Gaussian Neighborhood Descriptors for Brain Segmentation

Henrik Skibbe[1,3], Marco Reisert[2], Hans Burkhardt[1,3]
[1]Department of Computer Science, University of Freiburg, Germany
[2]Dept. of Diagnostic Radiology, Medical Physics, University Medical Center, Freiburg
[3]Center for Biological Signalling Studies (BIOSS), University of Freiburg *

Abstract

In this paper we introduce a novel way for describing and classifying high angular resolution diffusion-weighted magnetic-resonance images (HARDI) of the human brain. Our approach is capable to segment the brain images into gray matter (GM) and white matter (WM) tissue. For the segmentation a two step approach is suggested: The appearance of a training image is described locally at each voxel position in a rotation invariant manner. Then a classifier is trained and used for distinguishing between background (BG), GM and WM in unclassified images. In contrast to existing model-free methods we are not only using the raw measurements at each position, we also comprise neighboring measurements in a rotation invariant way. Experiments show that our method outperforms existing methods significantly. Furthermore, we show that our method gives also reasonable results for brains with pathologies like tumors.

1 Introduction

Diffusion weighted magnetic resonance imaging (DWI) plays a substantial role in neuroscience and clinical applications. One field of interest is the investigation of the neuronal fiber architecture located in the brain white matter connecting different regions in the brain (see e.g. [6]). The fibers them self cannot be recorded directly. However, the data is usually recorded using the high angular resolution diffusion imaging technique [10]. With that technique we obtain an image where each position is described by a diffusion tensor representing the brain's tissue. Having such a representation we gain a probability distribution representing both the probability that fibers are crossing an image position and in which direction and constellation they probably continue. For the analysis of the fiber structure a preprocessing step that identifies the WM within the image is required. In [8] a model-free approach is introduced for describing and classifying the data locally. A SVM classifier is trained and used to automatically segment the brain into regions of interest. The descriptors are spherical harmonic features widely used for describing and detecting objects in volumetric images in a rotation invariant manner (see e.g. [3]). What we propose here is a new algorithm that not only uses the diffusion-weighted signal of a single voxel, but also includes the neigboring voxels. This is done by making use of a family of operators known from spherical tensor calculus (see e.g. [5] for further readings). Such operators play an important role when realizing a fast voxel-wise description and detection of objects in huge volumetric images (see e.g. [9]). We further show that instead of a SVM (as proposed in [8]) the consideration of a random forest classifier [1] additionally increases the performance significantly.

2 Approach

We represent HARDI-images by a function $\mathbf{f} : \mathbb{R}^3 \times S_2 \to \mathbb{R}$, where S_2 denotes the unit-sphere in $\mathbb{R}^3$ (with respect to the Euclidean distance). This means at each voxel-position $\mathbf{x} \in \mathbb{R}^3$ we have a function on the unit-sphere $\mathbf{f}(\mathbf{x}, \mathbf{n}_i)$ representing the MR-signal (note that $\mathbf{n}_i \in \mathbb{R}^3, \|\mathbf{n}_i\| = 1$). One important fact is that the signal shows an even symmetry. More precisely, a fiber enters and leaves a voxel position in two opposite directions with the same probability. It follows that $\mathbf{f}(\mathbf{x}, \mathbf{n}_i) = \mathbf{f}(\mathbf{x}, -\mathbf{n}_i)$ (see figure 1 a)). In figure 2 a configuration for parallel and crossing fiber bundles are shown.

According to [8] we first decompose the signal at each voxel position into its different frequency components. For this we utilize the so called spherical harmonic basis functions $Y^\ell_m : S_2 \to \mathbb{C}$ (see e.g. [7] for further informations). Spherical harmonics can be considered as some kind of orthonormal Fourier basis on the sphere. With each index ℓ a certain frequency is represented by a set of functions which lower index m is ranging from $m = \{-\ell \ldots \ell\}$. An even index ℓ indicates a symmetric pattern, if odd it indicates an antisymmetric pattern (similar to the cos and sin functions in 1D). As mentioned before the signals are symmetric, hence we only need to consider spherical harmonics associated with an even index $\ell\%2 = 0$ to completely represent the signal in terms of spherical harmonics (in figure 1 b) spherical harmonics for $\ell = 0, 1, 2$ are depicted). Due to a limited recording time (a patient must avoid any movements for several minutes) the number of points evaluated on the sphere (fiber directions) is limited. Similar to [8] the number of directions of our data is in the range of 31-81 (see table 1). With that finite number of points (we call that number $N \in \mathbb{N}$) we can represent signals with frequencies up to order $\ell = 4$ without artifacts (see eq. 7) in [8]). The unitary discrete transformation mapping a discrete signal with frequency ℓ represented by a homogeneously distributed number of points on the sphere in terms of spherical harmonics is just a multiplication with a matrix $\mathbf{M}^\ell$. We exemplary show the transformation corresponding to $\ell = 2$ in figure 1 c). As shown in figure 1 c) the resulting vector $\mathbf{a}^\ell(\mathbf{x}) \in \mathbb{C}^\ell$ is the ℓ-frequency component of the signal represented in the spherical harmonic domain. For this we shortly

*This study was partly supported by the Excellence Initiative of the German Federal and State Governments (EXC 294). Marco Reisert is indebted to the "Baden-Württemberg Stiftung" for the support of this research project by the "Eliteprogramme for Postdocs". We thank Susanne Schnell, Constantin Anastosopoulus and Irina Mader for providing the data.

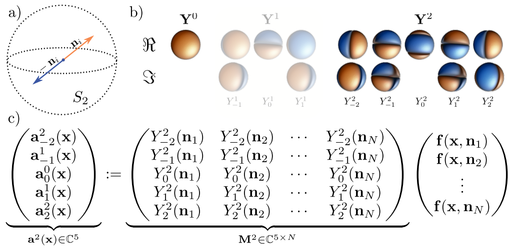


Figure 1. a) A probability distribution $\mathbf{f} : S_2 \to \mathbb{R}$ on a sphere represents the probability that fibers are crossing a voxel in a certain direction. This distribution is symmetric, hence $\mathbf{f}(\mathbf{n}_i) = \mathbf{f}(-\mathbf{n}_i)$. b) Spherical harmonic basis functions. Due to the symmetry of $\mathbf{f}$ we only need basis functions of even upper index to represent $\mathbf{f}$ in terms of spherical harmonics. c) Discrete transformation for transforming $\mathbf{f}$ into the harmonic domain.

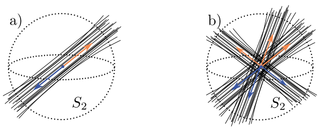


Figure 2. a) A parallel fiber configuration that is represented by the features $\|\mathbf{a}^2\|$, b) A fiber crossing that is represented by the features $\|\mathbf{a}^4\|$ (see eq. (1))

write

$$\mathbf{a}^\ell(\mathbf{x}) := \mathbf{M}^\ell \mathbf{f}(\mathbf{x})/N \quad , \tag{1}$$

where $\mathbf{a}^\ell(\mathbf{x}) \in \mathbb{C}^\ell$ is a vector valued expansion coefficient, $\mathbf{M}^\ell \in \mathbb{C}^{(2\ell+1)\times N}$ the unitary transformation matrix and $\mathbf{f}(\mathbf{x}) = (\mathbf{f}(\mathbf{x}, \mathbf{n}_1), \ldots, \mathbf{f}(\mathbf{x}, \mathbf{n}_N))^T$ are just the N measurements at voxel position $\mathbf{x}$ corresponding to N different directions. When omitting $\mathbf{x}$ and shortly writing $\mathbf{a}^\ell$ we consider $\mathbf{a}^\ell : \mathbb{R}^3 \to \mathbb{C}^{2\ell+1}$ as image of the expansion coefficients.

This transformation is performed for all voxel positions. We describe and classify the images using local image descriptors based on the spherical harmonic expansion coefficients $\mathbf{a}^\ell$. Similar to Cartesian Fourier analysis where the power spectrum is translation invariant, the power spectra of the coefficients $\mathbf{a}^\ell$ are rotation invariant (see e.g. [3]). In [8] these power-spectra are used in combination with an SVM-classifier to describe the image locally and then to segment the images into the areas of interest. Due to the limited number of sample points (here N) and the fact that the signal is symmetric the resulting feature vector only consists of the concatenation of three real-valued scalar-components, namely $\|\mathbf{a}^0(\mathbf{x})\|$ describing the isotropy of the distribution, $\|\mathbf{a}^2(\mathbf{x})\|$ representing the similarity to parallel fiber bundle (figure 2 a)) and $\|\mathbf{a}^4(\mathbf{x})\|$ representing crossings (figure 2 b)).

The proposed approach additionally includes information from neighboring voxels to compute local feature vectors. This is done in two aspects. On the one hand we consider different scales. More precisely, at each voxel position we consider an average distribution corresponding to a Gaussian windowed neighborhood of size $\sigma \in \mathbb{R}$, where $g(\sigma) := e^{-\mathbf{r}^T\mathbf{r}/(2\sigma)}$ is

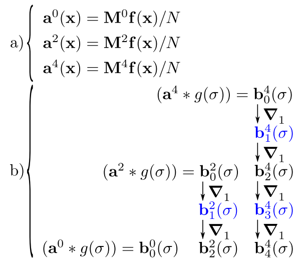


Figure 3. Work-flow of our algorithm. Step a): We first represent the signal in terms of spherical harmonics describing basic patterns of the fiber distributions (see figure 2). Step b): we generate further neighborhood descriptors describing the distributions in different scales more precisely. Rotation invariant descriptors are finally formed by computing and concatenating the power spectra of all expansion coefficients $\mathbf{a}^\ell$ and $\mathbf{b}_n^\ell$, respectively.

the 3D Gaussian function. What we obtain are expansion coefficients $\mathbf{b}_0^\ell(\mathbf{x}, \sigma) := (\mathbf{a}^\ell * g(\sigma))(\mathbf{x})$ having the same properties as the expansion coefficients $\mathbf{a}^\ell$ itself but also encoding the neighborhood. In addition to that we make use of so-called spherical tensor down-derivatives which we denote by $\boldsymbol{\nabla}_1$ (For further details we recommend [5]). When applying this operator voxel-wise up to ℓ times to expansion coefficients $\mathbf{b}_0^\ell(\mathbf{x}, \sigma) \in \mathbb{C}^{2\ell+1}$ we successively gain new expansion coefficients $\mathbf{b}_{n=\{0\ldots\ell\}}^\ell \in \mathbb{C}^{2(\ell-n)+1}$, with

$$\underbrace{\boldsymbol{\nabla}_1(\underbrace{\boldsymbol{\nabla}_1(\cdots \underbrace{\boldsymbol{\nabla}_1(\underbrace{\mathbf{a}^\ell * g(\sigma)}_{\mathbf{b}_0^\ell \in \mathbb{C}^{2\ell+1}})}_{\mathbf{b}_1^\ell \in \mathbb{C}^{2(\ell-1)+1}})}_{\mathbf{b}_{\ell-1}^\ell \in \mathbb{C}^3})}_{\mathbf{b}_\ell^\ell \in \mathbb{C}^1} \quad , \tag{2}$$

each describing certain characteristics like the local curvature of the smoothed fiber distribution (similar to an ordinary Taylor expansion). Due to the properties of $\boldsymbol{\nabla}_1$ [5] the power-spectra of all resulting coefficients are rotation invariant, too. Figure 3 illustrates the algorithm for a given scale. For representing the local fiber distribution we finally concatenate the power-spectra of all expansion coefficients and we get our feature image $\mathbf{c} : \mathbb{R}^3 \to \mathbb{R}^d$ with

$$\mathbf{c} := \{\|\mathbf{a}^0\|, ., \|\mathbf{a}^4\|, .., \|\mathbf{b}_0^0(\sigma_1)\|, .., \|\mathbf{b}_4^4(\sigma_2)\|, ..\} \quad , \tag{3}$$

where $d \in \mathbb{N}$ is the final descriptor dimension. Consider that each feature vector $\mathbf{c}(\mathbf{x})$ is rotation invariant with respect to rotation around $\mathbf{x}$.

3 Experiments

For our first experiments we use a database of the same size and being recorded using the same measurement parameters (resolution $2mm \times 2mm \times 2mm$ and b-value $1000mm^2/s$) as used in [8] containing 6 different measurements of healthy individuals. An overview of the images is given in table 1. For all our experiments we use the first dataset (data1) for training

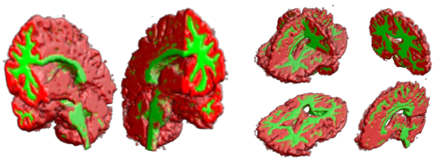


Figure 4. Isosurface showing the predictions for dataset 3 using GND and a RF classifier.

Table 1. Overview: The datasets. Data 1 is used for training.

	size	directions
data 1	$104 \times 104 \times 76$	81
data 2	$104 \times 104 \times 76$	31
data 3	$104 \times 104 \times 76$	81
data 4	$112 \times 112 \times 51$	81
data 5	$124 \times 124 \times 61$	61
data 6	$104 \times 104 \times 81$	61
defect 1	$104 \times 104 \times 69$	61
defect 2	$104 \times 104 \times 69$	61
defect 3	$104 \times 104 \times 69$	61

and the remaining images for evaluation. Consider that the datasets vary in size and number of diffusion-weighted directions. In our experiment we aim to segment the images into the classes background (BG), white matter (WM) and gray matter (GM). Ground truths for all datasets are generated by utilizing the SPM5 toolkit (Functional Imaging Laboratories 2005). This is done by registering the recorded image to a brain-atlas. Center slices of the resulting ground-truth images are shown in the first row of figure 5. For all images we compute features based on the spherical harmonic based descriptors proposed in [8] (we shortly call them SHD) and our descriptors describing the signals more precisely by involving Gaussian neighborhoods (we shortly call them GND). For the Gaussian neighborhoods we considered three different scales, namely $\sigma = 1.5, 3, 4.5$ leading to a descriptor with dimension $d = 30$ (3 power spectra of the ordinary spherical harmonic expansion coefficients $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3$ plus 3×9 new neighborhood descriptors computed for three different scales (see figure 3)).

As proposed in [8] we first use a support vector machine (SVM) with Gaussian kernel for training and testing. The feature vectors are normalized using the l_1 norm leading to the best performance. The parameters of the SVM are optimized on the training image using a grid search in combination with a cross validation [2]. For evaluation we use a PR-Graph generated by thresholding Platt's probabilistic output of the SVM [4]. The results showing the performance for detecting GM and WM (training on data 1, testing on data 2-6) are shown in figure 7(a). We also evaluated the performance on each dataset individually which is depicted in figure 7(c) (GM) and figure 7(b) (WM). We can observe, that the GND outperforms the SHD significantly. However, the performance is quite bad for the datasets 2,4 and 5. We believe that this is caused by over-fitting of the SVM to the given training dataset. Due to this we decide to additionally conduct an experiment using a random forest [1] (RF). We experienced that a RF is not only faster than a SVM with RBF kernel, it is also less likely to loose the ability of generalization due to an over-fitting to the training data. The RF classifier is based on a forest of decision trees,

Table 2. Equal Error Rate (smaller values correspond to a better performance)

		GND	SHD
RF	WM	**19%**	25%
	GM	**21%**	26%
SVM	WM	**24%**	26%
	GM	**33%**	57%

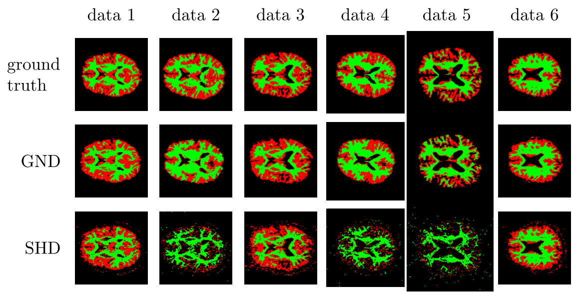


Figure 5. Ground truth shown together with the predictions of the random forest for both the Gaussian neighborhood descriptors (GND) and the spherical harmonic descriptors (SHD). (White matter (WM) is depicted in green, the gray matter (GM) in red)

each voting for a certain class. The final prediction is done by decision by majority. The parameters (number of trees in the forest and number of variables to split at each node in the decision trees) are determined experimentally on the training set by minimizing the OOB error rate [1]. For the GND the number of variables to split at each node (the parameter mtry) is set to two times the square root of the feature dim. The dimension of the descriptor of the SHD is quite small. For the best performance we use mtry=3 for the SHD. The number of trees is set to 1000 for both the SHD and the GND. For generating the PR-graphs we use the number of votes of the forest, voting for a certain class and we divide this number by the number of trees (1000). We obtain probability values representing how probable it is that a certain feature represents one of the three classes BG,GM or WM (centered Z-slices of such probability maps are depicted in figure 8). In figure 7(d) we show the results. Similar to the SVM scenario we do an evaluation on each dataset separately leading to the results in figure 7(f) (GM) and figure 7(e) (WM). The predictions (majority decision) of the RF are depicted together with the ground truth in figure 5. We can observe that the performance of both the SHD and the GND perform much better when using a RF than using a SVM (the equal error rate for both SVM and RF are listed in table 2).

We further use our algorithm (GND + RF) to create probability maps for three datasets (defect 1,2,3) containing measurements of individuals having different kinds of pathologies. The advantage of our approach is that it models the data locally. Hence it can deal with data where a significantly large area differs from the prototype. No ground truth was available However, figure 6 shows very promising results.

4 Conclusion

We presented a new application generating probability maps for the different regions of human brains from

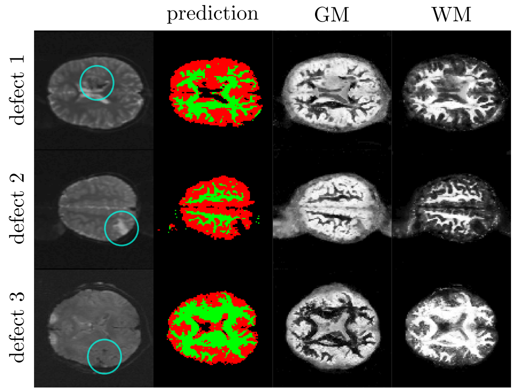
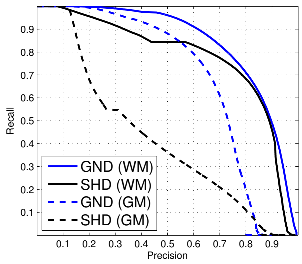
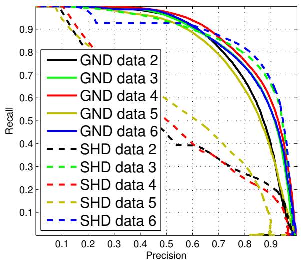


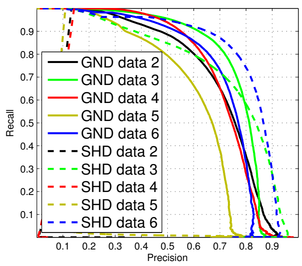
Figure 6. Predictions and probability maps for gray matter (GM) and white matter (WM) on the patient database.



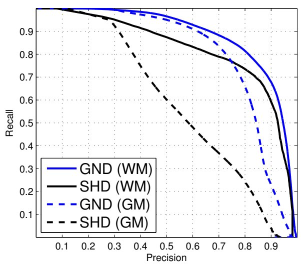
(a) PR-graph comparing the performance of the SHD and the GND using the Platt-probabilities of an SVM classifier.



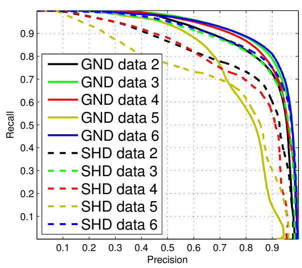
(b) Detecting WM: Performance evaluated for each dataset separately (SVM).



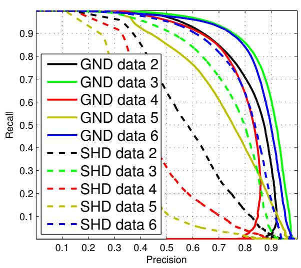
(c) Detecting GM: Performance evaluated for each dataset separately (SVM).



(d) PR-graph comparing the performance of the SHD and the GND using a RF classifier.



(e) Detecting WM: Performance evaluated for each dataset separately (RF).



(f) Detecting GM: Performance evaluated for each dataset separately (RF).

Figure 7. Comparing the performance of the SHD and the GND using both SVM and RF classifiers.

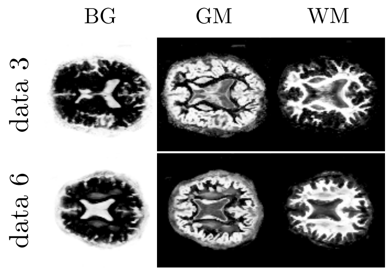


Figure 8. Probability map for background (BG), grey matter (GM) and white matter (WM).

HARDI in vivo data. For this we utilize spherical tensor algebra to rotationally invariant describe the local structures of the different areas of interest (the gray matter and the white matter) in a new manner. Probability maps representing the confidences for white and gray matter are generated by utilizing a random forest. Our experiments, where we consider a quantitative and qualitative evaluation, have shown that our method outperforms comparable existing model-free approaches significantly.

References

[1] Leo Breiman. Random forests. *Mach. Learn.*, 45:5–32, October 2001.

[2] Chih-Chung Chang and Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. Available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[3] Michael Kazhdan, Thomas Funkhouser, and Szymon Rusinkiewicz. Rotation invariant spherical harmonic representation of 3d shape descriptors. In *SGP '03: Proc. of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 156–164. Eurographics Association, 2003.

[4] John C. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *ADVANCES IN LARGE MARGIN CLASSIFIERS*, pages 61–74. MIT Press, 1999.

[5] M. Reisert and H. Burkhardt. Spherical tensor calculus for local adaptive filtering. In S. Aja-Fernández, R. de Luis García, D. Tao, and X. Li, editors, *Tensors in Image Processing and Computer Vision*. Springer, 2009.

[6] M. Reisert, I. Mader, C. Anastosoulus, M. Weigel, S. Schnell, and V.Kiselev. Global fiber reconstruction becomes practical. *Neuroimage*, 2010.

[7] M. Rose. *Elementary Theory of Angular Momentum.* Dover Publications, 1995.

[8] S. Schnell, D. Saur, B.W. Kreher, J. Hennig, H. Burkhardt, and V.G. Kiselev. Fully automated classification of hardi in vivo data using a support vector machine. *Neuroimage*, 46:642–651, 2009.

[9] Henrik Skibbe, Marco Reisert, Thorsten Schmidt, Klaus Palme, Olaf Ronneberger, and Hans Burkhardt. 3D object detection using a fast voxel-wise local spherical Fourier tensor transformation. In *Proc. of the DAGM*, pages 412–421, Darmstadt, Germany, 2010. LNCS, Springer.

[10] DS. Tuch, RM. Weisskoff, JW. Belliveau, and VJ. Wedeen. High angular resolution diffusion imaging of the human brain. In *Proceedings of the 7th Annual Meeting of the ISMRM*, Philadelphia,USA, 1999.