

Segmenting Thalamic Nuclei: What Can We Gain from HARDI?*

Thomas Schultz

Computation Institute, University of Chicago, Chicago IL, USA

Abstract. The contrast provided by diffusion MRI has been exploited repeatedly for in vivo segmentations of thalamic nuclei. This paper systematically investigates the benefits of high-angular resolution (HARDI) data for this purpose. An empirical analysis of clustering stability reveals a clear advantage of acquiring HARDI data at $b = 1000$ s/mm². However, based on stability arguments, as well as further visual and statistical evidence and theoretical insights about the impact of parameters, HARDI models such as the q-ball do not exhibit clear benefits over the standard diffusion tensor for thalamus segmentation at this b value.

1 Introduction

Since Wiegell et al. [1] demonstrated that the striations within the thalamic nuclei provide sufficient contrast in diffusion MRI to allow for an automated segmentation, several groups have proposed algorithmic methods to segment these nuclei based on the diffusion tensor (DTI) model [2–5]. More recently, segmentations of the thalamic nuclei have been achieved using high angular resolution (HARDI) data and the q-ball model [6, 7]. However, we are not aware of any studies that investigate the relative benefit from using HARDI models for this particular task, or provide guidelines on which model is most suitable for which measurement setup. It is the goal of our investigation to find out what can be gained from using HARDI data and specific HARDI models for this task.

Studying this question is essential to further refine methods for thalamus segmentation, but it is complicated by the fact that final segmentations do not only depend on the chosen model, but also on the algorithmic method and spatial regularization. Moreover, it is certainly possible to verify the overall plausibility of a segmentation by comparing it to known anatomy [1–7], but a lack of exact and reliable ground truth for individual subjects makes it difficult to argue about the relative validity in case of more subtle differences.

Therefore, we only turn to actual segmentations after a visual (Section 2.1) and statistical (Section 2.3) investigation of distance measures that result from different models, and a theoretical analysis of their parameters (Section 2.2). As

* I would like to thank Alfred Anwander (MPI CBS, Leipzig, Germany) for providing data and important feedback on this work. Gordon Kindlmann (University of Chicago) contributed through productive and inspiring discussions. This work has been supported by a fellowship within the DAAD Postdoc Program.

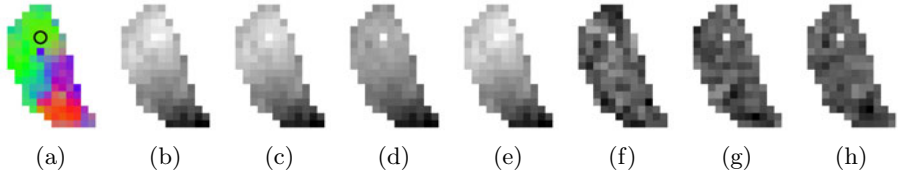


Fig. 1. Distances from a reference voxel (a) show a large agreement between the diffusion tensor deviatoric (b) and the q-ball model (c). The Sobolev norm (d) weights the higher orders (f)–(h) stronger than the second order (e).

an objective way of judging segmentations, we then perform an analysis of their stability under small perturbations of the input (Sections 2.4 and 3).

2 Evaluating Benefits of the Q-Ball Model

2.1 Distance Maps

Measures of dissimilarity between the values in different voxels are at the center of all segmentation approaches, so it is a natural first step to study their behavior.

Figure 1 shows an axial slice through the right thalamus in one of our subjects. Our data has 60 gradient directions at $b = 1000 \text{ s/mm}^2$, with 1.72 mm isotropic voxel size. Distances to a reference voxel, marked by a circle in (a), are linearly mapped to grayscale, from zero to the largest value in each map. To reduce the impact of partial voluming with the adjacent ventricles, we consider the diffusion tensor deviatoric $\tilde{\mathbf{D}} = \mathbf{D} - \text{tr}(\mathbf{D})/3\mathbf{I}$ instead of the full tensors \mathbf{D} (tr is matrix trace and \mathbf{I} is the identity).

The Frobenius norm $\|\mathbf{D}\| = \sqrt{\text{tr}(\mathbf{D}^T\mathbf{D})}$ of differences in $\tilde{\mathbf{D}}$ (Fig. 1 (b), range $[0, 7.91 \times 10^{-3} \text{ mm}^2/\text{s}]$) produces a similar map as the ℓ_2 norm of q-balls (c, $[0, 0.48]$), modeled by order-8 spherical harmonics (SHs). The Sobolev norm [7] (d, $[0, 0.81]$) is designed to be more sensitive to the alignment of peaks in the q-ball than the ℓ_2 norm. Its map reveals similar overall structure, with highest distances at the posterior end of the thalamus, but exhibits reduced contrast.

2.2 Influence of Model Parameters

We have used Laplace-Beltrami regularization for the q-balls [8], which introduces a regularization parameter ($\lambda = 0.008$). The Sobolev norm involves three additional parameters ($\alpha = 1$, $t = 0$, $\gamma = 0.21$). Even though we carefully calibrated these values according to the synthetic model proposed in [7], modified to match our experimental setup, it is informative to consider how our result depends on their exact choice.

The Sobolev norm amounts to a weighted ℓ_2 norm, in which higher orders are weighted more strongly [7]. This motivates mapping the ℓ_2 distances of the SH coefficients of orders 2 (Fig. 1 (e), $[0, 0.47]$), 4 (f, $[0, 0.062]$), 6 (g, $[0, 0.026]$) and 8 (h, $[0.011]$) separately. Since q-balls integrate to unity by definition [9], there is

no variability in order zero. All parameters of the Sobolev norm lead to squared distances that are linear combinations of the squared values from Fig. 1 (e)–(h).

We observed that the effect of different regularization parameters λ is also well-approximated by a re-weighting of different SH orders. This is explained by the fact that with Laplace-Beltrami regularization, the SH coefficient vector $\mathbf{c} \in \mathbb{R}^C$ is obtained from the diffusion-weighted signal vector $\mathbf{s} \in \mathbb{R}^N$ via

$$\mathbf{c} = (\mathbf{B}^T \mathbf{B} + \lambda \mathbf{L})^{-1} \mathbf{B}^T \mathbf{s}, \quad (1)$$

where the (i, j) th element of $\mathbf{B} \in \mathbb{R}^{N \times C}$ is computed by evaluating the SH basis function Y_j at the position (θ_i, ϕ_i) of gradient vector i . The diagonal matrix $\mathbf{L} \in \mathbb{R}^{C \times C}$ has entries $l_j^2(l_j + 1)^2$, l_j being the order of SH coefficient c_j . In the interest of space, the reader is referred to [8] for details on this notation.

Typical HARDI acquisition schemes achieve a near-uniform distribution of gradient directions over the (hemi-)sphere. In this case, up to a factor of $4\pi/N$, dot products between the columns of \mathbf{B} numerically approximate the integral

$$\langle S^{(1)}, S^{(2)} \rangle = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} S^{(1)}(\theta, \phi) S^{(2)}(\theta, \phi) \sin \theta d\phi d\theta, \quad (2)$$

which defines a scalar product for real-valued functions on the sphere. Given the orthonormality of the SH basis functions from [8] with respect to (2), this leads to the approximation $\mathbf{B}^T \mathbf{B} \approx N/(4\pi) \mathbf{I}$.

Consequently, $(\mathbf{B}^T \mathbf{B} + \lambda \mathbf{L})^{-1}$ in Eq. (1) is approximated by a diagonal matrix with entries $(N/(4\pi) + \lambda l_j^2(l_j + 1)^2)^{-1}$. Thus, the regularized $\tilde{c}_{l,m}$ can be approximated by scaling the unregularized $c_{l,m}$,

$$\tilde{c}_{l,m} \approx c_{l,m} \times \left(1 + \frac{4\pi \lambda l^2(l+1)^2}{N} \right)^{-1}. \quad (3)$$

In the ℓ_2 norm $\|\tilde{\mathbf{c}}\|$, this is reflected by a re-weighting of the terms $c_{l,m}$ by a function of SH order l and the regularization parameter λ .

Even replacing the q-balls by a different HARDI model, spherical deconvolution [10], would only amount to a re-weighting of different SH orders in the ℓ_2 norm. The reason is that linear deconvolution is performed by scaling the SH coefficients of the diffusion weighted signal by a function of SH order, which depends on the fiber response function [10]. Analytically evaluating the Funk-Radon transform, which relates the diffusion-weighted signal to the q-ball, reveals that it can be expressed in the same way [8].

These insights illustrate the high relevance of the maps in Fig. 1 (e)–(h) to our problem: They serve as linear building blocks of the distance maps that would result from q-ball or spherical deconvolution models, combined with either the ℓ_2 or the Sobolev norm. We cannot expect any structures that are not present in these maps to be brought forward by changing the parameters of these models.

Visually, the spatial structure of the second order map (e) largely agrees with the structure from the diffusion tensor deviatoric in (b). It is more difficult to discern structure in the higher orders. In this respect, the thalamus appears to differ from white matter, where higher orders carry meaningful structure in regions of fiber crossings, even at moderate b -values [11].

Table 1. A statistical investigation of distance measures (separately in the left/right part of the thalamus) supports the visual impression that distances at spherical harmonics order 2 expose a much clearer spatial structure than at higher orders.

	order 2	order 4	order 6	order 8
\bar{d}_A^2/\bar{d}_N^2	8.17/8.19	1.83/1.75	1.63/1.58	1.60/1.58
skewness of d_N^2	2.82/2.65	1.87/1.50	1.30/1.45	1.19/1.26

2.3 Distance Statistics

To quantify the visual impression that the distances at higher orders reveal much less spatial structure than distances at SH order two, we have computed average squared distances \bar{d}_A^2 between all pairs of voxels (in a full 3D thalamus mask, for the left and right side separately), and compared them to average squared distances \bar{d}_N^2 between face neighbors. If the voxels contained noise without any spatial structure, the expected ratio were $\bar{d}_A^2/\bar{d}_N^2 = 1$. Piecewise smooth or piecewise constant data with clear spatial structure should lead to $\bar{d}_A^2/\bar{d}_N^2 \gg 1$. Table 1 shows that \bar{d}_A^2/\bar{d}_N^2 is much larger for order 2 than it is for higher orders.

We have also considered the distribution of d_N^2 , the squared distances between face neighbors. In case of homogeneous regions that are separated by clear boundaries, we expect most distances to be small (within regions), with a heavy tail of much larger distances from boundaries. This should lead to a positive skew in the distribution. In fact, Table 1 reveals such a skew, most pronounced in the second order distances.

2.4 Stability Analysis

Even though the distance maps from DTI and q-balls look similar, we found that the resulting segmentations differ to some degree. In order to assess whether these differences are significant, we compute the adjusted RAND index [12] between the segmentations to quantify their overlap, and compare it to the overlap of segmentations from the same model, but slightly perturbed data.

It is widely accepted that successful segmentation of thalamic nuclei requires some sort of regularization that favors spatially connected structures. All methods we are aware of include such regularization, either by combining the data-based distance with a spatial distance [1, 3, 6, 7], through coupling forces in a level set framework [4], by Markovian relaxation [2], or by learning an atlas from the joint segmentation of multiple subjects [5].

Unfortunately, spatial regularization might lead to paradoxical results in a stability analysis: If a model does not reveal clear structure in the data, the resulting segmentation might be dominated by spatial information and consequently appear more stable than an alternative model that does reveal (somewhat fragile) structure. Such misleading results can be difficult to detect, even when checking the plausibility of the obtained segmentations.

To illustrate this, Fig. 2 (a) reproduces the segmentation presented by Grassi et al. [6] based on our own data. We use k -means with a linear combination of



Fig. 2. Given an individual segmentation result (a), it can be difficult to judge the relative influence of data-based (b) and spatial distances (c). The same slice is shown as in Fig. 1, rotated clockwise by 90° .

Euclidean spatial distance and the ℓ_2 distance between q-balls, giving the spatial distance the minimum weight required to obtain connected components. As in [6], an axial slice of the thalamus is segmented into a pulvinar, a lateral, a medial, and an anterior region. It is reassuring that a very similar result is obtained without spatial regularization (Fig. 2 (b)). However, even if we do not make any use of the data (Fig. 2 (c)), spatial distances alone lead to a segmentation that looks similar enough that it might pass a superficial test of plausibility.

Therefore, we run k -means based on the diffusion information alone, on the right thalamus (in 3D) and for different numbers of regions (2–7). Since the result of k -means may depend on the initialization, we perform each clustering ten times, from randomly selected seeds, and keep the result with the lowest sum of squared distances of data points to their cluster centers. In order to create slightly perturbed model fits without changing the noise characteristics of the data, we obtained 150 perturbed models via weighted least squares, where the weights of the 60 available directions were sampled uniformly at random from $[0, 1]$. The average relative distance $\|\mathbf{D}_p - \mathbf{D}\|/\|\mathbf{D}\|$ of a diffusion tensor \mathbf{D}_p that has been perturbed in this way to its unperturbed version \mathbf{D} is around 3%.

The solid lines in Fig. 3 (a) plot the mean adjusted RAND index when comparing segmentations of perturbed data to their unperturbed counterparts. Results from the diffusion tensor are marked with circles, q-ball with the ℓ_2 norm by squares, and q-ball with Sobolev norm by triangles. None of the models consistently produced more stable results than the others, but we noticed a steep drop in stability for more than five regions, especially when using q-balls.

The dashed lines in the same Figure plot the mean adjusted RAND when comparing segmentations from the same weighting of gradients, but different models. The average overlap of DTI vs. q-ball with ℓ_2 norm (squares) and DTI vs. q-ball with Sobolev norm (triangles) on the same data is consistently larger than the overlap when using the same model on perturbed vs. unperturbed data.

This effect is statistically significant. For each set of random weights, we compared the overlap within the same model (perturbed vs. unperturbed) to the overlap across models (DTI vs. q-ball). For all numbers of regions, we found the latter to be greater in so many cases that a one-sided binomial test rejects, with $p < 0.01$ (the largest p was $p = 5.5 \times 10^{-5}$), the null hypothesis that the within-model overlap and the across-model overlap are equally likely to be greater than each other. We conclude that segmentations change significantly more under slight perturbations of the input than when switching between diffusion tensor and q-ball models.

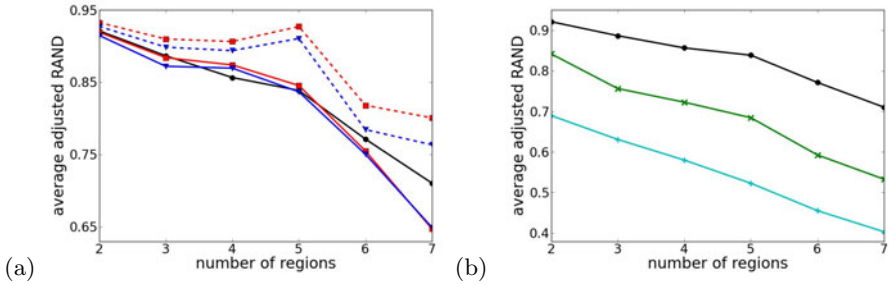


Fig. 3. Solid lines in (a) show segmentation stability under slight perturbations of the input (DTI: black circles, q-ball with ℓ_2 norm: red squares, Sobolev norm: blue triangles). It is consistently lower than the average similarity between results from DTI and q-ball on the same data (dashed lines). However, (b) shows that 60 directions (black circles) improve stability considerably when compared to 40 (green x-es) or 20 directions (cyan + signs), even with the DTI model.

2.5 Segmentation Based on Principal Directions

In their spectral method for thalamus segmentation, Ziyang et al. [2] use angular differences between principal eigenvectors as an alternative to the full diffusion tensor information. We have explored a similar strategy by extracting q-ball maxima. An initial estimate of the maximum has been obtained from a dense sampling on the sphere, and the result has been refined via gradient ascent.

Based on visual inspection of RGB maps (Fig. 4 (a) and (b)), the resulting overall structure appears similar. Within the thalamus mask, the average angular difference (AAD) between both estimates is around 10° . However, principal directions from q-balls are less stable under the perturbation from the previous Section (AAD 6.4°) than those estimated from the diffusion tensor (AAD 4.9°).

In order to average directions in the k -means algorithm, we added their outer products and took the principal eigenvector of the result. In general, we found segmentations that use principal directions to be less stable than those based on the deviatoric of the diffusion tensor, in particular when the estimate was derived from the q-ball model. Interpretation of Fig. 4 (c) is analogous to Fig. 3; stars indicate use of the principal direction from DTI, squares from q-ball. The stability of the deviatoric (circles) is repeated from Fig. 3 for reference.

3 Evaluating Benefits of Acquiring HARDI Data

Since our results in the previous Section did not demonstrate a clear advantage of using the q-ball model, it is natural to ask if it is even worthwhile to acquire 60 gradient directions at $b = 1000 \text{ s/mm}^2$ when the goal is not tractography, but segmentation of the thalamic nuclei. To investigate this question, we have repeated the stability analysis based on the diffusion tensor model and subsampled sets of 40 and 20 gradient directions, respectively. In order to retain a reasonable

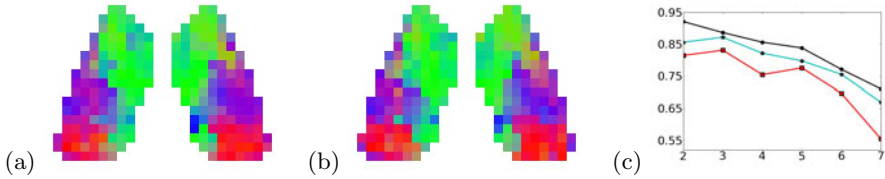


Fig. 4. Principal directions estimated from diffusion tensors (a) and q-balls (b) largely agree. However, (c) shows that segmentations based on principal directions from DTI (cyan stars) and q-ball (red squares) are less stable than ones from the diffusion tensor deviatoric (black circles).

gradient distribution, we generated sets of 40 and 20 directions based on electrostatic repulsion [13], and kept the original measurements that were closest to these directions.

Fig. 3 (b) compares the average overlap of segmentations from perturbed data (random weights of the gradients) to a segmentation based on the same set of gradients with unit weights. The original dataset is marked with circles (and is identical to Fig. 3 (a)), the subsampled data is marked with x (40 directions) and + signs (20 directions). It is obvious that stability of the segmentation suffers considerably when using fewer directions.

4 Conclusion

While it is widely accepted that HARDI models provide useful additional information within the white matter [11, 14], it has not been studied systematically what they can contribute to the segmentation of gray matter structures such as the thalamus. We have shed light on this interesting question from different perspectives, using a mixture of visual and statistical methods, theoretical reasoning about the impact of parameters, and empirical stability analysis.

For the data available to us ($b = 1000 \text{ s/mm}^2$), we conclude that segmentations of thalamic nuclei clearly benefit from acquiring data at high angular resolution, but that this is mainly due to an improved estimate of the information present in the standard diffusion tensor model. We found no clear benefit from switching to more complex models, such as q-ball, where most of the useful information seems to reside in the second-order spherical harmonics coefficients. This finding might not carry over to much higher b -values, as they have been used in [6, 7]. The methods proposed in our work could be used to test for potential benefits of the q-ball model when combined with such measurement setups.

We have only considered segmentations based on local diffusion properties. A different strategy for segmenting thalamic nuclei has involved a probabilistic tracking of the corticothalamic/thalamocortical connections [15, 16]. Since significant advantages of using HARDI models have been reported when tracking non-dominant fibers, even at $b = 1000 \text{ s/mm}^2$ [14], the relative benefit of tractography-based over purely local methods merits further investigation.

References

1. Wiegell, M.R., Tuch, D., Larsson, H.B., Wedeen, V.J.: Automatic segmentation of thalamic nuclei from diffusion tensor magnetic resonance imaging. *NeuroImage* 19, 391–401 (2003)
2. Ziyang, U., Tuch, D., Westin, C.-F.: Segmentation of thalamic nuclei from DTI using spectral clustering. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) *MICCAI 2006*. LNCS, vol. 4191, pp. 807–814. Springer, Heidelberg (2006)
3. Duan, Y., Li, X., Xi, Y.: Thalamus segmentation from diffusion tensor magnetic resonance imaging. *Int. J. Biomed. Imaging* 2007 (2007)
4. Jonasson, L., Hagmann, P., Pollo, C., Bresson, X., Wilson, C.R., Meuli, R., Thiran, J.P.: A level set method for segmentation of the thalamus and its nuclei in DT-MRI. *Signal Processing* 87(2), 309–321 (2007)
5. Ziyang, U., Westin, C.-F.: Joint segmentation of thalamic nuclei from a population of diffusion tensor MR images. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I*. LNCS, vol. 5241, pp. 279–286. Springer, Heidelberg (2008)
6. Grassi, A., Cammoun, L., Pollo, C., Hagmann, P., Meuli, R., Thiran, J.P.: Thalamic nuclei clustering on high angular resolution diffusion images. In: *Proc. Int. Soc. Magn. Reson. Med.*, p. 1777 (2008)
7. Brunenberg, E., Duits, R., ter Haar Romeny, B., Platel, B.: A sobolev norm based distance measure for HARDI clustering. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) *MICCAI 2010*. LNCS, vol. 6361, pp. 175–182. Springer, Heidelberg (2010)
8. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical Q-Ball imaging. *Magn. Reson. Med.* 58, 497–510 (2007)
9. Tuch, D.S.: Q-Ball imaging. *Magn. Reson. Med.* 52, 1358–1372 (2004)
10. Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* 23, 1176–1185 (2004)
11. Alexander, D.C., Barker, G.J., Arridge, S.R.: Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magn. Reson. Med.* 48, 331–340 (2002)
12. Hubert, L., Arabie, P.: Comparing partitions. *J. Classif.* 2, 193–218 (1985)
13. Jones, D.K., Horsfield, M.A., Simmons, A.: Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn. Reson. Med.* 42, 515–525 (1999)
14. Behrens, T.E.J., Johansen-Berg, H., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* 34, 144–155 (2007)
15. Behrens, T.E.J., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C.A.M., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat. Neurosci.* 6(7), 750–757 (2003)
16. Johansen-Berg, H., Behrens, T.E.J., Sillery, E., Ciccarelli, O., Thompson, A.J., Smith, S.M., Matthews, P.M.: Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Cereb. Cortex* 15(1), 31–39 (2005)