numerische Mathematik)

$$min\|E|n\|_F^2 \quad \text{subject to} \quad g = (H+E)f + \frac{n}{\gamma}$$

- Minimum is obtained as the minimum singular value of $[H, \gamma g]$
- For flexibility use Rayleigh quotient formulation

$$min_f \frac{\|Hf - g\|_2^2}{1 + \gamma^2 \|f\|_2^2}$$

- $\gamma = 0$ is the LS problem
- $\gamma = 1$ is the standard TLS problem

Scaled TLS: different noise levels

- Theory (Paige and Strakos, Numerische Mathematik)

$$min\|E|n\|_F^2 \quad \text{subject to} \quad g = (H+E)f + \frac{n}{\gamma}$$

- Minimum is obtained as the minimum singular value of $[H, \gamma g]$
- For flexibility use Rayleigh quotient formulation

$$min_f \frac{\|Hf - g\|_2^2}{1 + \gamma^2 \|f\|_2^2}$$

- $\gamma = 0$ is the LS problem
- $\gamma = 1$ is the standard TLS problem
- γ accounts for different noise levels in H and g.

Scaled Total Least Squares with Regularization

- Regularize scaled RQ:

$$min_f \frac{\|Hf - g\|_2^2}{1 + \gamma^2 \|f\|_2^2} + \lambda R(f)$$

Scaled Total Least Squares with Regularization

- Regularize scaled RQ:

$$min_f \frac{\|Hf - g\|_2^2}{1 + \gamma^2 \|f\|_2^2} + \lambda R(f)$$

- Permits careful investigation of effect of noise levels in H and g.

Scaled Total Least Squares with Regularization

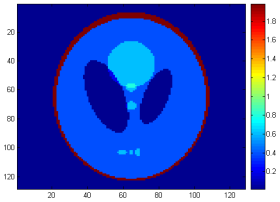
- Regularize scaled RQ:

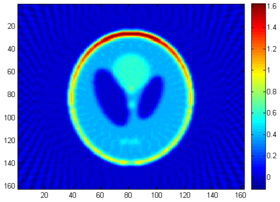
$$min_f \frac{\|Hf - g\|_2^2}{1 + \gamma^2 \|f\|_2^2} + \lambda R(f)$$

- Permits careful investigation of effect of noise levels in H and g.
- Which is greater , the error in the PSF or the error in the measured data?

Test Problem Noisy Shepp Logan Phantom

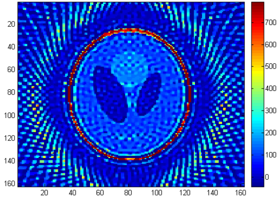
- Use 128x128 Shepp Logan Phantom
  - Blur with Gaussian $h(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{|x|^2}{2\sigma^2}}$ with $\sigma = 1.5$ (6mm half width)
  - Take forward Radon transform with 45 angles
  - Add Poisson Noise to sinogram
  - Transform back, with filtered back projection



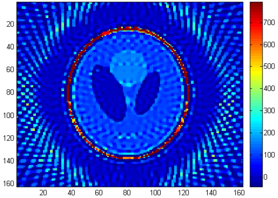


Deconvolving the Shepp-Logan Phantom

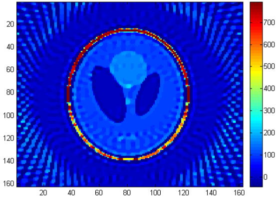
- Gauss PSF with $\sigma = 2$ and TV regularization



$\gamma^2 = 0$



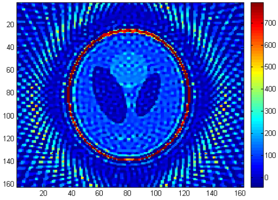
$\gamma^2 = 7.3e - 9$



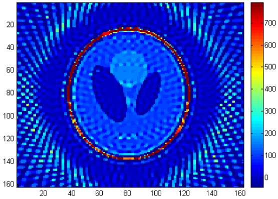
$\gamma^2 = 3.7e - 7$

Deconvolving the Shepp-Logan Phantom

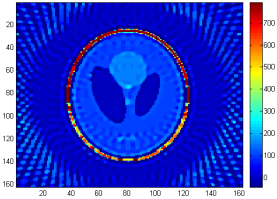
- Gauss PSF with $\sigma = 2$ and TV regularization



$\gamma^2 = 0$



$\gamma^2 = 7.3e - 9$

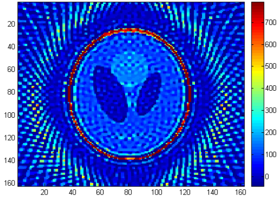


$\gamma^2 = 3.7e - 7$

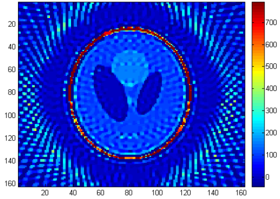
- Scaling shows how to improve impact of badly chosen PSF.

Deconvolving the Shepp-Logan Phantom

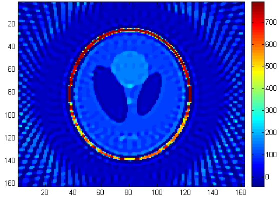
- Gauss PSF with $\sigma = 2$ and TV regularization



$\gamma^2 = 0$



$\gamma^2 = 7.3e - 9$

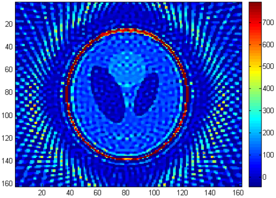


$\gamma^2 = 3.7e - 7$

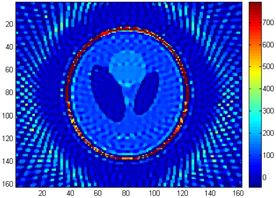
- Scaling shows how to improve impact of badly chosen PSF.
- Scaling $\|Hf\| = \|g\| = 1$ (Notice for given image $\|g\|^2 \approx 10^9$)

Deconvolving the Shepp-Logan Phantom

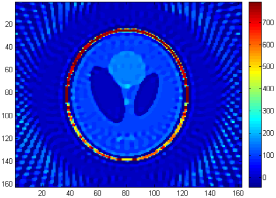
- Gauss PSF with $\sigma = 2$ and TV regularization



$\gamma^2 = 0$



$\gamma^2 = 7.3e - 9$

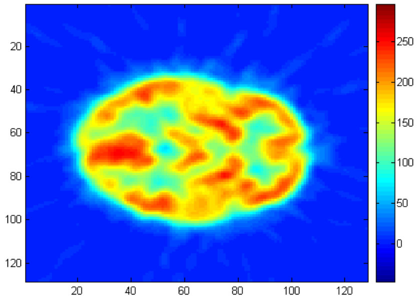


$\gamma^2 = 3.7e - 7$

- Scaling shows how to improve impact of badly chosen PSF.
- Scaling $\|Hf\| = \|g\| = 1$ (Notice for given image $\|g\|^2 \approx 10^9$)
- For scaled problem γ^2 is 1, 50, resp.

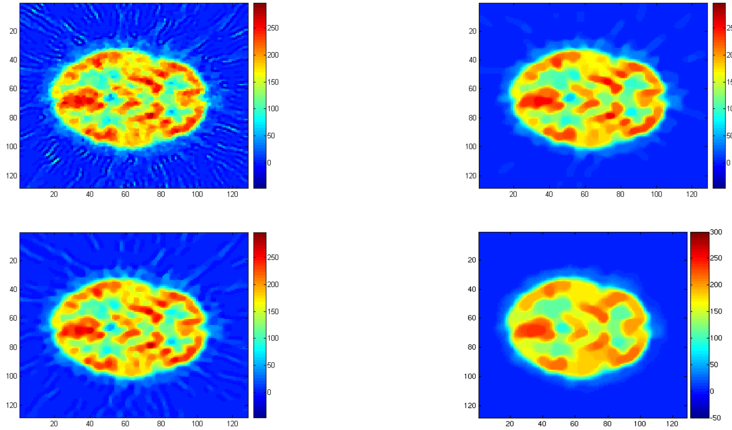
Real PET data

- Real PET data

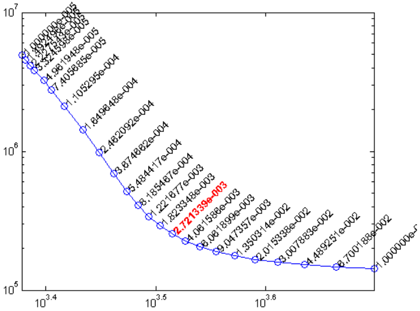


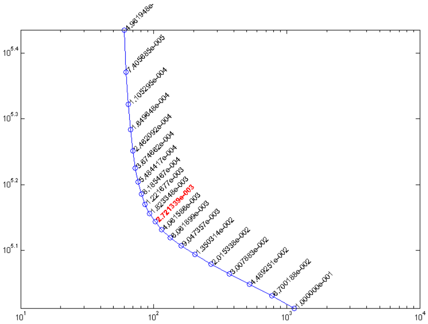
Numerical Results Scaled RTVTLS

PSF $6mm$ half width Gaussian, $\gamma = 0$ (LS), $3.7e-7$, $1e-5$ and $1e-4$ (top to bottom and left to right 0, 50, $1.4e3$, $1.4e4$)



L-curve





- Notice optimal λ by L-curve is similar for LS and Scaled TLS.

Observations

- RTLS with TV handles inexact PSFs better then simple RLS

Observations

- RTLS with TV handles inexact PSFs better then simple RLS
- Scaled RTVTLS with parameter γ in

$$min_f \frac{\|Hf - g\|_2^2}{1 + \gamma^2 \|f\|_2^2} + \lambda R(f)$$

allows further tuning in case of an unknown PSF

Observations

- RTLS with TV handles inexact PSFs better then simple RLS
- Scaled RTVTLS with parameter γ in

$$min_f \frac{\|Hf - g\|_2^2}{1 + \gamma^2 \|f\|_2^2} + \lambda R(f)$$

  allows further tuning in case of an unknown PSF
- Iterations are expensive.

Conclusions and Future Work

- Further investigation of RTVTLS and relation to RTVLS (also with scaling)

Conclusions and Future Work

- Further investigation of RTVTLS and relation to RTVLS (also with scaling)
- Improve efficiency of algorithms (methods of Guo and Renaut)

Conclusions and Future Work

- Further investigation of RTVTLS and relation to RTVLS (also with scaling)
- Improve efficiency of algorithms (methods of Guo and Renaut)
- Further interaction with medical consultants for impact and direction of the work.

Conclusions and Future Work

- Further investigation of RTVTLS and relation to RTVLS (also with scaling)
- Improve efficiency of algorithms (methods of Guo and Renaut)
- Further interaction with medical consultants for impact and direction of the work.
- What can be achieved with wavelets- see Wolfgang Stefan pm.

Acknowledgments

- Hongbin Guo (Total Least Squares)
- Haewon Nam and Kewei Chen for the data and discussions on PET imaging
- Supported by: Arizona Alzheimer's Research Center and NIH NIBIB