Sparse Modeling for Prospective Head Motion Correction

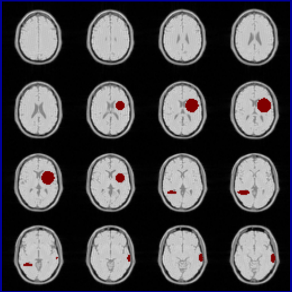
Daniel S. Weller

University of Michigan

September 26, 2013

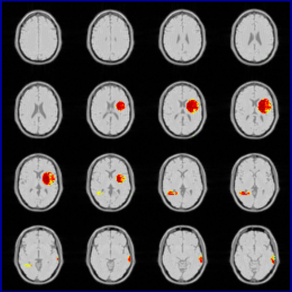
Motivation

- Head motion reduces the sensitivity of functional magnetic resonance imaging (fMRI) experiments.
- I propose a motion compensation method that uses just the data collected during a conventional fMRI.



True correlations (red)

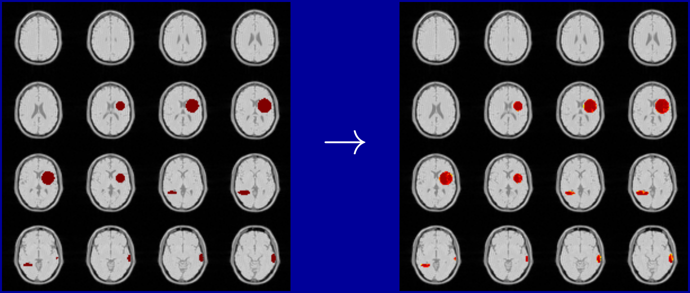
$\longrightarrow$



After head motion

Motivation

- Head motion reduces the sensitivity of functional magnetic resonance imaging (fMRI) experiments.
- I propose a motion compensation method that uses just the data collected during a conventional fMRI.



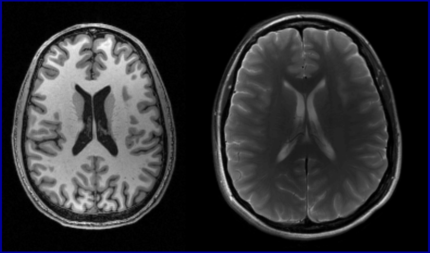
True correlations (red) → With proposed correction

Outline

- Magnetic Resonance Imaging
- Functional MRI
- Head Motion in fMRI
- Proposed Method for Prospective Correction
- Real Time Implementation
- Simulation Results
- Conclusions and Future Work
- Summary

What is MRI?

MRI Scanner

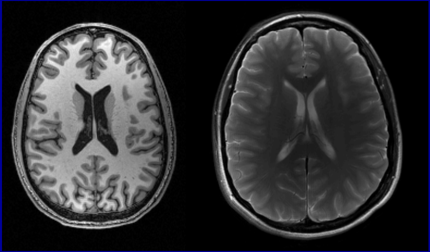


Axial brain images show different soft tissue contrast.

- Magnetic resonance imaging (MRI) provides excellent soft tissue contrast.
- Densities and magnetic properties of particles tell us a lot about organ structure and composition.
- Dynamic/functional MRI time series capture organ function.

What is MRI?

MRI Scanner

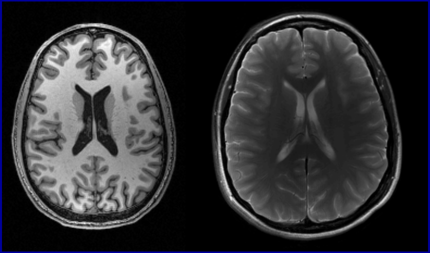


Axial brain images show different soft tissue contrast.

- Magnetic resonance imaging (MRI) provides excellent soft tissue contrast.
- Densities and magnetic properties of particles tell us a lot about organ structure and composition.
- Dynamic/functional MRI time series capture organ function.

What is MRI?

MRI Scanner

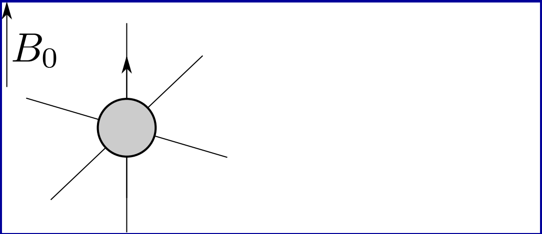


Axial brain images show different soft tissue contrast.

- Magnetic resonance imaging (MRI) provides excellent soft tissue contrast.
- Densities and magnetic properties of particles tell us a lot about organ structure and composition.
- Dynamic/functional MRI time series capture organ function.

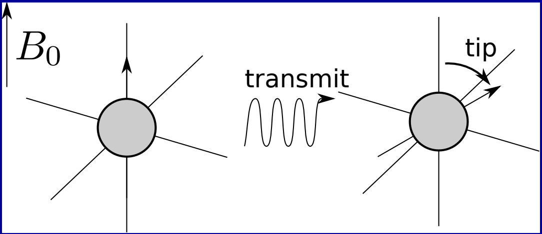
What does MRI measure?

- Certain particles have magnetic moments (spins), which tend to align with a strong external magnetic field B_0.
- When excited by a radiofrequency (RF) pulse, these spins tip away from the main field.
- They precess at the Larmor frequency γB_0, proportional to the main field, as they return to equilibrium (relax).
- These spins induce an emf in a nearby receiver loop coil.



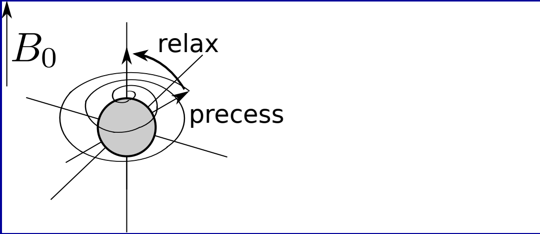
What does MRI measure?

- Certain particles have magnetic moments (spins), which tend to align with a strong external magnetic field B_0.
- When excited by a radiofrequency (RF) pulse, these spins tip away from the main field.
- They precess at the Larmor frequency γB_0, proportional to the main field, as they return to equilibrium (relax).
- These spins induce an emf in a nearby receiver loop coil.



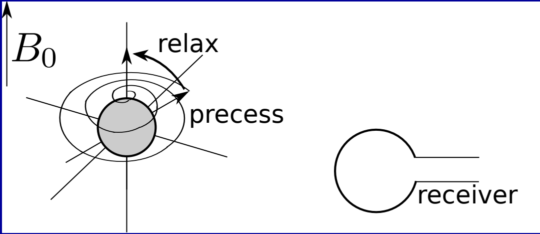
What does MRI measure?

- Certain particles have magnetic moments (spins), which tend to align with a strong external magnetic field B_0.
- When excited by a radiofrequency (RF) pulse, these spins tip away from the main field.
- They precess at the Larmor frequency γB_0, proportional to the main field, as they return to equilibrium (relax).
- These spins induce an emf in a nearby receiver loop coil.



What does MRI measure?

- Certain particles have magnetic moments (spins), which tend to align with a strong external magnetic field B_0.
- When excited by a radiofrequency (RF) pulse, these spins tip away from the main field.
- They precess at the Larmor frequency γB_0, proportional to the main field, as they return to equilibrium (relax).
- These spins induce an emf in a nearby receiver loop coil.



Generating an MR image

- The received signal superimposes contributions from all the precessing spins in the excited slice S:

$$s(t) = e^{-j\gamma B_0 t} \iint_S m(x,y)\, dx\, dy.$$

- Gradient coils control additional spatially varying magnetic fields G_x, G_y to acquire images: $B_z = B_0 + G_x x + G_y y$.
- The received signal is called k-space:

$$e^{j\gamma B_0 t} s(t) = \iint_S m(x,y) e^{-j\gamma\left(\int_0^t G_x(\tau)x + G_y(\tau)y\, d\tau\right)}\, dx\, dy.$$

- By adjusting the gradients, we effectively sample k-space.
- Applying z-gradient G_z during excitation selects a slice.

Generating an MR image

- The received signal superimposes contributions from all the precessing spins in the excited slice S:

$$s(t) = e^{-j\gamma B_0 t} \iint_S m(x, y)\, dx\, dy.$$

- Gradient coils control additional spatially varying magnetic fields G_x, G_y to acquire images: $B_z = B_0 + G_x x + G_y y$.
- The received signal is called k-space:

$$e^{j\gamma B_0 t} s(t) = \iint_S m(x, y) e^{-j\gamma(\int_0^t G_x(\tau)x + G_y(\tau)y\, d\tau)}\, dx\, dy.$$

- By adjusting the gradients, we effectively sample k-space.
- Applying z-gradient G_z during excitation selects a slice.

Generating an MR image

- The received signal superimposes contributions from all the precessing spins in the excited slice S:

$$s(t) = e^{-j\gamma B_0 t} \iint_S m(x,y)\, dx\, dy.$$

- Gradient coils control additional spatially varying magnetic fields G_x, G_y to acquire images: $B_z = B_0 + G_x x + G_y y$.
- The received signal is called k-space:

$$e^{j\gamma B_0 t} s(t) = \iint_S m(x,y) e^{-j\gamma\left(\int_0^t G_x(\tau)x + G_y(\tau)y\, d\tau\right)}\, dx\, dy.$$

- By adjusting the gradients, we effectively sample k-space.
- Applying z-gradient G_z during excitation selects a slice.

Generating an MR image

- The received signal superimposes contributions from all the precessing spins in the excited slice S:

$$s(t) = e^{-j\gamma B_0 t} \iint_S m(x,y)\, dx\, dy.$$

- Gradient coils control additional spatially varying magnetic fields G_x, G_y to acquire images: $B_z = B_0 + G_x x + G_y y$.
- The received signal is called k-space:

$$e^{j\gamma B_0 t} s(t) = \iint_S m(x,y) e^{-j\gamma\left(\int_0^t G_x(\tau)x + G_y(\tau)y\, d\tau\right)}\, dx\, dy.$$

- By adjusting the gradients, we effectively sample k-space.
- Applying z-gradient G_z during excitation selects a slice.

Generating an MR image

- The received signal superimposes contributions from all the precessing spins in the excited slice S:

$$s(t) = e^{-j\gamma B_0 t} \iint_S m(x,y)\, dx\, dy.$$

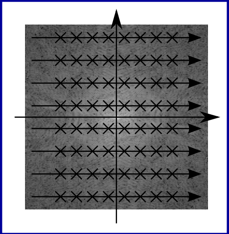
- Gradient coils control additional spatially varying magnetic fields G_x, G_y to acquire images: $B_z = B_0 + G_x x + G_y y$.
- The received signal is called k-space:

$$e^{j\gamma B_0 t} s(t) = \iint_S m(x,y) e^{-j\gamma \left(\int_0^t G_x(\tau) x + G_y(\tau) y \, d\tau\right)}\, dx\, dy.$$

- By adjusting the gradients, we effectively sample k-space.
- Applying z-gradient G_z during excitation selects a slice.

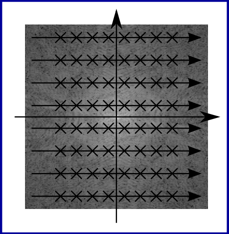
Accelerating MR imaging

- The sampling process is sequential, relatively slow.
- We must wait between excitations for the signal to relax.
- To accelerate imaging:
  - increase sample spacing, causing aliasing,
  - reduce the sampling extent, decreasing resolution,
  - sample half of k-space, assuming a real-valued image, or
  - traverse more of k-space during each shot



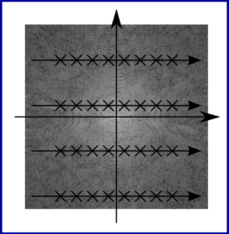
Accelerating MR imaging

- The sampling process is sequential, relatively slow.
- We must wait between excitations for the signal to relax.
- To accelerate imaging:
  - increase sample spacing, causing aliasing,
  - reduce the sampling extent, decreasing resolution,
  - sample half of k-space, assuming a real-valued image, or
  - traverse more of k-space during each shot



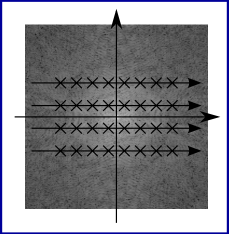
Accelerating MR imaging

- The sampling process is sequential, relatively slow.
- We must wait between excitations for the signal to relax.
- To accelerate imaging:
  - increase sample spacing, causing aliasing,
  - reduce the sampling extent, decreasing resolution,
  - sample half of k-space, assuming a real-valued image, or
  - traverse more of k-space during each shot



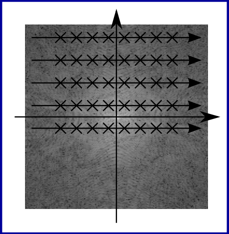
Accelerating MR imaging

- The sampling process is sequential, relatively slow.
- We must wait between excitations for the signal to relax.
- To accelerate imaging:
  - increase sample spacing, causing aliasing,
  - reduce the sampling extent, decreasing resolution,
  - sample half of k-space, assuming a real-valued image, or
  - traverse more of k-space during each shot



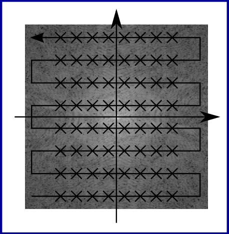
Accelerating MR imaging

- The sampling process is sequential, relatively slow.
- We must wait between excitations for the signal to relax.
- To accelerate imaging:
  - increase sample spacing, causing aliasing,
  - reduce the sampling extent, decreasing resolution,
  - sample half of k-space, assuming a real-valued image, or
  - traverse more of k-space during each shot



Accelerating MR imaging

- The sampling process is sequential, relatively slow.
- We must wait between excitations for the signal to relax.
- To accelerate imaging:
  - increase sample spacing, causing aliasing,
  - reduce the sampling extent, decreasing resolution,
  - sample half of k-space, assuming a real-valued image, or
  - traverse more of k-space during each shot



Contrast and relaxation

- The signal model assumed $m(x, y)$ is constant over time.
- However, the spin magnitude varies over time due to relaxation:
  - longitudinal component recovers ($\propto 1 - e^{-t/T_1}$), and
  - transverse component decays ($\propto e^{-t/T_2}$)
- Signal dephasing due to local field inhomogeneity also manifests as transverse relaxation (time constant T_2'):
  - total transverse relaxation rate $1/T_2^* = 1/T_2 + 1/T_2'$
- Repetition time T_R is time between excitations; shorter T_R $\Rightarrow$ greater T_1 contrast.
- Echo time T_E is time between excitation and sampling center of k-space; longer $T_E \Rightarrow$ greater T_2^* contrast.

Contrast and relaxation

- The signal model assumed $m(x, y)$ is constant over time.
- However, the spin magnitude varies over time due to relaxation:
  - longitudinal component recovers ($\propto 1 - e^{-t/T_1}$), and
  - transverse component decays ($\propto e^{-t/T_2}$)
- Signal dephasing due to local field inhomogeneity also manifests as transverse relaxation (time constant T_2'):
  - total transverse relaxation rate $1/T_2^* = 1/T_2 + 1/T_2'$
- Repetition time T_R is time between excitations; shorter T_R $\Rightarrow$ greater T_1 contrast.
- Echo time T_E is time between excitation and sampling center of k-space; longer $T_E \Rightarrow$ greater T_2^* contrast.

Contrast and relaxation

- The signal model assumed $m(x, y)$ is constant over time.
- However, the spin magnitude varies over time due to relaxation:
  - longitudinal component recovers ($\propto 1 - e^{-t/T_1}$), and
  - transverse component decays ($\propto e^{-t/T_2}$)
- Signal dephasing due to local field inhomogeneity also manifests as transverse relaxation (time constant T_2'):
  - total transverse relaxation rate $1/T_2^* = 1/T_2 + 1/T_2'$
- Repetition time T_R is time between excitations; shorter T_R $\Rightarrow$ greater T_1 contrast.
- Echo time T_E is time between excitation and sampling center of k-space; longer $T_E \Rightarrow$ greater T_2^* contrast.

Contrast and relaxation

- The signal model assumed $m(x, y)$ is constant over time.
- However, the spin magnitude varies over time due to relaxation:
  - longitudinal component recovers ($\propto 1 - e^{-t/T_1}$), and
  - transverse component decays ($\propto e^{-t/T_2}$)
- Signal dephasing due to local field inhomogeneity also manifests as transverse relaxation (time constant T_2'):
  - total transverse relaxation rate $1/T_2^* = 1/T_2 + 1/T_2'$
- Repetition time T_R is time between excitations; shorter T_R $\Rightarrow$ greater T_1 contrast.
- Echo time T_E is time between excitation and sampling center of k-space; longer $T_E \Rightarrow$ greater T_2^* contrast.

Contrast and relaxation

- The signal model assumed $m(x, y)$ is constant over time.
- However, the spin magnitude varies over time due to relaxation:
  - longitudinal component recovers ($\propto 1 - e^{-t/T_1}$), and
  - transverse component decays ($\propto e^{-t/T_2}$)
- Signal dephasing due to local field inhomogeneity also manifests as transverse relaxation (time constant T_2'):
  - total transverse relaxation rate $1/T_2^* = 1/T_2 + 1/T_2'$
- Repetition time T_R is time between excitations; shorter T_R $\Rightarrow$ greater T_1 contrast.
- Echo time T_E is time between excitation and sampling center of k-space; longer $T_E \Rightarrow$ greater T_2^* contrast.

Outline

- Magnetic Resonance Imaging
- Functional MRI
- Head Motion in fMRI
- Proposed Method for Prospective Correction
- Real Time Implementation
- Simulation Results
- Conclusions and Future Work
- Summary

The BOLD effect for fMRI

- Neuronal impulses are too short, weak to register using conventional MRI.
- Instead, indirectly measure activity by tracking blood oxygen metabolism.
- Hemoglobin has two states:
  - with bound oxygen is diamagnetic, and
  - without bound oxygen is paramagnetic ($\uparrow 1/T_2^*$)
- BOLD contrast results from increased blood flow, and increased metabolism:
  - $\uparrow$ blood flow $\Rightarrow \downarrow$ deoxygenated blood $\Rightarrow \downarrow 1/T_2^*$, and
  - $\uparrow$ metabolism $\Rightarrow \uparrow$ deoxygenated blood $\Rightarrow \uparrow 1/T_2^*$

The BOLD effect for fMRI

- Neuronal impulses are too short, weak to register using conventional MRI.
- Instead, indirectly measure activity by tracking blood oxygen metabolism.
- Hemoglobin has two states:
  - with bound oxygen is diamagnetic, and
  - without bound oxygen is paramagnetic ($\uparrow 1/T_2^*$)
- BOLD contrast results from increased blood flow, and increased metabolism:
  - $\uparrow$ blood flow $\Rightarrow \downarrow$ deoxygenated blood $\Rightarrow \downarrow 1/T_2^*$, and
  - $\uparrow$ metabolism $\Rightarrow \uparrow$ deoxygenated blood $\Rightarrow \uparrow 1/T_2^*$

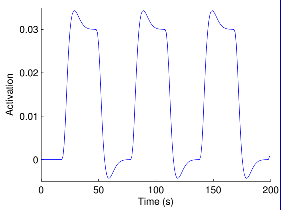
The BOLD effect for fMRI

- Neuronal impulses are too short, weak to register using conventional MRI.
- Instead, indirectly measure activity by tracking blood oxygen metabolism.
- Hemoglobin has two states:
  - with bound oxygen is diamagnetic, and
  - without bound oxygen is paramagnetic ($\uparrow 1/T_2^*$)
- BOLD contrast results from increased blood flow, and increased metabolism:
  - $\uparrow$ blood flow $\Rightarrow \downarrow$ deoxygenated blood $\Rightarrow \downarrow 1/T_2^*$, and
  - $\uparrow$ metabolism $\Rightarrow \uparrow$ deoxygenated blood $\Rightarrow \uparrow 1/T_2^*$

The BOLD effect for fMRI

- Neuronal impulses are too short, weak to register using conventional MRI.
- Instead, indirectly measure activity by tracking blood oxygen metabolism.
- Hemoglobin has two states:
  - with bound oxygen is diamagnetic, and
  - without bound oxygen is paramagnetic ($\uparrow 1/T_2^*$)
- BOLD contrast results from increased blood flow, and increased metabolism:
  - $\uparrow$ blood flow $\Rightarrow \downarrow$ deoxygenated blood $\Rightarrow \downarrow 1/T_2^*$, and
  - $\uparrow$ metabolism $\Rightarrow \uparrow$ deoxygenated blood $\Rightarrow \uparrow 1/T_2^*$

Modeling the BOLD signal

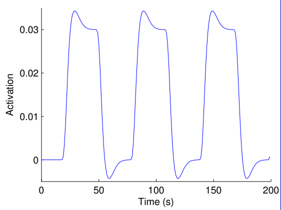


Activations from block design convolved with canonical hrf from SPM8[1]

- The functional MRI process is modeled as an LTI system with an impulse response known as the hemodynamic response function (hrf).
- In reality, the hrf varies spatially and over time.
- It changes from subject-to-subject and scan-to-scan.
- Software like SPM8[1] use basis functions for the hrf.

[1] `http://www.fil.ion.ucl.ac.uk/spm/`

Modeling the BOLD signal

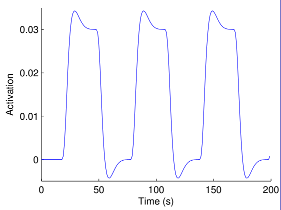


Activations from block design convolved with canonical hrf from SPM8[1]

- The functional MRI process is modeled as an LTI system with an impulse response known as the hemodynamic response function (hrf).
- In reality, the hrf varies spatially and over time.
- It changes from subject-to-subject and scan-to-scan.
- Software like SPM8[1] use basis functions for the hrf.

[1] http://www.fil.ion.ucl.ac.uk/spm/

Modeling the BOLD signal

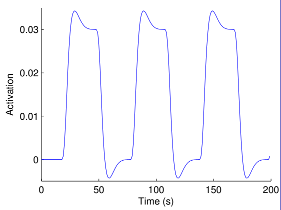


Activations from block design convolved with canonical hrf from SPM8[1]

- The functional MRI process is modeled as an LTI system with an impulse response known as the hemodynamic response function (hrf).
- In reality, the hrf varies spatially and over time.
- It changes from subject-to-subject and scan-to-scan.
- Software like SPM8[1] use basis functions for the hrf.

[1] http://www.fil.ion.ucl.ac.uk/spm/

Modeling the BOLD signal



Activations from block design convolved with canonical hrf from SPM8[1]

- The functional MRI process is modeled as an LTI system with an impulse response known as the hemodynamic response function (hrf).
- In reality, the hrf varies spatially and over time.
- It changes from subject-to-subject and scan-to-scan.
- Software like SPM8[1] use basis functions for the hrf.

[1] http://www.fil.ion.ucl.ac.uk/spm/

Functional MRI analysis

- To identify activated brain regions, we correlate the ideal time series activations against the time series for each voxel in the brain.
- This analysis yields the general linear model (GLM):

$$\underbrace{\boldsymbol{Y}}_{\text{data}} = \underbrace{\begin{bmatrix} \boldsymbol{G} & \mathbf{1} \end{bmatrix}}_{\text{regressors}} \underbrace{\boldsymbol{\beta}}_{\text{weights}} + \underbrace{\boldsymbol{\varepsilon}}_{\text{error}}.$$

- Software packages like SPM8 implement this analysis and provide visualization tools.

Functional MRI analysis

- To identify activated brain regions, we correlate the ideal time series activations against the time series for each voxel in the brain.
- This analysis yields the general linear model (GLM):

$$\underbrace{Y}_{\text{data}} = \underbrace{\begin{bmatrix} G & 1 \end{bmatrix}}_{\text{regressors}} \underbrace{\beta}_{\text{weights}} + \underbrace{\varepsilon}_{\text{error}}.$$

- Software packages like SPM8 implement this analysis and provide visualization tools.

Functional MRI analysis

- To identify activated brain regions, we correlate the ideal time series activations against the time series for each voxel in the brain.
- This analysis yields the general linear model (GLM):

$$\underbrace{\boldsymbol{Y}}_{\text{data}} = \underbrace{\begin{bmatrix} \boldsymbol{G} & \mathbf{1} \end{bmatrix}}_{\text{regressors}} \underbrace{\boldsymbol{\beta}}_{\text{weights}} + \underbrace{\varepsilon}_{\text{error}}.$$

- Software packages like SPM8 implement this analysis and provide visualization tools.

Pre-processing fMRI data

Conventionally, data are pre-processed in the following ways:

1. Slice-by-slice acquisitions are time-shifted to account for different slice timings.
2. Head motion is estimated and used to register volumes to the first/middle volume of the time series.
3. Time series can be realigned to a separate reference volume for visualization purposes.
4. Volumes are normalized to a fixed coordinate system (e.g., Tailarach) in group studies.
5. Data can be smoothed/blurred using a Gaussian kernel to reduce noise, cross-subject variance.

Pre-processing fMRI data

Conventionally, data are pre-processed in the following ways:

1. Slice-by-slice acquisitions are time-shifted to account for different slice timings.
2. Head motion is estimated and used to register volumes to the first/middle volume of the time series.
3. Time series can be realigned to a separate reference volume for visualization purposes.
4. Volumes are normalized to a fixed coordinate system (e.g., Tailarach) in group studies.
5. Data can be smoothed/blurred using a Gaussian kernel to reduce noise, cross-subject variance.

Pre-processing fMRI data

Conventionally, data are pre-processed in the following ways:

1. Slice-by-slice acquisitions are time-shifted to account for different slice timings.
2. Head motion is estimated and used to register volumes to the first/middle volume of the time series.
3. Time series can be realigned to a separate reference volume for visualization purposes.
4. Volumes are normalized to a fixed coordinate system (e.g., Tailarach) in group studies.
5. Data can be smoothed/blurred using a Gaussian kernel to reduce noise, cross-subject variance.

Pre-processing fMRI data

Conventionally, data are pre-processed in the following ways:

1. Slice-by-slice acquisitions are time-shifted to account for different slice timings.
2. Head motion is estimated and used to register volumes to the first/middle volume of the time series.
3. Time series can be realigned to a separate reference volume for visualization purposes.
4. Volumes are normalized to a fixed coordinate system (e.g., Tailarach) in group studies.
5. Data can be smoothed/blurred using a Gaussian kernel to reduce noise, cross-subject variance.

Pre-processing fMRI data

Conventionally, data are pre-processed in the following ways:

1. Slice-by-slice acquisitions are time-shifted to account for different slice timings.
2. Head motion is estimated and used to register volumes to the first/middle volume of the time series.
3. Time series can be realigned to a separate reference volume for visualization purposes.
4. Volumes are normalized to a fixed coordinate system (e.g., Tailarach) in group studies.
5. Data can be smoothed/blurred using a Gaussian kernel to reduce noise, cross-subject variance.

Pre-processing fMRI data

Conventionally, data are pre-processed in the following ways:

1. Slice-by-slice acquisitions are time-shifted to account for different slice timings.
2. Head motion is estimated and used to register volumes to the first/middle volume of the time series.
3. Time series can be realigned to a separate reference volume for visualization purposes.
4. Volumes are normalized to a fixed coordinate system (e.g., Tailarach) in group studies.
5. Data can be smoothed/blurred using a Gaussian kernel to reduce noise, cross-subject variance.

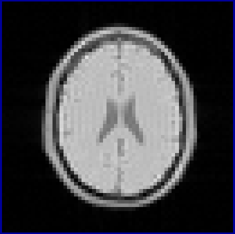
Our single-subject simulations require only step #2.

Outline

- Magnetic Resonance Imaging
- Functional MRI
- Head Motion in fMRI
- Proposed Method for Prospective Correction
- Real Time Implementation
- Simulation Results
- Conclusions and Future Work
- Summary

Head motion in fMRI

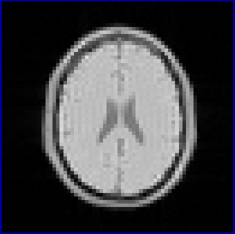
- Head motion causes several problems in brain imaging:
    - Misaligned time series have inconsistent brain coordinates.
    - Motion distorts the main field and alters the dephasing rate.
    - Nonuniform excitation from inter-slice motion introduces “spin history” effects.



Simulated brain with motion.

Head motion in fMRI

- Head motion causes several problems in brain imaging:
  - Misaligned time series have inconsistent brain coordinates.
  - Motion distorts the main field and alters the dephasing rate.
  - Nonuniform excitation from inter-slice motion introduces "spin history" effects.



Simulated brain with motion.

Head motion in fMRI

- Head motion causes several problems in brain imaging:
  - Misaligned time series have inconsistent brain coordinates.
  - Motion distorts the main field and alters the dephasing rate.
  - Nonuniform excitation from inter-slice motion introduces "spin history" effects.

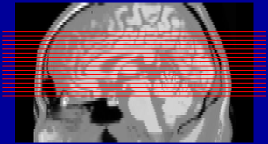
Figure 9 in [1]
URL: `http://dx.doi.org/10.1002/mrm.24314`

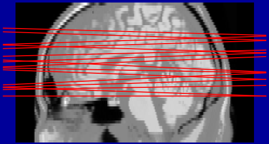
Head motion causes field map distortions[1].

[1] J Maclaren et al., Mag. Res. Med., 69(3), 2013.

Head motion in fMRI

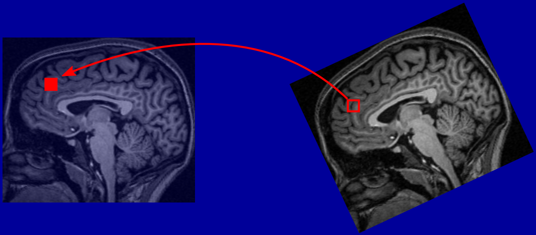
- Head motion causes several problems in brain imaging:
  - Misaligned time series have inconsistent brain coordinates.
  - Motion distorts the main field and alters the dephasing rate.
  - Nonuniform excitation from inter-slice motion introduces "spin history" effects.





Slice selection without (left) and with (right) motion.

Head motion correction



Retrospective image registration/interpolation.

- Retrospective correction (post-scan):
  - registration: spatial interpolation of reconstructed images
  - motion estimates as nuisance regressors in the GLM
- Prospective correction (during the scan):
  - slice prescription, k-space trajectory adjusted between frames or slices
  - corrupted data re-scanned (rare in fMRI)

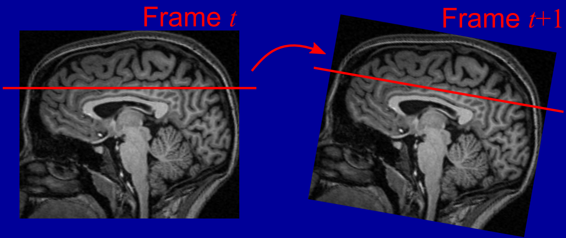
Head motion correction

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{G} & \mathbf{1} \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} -\widehat{\boldsymbol{\alpha}}^{0}- \\ \vdots \\ -\widehat{\boldsymbol{\alpha}}^{N_F-1}- \end{bmatrix} \boldsymbol{\beta}_{\text{motion}} + \boldsymbol{\varepsilon}.$$

Nuisance regressors in the GLM.

- Retrospective correction (post-scan):
  - registration: spatial interpolation of reconstructed images
  - motion estimates as nuisance regressors in the GLM
- Prospective correction (during the scan):
  - slice prescription, k-space trajectory adjusted between frames or slices
  - corrupted data re-scanned (rare in fMRI)

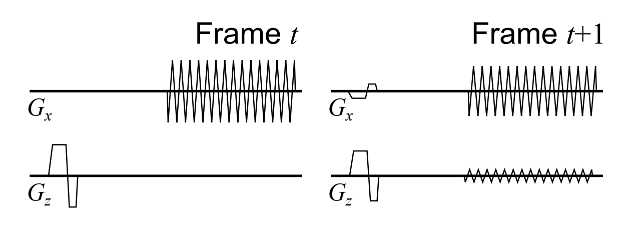
Head motion correction



Gradients, slice prescriptions adjusted.

- Retrospective correction (post-scan):
  - registration: spatial interpolation of reconstructed images
  - motion estimates as nuisance regressors in the GLM
- Prospective correction (during the scan):
  - slice prescription, k-space trajectory adjusted between frames or slices
  - corrupted data re-scanned (rare in fMRI)

Head motion correction



Gradients, slice prescriptions adjusted.

- Retrospective correction (post-scan):
  - registration: spatial interpolation of reconstructed images
  - motion estimates as nuisance regressors in the GLM
- Prospective correction (during the scan):
  - slice prescription, k-space trajectory adjusted between frames or slices
  - corrupted data re-scanned (rare in fMRI)

Prospective correction methods

- Retrospective methods are suboptimal.
  - registration: interpolation smears activations across voxels
  - regression: activations mistaken for task-correlated motion
- Prospective motion estimation methods include:
  - external tracking: markers, cameras, etc.
  - navigational techniques using MRI data

Prospective correction methods

- Retrospective methods are suboptimal.
  - registration: interpolation smears activations across voxels
  - regression: activations mistaken for task-correlated motion
- Prospective motion estimation methods include:
  - external tracking: markers, cameras, etc.
  - navigational techniques using MRI data

Prospective correction methods

- Retrospective methods are suboptimal.
  - registration: interpolation smears activations across voxels
  - regression: activations mistaken for task-correlated motion
- Prospective motion estimation methods include:
  - external tracking: markers, cameras, etc.
  - navigational techniques using MRI data

Figure 2b in [1]
URL: `http://dx.doi.org/10.1002/mrm.24314`

Methods for prospective motion estimation[1].

[1] J Maclaren et al., Mag. Res. Med., 69(3), 2013.

Navigation

- Image-based navigators align the MR time series images.
- Fast k-space navigators have static contrast, versus time-varying BOLD contrast found in fMRI data.
- With multi-channel receivers, relative intensity changes in received signals (called FID signals) can signify motion.

Figure 2a in [1]
URL: `http://dx.doi.org/10.1002/mrm.24314`

K-space, image-space, and FID navigator methods[1].

[1] J Maclaren et al., Mag. Res. Med., 69(3), 2013.

Navigation

- Image-based navigators align the MR time series images.
- Fast k-space navigators have static contrast, versus time-varying BOLD contrast found in fMRI data.
- With multi-channel receivers, relative intensity changes in received signals (called FID signals) can signify motion.

Figure 2a in [1]
URL: `http://dx.doi.org/10.1002/mrm.24314`

K-space, image-space, and FID navigator methods[1].

[1] J Maclaren et al., Mag. Res. Med., 69(3), 2013.

Navigation

- Image-based navigators align the MR time series images.
- Fast k-space navigators have static contrast, versus time-varying BOLD contrast found in fMRI data.
- With multi-channel receivers, relative intensity changes in received signals (called FID signals) can signify motion.

Figure 2a in [1]
URL: `http://dx.doi.org/10.1002/mrm.24314`

K-space, image-space, and FID navigator methods[1].

[1] J Maclaren et al., Mag. Res. Med., 69(3), 2013.

Outline

- Magnetic Resonance Imaging
- Functional MRI
- Head Motion in fMRI
- Proposed Method for Prospective Correction
- Real Time Implementation
- Simulation Results
- Conclusions and Future Work
- Summary

Related work

- Image-space navigators were used with a linear motion model[1].
- Two-dimensional k-space navigators were used with an extended Kalman filter[2].
- Sparse residual models were employed in retrospective joint reconstruction/registration methods[3,4].

[1] S Thesen et al., Mag. Res. Med., 44(3), 2000.

Related work

- Image-space navigators were used with a linear motion model[1].
- Two-dimensional k-space navigators were used with an extended Kalman filter[2].
- Sparse residual models were employed in retrospective joint reconstruction/registration methods[3,4].

[1] S Thesen et al., Mag. Res. Med., 44(3), 2000.
[2] N White et al., Mag. Res. Med., 63(1), 2010.

Related work

- Image-space navigators were used with a linear motion model[1].
- Two-dimensional k-space navigators were used with an extended Kalman filter[2].
- Sparse residual models were employed in retrospective joint reconstruction/registration methods[3,4].

[1] S Thesen et al., Mag. Res. Med., 44(3), 2000.
[2] N White et al., Mag. Res. Med., 63(1), 2010.
[3] H Jung et al., Mag. Res. Med., 61(1), 2009.
[4] MS Asif et al., Mag. Res. Med., in press.

Related work

- Image-space navigators were used with a linear motion model[1].
- Two-dimensional k-space navigators were used with an extended Kalman filter[2].
- Sparse residual models were employed in retrospective joint reconstruction/registration methods[3,4].

- I propose using image-space navigators and combining a sparse residual model with Kalman-like filtering[5].

[1] S Thesen et al., Mag. Res. Med., 44(3), 2000.
[2] N White et al., Mag. Res. Med., 63(1), 2010.
[3] H Jung et al., Mag. Res. Med., 61(1), 2009.
[4] MS Asif et al., Mag. Res. Med., in press.
[5] DSW et al., Proc. SPIE Wavelets and Sparsity XV, in press.

Coordinate systems

Consider images $\boldsymbol{x}^t|_{t=0,\ldots,N_F-1}$ in 3 coordinate systems:

1. $\boldsymbol{x}_0^t$ has the same physical coordinate system for all t.

2. $\boldsymbol{x}_{\text{reg}}^t = \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^t)^{(-1)})\boldsymbol{x}_0^t$ is registered to $\boldsymbol{x}^0$ with estimate $\widehat{\boldsymbol{\alpha}}^t$.

3. Measure $\boldsymbol{x}_{\text{meas}}^t = \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)})\boldsymbol{x}_0^t$ prospectively corrected using the previous frame's motion.

$t = 0$

Fixed Registered Measured

$t = 1$

$t = 2$

Coordinate systems

Consider images $\boldsymbol{x}^t|_{t=0,\ldots,N_F-1}$ in 3 coordinate systems:

1. $\boldsymbol{x}_0^t$ has the same physical coordinate system for all t.
   - true motion $\boldsymbol{\alpha}^t$: rigid head motion from frame 0 to t
2. $\boldsymbol{x}_{\text{reg}}^t = \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^t)^{(-1)})\boldsymbol{x}_0^t$ is registered to $\boldsymbol{x}^0$ with estimate $\widehat{\boldsymbol{\alpha}}^t$.
3. Measure $\boldsymbol{x}_{\text{meas}}^t = \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)})\boldsymbol{x}_0^t$ prospectively corrected using the previous frame's motion.

$t = 0$

Fixed Registered Measured

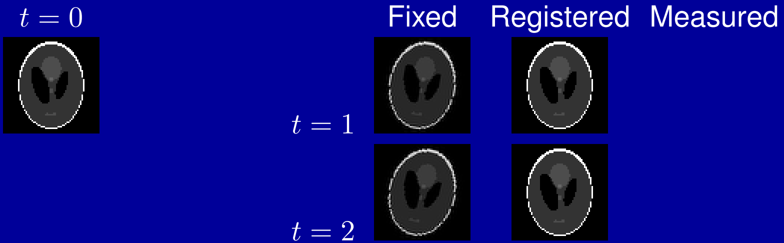
$t = 1$

$t = 2$

Coordinate systems

Consider images $\boldsymbol{x}^t|_{t=0,\ldots,N_F-1}$ in 3 coordinate systems:

1. $\boldsymbol{x}_0^t$ has the same physical coordinate system for all t.
   - true motion $\boldsymbol{\alpha}^t$: rigid head motion from frame 0 to t
2. $\boldsymbol{x}_{\text{reg}}^t = \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^t)^{(-1)})\boldsymbol{x}_0^t$ is registered to $\boldsymbol{x}^0$ with estimate $\widehat{\boldsymbol{\alpha}}^t$.
   - $\boldsymbol{\alpha}^{(-1)}$ is the inverse motion of $\boldsymbol{\alpha}$
   - $\boldsymbol{T}(\boldsymbol{\alpha}_n, \ldots, \boldsymbol{\alpha}_1)\boldsymbol{x}$ transforms $\boldsymbol{x}$ by motions $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$
3. Measure $\boldsymbol{x}_{\text{meas}}^t = \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)})\boldsymbol{x}_0^t$ prospectively corrected using the previous frame's motion.



Coordinate systems

Consider images $\boldsymbol{x}^t|_{t=0,\ldots,N_F-1}$ in 3 coordinate systems:

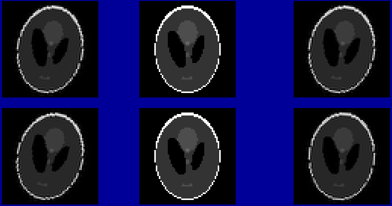
1. $\boldsymbol{x}_0^t$ has the same physical coordinate system for all t.
   - true motion $\boldsymbol{\alpha}^t$: rigid head motion from frame 0 to t
2. $\boldsymbol{x}_{\text{reg}}^t = \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^t)^{(-1)})\boldsymbol{x}_0^t$ is registered to $\boldsymbol{x}^0$ with estimate $\widehat{\boldsymbol{\alpha}}^t$.
   - $\boldsymbol{\alpha}^{(-1)}$ is the inverse motion of $\boldsymbol{\alpha}$
   - $\boldsymbol{T}(\boldsymbol{\alpha}_n, \ldots, \boldsymbol{\alpha}_1)\boldsymbol{x}$ transforms $\boldsymbol{x}$ by motions $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$
3. Measure $\boldsymbol{x}_{\text{meas}}^t = \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)})\boldsymbol{x}_0^t$ prospectively corrected using the previous frame's motion.

$t = 0$

Fixed Registered Measured

$t = 1$

$t = 2$



Measurement model

- Sample k-space in measurement coordinates:

$$\boldsymbol{d}^t = \boldsymbol{\mathcal{F}} \boldsymbol{x}^t_{\text{meas}} + \boldsymbol{n}^t,$$

where $\mathcal{F}$ is the Discrete Fourier Transform (DFT), and $\boldsymbol{n}^t$ is iid zero-mean complex Gaussian with variance σ^2.

- To estimate $\boldsymbol{\alpha}^t$, relate $\boldsymbol{x}^t_{\text{meas}}$ to registered coordinates:

- Define the residual image $\boldsymbol{s}^t$ in measurement coordinates:

- Model $\boldsymbol{s}^t$ as sparse, which is equivalent to using the SAD registration cost function:

Measurement model

- Sample k-space in measurement coordinates:
$$\boldsymbol{d}^t = \boldsymbol{\mathcal{F}} \boldsymbol{x}^t_{\text{meas}} + \boldsymbol{n}^t,$$
where $\boldsymbol{\mathcal{F}}$ is the Discrete Fourier Transform (DFT), and
$\boldsymbol{n}^t$ is iid zero-mean complex Gaussian
with variance σ^2.
- To estimate $\boldsymbol{\alpha}^t$, relate $\boldsymbol{x}^t_{\text{meas}}$ to registered coordinates:
$$\boldsymbol{d}^t = \boldsymbol{\mathcal{F}} \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t) \boldsymbol{x}^t_{\text{reg}} + \boldsymbol{n}^t.$$
- Define the residual image $\boldsymbol{s}^t$ in measurement coordinates:
- Model $\boldsymbol{s}^t$ as sparse, which is equivalent to using the SAD registration cost function:

Measurement model

- Sample k-space in measurement coordinates:

$$\boldsymbol{d}^t = \boldsymbol{\mathcal{F}}\boldsymbol{x}^t_{\text{meas}} + \boldsymbol{n}^t,$$

where $\boldsymbol{\mathcal{F}}$ is the Discrete Fourier Transform (DFT), and $\boldsymbol{n}^t$ is iid zero-mean complex Gaussian with variance σ^2.

- To estimate $\boldsymbol{\alpha}^t$, relate $\boldsymbol{x}^t_{\text{meas}}$ to registered coordinates:

$$\boldsymbol{d}^t = \boldsymbol{\mathcal{F}}\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)\boldsymbol{x}^t_{\text{reg}} + \boldsymbol{n}^t.$$

- Define the residual image $\boldsymbol{s}^t$ in measurement coordinates:

$$\boldsymbol{s}^t = \boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)\boldsymbol{x}^{t-1}_{\text{reg}}.$$

- Model $\boldsymbol{s}^t$ as sparse, which is equivalent to using the SAD registration cost function:

Measurement model

- Sample k-space in measurement coordinates:

$$\boldsymbol{d}^t = \boldsymbol{\mathcal{F}} \boldsymbol{x}^t_{\text{meas}} + \boldsymbol{n}^t,$$

where $\boldsymbol{\mathcal{F}}$ is the Discrete Fourier Transform (DFT), and
$\boldsymbol{n}^t$ is iid zero-mean complex Gaussian
with variance σ^2.

- To estimate $\boldsymbol{\alpha}^t$, relate $\boldsymbol{x}^t_{\text{meas}}$ to registered coordinates:

$$\boldsymbol{d}^t = \boldsymbol{\mathcal{F}} \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t) \boldsymbol{x}^t_{\text{reg}} + \boldsymbol{n}^t.$$

- Define the residual image $\boldsymbol{s}^t$ in measurement coordinates:

$$\boldsymbol{s}^t = \boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t) \boldsymbol{x}^{t-1}_{\text{reg}}.$$

- Model $\boldsymbol{s}^t$ as sparse, which is equivalent to using the SAD registration cost function:

$$C(\boldsymbol{\alpha}) = \| \boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}) \boldsymbol{x}^{t-1}_{\text{reg}} \|_1.$$

Kalman-like motion filtering

- To enforce smoothness, we model motion $\{\boldsymbol{\alpha}^t\}$ as a first-order random walk:

$$\boldsymbol{\alpha}^t = \boldsymbol{\alpha}^{t-1} + \boldsymbol{a}^t,$$

  where innovations $\boldsymbol{a}^t$ are iid (over time) Normal$(\mathbf{0}, \boldsymbol{Q})$.
- Subtracting the sparse residual image component of the data yield Kalman filter-like measurements of the motion:
- Knowing $\boldsymbol{s}^t$ yields an extended Kalman filter:

Kalman-like motion filtering

- To enforce smoothness, we model motion $\{\boldsymbol{\alpha}^t\}$ as a first-order random walk:
$$\boldsymbol{\alpha}^t = \boldsymbol{\alpha}^{t-1} + \boldsymbol{a}^t,$$
where innovations $\boldsymbol{a}^t$ are iid (over time) Normal$(\mathbf{0}, \boldsymbol{Q})$.
- Subtracting the sparse residual image component of the data yield Kalman filter-like measurements of the motion:
$$\boldsymbol{d}^t - \boldsymbol{\mathcal{F}} \boldsymbol{s}^t = \boldsymbol{\mathcal{F}} \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t) \boldsymbol{x}_{\text{reg}}^{t-1} + \boldsymbol{n}^t.$$
- Knowing $\boldsymbol{s}^t$ yields an extended Kalman filter:

Kalman-like motion filtering

- To enforce smoothness, we model motion $\{\boldsymbol{\alpha}^t\}$ as a first-order random walk:
$$\boldsymbol{\alpha}^t = \boldsymbol{\alpha}^{t-1} + \boldsymbol{a}^t,$$
where innovations $\boldsymbol{a}^t$ are iid (over time) Normal$(\mathbf{0}, \boldsymbol{Q})$.
- Subtracting the sparse residual image component of the data yield Kalman filter-like measurements of the motion:
$$\boldsymbol{d}^t - \boldsymbol{\mathcal{F}} \boldsymbol{s}^t = \boldsymbol{\mathcal{F}} \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t) \boldsymbol{x}_{\text{reg}}^{t-1} + \boldsymbol{n}^t.$$
- Knowing $\boldsymbol{s}^t$ yields an extended Kalman filter:
$$\widehat{\boldsymbol{\alpha}}^t = \arg\min_{\boldsymbol{\alpha}} \tfrac{1}{2\sigma^2} \|\boldsymbol{\mathcal{F}} \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}) \boldsymbol{x}_{\text{reg}}^{t-1} - (\boldsymbol{d}^t - \boldsymbol{\mathcal{F}} \boldsymbol{s}^t)\|_2^2 + \tfrac{1}{2} \|\boldsymbol{\alpha} - \widehat{\boldsymbol{\alpha}}^{t-1}\|^2_{(\boldsymbol{P}_{\boldsymbol{\alpha}}^{t|t-1})^{-1}},$$
where $\boldsymbol{P}_{\boldsymbol{\alpha}}^{t|t-1}$ is the filter's prediction error covariance.

Joint registration and reconstruction

Jointly minimize $f(\boldsymbol{x}^t_{\text{meas}}, \boldsymbol{\alpha}^t)$ to estimate motion $\widehat{\boldsymbol{\alpha}}^t$:

$$
\begin{aligned}
f(\boldsymbol{x}^t_{\text{meas}}, \boldsymbol{\alpha}^t) = \tfrac{1}{2\sigma^2} \|\boldsymbol{\mathcal{F}} \boldsymbol{x}^t_{\text{meas}} - \boldsymbol{d}^t\|_2^2 \\
+ \lambda \|\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t) \boldsymbol{x}^{t-1}_{\text{reg}}\|_1 \\
+ \tfrac{1}{2} \|\boldsymbol{\alpha}^t - \widehat{\boldsymbol{\alpha}}^{t-1}\|^2_{(\boldsymbol{P}_{\boldsymbol{\alpha}}^{t|t-1})^{-1}}
\end{aligned}
$$

The objective function has three parts:

Joint registration and reconstruction

Jointly minimize $f(\boldsymbol{x}^t_{\text{meas}}, \boldsymbol{\alpha}^t)$ to estimate motion $\widehat{\boldsymbol{\alpha}}^t$:

$$
\begin{aligned}
f(\boldsymbol{x}^t_{\text{meas}}, \boldsymbol{\alpha}^t) = {} & \frac{1}{2\sigma^2} \|\boldsymbol{\mathcal{F}} \boldsymbol{x}^t_{\text{meas}} - \boldsymbol{d}^t\|_2^2 \\
& + \lambda \|\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t) \boldsymbol{x}^{t-1}_{\text{reg}}\|_1 \\
& + \frac{1}{2} \|\boldsymbol{\alpha}^t - \widehat{\boldsymbol{\alpha}}^{t-1}\|^2_{(\boldsymbol{P}^{t|t-1}_{\boldsymbol{\alpha}})^{-1}}
\end{aligned}
$$

The objective function has three parts:

1. the data fidelity term,
   - equivalent to $\frac{1}{2\sigma^2} \|\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{\mathcal{F}}^{-1} \boldsymbol{d}^t\|_2^2$ (with unitary DFT);

Joint registration and reconstruction

Jointly minimize $f(\boldsymbol{x}^t_{\text{meas}}, \boldsymbol{\alpha}^t)$ to estimate motion $\widehat{\boldsymbol{\alpha}}^t$:

$$\begin{aligned} f(\boldsymbol{x}^t_{\text{meas}}, \boldsymbol{\alpha}^t) = {} & \tfrac{1}{2\sigma^2} \|\boldsymbol{\mathcal{F}} \boldsymbol{x}^t_{\text{meas}} - \boldsymbol{d}^t\|_2^2 \\ & + \lambda \|\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t) \boldsymbol{x}^{t-1}_{\text{reg}}\|_1 \\ & + \tfrac{1}{2} \|\boldsymbol{\alpha}^t - \widehat{\boldsymbol{\alpha}}^{t-1}\|^2_{(\boldsymbol{P}^{t|t-1}_{\boldsymbol{\alpha}})^{-1}} \end{aligned}$$

The objective function has three parts:

1. the data fidelity term,
   - equivalent to $\frac{1}{2\sigma^2} \|\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{\mathcal{F}}^{-1} \boldsymbol{d}^t\|_2^2$ (with unitary DFT);
2. the registration cost function, which also regularizes $\boldsymbol{x}^t_{\text{meas}}$;

Joint registration and reconstruction

Jointly minimize $f(\boldsymbol{x}^t_{\text{meas}}, \boldsymbol{\alpha}^t)$ to estimate motion $\widehat{\boldsymbol{\alpha}}^t$:

$$
\begin{aligned}
f(\boldsymbol{x}^t_{\text{meas}}, \boldsymbol{\alpha}^t) = \tfrac{1}{2\sigma^2}\|\boldsymbol{\mathcal{F}}\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{d}^t\|_2^2 \\
+ \lambda\|\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)\boldsymbol{x}^{t-1}_{\text{reg}}\|_1 \\
+ \tfrac{1}{2}\|\boldsymbol{\alpha}^t - \widehat{\boldsymbol{\alpha}}^{t-1}\|^2_{(\boldsymbol{P}_{\boldsymbol{\alpha}}^{t|t-1})^{-1}}
\end{aligned}
$$

The objective function has three parts:

1. the data fidelity term,
   - equivalent to $\frac{1}{2\sigma^2}\|\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{\mathcal{F}}^{-1}\boldsymbol{d}^t\|_2^2$ (with unitary DFT);
2. the registration cost function, which also regularizes $\boldsymbol{x}^t_{\text{meas}}$;
3. the Kalman consistency term, promoting smoothness.

Outline

- Magnetic Resonance Imaging
- Functional MRI
- Head Motion in fMRI
- Proposed Method for Prospective Correction
- Real Time Implementation
- Simulation Results
- Conclusions and Future Work
- Summary

Variable splitting

- The residual image is a natural split variable:
$$\boldsymbol{s}^t = \boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}(\boldsymbol{x}^{t-1}_{\text{reg}}; (\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t).$$
- The augmented Lagrangian has scaled dual $\boldsymbol{u}$, penalty μ:
- The choice of penalty $\mu > 0$ controls the overall convergence rate, by trading off minimizing the objective function and satisfying the variable-split constraint.

Variable splitting

- The residual image is a natural split variable:
$$\boldsymbol{s}^t = \boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}(\boldsymbol{x}^{t-1}_{\text{reg}}; (\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t).$$
- The augmented Lagrangian has scaled dual $\boldsymbol{u}$, penalty μ:

$$\begin{aligned} AL(\boldsymbol{x}^t_{\text{meas}}, \boldsymbol{\alpha}^t, \boldsymbol{s}^t; \boldsymbol{u}) = {} & \tfrac{1}{2\sigma^2}\|\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{\mathcal{F}}^{-1}\boldsymbol{d}^t\|_2^2 + \lambda\|\boldsymbol{s}^t\|_1 \\ & + \tfrac{\mu}{2}\|\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)\boldsymbol{x}^{t-1}_{\text{reg}} - \boldsymbol{s}^t + \boldsymbol{u}\|_2^2 \\ & + \tfrac{1}{2}\|\boldsymbol{\alpha}^t - \widehat{\boldsymbol{\alpha}}^{t-1}\|^2_{(\boldsymbol{P}^{t|t-1}_{\boldsymbol{\alpha}})^{-1}}. \end{aligned}$$

- The choice of penalty $\mu > 0$ controls the overall convergence rate, by trading off minimizing the objective function and satisfying the variable-split constraint.

Linearizing the motion transform

- The transform $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)$ is nonlinear and nonconvex in $\boldsymbol{\alpha}^t$.
- Assuming smooth motion, $\boldsymbol{\alpha}^t$ is close to $\boldsymbol{\alpha}^{t-1}$, so initializing with $\widehat{\boldsymbol{\alpha}}^{t-1}$ likely yields a global minimum.
- Linearize the transform $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)$ around $\boldsymbol{\alpha}^t = \widehat{\boldsymbol{\alpha}}^{t-1}$, and define $\widehat{\boldsymbol{a}}^t = \boldsymbol{\alpha}^t - \widehat{\boldsymbol{\alpha}}^{t-1}$:
- The resulting approximation to the augmented Lagrangian is convex.

Linearizing the motion transform

- The transform $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)$ is nonlinear and nonconvex in $\boldsymbol{\alpha}^t$.
- Assuming smooth motion, $\boldsymbol{\alpha}^t$ is close to $\boldsymbol{\alpha}^{t-1}$, so initializing with $\widehat{\boldsymbol{\alpha}}^{t-1}$ likely yields a global minimum.
- Linearize the transform $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)$ around $\boldsymbol{\alpha}^t = \widehat{\boldsymbol{\alpha}}^{t-1}$, and define $\widehat{\boldsymbol{a}}^t = \boldsymbol{\alpha}^t - \widehat{\boldsymbol{\alpha}}^{t-1}$:
- The resulting approximation to the augmented Lagrangian is convex.

Linearizing the motion transform

- The transform $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)$ is nonlinear and nonconvex in $\boldsymbol{\alpha}^t$.
- Assuming smooth motion, $\boldsymbol{\alpha}^t$ is close to $\boldsymbol{\alpha}^{t-1}$, so initializing with $\widehat{\boldsymbol{\alpha}}^{t-1}$ likely yields a global minimum.
- Linearize the transform $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)$ around $\boldsymbol{\alpha}^t = \widehat{\boldsymbol{\alpha}}^{t-1}$, and define $\widehat{\boldsymbol{a}}^t = \boldsymbol{\alpha}^t - \widehat{\boldsymbol{\alpha}}^{t-1}$:

$$\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)\boldsymbol{x}_{\text{reg}}^{t-1} \approx \boldsymbol{x}_{\text{reg}}^{t-1} + \boldsymbol{J_T}\widehat{\boldsymbol{a}}^t,$$

  where $\boldsymbol{J_T}$ is the Jacobian matrix of $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)\boldsymbol{x}_{\text{reg}}^{t-1}$ evaluated at $\boldsymbol{\alpha}^t = \widehat{\boldsymbol{\alpha}}^{t-1}$.
- The resulting approximation to the augmented Lagrangian is convex.

Linearizing the motion transform

- The transform $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)$ is nonlinear and nonconvex in $\boldsymbol{\alpha}^t$.
- Assuming smooth motion, $\boldsymbol{\alpha}^t$ is close to $\boldsymbol{\alpha}^{t-1}$, so initializing with $\widehat{\boldsymbol{\alpha}}^{t-1}$ likely yields a global minimum.
- Linearize the transform $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)$ around $\boldsymbol{\alpha}^t = \widehat{\boldsymbol{\alpha}}^{t-1}$, and define $\widehat{\boldsymbol{a}}^t = \boldsymbol{\alpha}^t - \widehat{\boldsymbol{\alpha}}^{t-1}$:

$$\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)\boldsymbol{x}_{\text{reg}}^{t-1} \approx \boldsymbol{x}_{\text{reg}}^{t-1} + \boldsymbol{J}_{\boldsymbol{T}}\widehat{\boldsymbol{a}}^t,$$

  where $\boldsymbol{J}_{\boldsymbol{T}}$ is the Jacobian matrix of $\boldsymbol{T}((\widehat{\boldsymbol{\alpha}}^{t-1})^{(-1)}, \boldsymbol{\alpha}^t)\boldsymbol{x}_{\text{reg}}^{t-1}$ evaluated at $\boldsymbol{\alpha}^t = \widehat{\boldsymbol{\alpha}}^{t-1}$.
- The resulting approximation to the augmented Lagrangian is convex.

Alternating minimization

Use alternating directions method of multipliers (ADMM)[1,2].
Each iteration consists of three steps:

1. Update $\boldsymbol{x}^t_{\text{meas}}, \widehat{\boldsymbol{a}}^t$ together (least-squares problem):

$$\{\boldsymbol{x}^t_{\text{meas}}, \widehat{\boldsymbol{a}}^t\} \leftarrow \arg\min_{\boldsymbol{x},\widehat{\boldsymbol{a}}} \tfrac{1}{2\sigma^2}\|\boldsymbol{x} - \boldsymbol{\mathcal{F}}^{-1}\boldsymbol{d}^t\|_2^2 + \tfrac{1}{2}\|\widehat{\boldsymbol{a}}\|^2_{(\boldsymbol{P}_{\boldsymbol{\alpha}}^{t|t-1})^{-1}}$$
$$+ \tfrac{\mu}{2}\|\boldsymbol{x} - \boldsymbol{J}_{\boldsymbol{T}}\widehat{\boldsymbol{a}} - (\boldsymbol{x}^{t-1}_{\text{reg}} + \boldsymbol{s}^t - \boldsymbol{u})\|_2^2.$$

2. Update $\boldsymbol{s}^t$ using shrinkage:

3. Update scaled dual variable:

[1] R Glowinski and A Marrocco, Inf. Rech. Oper., R-2, 1975.
[2] D Gabay and B Mercier, Comput. Math. Appl., 2(1), 1976.

Alternating minimization

Use alternating directions method of multipliers (ADMM)[1,2].
Each iteration consists of three steps:

1. Update $\boldsymbol{x}^t_{\text{meas}}, \widehat{\boldsymbol{a}}^t$ together (least-squares problem):

$$\{\boldsymbol{x}^t_{\text{meas}}, \widehat{\boldsymbol{a}}^t\} \leftarrow \arg\min_{\boldsymbol{x},\widehat{\boldsymbol{a}}} \tfrac{1}{2\sigma^2}\|\boldsymbol{x} - \boldsymbol{\mathcal{F}}^{-1}\boldsymbol{d}^t\|_2^2 + \tfrac{1}{2}\|\widehat{\boldsymbol{a}}\|^2_{(\boldsymbol{P}^{t|t-1}_{\boldsymbol{\alpha}})^{-1}} + \tfrac{\mu}{2}\|\boldsymbol{x} - \boldsymbol{J_T}\widehat{\boldsymbol{a}} - (\boldsymbol{x}^{t-1}_{\text{reg}} + \boldsymbol{s}^t - \boldsymbol{u})\|_2^2.$$

2. Update $\boldsymbol{s}^t$ using shrinkage:

$$\boldsymbol{s}^t \leftarrow \arg\min_{\boldsymbol{s}} \lambda\|\boldsymbol{s}\|_1 + \tfrac{\mu}{2}\|\boldsymbol{s} - (\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{J_T}\widehat{\boldsymbol{a}}^t - \boldsymbol{x}^{t-1}_{\text{reg}} + \boldsymbol{u})\|_2^2.$$

3. Update scaled dual variable:

[1] R Glowinski and A Marrocco, Inf. Rech. Oper., R-2, 1975.
[2] D Gabay and B Mercier, Comput. Math. Appl., 2(1), 1976.

Alternating minimization

Use alternating directions method of multipliers (ADMM)[1,2].
Each iteration consists of three steps:

1. Update $\boldsymbol{x}^t_{\text{meas}}, \widehat{\boldsymbol{a}}^t$ together (least-squares problem):

$$\{\boldsymbol{x}^t_{\text{meas}}, \widehat{\boldsymbol{a}}^t\} \leftarrow \arg\min_{\boldsymbol{x},\widehat{\boldsymbol{a}}} \tfrac{1}{2\sigma^2}\|\boldsymbol{x} - \boldsymbol{\mathcal{F}}^{-1}\boldsymbol{d}^t\|_2^2 + \tfrac{1}{2}\|\widehat{\boldsymbol{a}}\|^2_{(\boldsymbol{P}_{\boldsymbol{\alpha}}^{t|t-1})^{-1}}$$
$$+ \tfrac{\mu}{2}\|\boldsymbol{x} - \boldsymbol{J}_{\boldsymbol{T}}\widehat{\boldsymbol{a}} - (\boldsymbol{x}^{t-1}_{\text{reg}} + \boldsymbol{s}^t - \boldsymbol{u})\|_2^2.$$

2. Update $\boldsymbol{s}^t$ using shrinkage:

$$\boldsymbol{s}^t \leftarrow \arg\min_{\boldsymbol{s}} \lambda\|\boldsymbol{s}\|_1 + \tfrac{\mu}{2}\|\boldsymbol{s} - (\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{J}_{\boldsymbol{T}}\widehat{\boldsymbol{a}}^t - \boldsymbol{x}^{t-1}_{\text{reg}} + \boldsymbol{u})\|_2^2.$$

3. Update scaled dual variable:

$$\boldsymbol{u} \leftarrow \boldsymbol{u} + (\boldsymbol{x}^t_{\text{meas}} - \boldsymbol{J}_{\boldsymbol{T}}\widehat{\boldsymbol{a}}^t - \boldsymbol{x}^{t-1}_{\text{reg}} - \boldsymbol{s}^t).$$

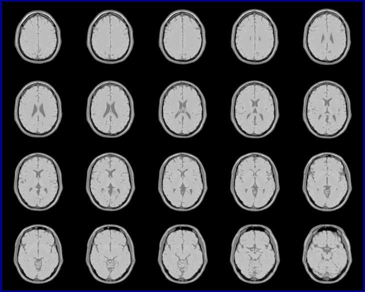
[1] R Glowinski and A Marrocco, Inf. Rech. Oper., R-2, 1975.
[2] D Gabay and B Mercier, Comput. Math. Appl., 2(1), 1976.

Outline

- Magnetic Resonance Imaging
- Functional MRI
- Head Motion in fMRI
- Proposed Method for Prospective Correction
- Real Time Implementation
- Simulation Results
- Conclusions and Future Work
- Summary

Experimental design

- T_2^*-weighted Brainweb[1] phantom ($1 \times 1 \times 3$ mm resolution)
- Simulated activations, head motion on high-resolution phantom for 200 frames (TR $= 1$ s)
- Sampled k-space for $16 + 4$ slices at $4 \times 4 \times 3$ mm resolution (64×64 samples/slice) with 40 dB SNR

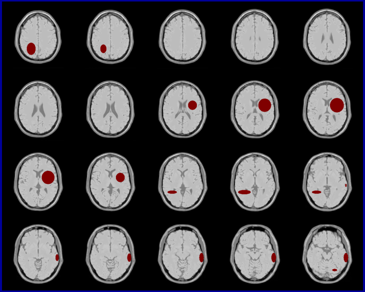


$16 + 4$ high-resolution slices

[1] RKS Kwan et al., IEEE Trans. Med. Imag., 18(11), 1999.

Experimental design

- T_2^*-weighted Brainweb[1] phantom ($1 \times 1 \times 3$ mm resolution)
- Simulated activations, head motion on high-resolution phantom for 200 frames (TR $= 1$ s)
- Sampled k-space for $16 + 4$ slices at $4 \times 4 \times 3$ mm resolution (64×64 samples/slice) with 40 dB SNR

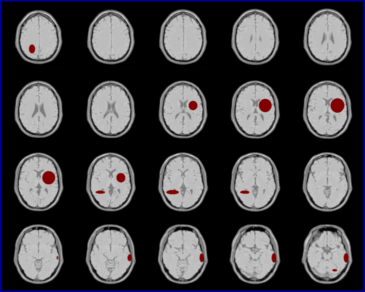


High-resolution slices + activations (red)

[1] RKS Kwan et al., IEEE Trans. Med. Imag., 18(11), 1999.

Experimental design

- T_2^*-weighted Brainweb[1] phantom ($1 \times 1 \times 3$ mm resolution)
- Simulated activations, head motion on high-resolution phantom for 200 frames (TR $= 1$ s)
- Sampled k-space for $16 + 4$ slices at $4 \times 4 \times 3$ mm resolution (64×64 samples/slice) with 40 dB SNR

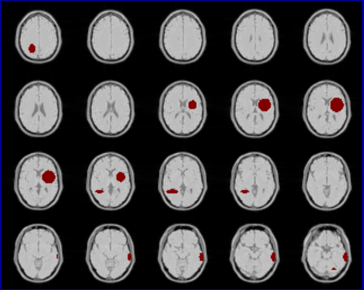


High-resolution slices + motion

[1] RKS Kwan et al., IEEE Trans. Med. Imag., 18(11), 1999.

Experimental design

- T_2^*-weighted Brainweb[1] phantom ($1 \times 1 \times 3$ mm resolution)
- Simulated activations, head motion on high-resolution phantom for 200 frames (TR $= 1$ s)
- Sampled k-space for $16 + 4$ slices at $4 \times 4 \times 3$ mm resolution (64×64 samples/slice) with 40 dB SNR

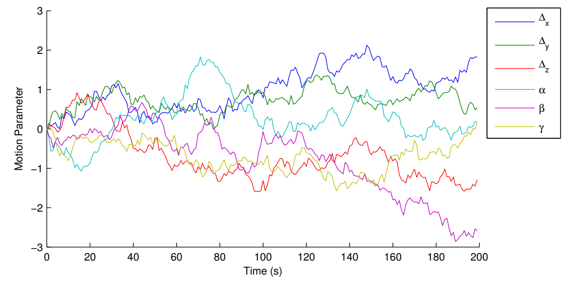


$4 \times 4 \times 3$ mm slices + motion

[1] RKS Kwan et al., IEEE Trans. Med. Imag., 18(11), 1999.

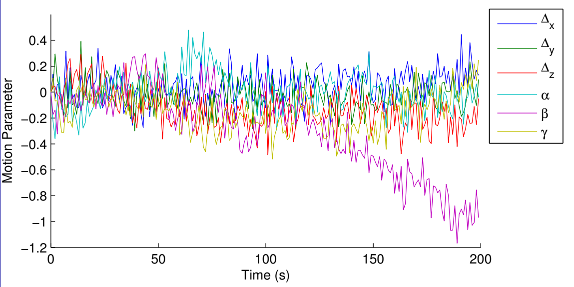
Simulated motion

- Rigid-body motion is described using six parameters:
    1. Δ_x: right-to-left translation (mm)
    2. Δ_y: anterior-to-posterior translation (mm)
    3. Δ_z: superior-to-inferior translation (mm)
    4. α: axial (xy-)plane rotation (degrees)
    5. β: coronal (xz-)plane rotation (degrees)
    6. γ: sagittal (yz-)plane rotation (degrees)



Simulated rigid-body motion

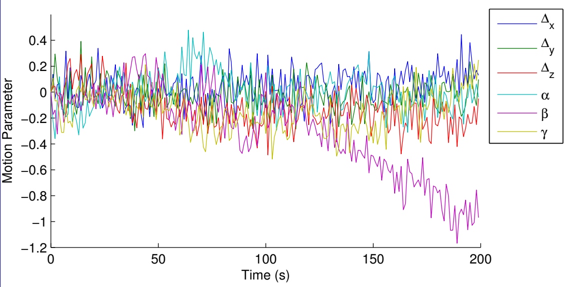
Prospective motion correction



Residual motion after prospective correction

- Residual motion is substantially reduced:
  Uncorrected: tx $= 1.6 \pm 0.53$ mm, rot $= 1.4 \pm 0.68$ deg.
  Residual: tx $= 0.25 \pm 0.10$ mm, rot $= 0.45 \pm 0.11$ deg.
- Retrospective registration can mitigate this residual motion.

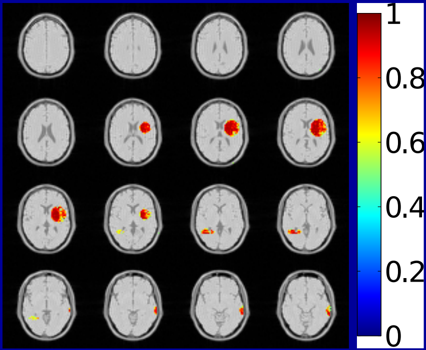
Prospective motion correction



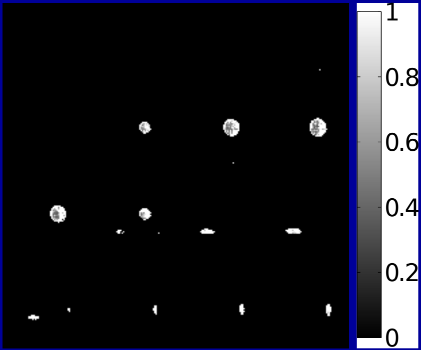
Residual motion after prospective correction

- Residual motion is substantially reduced:
  Uncorrected: tx $= 1.6 \pm 0.53$ mm, rot $= 1.4 \pm 0.68$ deg.
  Residual: tx $= 0.25 \pm 0.10$ mm, rot $= 0.45 \pm 0.11$ deg.
- Retrospective registration can mitigate this residual motion.

Time-series correlation maps – no correction

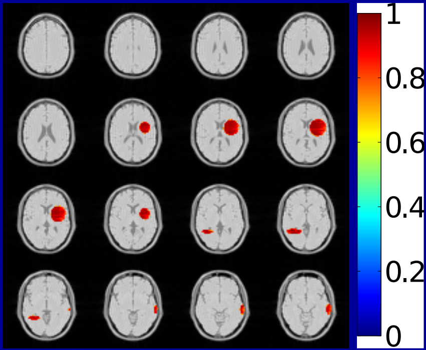


Correlation Map

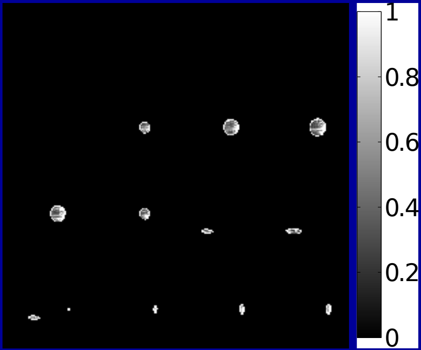


Differences $(10\times)$

Time-series correlation maps – retrospective

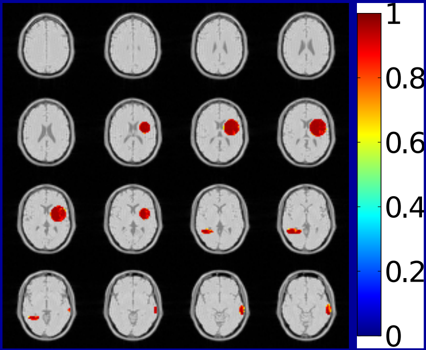


Correlation Map

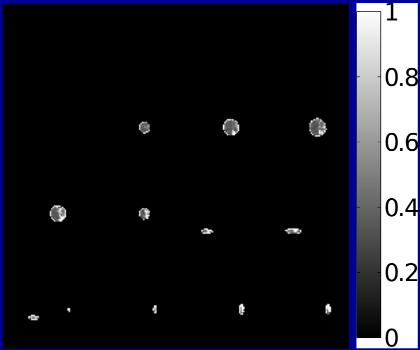


Differences $(10\times)$

Time-series correlation maps – prospective

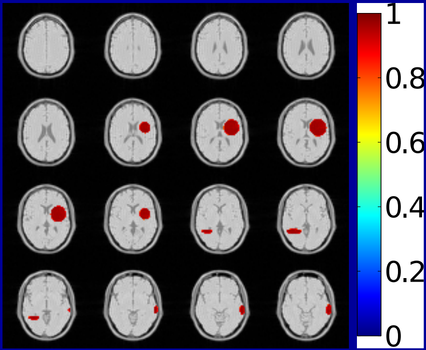


Correlation Map

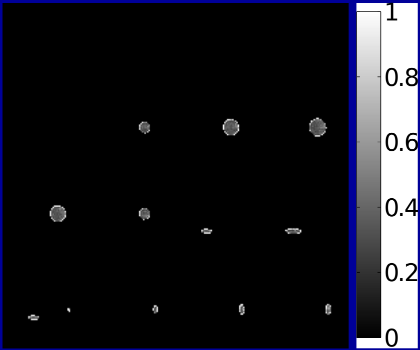


Differences $(10\times)$

Time-series correlation maps – both corrections

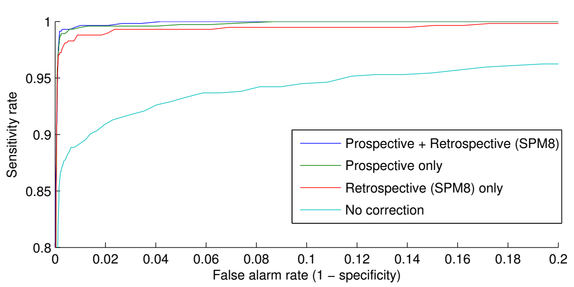


Correlation Map



Differences $(10\times)$

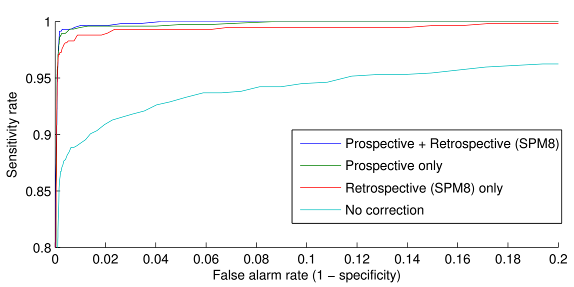
Performance analysis of activation maps



Receiver operating characteristic (ROC) curves for correlation analysis

- Prospective correction improves sensitivity and specificity
- Spatial interpolation may be responsible for reduced sensitivity with just retrospective correction

Performance analysis of activation maps



Receiver operating characteristic (ROC) curves for correlation analysis

- Prospective correction improves sensitivity and specificity
- Spatial interpolation may be responsible for reduced sensitivity with just retrospective correction

Simulation with unknown parameters

- Estimate σ^2 using sample variance of noise-only data.
- Calibrate for 2-D EPI Nyquist ghost correction[1]:
- Estimate the innovation covariance Q:

Simulation with unknown parameters

- Estimate σ^2 using sample variance of noise-only data.
- Calibrate for 2-D EPI Nyquist ghost correction[1]:
  - use 2-D reference scans (forwards & backwards)
  - transform linear terms for prospective correction
- Estimate the innovation covariance Q:

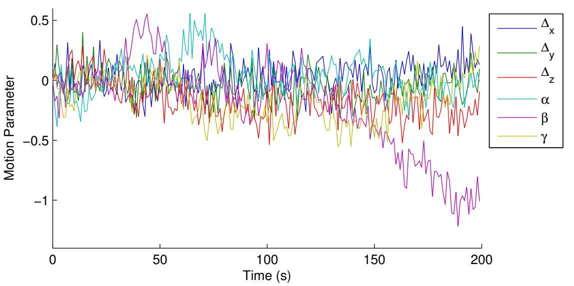
[1] N Chen and AM Wyrwicz, Mag. Res. Med., 51(6), 2004.

Simulation with unknown parameters

- Estimate σ^2 using sample variance of noise-only data.
- Calibrate for 2-D EPI Nyquist ghost correction[1]:
  - use 2-D reference scans (forwards & backwards)
  - transform linear terms for prospective correction
- Estimate the innovation covariance Q:
  - initialize to a large value
  - update the sample covariance using estimated $\widehat{a}$'s
  - since Q is time-varying, update using just last ten frames

[1] N Chen and AM Wyrwicz, Mag. Res. Med., 51(6), 2004.

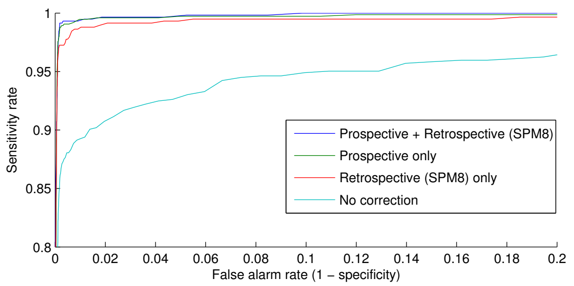
Simulation with unknown parameters



Residual motion after prospective correction

- Residual motion is nearly the same as before:
  Uncorrected: tx $= 1.6 \pm 0.53$ mm, rot $= 1.4 \pm 0.68$ deg.
  Residual: tx $= 0.28 \pm 0.11$ mm, rot $= 0.47 \pm 0.26$ deg.
- Prospective correction remains effective at improving sensitivity, specificity.

Simulation with unknown parameters

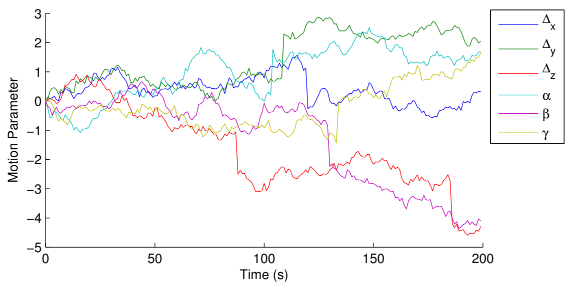


Receiver operating characteristic (ROC) curves for correlation analysis

- Residual motion is nearly the same as before:
  Uncorrected: tx $= 1.6 \pm 0.53$ mm, rot $= 1.4 \pm 0.68$ deg.
  Residual: tx $= 0.28 \pm 0.11$ mm, rot $= 0.47 \pm 0.26$ deg.
- Prospective correction remains effective at improving sensitivity, specificity.

Correcting for impulsive motion

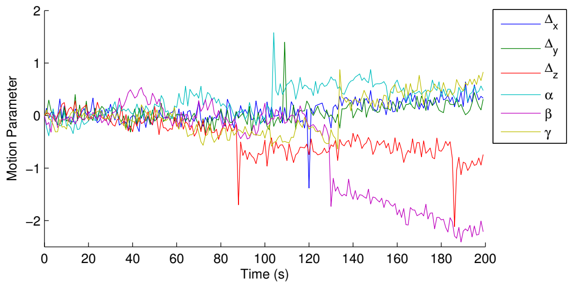
- Impulsive motion is significant over a short duration.
- I generated 1 s impulses of ± 1.5 mm/degrees per second occurring every 100 s, on average.
- The residual motion effects are mainly short-lived.
- The improvement in sensitivity of prospective correction remains significant.



True motion including simulated impulses

Correcting for impulsive motion

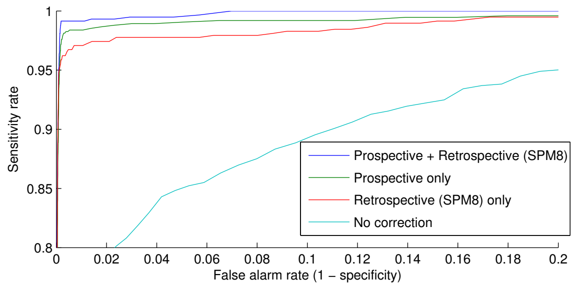
- Impulsive motion is significant over a short duration.
- I generated 1 s impulses of ± 1.5 mm/degrees per second occurring every 100 s, on average.
- The residual motion effects are mainly short-lived.
- The improvement in sensitivity of prospective correction remains significant.



Residual motion after prospective correction

Correcting for impulsive motion

- Impulsive motion is significant over a short duration.
- I generated 1 s impulses of ± 1.5 mm/degrees per second occurring every 100 s, on average.
- The residual motion effects are mainly short-lived.
- The improvement in sensitivity of prospective correction remains significant.



Receiver operating characteristic (ROC) curves for correlation analysis

Outline

- Magnetic Resonance Imaging
- Functional MRI
- Head Motion in fMRI
- Proposed Method for Prospective Correction
- Real Time Implementation
- Simulation Results
- Conclusions and Future Work
- Summary

Conclusions

- The residual motion is much smaller than absolute motion.
- Correlation maps show fewer errors and mis-classifications.
- The proposed method is more statistically robust than standard retrospective registration.
- The algorithm remains effective when noise/innovation statistics are unknown or impulsive motion is added.
- Limitations include:

Conclusions

- The residual motion is much smaller than absolute motion.
- Correlation maps show fewer errors and mis-classifications.
- The proposed method is more statistically robust than standard retrospective registration.
- The algorithm remains effective when noise/innovation statistics are unknown or impulsive motion is added.
- Limitations include:

Conclusions

- The residual motion is much smaller than absolute motion.
- Correlation maps show fewer errors and mis-classifications.
- The proposed method is more statistically robust than standard retrospective registration.
- The algorithm remains effective when noise/innovation statistics are unknown or impulsive motion is added.
- Limitations include:

Conclusions

- The residual motion is much smaller than absolute motion.
- Correlation maps show fewer errors and mis-classifications.
- The proposed method is more statistically robust than standard retrospective registration.
- The algorithm remains effective when noise/innovation statistics are unknown or impulsive motion is added.
- Limitations include:

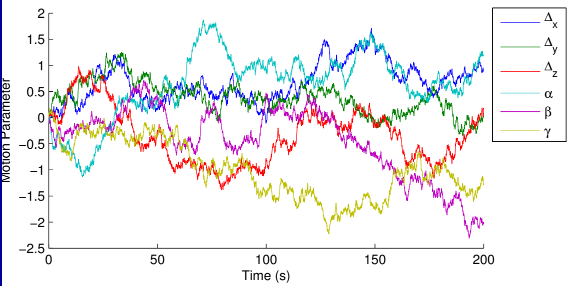
Conclusions

- The residual motion is much smaller than absolute motion.
- Correlation maps show fewer errors and mis-classifications.
- The proposed method is more statistically robust than standard retrospective registration.
- The algorithm remains effective when noise/innovation statistics are unknown or impulsive motion is added.
- Limitations include:
  - motion assumed constant over a TR
  - ignored other time-varying effects such as scanner (B_0) drift, susceptibility variations, and physiological signals

Slice-by-slice motion correction

- Motion will be different for each slice in a slice-by-slice acquisition.
- Prospective correction reduces residual per-slice motion:

Uncorrected: tx $= 1.1 \pm 0.27$ mm, rot $= 1.5 \pm 0.64$ deg.
Residual: tx $= 0.30 \pm 0.11$ mm, rot $= 0.47 \pm 0.20$ deg.

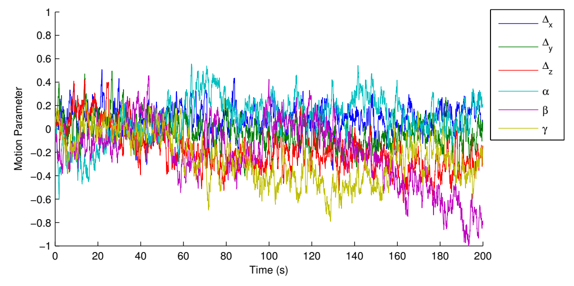


True slice-by-slice motion

Slice-by-slice motion correction

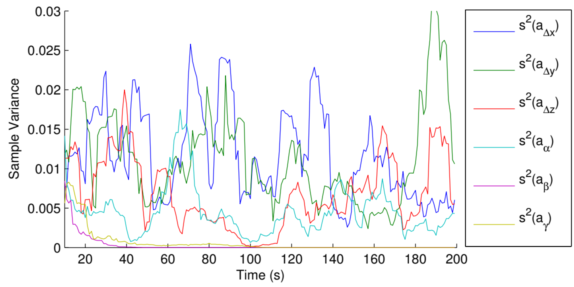
- Motion will be different for each slice in a slice-by-slice acquisition.
- Prospective correction reduces residual per-slice motion:

Uncorrected: tx $= 1.1 \pm 0.27$ mm, rot $= 1.5 \pm 0.64$ deg.
Residual: tx $= 0.30 \pm 0.11$ mm, rot $= 0.47 \pm 0.20$ deg.



Residual motion after prospective correction (known Q)

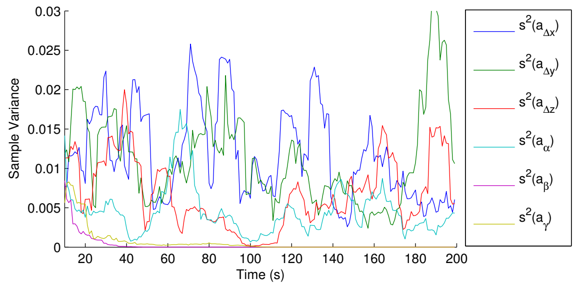
Slice-by-slice motion correction



Estimates of innovation sample variances (diagonal of $\boldsymbol{Q}$)

- Estimates of $\boldsymbol{Q}$ are less stable with slice-by-slice motion.
- Innovation variances for out-of-plane rotations β, $\gamma \downarrow 0$.
- In turn, $(\boldsymbol{P}_{\boldsymbol{\alpha}}^{t|t-1})^{-1}$ becomes arbitrarily large, yielding motion estimates mostly ignoring the data.

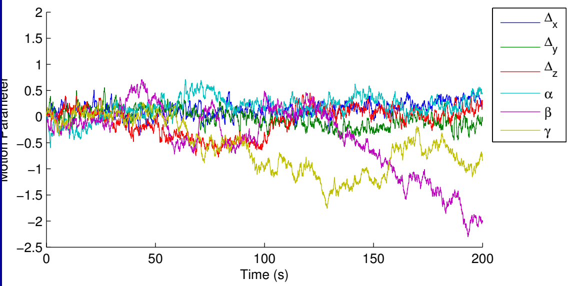
Slice-by-slice motion correction



Estimates of innovation sample variances (diagonal of $\boldsymbol{Q}$)

- Estimates of $\boldsymbol{Q}$ are less stable with slice-by-slice motion.
- Innovation variances for out-of-plane rotations $\beta, \gamma \downarrow 0$.
- In turn, $(\boldsymbol{P}_{\boldsymbol{\alpha}}^{t|t-1})^{-1}$ becomes arbitrarily large, yielding motion estimates mostly ignoring the data.

Slice-by-slice motion correction



Residual motion after prospective correction (estimated $\boldsymbol{Q}$)

- Estimates of Q are less stable with slice-by-slice motion.
- Innovation variances for out-of-plane rotations β, $\gamma \downarrow 0$.
- In turn, $(\boldsymbol{P}_{\boldsymbol{\alpha}}^{t|t-1})^{-1}$ becomes arbitrarily large, yielding motion estimates mostly ignoring the data.

Slice-by-slice motion correction

- Enforcing a minimum threshold on the innovation sample variances would mitigate the effect of poor estimates of $\boldsymbol{Q}$ on future motion estimates.
- Alternatively, we can extend our Kalman filter model to account for slice-by-slice motion:

$$\begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_{N_S} \end{bmatrix}^t = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_{N_S} \end{bmatrix}^{t-1} + \begin{bmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_{N_S} \end{bmatrix}^t, \text{ where}$$

$$\begin{bmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_{N_S} \end{bmatrix}^t \sim \text{Normal}\left(\boldsymbol{0}, \begin{bmatrix} N_S & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & N_S \end{bmatrix} \otimes \boldsymbol{Q} \right).$$

Slice-by-slice motion correction

- Enforcing a minimum threshold on the innovation sample variances would mitigate the effect of poor estimates of Q on future motion estimates.
- Alternatively, we can extend our Kalman filter model to account for slice-by-slice motion:

$$\begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_{N_S} \end{bmatrix}^t = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_{N_S} \end{bmatrix}^{t-1} + \begin{bmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_{N_S} \end{bmatrix}^t, \text{ where}$$

$$\begin{bmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_{N_S} \end{bmatrix}^t \sim \text{Normal}\left(\mathbf{0}, \begin{bmatrix} N_S & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & N_S \end{bmatrix} \otimes \boldsymbol{Q}\right).$$

Accounting for other time-varying signals

- Other time-varying components of the BOLD signal include:
  - scanner drift, which has a global effect on T_2^*,
  - breathing-induced global modulation of the main field, T_2^*,
  - and cardiac pulsatility, which varies blood flow, especially near the cerebral arteries and ventricles.
- The global effects are spatially smooth.
- Introducing a wavelet transform can isolate drift and respiratory changes to the approximation coefficients.
- We propose performing registration using just the detail coefficients.

Accounting for other time-varying signals

- Other time-varying components of the BOLD signal include:
  - scanner drift, which has a global effect on T_2^*,
  - breathing-induced global modulation of the main field, T_2^*,
  - and cardiac pulsatility, which varies blood flow, especially near the cerebral arteries and ventricles.
- The global effects are spatially smooth.
- Introducing a wavelet transform can isolate drift and respiratory changes to the approximation coefficients.
- We propose performing registration using just the detail coefficients.

Accounting for other time-varying signals

- Other time-varying components of the BOLD signal include:
  - scanner drift, which has a global effect on T_2^*,
  - breathing-induced global modulation of the main field, T_2^*,
  - and cardiac pulsatility, which varies blood flow, especially near the cerebral arteries and ventricles.
- The global effects are spatially smooth.
- Introducing a wavelet transform can isolate drift and respiratory changes to the approximation coefficients.
- We propose performing registration using just the detail coefficients.

Accounting for other time-varying signals

- Other time-varying components of the BOLD signal include:
  - scanner drift, which has a global effect on T_2^*,
  - breathing-induced global modulation of the main field, T_2^*,
  - and cardiac pulsatility, which varies blood flow, especially near the cerebral arteries and ventricles.
- The global effects are spatially smooth.
- Introducing a wavelet transform can isolate drift and respiratory changes to the approximation coefficients.
- We propose performing registration using just the detail coefficients.

Accounting for other time-varying signals

- Other time-varying components of the BOLD signal include:
  - scanner drift, which has a global effect on T_2^*,
  - breathing-induced global modulation of the main field, T_2^*,
  - and cardiac pulsatility, which varies blood flow, especially near the cerebral arteries and ventricles.
- The global effects are spatially smooth.
- Introducing a wavelet transform can isolate drift and respiratory changes to the approximation coefficients.
- We propose performing registration using just the detail coefficients.

Accounting for other time-varying signals

- Other time-varying components of the BOLD signal include:
  - scanner drift, which has a global effect on T_2^*,
  - breathing-induced global modulation of the main field, T_2^*,
  - and cardiac pulsatility, which varies blood flow, especially near the cerebral arteries and ventricles.
- The global effects are spatially smooth.
- Introducing a wavelet transform can isolate drift and respiratory changes to the approximation coefficients.
- We propose performing registration using just the detail coefficients.

Outline

- Magnetic Resonance Imaging
- Functional MRI
- Head Motion in fMRI
- Proposed Method for Prospective Correction
- Real Time Implementation
- Simulation Results
- Conclusions and Future Work
- Summary

Summary

- Functional MRI tracks brain function during tasks by using the signal variation generated by metabolizing hemoglobin.
- Prospective correction can improve statistical sensitivity of fMRI time series acquired in the presence of head motion.
- The proposed method outperforms retrospective registration in simulated data, even with impulsive motion and unknown model parameters.
- Future work will make the proposed method robust to inter-slice motion and global image time variations.
- I am also preparing to evaluate my proposed method in real fMRI studies.

Summary

- Functional MRI tracks brain function during tasks by using the signal variation generated by metabolizing hemoglobin.
- Prospective correction can improve statistical sensitivity of fMRI time series acquired in the presence of head motion.
- The proposed method outperforms retrospective registration in simulated data, even with impulsive motion and unknown model parameters.
- Future work will make the proposed method robust to inter-slice motion and global image time variations.
- I am also preparing to evaluate my proposed method in real fMRI studies.

Summary

- Functional MRI tracks brain function during tasks by using the signal variation generated by metabolizing hemoglobin.
- Prospective correction can improve statistical sensitivity of fMRI time series acquired in the presence of head motion.
- The proposed method outperforms retrospective registration in simulated data, even with impulsive motion and unknown model parameters.
- Future work will make the proposed method robust to inter-slice motion and global image time variations.
- I am also preparing to evaluate my proposed method in real fMRI studies.

Summary

- Functional MRI tracks brain function during tasks by using the signal variation generated by metabolizing hemoglobin.
- Prospective correction can improve statistical sensitivity of fMRI time series acquired in the presence of head motion.
- The proposed method outperforms retrospective registration in simulated data, even with impulsive motion and unknown model parameters.
- Future work will make the proposed method robust to inter-slice motion and global image time variations.
- I am also preparing to evaluate my proposed method in real fMRI studies.

Summary

- Functional MRI tracks brain function during tasks by using the signal variation generated by metabolizing hemoglobin.
- Prospective correction can improve statistical sensitivity of fMRI time series acquired in the presence of head motion.
- The proposed method outperforms retrospective registration in simulated data, even with impulsive motion and unknown model parameters.
- Future work will make the proposed method robust to inter-slice motion and global image time variations.
- I am also preparing to evaluate my proposed method in real fMRI studies.

Questions?

Thank you for your attention.

Acknowledgments:

- Jeff Fessler and Doug Noll
- NIH F32 EB015914 and P01 CA087634