Introduction to diffusion-weighted MRI

Anastasia Yendiki

HMS/MGH/MIT Athinoula A. Martinos Center for Biomedical Imaging
White-matter imaging

- Axons measure ~\(\mu\)m in width
- They group together in bundles that traverse the white matter
- We cannot image individual axons but we can image bundles with diffusion MRI
- Useful in studying neurodegenerative diseases, stroke, aging, development...

From Gray’s Anatomy: IX. Neurology

From the National Institute on Aging
Diffusion in brain tissue

- Differentiate tissues based on the diffusion (random motion) of water molecules within them

- Gray matter: Diffusion is unrestricted ⇒ isotropic

- White matter: Diffusion is restricted ⇒ anisotropic
Properties of diffusion

- At every voxel we want to know:
  - Is this in white matter?
  - If yes, what pathway(s) is it part of?
    - What is the orientation of diffusion?
    - What is the magnitude of diffusion?
- A grayscale image cannot capture all this!
Diffusion MRI (dMRI)

- Magnetic resonance imaging can provide “diffusion encoding”
- Magnetic field strength is varied by gradients in different directions
- Image intensity is attenuated depending on water diffusion in each direction
- Compare with baseline images to infer on diffusion process
**Diffusion encoding in MRI**

- Apply two gradient pulses in some direction $y$:

  - **Case 1:** If spins aren’t diffusing
    
    $y = y_1, y_2 \rightarrow y = y_1, y_2$

    - No displacement in $y \Rightarrow$
    - No dephasing $\Rightarrow$
    - No net signal change
Diffusion encoding in MRI

- Apply two gradient pulses:

\[
\begin{array}{c}
s_0^\circ \quad G_y \\ 180^\circ \quad G_y \\
\end{array}
\]

- Case 2: If spins are diffusing

\[
y = y_1, y_2 \quad \rightarrow \quad y = y_1 + \Delta y_1, y_2 + \Delta y_2
\]

Displacement in $y \Rightarrow$ Dephasing $\Rightarrow$ Signal attenuation
Choice 1: Gradient directions

- True diffusion direction $||$ Applied gradient direction
  $\Rightarrow$ Maximum attenuation

  ![Diffusion-encoding gradient $g$](image)
  Diffusion detected

- True diffusion direction $\perp$ Applied gradient direction
  $\Rightarrow$ No attenuation

  ![Diffusion-encoding gradient $g$](image)
  Diffusion not detected

- To capture all diffusion directions well, gradient directions should cover 3D space uniformly

  ![Diffusion-encoding gradient $g$](image)
  Diffusion partly detected

Why’n’how | Intro to diffusion-weighted MRI
How many directions?

- Acquiring data with more gradient directions leads to:
  + More reliable estimation of diffusion measures
  - Increased imaging time ⇒ Subject discomfort, more susceptible to artifacts due to motion, respiration, etc.

- DTI:
  - Six directions is the minimum
  - Usually a few 10’s of directions
  - Diminishing returns after a certain number [Jones, 2004]

- HARDI/DSI:
  - Usually a few 100’s of directions
Choice 2: b-value

The b-value depends on acquisition parameters:

\[ b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3) \]

- \( \gamma \) the gyromagnetic ratio
- \( G \) the strength of the diffusion-encoding gradient
- \( \delta \) the duration of each diffusion-encoding pulse
- \( \Delta \) the interval b/w diffusion-encoding pulses
How high b-value?

• Increasing the b-value leads to:
  + Increased contrast b/w areas of higher and lower diffusivity in principle
  – Decreased signal-to-noise ratio ⇒ Less reliable estimation of diffusion measures in practice

• DTI: b ~ 1000 sec/mm²
• HARDI/DSI: b ~ 10,000 sec/mm²

• Data can be acquired at multiple b-values for trade-off
• Repeat acquisition and average to increase signal-to-noise ratio
Looking at the data

A diffusion data set consists of:

- A set of non-diffusion-weighted a.k.a. “baseline” a.k.a. “low-b” images (b-value = 0)
- A set of diffusion-weighted (DW) images acquired with different gradient directions $g_1$, $g_2$, ... and b-value > 0
- The diffusion-weighted images have lower intensity values
## Models of diffusion

<table>
<thead>
<tr>
<th>Model</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diffusion spectrum:</strong></td>
<td>DTK (DSI option)</td>
</tr>
<tr>
<td>Full distribution of orientation and magnitude</td>
<td></td>
</tr>
<tr>
<td><strong>Orientation distribution function (ODF):</strong></td>
<td>DTK (Q-ball option)</td>
</tr>
<tr>
<td>No magnitude info</td>
<td></td>
</tr>
<tr>
<td><strong>Ball-and-stick:</strong></td>
<td>FSL (bedpost)</td>
</tr>
<tr>
<td>Orientation and magnitude for up to N anisotropic compartments (default N=2)</td>
<td></td>
</tr>
<tr>
<td><strong>Tensor:</strong></td>
<td>DTK (DTI option)</td>
</tr>
<tr>
<td>Single orientation and magnitude</td>
<td>FSL (dtifit)</td>
</tr>
</tbody>
</table>
Tensors

- One way to express diffusion is by a tensor $D$.

- A tensor is a 3x3 symmetric, positive-definite matrix:

$$D = \begin{bmatrix}
  d_{11} & d_{12} & d_{13} \\
  d_{12} & d_{22} & d_{23} \\
  d_{13} & d_{23} & d_{33}
\end{bmatrix}$$

- $D$ is symmetric 3x3 $\Rightarrow$ It has 6 unique elements.

- Suffices to estimate the upper (lower) triangular part.
Tensor: Eigenvalues & eigenvectors

- The matrix $D$ is positive-definite $\Rightarrow$
  - It has 3 real, positive eigenvalues $\lambda_1, \lambda_2, \lambda_3 > 0$.
  - It has 3 orthogonal eigenvectors $e_1, e_2, e_3$.

$$D = \lambda_1 e_1 \cdot e_1' + \lambda_2 e_2 \cdot e_2' + \lambda_3 e_3 \cdot e_3'$$

**Eigenvalue**

**Eigenvector**
Tensor: Physical interpretation

- Eigenvectors express diffusion direction
- Eigenvalues express diffusion magnitude

Isotropic diffusion:
\[ \lambda_1 \approx \lambda_2 \approx \lambda_3 \]

Anisotropic diffusion:
\[ \lambda_1 \gg \lambda_2 \approx \lambda_3 \]
Tensor: Summary measures

- Mean diffusivity (MD): Mean of the 3 eigenvalues
  \[ \text{MD}(j) = \frac{\lambda_1(j) + \lambda_2(j) + \lambda_3(j)}{3} \]

- Fractional anisotropy (FA): Variance of the 3 eigenvalues, normalized so that \(0 \leq \text{FA} \leq 1\)
  \[ \text{FA}(j)^2 = \frac{3}{2} \frac{[\lambda_1(j)-\text{MD}(j)]^2 + [\lambda_2(j)-\text{MD}(j)]^2 + [\lambda_3(j)-\text{MD}(j)]^2}{\lambda_1(j)^2 + \lambda_2(j)^2 + \lambda_3(j)^2} \]
Tensor: More summary measures

- Axial diffusivity: Greatest eigenvalue
  
  \[ \text{AD}(j) = \lambda_1(j) \]

- Radial diffusivity: Average of 2 lesser eigenvalues
  
  \[ \text{RD}(j) = \frac{\lambda_2(j) + \lambda_3(j)}{2} \]

- These measures are used in combination to interpret group differences
Tensor: Visualization

Image:
An intensity value at each voxel

Tensor map:
A tensor at each voxel

Direction of eigenvector corresponding to greatest eigenvalue
Tensor: Visualization

Image:
An intensity value at each voxel

Tensor map:
A tensor at each voxel

Direction of eigenvector corresponding to greatest eigenvalue
Red: L-R, Green: A-P, Blue: I-S
Data analysis steps

- **Pre-process images**
  - FSL: eddy_correct, rotate_bvecs

- **Fit a diffusion model at every voxel**
  - DTK: DSI, Q-ball, or DTI
  - FSL: Ball-and-stick (bedpost) or DTI (dtifit)

- **Compute measures of anisotropy/diffusivity and compare them between populations**
  - Voxel-based, ROI-based, or tract-based statistical analysis

- **For tract-based: Reconstruct pathways**
  - DTK: Deterministic tractography using DSI, Q-ball, or DTI model
  - FSL: Probabilistic tractography (probtrack) using ball-and-stick model
Tractography studies

• Exploratory tractography:
  – Example: *Show me all regions that the motor cortex is connected to*
  – Seed region can be anatomically defined (motor cortex) or functionally defined (region activated in an fMRI finger-tapping task)

• Tractography of known pathways:
  – Example: *Show me the corticospinal tract*
  – Use prior anatomical knowledge of the pathway’s terminations and trajectory (connects motor cortex and brainstem through capsule)
Tractography methods

- Every tractography method can be characterized by:
  
  - Which model of diffusion does it use?
    - Representation of local orientation of diffusion at every voxel that is fit to image data (tensor, ball-and-stick, etc.)
  
  - Is it deterministic or probabilistic?
    - Deterministic estimates only the most likely orientation
    - Probabilistic also estimates the uncertainty around that
  
  - Is it local or global?
    - Local fits the pathway to the data one step at a time
    - Global fits the entire pathway at once
Deterministic vs. probabilistic

- **Deterministic methods** give you an estimate of model parameters

- **Probabilistic methods** give you the uncertainty (probability distribution) of the estimate
Deterministic vs. probabilistic

Deterministic tractography:
One streamline per seed voxel

Probabilistic tractography:
Multiple streamline samples per seed voxel (drawn from probability distribution)
Deterministic vs. probabilistic

**Deterministic tractography:**
One streamline per seed voxel

**Probabilistic tractography:**
A probability distribution
(sum of all streamline samples from all seed voxels)
Local vs. global

Local tractography:
Fits pathway step-by-step, using local diffusion orientation at each step

Global tractography:
Fits the entire pathway, using diffusion orientation at all voxels along pathway length
Local tractography

- Best suited for exploratory tractography studies
- All connections from a seed region, not constrained to a specific target region
- How do we isolate a specific white-matter pathway?
  - Thresholding?
  - Intermediate masks?
- Non-dominant connections are hard to reconstruct
- Results are not symmetric between “seed” and “target” regions
- Sensitive to areas of high local uncertainty in orientation (e.g., pathway crossings), errors propagate from those areas
Global tractography

- Best suited for reconstruction of known white-matter pathways
- Constrained to connection of two specific end regions
- Not sensitive to areas of high local uncertainty in orientation, integrates over entire pathway
- Symmetric between “seed” and “target” regions

- Computationally expensive: Need to search through a large solution space of all possible connections between two regions
Tractography takes time!

- Get whole-brain tract solutions, edit manually
- Use knowledge of anatomy to isolate specific pathways
Tractography for busy people

- Integrate knowledge of anatomy into algorithm to reconstruct pathways automatically without manual labeling

- TRActs Constrained by UnderLying Anatomy (TRACULA)
- Available in FreeSurfer

Yendiki et al., 2011
TRACULA

- Automated reconstruction of 18 major white-matter pathways
- Global probabilistic tractography with the ball-and-stick model
- No manual definition of seed ROIs needed

- Use prior information on pathway anatomy:
  - Atlas contains manually labeled pathways and FreeSurfer segmentations of training subjects
  - Learn neighboring anatomical labels along pathway

Yendiki et al., 2011
TRACULA outputs

- Posterior probability distribution of each pathway given data (3D volume)

- Tract-based diffusion measures (FA, MD, RD, AD, etc):
  - Average values over entire pathway
  - Average values at each intersection along a pathway
Head motion in diffusion MRI

- Head motion during a dMRI scan can lead to:
  - **Misalignment** between consecutive DWI volumes in the series
  - **Attenuation** in the intensities of a single DWI volume/slice, if the motion occurred during the diffusion-encoding gradient pulse
  - The former can be corrected with rigid registration, *the latter can’t*

- Conventional EPI sequences for dMRI ignore the problem
  - If motion in several directions ⇒ underestimation of anisotropy
  - False positives in group studies where one group moves more
  - Effects more severe when higher $b$-values, more directions acquired
Motion in a dMRI group study

- 50 children with autism spectrum disorder (ASD) and 62 typically developing children (TD), ages 5-12
- 165 total scans (some retest)
- DWI: 3T, 2mm isotropic, 30 directions, $b=700$ s/mm$^2$
- Translation, rotation, and intensity drop-out due to motion
- Outlier scans excluded
- Pathways reconstructed with TRACULA

Data courtesy of Dr. Nancy Kanwisher and Ellison autism study
ASD vs. TD

Differences in dMRI measures between groups with low differences in head motion

Differences in dMRI measures between groups with high differences in head motion

Yendiki et al., 2014
TD vs. TD

Yendiki et al., 2014

Differences in dMRI measures between groups with **low differences in head motion**

Differences in dMRI measures between groups with **high differences in head motion**
Motion compensation strategies

• Retrospective:
  – Registration-based
    • Does not correct for intensity drop-out, less robust at high b-values
  – Outlier removal
    • Must have redundancy in data, remove comparably from every group
  – Nuisance regressors
  – Motion matching between groups

• Prospective:
  – Motion-compensated sequences
  – Accelerated sequences
Overview

- Diffusion MRI measures the preferential orientation and the anisotropy of water diffusion at each voxel

- **Data acquisition:** Diffusion gradient directions, b-values

- **Data analysis:**
  - Diffusion model (tensor, ball-and-stick, ODF, etc.)
  - Voxel-based vs. tract-based group studies
  - Tractography:
    - Exploratory vs. reconstruction of known pathways
    - Deterministic vs. probabilistic
    - Local vs. global

- Beware of *head motion!*