Fast and Accurate HARDI and its Application to Neurological Diagnosis

Dr. Oleg Michailovich

Department of Electrical and Computer Engineering
University of Waterloo

June 21, 2011

Outline

Diffusion MRI

The *average diffusion propagator* $p(\mathbf{r})$ quantifies the probability of a spin to relocate to position $\mathbf{r} + d\mathbf{r}$ at experimental time τ.

$$p(\mathbf{r}) = \int_{\mathbb{R}^3} s(\mathbf{q}) \, e^{-2\pi i (\mathbf{q} \cdot \mathbf{r})} d\mathbf{q}, \quad \mathbf{r} \in \mathbb{R}^3, \mathbf{q} \in \mathbb{R}^3$$

where

- $s(\mathbf{q})$ is a normalized diffusion signal
- $\mathbf{q}$ is the wavevector

Diffusion MRI

The *average diffusion propagator* $p(\mathbf{r})$ quantifies the probability of a spin to relocate to position $\mathbf{r} + d\mathbf{r}$ at experimental time τ.

$$p(\mathbf{r}) = \int_{\mathbb{R}^3} s(\mathbf{q})\, e^{-2\pi i\, (\mathbf{q} \cdot \mathbf{r})} d\mathbf{q}, \quad \mathbf{r} \in \mathbb{R}^3, \mathbf{q} \in \mathbb{R}^3$$

where

- $s(\mathbf{q})$ is a normalized diffusion signal
- $\mathbf{q}$ is the wavevector

In **H**igh **A**ngular **R**esolution **D**iffusion **I**maging, the estimation of $p(\mathbf{r})$ is superseded by estimating its radial projection known as an *orientation distribution function* (ODF):

$$Q(\mathbf{u}) = \int_0^\infty p(\alpha\, \mathbf{u})\, \alpha^2\, d\alpha, \quad \mathbf{u} \in \mathbb{S}^2.$$

In this case, $s(\mathbf{q})$ can be restricted to a spherical shell.

HARDI vs. DTI

Diffusion tensor imaging (DTI) is a well-established tool of diagnostic dMRI.

HARDI vs. DTI

Diffusion tensor imaging (DTI) is a well-established tool of diagnostic dMRI.

DTI

1. Unimodal Gaussian diffusion model is assumed.
2. Typical number of diffusion - encoding gradients is about 20 - 30.
3. Scan duration is about 5 - 10 mins.
4. *Incapable of discriminating multimodal diffusion flows within a voxel*.

HARDI vs. DTI

Diffusion tensor imaging (DTI) is a well-established tool of diagnostic dMRI.

DTI

1. Unimodal Gaussian diffusion model is assumed.
2. Typical number of diffusion - encoding gradients is about 20 - 30.
3. Scan duration is about 5 - 10 mins.
4. *Incapable of discriminating multimodal diffusion flows within a voxel*.

HARDI

1. No assumptions on the diffusion model are made.
2. Typical number of diffusion - encoding gradients is about 80 - 100.
3. Scan duration is about 20 - 30 mins.
4. *Can be used to delineate multimodal diffusion flows within a voxel*.
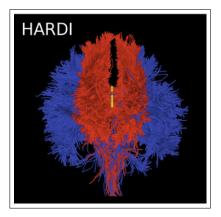
HARDI vs. DTI: Fibre tractography

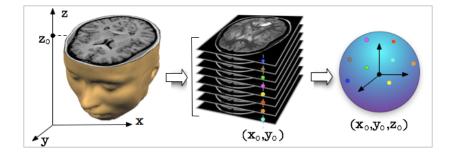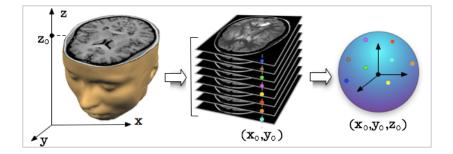**From**: J. Malcolm, O. Michailovich, S. Bouix, C.-F. Westin, A. Tannenbaum, M. Shenton and Y. Rathi, "A filtered approach to neural tractography using the Watson directional function," Medical Image Analysis, 14(1), 2009.

Thus, with $\mathbf{r} = (x, y, z) \in \Omega \subset \mathbb{R}^3$, HARDI measurements can be modelled as $s(\mathbf{u}, \mathbf{r}) : \mathbb{S}^2 \times \Omega \to \mathbb{R}^+$.

Thus, with $\mathbf{r} = (x, y, z) \in \Omega \subset \mathbb{R}^3$, HARDI measurements can be modelled as $s(\mathbf{u}, \mathbf{r}) : \mathbb{S}^2 \times \Omega \to \mathbb{R}^+$.

Fundamental practical limitation

The practical value of HARDI is greatly impaired by the problem of prohibitively long acquisition times.

The main questions to be addressed are:

Research questions

The main questions to be addressed are:

1. Are there tools of mathematical imaging which could be used to speed up HARDI acquisition so as to make its requirements comparable to those of DTI?

Research questions

The main questions to be addressed are:

1. Are there tools of mathematical imaging which could be used to speed up HARDI acquisition so as to make its requirements comparable to those of DTI?

2. To what extent the resulting approximation errors can affect the accuracy of subsequent diagnostic inference?

Multifiber mixture model

At each $\mathbf{r} \in \Omega$, the HARDI signal can be modelled as

$$s(\mathbf{u}, \mathbf{r}) = \sum_{i=1}^{M} \alpha_i(\mathbf{r}) \overbrace{\exp\left\{-b\left(\mathbf{u}^T D_i(\mathbf{r})\,\mathbf{u}\right)\right\}}^{s_i(\mathbf{u},\mathbf{r})}$$

Multifiber mixture model

At each $\mathbf{r} \in \Omega$, the HARDI signal can be modelled as

$$s(\mathbf{u}, \mathbf{r}) = \sum_{i=1}^{M} \alpha_i(\mathbf{r}) \overbrace{\exp\left\{-b\left(\mathbf{u}^T D_i(\mathbf{r})\,\mathbf{u}\right)\right\}}^{s_i(\mathbf{u}, \mathbf{r})}$$
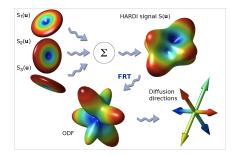
Multifiber mixture model

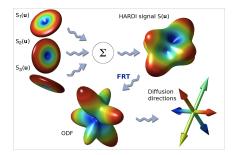At each $\mathbf{r} \in \Omega$, the HARDI signal can be modelled as

$$s(\mathbf{u}, \mathbf{r}) = \sum_{i=1}^{M} \alpha_i(\mathbf{r}) \overbrace{\exp\left\{-b\left(\mathbf{u}^T D_i(\mathbf{r})\,\mathbf{u}\right)\right\}}^{s_i(\mathbf{u},\mathbf{r})}$$

Observation

The energy of $s_i(\mathbf{u})$ is supported alongside the *great circles* of $\mathbb{S}^2$.

Spherical ridgelets

Central requirement

The $\mathbb{L}_2$-energy of HARDI signals can be efficiently "encoded" in terms of representation atoms, whose $\mathbb{L}_2$-energy is concentrated alongside the great circles of $\mathbb{S}^2$.

Spherical ridgelets

Central requirement

The $\mathbb{L}_2$-energy of HARDI signals can be efficiently "encoded" in terms of representation atoms, whose $\mathbb{L}_2$-energy is concentrated alongside the great circles of $\mathbb{S}^2$.

PLANAR RIDGELETS

SPHERICAL RIDGELETS

Spherical ridgelets: construction

Let $\kappa(x) = \exp\{-\rho\, x\,(x+1)\}$ be a Gaussian function, which we use to define:

$$\kappa_j(x) = \kappa(2^{-j}x) = \exp\left\{-\rho\,\frac{x}{2^j}\left(\frac{x}{2^j}+1\right)\right\}, \quad j \in \mathbb{N}.$$

Spherical ridgelets: construction

Let $\kappa(x) = \exp\{-\rho\, x\, (x+1)\}$ be a Gaussian function, which we use to define:

$$\kappa_j(x) = \kappa(2^{-j}x) = \exp\left\{-\rho\,\frac{x}{2^j}\left(\frac{x}{2^j}+1\right)\right\}, \quad j \in \mathbb{N}.$$

The *Gauss-Weierstrass scaling function* $\chi_{j,\mathbf{v}} : \mathbb{S}^2 \to \mathbb{R}$ at resolution $j \in \mathbb{N}$ and orientation $\mathbf{v} \in \mathbb{S}^2$ is defined as:

$$\chi_{j,\mathbf{v}}(\mathbf{u}) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\, \kappa_j(n)\, P_n(\mathbf{u} \cdot \mathbf{v}), \quad \forall \mathbf{u} \in \mathbb{S}^2,$$

where P_n is the Legendre polynomial of order n.

Spherical ridgelets: construction

Let $\kappa(x) = \exp\{-\rho\,x\,(x+1)\}$ be a Gaussian function, which we use to define:

$$\kappa_j(x) = \kappa(2^{-j}x) = \exp\left\{-\rho\,\frac{x}{2^j}\left(\frac{x}{2^j}+1\right)\right\}, \quad j \in \mathbb{N}.$$

The *Gauss-Weierstrass scaling function* $\chi_{j,\mathbf{v}} : \mathbb{S}^2 \to \mathbb{R}$ at resolution $j \in \mathbb{N}$ and orientation $\mathbf{v} \in \mathbb{S}^2$ is defined as:

$$\chi_{j,\mathbf{v}}(\mathbf{u}) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\,\kappa_j(n)\,P_n(\mathbf{u}\cdot\mathbf{v}), \quad \forall \mathbf{u} \in \mathbb{S}^2,$$

where P_n is the Legendre polynomial of order n.

Finally, the *spherical ridgelets* $\psi_{j,\mathbf{v}}$ are obtained from $\chi_{j,\mathbf{v}}$ according to:

$$\psi_{j,\mathbf{v}} = \frac{1}{2\pi}\mathcal{R}\left\{\chi_{j+1,\mathbf{v}} - \chi_{j,\mathbf{v}}\right\},$$

where $\mathcal{R}$ denotes the Funk-Radon transform.

Spherical ridgelets: construction (cont.)

Theorem

The semi-discrete set of spherical ridgelets $\{\psi_{j,\mathbf{v}}\}_{j\in\mathbb{N},\mathbf{v}\in\mathbb{S}^2}$ is a frame for the subspace $\mathcal{S}\in\mathbb{L}^2(\mathbb{S}^2)$ of symmetric spherical functions.

Spherical ridgelets: construction (cont.)

Theorem

The semi-discrete set of spherical ridgelets $\{\psi_{j,\mathbf{v}}\}_{j \in \mathbb{N}, \mathbf{v} \in \mathbb{S}^2}$ is a frame for the subspace $\mathcal{S} \in \mathbb{L}^2(\mathbb{S}^2)$ of symmetric spherical functions.

In practical computations, given a set of K diffusion-encoding orientations $\{\mathbf{u}_k\}_{k=1}^{K}$, the ridgelet frame is discretized to result in:

$$s(\mathbf{r}) = \Psi c(\mathbf{r}) + e(\mathbf{r}), \quad \forall \mathbf{r} \in \Omega,$$

Spherical ridgelets: construction (cont.)

Theorem

The semi-discrete set of spherical ridgelets $\{\psi_{j,\mathbf{v}}\}_{j\in\mathbb{N},\mathbf{v}\in\mathbb{S}^2}$ is a frame for the subspace $\mathcal{S}\in\mathbb{L}^2(\mathbb{S}^2)$ of symmetric spherical functions.

In practical computations, given a set of K diffusion-encoding orientations $\{\mathbf{u}_k\}_{k=1}^K$, the ridgelet frame is discretized to result in:

$$s(\mathbf{r}) = \Psi c(\mathbf{r}) + e(\mathbf{r}), \quad \forall \mathbf{r} \in \Omega,$$

where

- $s(\mathbf{r}) = [s(\mathbf{u}_1, \mathbf{r}), s(\mathbf{u}_2, \mathbf{r}), \ldots, s(\mathbf{u}_K, \mathbf{r})]^T$
- Ψ is the ridgelet (dictionary) matrix
- $c(\mathbf{r})$ is the vector of ridgelet representation coefficients
- $e(\mathbf{r})$ is a noise term

Reconstruction of HARDI signals

Bad news

- Noise contamination is typically severe.
- # ridgelets $\gg K$

Reconstruction of HARDI signals

Bad news

- Noise contamination is typically severe.
- # ridgelets $\gg K$

Good news

- The spherical ridgelets have a low coherence w.r.t. the Dirac sampling functions ($\mu \approx 0.56$).
- Spherical ridgelets provide sparse representation of HARDI signals (only $6 \div 8$ ridgelets are needed on average).

Reconstruction of HARDI signals

Bad news

- Noise contamination is typically severe.
- $\#$ ridgelets $\gg K$

Good news

- The spherical ridgelets have a low coherence w.r.t. the Dirac sampling functions ($\mu \approx 0.56$).
- Spherical ridgelets provide sparse representation of HARDI signals (only $6 \div 8$ ridgelets are needed on average).

Therefore, one can try to recover $c(\mathbf{r})$ through:

$$\min_{c(\mathbf{r})} \|c(\mathbf{r})\|_1$$
$$\text{s.t.} \quad \|\Psi c(\mathbf{r}) - s(\mathbf{r})\|_2 \leq \epsilon$$
$$\Psi c(\mathbf{r}) \succeq 0$$

which needs to be solved at each $\mathbf{r} \in \Omega$ independently.

Composite compressed sensing

General idea

Combine the sparse constraints in the diffusion domain (**u**) with a smoothness constraint in the spatial domain (**r**).

Composite compressed sensing

General idea

Combine the sparse constraints in the diffusion domain ($\mathbf{u}$) with a smoothness constraint in the spatial domain ($\mathbf{r}$).

To this end, we define:

$$\bar{s} = [s(\mathbf{r}_1), s(\mathbf{r}_2), \ldots, s(\mathbf{r}_{\#\Omega})]^T$$

$$\bar{c} = [c(\mathbf{r}_1), c(\mathbf{r}_2), \ldots, c(\mathbf{r}_{\#\Omega})]^T$$

$$\mathcal{A} = \left[\begin{array}{cccc} \Psi & 0 & \ldots & 0 \\ 0 & \Psi & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \Psi \end{array}\right]$$

and

$$\|\mathcal{A}\bar{c}\|_{TV} = \sum_{k=0}^{K-1} \|\mathcal{D}_k\{\mathcal{A}\bar{c}\}\|_{TV}$$

where $\mathcal{D}_k\{\bar{s}\}[n] = \bar{s}[Kn + k]$ is a subsampling operator.

Composite compressed sensing (cont.)

The spatially constrained CS problem can be now defined as

$$\min_{\bar{c}} \left\{ \|\bar{c}\|_1 + \mu \|\mathcal{A}\bar{c}\|_{TV} \right\}$$
$$\text{s.t.} \quad \|\mathcal{A}\bar{c} - \bar{s}\|_2 \leq \epsilon$$
$$\mathcal{A}\bar{c} \succeq 0$$

Composite compressed sensing (cont.)

The spatially constrained CS problem can be now defined as

$$\min_{\bar{c}} \left\{ \|\bar{c}\|_1 + \mu \|\mathcal{A}\bar{c}\|_{TV} \right\}$$
$$\text{s.t.} \quad \|\mathcal{A}\bar{c} - \bar{s}\|_2 \leq \epsilon$$
$$\mathcal{A}\bar{c} \succeq 0$$

or, equivalently, in the Lagrangian form as

$$\min_{\bar{c}} \left\{ \frac{1}{2} \|\mathcal{A}\bar{c} - \bar{s}\|_2^2 + \lambda \|\bar{c}\|_1 + \mu \|\mathcal{A}\bar{c}\|_{TV} + i_C(\mathcal{A}\bar{c}) \right\}.$$

Composite compressed sensing (cont.)

The spatially constrained CS problem can be now defined as

$$\min_{\bar{c}} \left\{ \|\bar{c}\|_1 + \mu\|\mathcal{A}\bar{c}\|_{TV} \right\}$$
$$\text{s.t.} \quad \|\mathcal{A}\bar{c} - \bar{s}\|_2 \leq \epsilon$$
$$\mathcal{A}\bar{c} \succeq 0$$

or, equivalently, in the Lagrangian form as

$$\min_{\bar{c}} \left\{ \frac{1}{2}\|\mathcal{A}\bar{c} - \bar{s}\|_2^2 + \lambda\|\bar{c}\|_1 + \mu\|\mathcal{A}\bar{c}\|_{TV} + i_C(\mathcal{A}\bar{c}) \right\}.$$

Alternatively, one can solve

$$\min_{\bar{c},\bar{u}} \left\{ \frac{1}{2}\|\bar{u} - \bar{s}\|_2^2 + \lambda\|\bar{c}\|_1 + \mu\|\bar{u}\|_{TV} + i_C(\bar{u}) \right\}$$
$$\text{s.t.} \quad \mathcal{A}\bar{c} = \bar{u}$$

Solution by means of ADMM

The above problem can be solved iteratively as

$$
\left(\bar{u}^{t+1}, \bar{c}^{t+1}\right) =
$$
$$
\arg\min_{\bar{c}, \bar{u}} \left\{ \frac{1}{2}\|\bar{u} - \bar{s}\|_2^2 + \lambda \|\bar{c}\|_1 + \mu \|\bar{u}\|_{TV} + \frac{\gamma}{2}\|\bar{u} - \mathcal{A}\bar{c} - p^t\|_2^2 \right\}
$$
$$
p^{t+1} = p^t + \left(\mathcal{A}\bar{c}^{t+1} - \bar{u}^{t+1}\right)
$$

Solution by means of ADMM

The above problem can be solved iteratively as

$$
\left(\bar{u}^{t+1}, \bar{c}^{t+1}\right) =
$$
$$
\arg\min_{\bar{c},\bar{u}} \left\{ \frac{1}{2}\|\bar{u} - \bar{s}\|_2^2 + \lambda\,\|\bar{c}\|_1 + \mu\,\|\bar{u}\|_{TV} + \frac{\gamma}{2}\|\bar{u} - \mathcal{A}\bar{c} - p^t\|_2^2 \right\}
$$
$$
p^{t+1} = p^t + \left(\mathcal{A}\bar{c}^{t+1} - \bar{u}^{t+1}\right)
$$

Finally, splitting the variables results in

$$
\text{Step 1: } \bar{c}^{t+1} = \arg\min_{\bar{c}} \left\{ \frac{1}{2}\|\mathcal{A}\bar{c} - \bar{d}^t\|_2^2 + \alpha\|\bar{c}\|_1 \right\}
$$
$$
\text{Step 2: } \bar{u}^{t+1} = \arg\min_{\bar{c}} \left\{ \frac{1}{2}\|\bar{u} - \bar{d}^{t+1}\|_2^2 + \beta\|\bar{u}\|_{TV} \right\}
$$
$$
\text{Step 3: } p^{t+1} = p^t + \left(\mathcal{A}\bar{c}^{t+1} - \bar{u}^{t+1}\right)
$$

Numerical aspects

- Step 1 is *separable in the diffusion variable*, and hence it can be executed in a "voxel-by-voxel" manner (using, e.g., FISTA)

Numerical aspects

- Step 1 is *separable in the diffusion variable*, and hence it can be executed in a "voxel-by-voxel" manner (using, e.g., FISTA)
- Step 2 is *separable in the spatial variable*, and hence it can be executed in a "slice-by-slice" manner (using, e.g., Chambolle's algorithm).

Numerical aspects
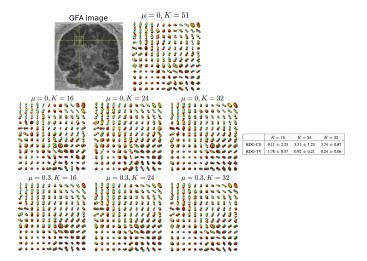
- Step 1 is *separable in the diffusion variable*, and hence it can be executed in a "voxel-by-voxel" manner (using, e.g., FISTA)
- Step 2 is *separable in the spatial variable*, and hence it can be executed in a "slice-by-slice" manner (using, e.g., Chambolle's algorithm).
- The overall computational load scales proportionally with the number of processing cores.

In vivo experiments

GFA image

$\mu = 0, K = 51$

$\mu = 0, K = 16$ $\mu = 0, K = 24$ $\mu = 0, K = 32$

$\mu = 0.3, K = 16$ $\mu = 0.3, K = 24$ $\mu = 0.3, K = 32$

	$K = 16$	$K = 24$	$K = 32$
RDG-CS	9.11 ± 2.23	5.31 ± 1.23	3.74 ± 0.87
RDG-TV	1.78 ± 0.37	0.92 ± 0.21	0.24 ± 0.06

O. Michailovich, Y. Rathi and S. Dolui, "Spatially regularized compressed sensing for high angular resolution diffusion imaging," IEEE Transactions on Medical Imaging, 30(5), 2011.

ODF revised

• Given a recovered signal $s(\mathbf{u}, \mathbf{r})$, its corresponding ODF $Q(\mathbf{u}, \mathbf{r})$ can be computed as (Aganj *et al*, 2010):

$$Q(\mathbf{u}, \mathbf{r}) = \frac{1}{4\pi} + \frac{1}{16\pi^2} \mathcal{R} \left\{ \nabla_b^2 \ln(-\ln s(\mathbf{u}, \mathbf{r})) \right\}$$

where ∇_b^2 stands for the Laplace-Beltrami operator.

ODF revised

• Given a recovered signal $s(\mathbf{u}, \mathbf{r})$, its corresponding ODF $Q(\mathbf{u}, \mathbf{r})$ can be computed as (Aganj $et~al$, 2010):

$$Q(\mathbf{u}, \mathbf{r}) = \frac{1}{4\pi} + \frac{1}{16\pi^2} \mathcal{R} \left\{ \nabla_b^2 \ln(-\ln s(\mathbf{u}, \mathbf{r})) \right\}$$

where ∇_b^2 stands for the Laplace-Beltrami operator.

• In practice, $\mathbf{r}$ belongs to a discrete set Ω:

$$\Omega = \left\{ \mathbf{r}_i = (x_i, y_i, z_i) \in \mathbb{R}^3 \mid i \in \mathcal{I} \right\}.$$

ODF revised

• Given a recovered signal $s(\mathbf{u}, \mathbf{r})$, its corresponding ODF $Q(\mathbf{u}, \mathbf{r})$ can be computed as (Aganj *et al*, 2010):

$$Q(\mathbf{u}, \mathbf{r}) = \frac{1}{4\pi} + \frac{1}{16\pi^2} \mathcal{R} \left\{ \nabla_b^2 \ln(-\ln s(\mathbf{u}, \mathbf{r})) \right\}$$

where ∇_b^2 stands for the Laplace-Beltrami operator.

• In practice, $\mathbf{r}$ belongs to a discrete set Ω:

$$\Omega = \left\{ \mathbf{r}_i = (x_i, y_i, z_i) \in \mathbb{R}^3 \mid i \in \mathcal{I} \right\}.$$

• Denoting $Q(\mathbf{u}, \mathbf{r}_i) = Q_i(\mathbf{u})$, we refer to the pair

$$\mathbf{D}_\Omega = (\{\mathbf{r}_i\}_{i \in \mathcal{I}}, \{Q_i(\mathbf{u})\}_{i \in \mathcal{I}})$$

as a *directional diffusion structure* (DDS).

Graph representation of DDS

• Formally, the DDS $\mathbf{D}_\Omega$ is a subset of the probability manifold $\mathbb{R}^3 \times \mathcal{P}$, where $\mathcal{P}$ is a set of probability densities on $\mathbb{S}^2$.

Graph representation of DDS

• Formally, the DDS $\mathbf{D}_\Omega$ is a subset of the probability manifold $\mathbb{R}^3 \times \mathcal{P}$, where $\mathcal{P}$ is a set of probability densities on $\mathbb{S}^2$.

• $\mathbf{D}_\Omega$ is a high-dimensional space, and it is therefore imperative to find a lower dimensional representation of $\mathbf{D}_\Omega$ in *a linear metric space*.

Graph representation of DDS

• Formally, the DDS $\mathbf{D}_\Omega$ is a subset of the probability manifold $\mathbb{R}^3 \times \mathcal{P}$, where $\mathcal{P}$ is a set of probability densities on $\mathbb{S}^2$.

• $\mathbf{D}_\Omega$ is a high-dimensional space, and it is therefore imperative to find a lower dimensional representation of $\mathbf{D}_\Omega$ in *a linear metric space*.

• Such a representation should:

1. Preserve the directivity information contained in $\mathbf{D}_\Omega$
2. Be invariant under Euclidean transformations

Graph representation of DDS

• Formally, the DDS $\mathbf{D}_\Omega$ is a subset of the probability manifold $\mathbb{R}^3 \times \mathcal{P}$, where $\mathcal{P}$ is a set of probability densities on $\mathbb{S}^2$.

• $\mathbf{D}_\Omega$ is a high-dimensional space, and it is therefore imperative to find a lower dimensional representation of $\mathbf{D}_\Omega$ in *a linear metric space*.

• Such a representation should:

1. Preserve the directivity information contained in $\mathbf{D}_\Omega$
2. Be invariant under Euclidean transformations

• To this end, we first transform $\mathbf{D}_\Omega$ into a discrete metrizable manifold.

Graph representation of DDS (cont.)

• $\mathbf{D}_\Omega$ can be associated with an undirected weighted graph $G_\omega = (V, E)$, with no self-loops and no multiple edges.

Graph representation of DDS (cont.)

• $\mathbf{D}_\Omega$ can be associated with an undirected weighted graph $G_\omega = (V, E)$, with no self-loops and no multiple edges.

• Each vertex $v_i \in V$ of G_ω is related to $\mathbf{r}_i \in \Omega$, while the connectivity on G_ω is defined by means of its weights:

$$
\omega_{i,j} = \begin{cases}
d_Q(Q_i, Q_j), & \text{if } \|\mathbf{r}_i - \mathbf{r}_j\|_2 = \Delta \\
\sqrt{2}\, d_Q(Q_i, Q_j), & \text{if } \|\mathbf{r}_i - \mathbf{r}_j\|_2 = \sqrt{2}\Delta \\
+\infty, & \text{otherwise}
\end{cases}
$$

where Δ is a spatial resolution parameter.

Graph representation of DDS (cont.)

• $\mathbf{D}_\Omega$ can be associated with an undirected weighted graph $G_\omega = (V, E)$, with no self-loops and no multiple edges.

• Each vertex $v_i \in V$ of G_ω is related to $\mathbf{r}_i \in \Omega$, while the connectivity on G_ω is defined by means of its weights:

$$\omega_{i,j} = \begin{cases} d_Q(Q_i, Q_j), & \text{if } \|\mathbf{r}_i - \mathbf{r}_j\|_2 = \Delta \\ \sqrt{2}\, d_Q(Q_i, Q_j), & \text{if } \|\mathbf{r}_i - \mathbf{r}_j\|_2 = \sqrt{2}\Delta \\ +\infty, & \text{otherwise} \end{cases}$$

where Δ is a spatial resolution parameter.

• Since physiologically significant regions within the brain are topologically connected, G_ω can be endowed with a metric:

$$d_{G_\omega}(u, v) = \inf \{l(u, v)\},$$

where $l(u, v) = \sum_{k=1}^{n} \omega_{k-1,k}$ is the length of a path connecting u and v.

• To finalize the definition of G_ω as a metric space, the distance $d_Q(Q_i, Q_j)$ is defined to be the Jensen-Shannon divergence (aka information radius):

$$d_\psi(Q_i, Q_j) =$$
$$= \left[\frac{1}{2} \int_{\mathbb{S}^2} Q_i(\mathbf{u}) \ln \frac{2\,Q_i(\mathbf{u})}{Q_i(\mathbf{u}) + Q_j(\mathbf{u})} \, d\eta(\mathbf{u}) + \right.$$
$$\left. + \frac{1}{2} \int_{\mathbb{S}^2} Q_j(\mathbf{u}) \ln \frac{2\,Q_j(\mathbf{u})}{Q_i(\mathbf{u}) + Q_j(\mathbf{u})} \, d\eta(\mathbf{u}) \right]^{1/2}$$

• To finalize the definition of G_ω as a metric space, the distance $d_Q(Q_i, Q_j)$ is defined to be the Jensen-Shannon divergence (aka information radius):

$$d_\psi(Q_i, Q_j) =$$
$$= \left[\frac{1}{2} \int_{\mathbb{S}^2} Q_i(\mathbf{u}) \ln \frac{2 Q_i(\mathbf{u})}{Q_i(\mathbf{u}) + Q_j(\mathbf{u})} \, d\eta(\mathbf{u}) + \right.$$
$$\left. + \frac{1}{2} \int_{\mathbb{S}^2} Q_j(\mathbf{u}) \ln \frac{2 Q_j(\mathbf{u})}{Q_i(\mathbf{u}) + Q_j(\mathbf{u})} \, d\eta(\mathbf{u}) \right]^{1/2}$$

• Note that the Jensen-Shannon divergence defines a metric on the space of spherical probability densities $\mathcal{P}$.

Isometric embedding via MDS

• Given a DDS $\mathbf{D}_\Omega$ and its associated graph representation G_ω, one can use d_{G_ω} to compute the (geodesic) distances between every pair of vertices in G_ω.

Isometric embedding via MDS

• Given a DDS $\mathbf{D}_\Omega$ and its associated graph representation G_ω, one can use d_{G_ω} to compute the (geodesic) distances between every pair of vertices in G_ω.

• These distances can be arranged into an $N \times N$ matrix

$$\delta = \{\delta_{i,j} = d_{G_\omega}(v_i, v_j)\}$$

with $N = \#\Omega$.

Isometric embedding via MDS

• Given a DDS $\mathbf{D}_\Omega$ and its associated graph representation G_ω, one can use d_{G_ω} to compute the (geodesic) distances between every pair of vertices in G_ω.

• These distances can be arranged into an $N \times N$ matrix

$$\delta = \{\delta_{i,j} = d_{G_\omega}(v_i, v_j)\}$$

with $N = \#\Omega$.

• A lower dimensional representation of $\mathbf{D}_\Omega$ can be found by means of *multidimensional scaling* (MDS) which finds a configuration of N points $X_N = \{\mathbf{t}_k\}_{k=1}^N$ in, e.g., $\mathbb{R}^d$ that minimizes the stress

$$\mathcal{E}(X_N) = \sum_{i<j}(\|\mathbf{t}_i - \mathbf{t}_j\|^2 - \delta_{i,j})^2$$

Isometric embedding via MDS

• Given a DDS $\mathbf{D}_\Omega$ and its associated graph representation G_ω, one can use d_{G_ω} to compute the (geodesic) distances between every pair of vertices in G_ω.

• These distances can be arranged into an $N \times N$ matrix

$$\delta = \{\delta_{i,j} = d_{G_\omega}(v_i, v_j)\}$$

with $N = \#\Omega$.

• A lower dimensional representation of $\mathbf{D}_\Omega$ can be found by means of *multidimensional scaling* (MDS) which finds a configuration of N points $X_N = \{\mathbf{t}_k\}_{k=1}^N$ in, e.g., $\mathbb{R}^d$ that minimizes the stress

$$\mathcal{E}(X_N) = \sum_{i<j}(\|\mathbf{t}_i - \mathbf{t}_j\|^2 - \delta_{i,j})^2$$

• In this work, we use $d = 3$.

Euclidean invariant biomarkers

• The low-dimensional representation X_N of a DDS can be characterized by its moments

$$\xi_{p,q,r} = \frac{1}{N} \sum_{k=1}^{N} (\mathbf{t}_k^1)^p (\mathbf{t}_k^2)^q (\mathbf{t}_k^3)^r,$$

with $p+q+r \leq P$ and the first-order moments $\xi_{0,0,1}$, $\xi_{0,1,0}$ and $\xi_{1,1,0}$ set to zero.

Euclidean invariant biomarkers

- The low-dimensional representation X_N of a DDS can be characterized by its moments

$$\xi_{p,q,r} = \frac{1}{N} \sum_{k=1}^{N} (\mathbf{t}_k^1)^p (\mathbf{t}_k^2)^q (\mathbf{t}_k^3)^r,$$

with $p+q+r \leq P$ and the first-order moments $\xi_{0,0,1}$, $\xi_{0,1,0}$ and $\xi_{1,1,0}$ set to zero.

- Given two DDSs $\mathbf{D}_\Omega^1$ and $\mathbf{D}_\Omega^2$ and their corresponding moments $\{\xi_{p,q,r}^1\}$ and $\{\xi_{p,q,r}^2\}$, the distance between these structures can be defined to be:

$$d_{\mathbf{D}}(\mathbf{D}_\Omega^1, \mathbf{D}_\Omega^2) = \sqrt{\sum_{p+q+r \leq P} \left| \xi_{p,q,r}^1 - \xi_{p,q,r}^2 \right|^2}$$

Euclidean invariant biomarkers

• The low-dimensional representation X_N of a DDS can be characterized by its moments

$$\xi_{p,q,r} = \frac{1}{N} \sum_{k=1}^{N} (\mathbf{t}_k^1)^p (\mathbf{t}_k^2)^q (\mathbf{t}_k^3)^r,$$

with $p+q+r \leq P$ and the first-order moments $\xi_{0,0,1}$, $\xi_{0,1,0}$ and $\xi_{1,1,0}$ set to zero.

• Given two DDSs $\mathbf{D}_\Omega^1$ and $\mathbf{D}_\Omega^2$ and their corresponding moments $\{\xi_{p,q,r}^1\}$ and $\{\xi_{p,q,r}^2\}$, the distance between these structures can be defined to be:

$$d_{\mathbf{D}}(\mathbf{D}_\Omega^1, \mathbf{D}_\Omega^2) = \sqrt{\sum_{p+q+r \leq P} \left| \xi_{p,q,r}^1 - \xi_{p,q,r}^2 \right|^2}$$

• In this work, P is set to be equal to 3, which results in a total of 16 moments of the 2nd and 3rd orders.
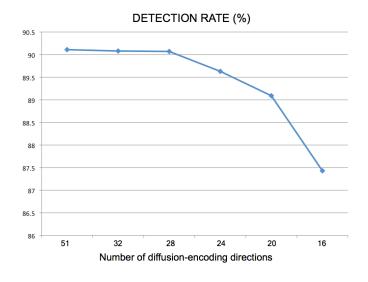
Diagnosis of FES

1. The subject pool consisted of 20 FE patients (16 males, 4 females, average age: 21.21 ± 4.56 years) and 20 normal controls (15 males, 5 females, average age: 22.47 ± 3.48 years) with the p-value for age being 0.34.

Diagnosis of FES

1. The subject pool consisted of 20 FE patients (16 males, 4 females, average age: 21.21 ± 4.56 years) and 20 normal controls (15 males, 5 females, average age: 22.47 ± 3.48 years) with the *p*-value for age being 0.34.

2. HARDI data was acquired using a 3T MRI scanner with spatial resolution of $1.667 \times 1.667 \times 1.7$ mm^3 and a *b*-value of 900 s/mm^2.

Diagnosis of FES

1. The subject pool consisted of 20 FE patients (16 males, 4 females, average age: 21.21 ± 4.56 years) and 20 normal controls (15 males, 5 females, average age: 22.47 ± 3.48 years) with the p-value for age being 0.34.

2. HARDI data was acquired using a 3T MRI scanner with spatial resolution of $1.667 \times 1.667 \times 1.7$ mm^3 and a b-value of 900 s/mm^2.

3. For each subject, expert segmentation was carried out to identify the brain regions corresponding to the left (Ω_L^i) and right (Ω_R^i) centrum semiovale (with $1 \leq i \leq 20$).

Diagnosis of FES

1. The subject pool consisted of 20 FE patients (16 males, 4 females, average age: 21.21 ± 4.56 years) and 20 normal controls (15 males, 5 females, average age: 22.47 ± 3.48 years) with the p-value for age being 0.34.

2. HARDI data was acquired using a 3T MRI scanner with spatial resolution of $1.667 \times 1.667 \times 1.7$ mm^3 and a b-value of 900 s/mm^2.

3. For each subject, expert segmentation was carried out to identify the brain regions corresponding to the left (Ω_L^i) and right (Ω_R^i) centrum semiovale (with $1 \leq i \leq 20$).

4. For each subject in the FE and normal control groups, the pairwise distances between $X_{N_i}^L$ and $X_{N_i}^R$ were computed and used as a diagnostic biomarker.

DETECTION RATE (%)

Number of diffusion-encoding directions

Conclusions

- The sparsifying properties of spherical ridgelets are crucial for the CS-based reconstruction of HARDI signals.
- Adding the spatial regularization and positivity constraints can substantially improve the reconstruction accuracy.
- HARDI can be performed using as few diffusion gradients as it is required by a standard DTI (i.e. $K = 20 \div 30$).
- A new method for low-dimensional representation of HARDI signals based on isometric embedding was formulated.
- It was demonstrated that the performance of the method remains reliable, when the sampling density is reduced by a factor of 4 with respect to its conventional value.