Gaussian Neighborhood Descriptors for Brain Segmentation

Henrik Skibbe\textsuperscript{1,3}, Marco Reisert\textsuperscript{2}, Hans Burkhardt\textsuperscript{1,3}

\textsuperscript{1}Department of Computer Science, University of Freiburg, Germany
\textsuperscript{2}Dept. of Diagnostic Radiology, Medical Physics, University Medical Center, Freiburg
\textsuperscript{3}Center for Biological Signalling Studies (BIOS), 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. [9]). 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 neighboring 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 \( f : \mathbb{R}^3 \times S_2 \rightarrow \mathbb{R} \), where \( S_2 \) denotes the unit-sphere in \( \mathbb{R}^3 \) (with respect to the Euclidean distance). This means at each voxel-position \( x \in \mathbb{R}^3 \) we have a function on the unit-sphere \( f(x, n) \) representing the MR-signal (note that \( n \in \mathbb{R}^3, ||n|| = 1 \)). 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 \( f(x, n) = f(x, -n) \) (see figure 1a). 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 \( Y_m^\ell : S_2 \rightarrow \mathbb{C} \) (see e.g. [8] for further informations). Spherical harmonics can be considered as some kind of orthonormal Fourier basis on the sphere. With each index \( \ell \) a certain frequency is represented by a set of functions which lower index \( m \) is ranging from \( m = \{-\ell, \ldots, \ell\} \). An even index \( \ell \) 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 \( \ell \). Without artifacts (see eq. 7) in [8] 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 \( \ell \) 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 1c. As shown in figure 1c the resulting vector \( \mathbf{a}^\ell(x) \in \mathbb{C}^\ell \) is the \( \ell \)-frequency component of the signal represented in the spherical harmonic domain. For this we shortly
Figure 1. a) A probability distribution $f : S_2 \rightarrow \mathbb{R}$ on a sphere represents the probability that fibers are crossing a voxel in a certain direction. This distribution is symmetric, hence $f(n_i) = f(-n_i)$. b) Spherical harmonic basis functions. Due to the symmetry of $f$ we only need basis functions of even upper index to represent $f$ in terms of spherical harmonics. c) Discrete transformation for transforming $f$ into the harmonic domain.

Write

$$a^\ell(x) := M^\ell f(x)/N ,$$

where $a^\ell(x) \in \mathbb{C}^\ell$ is a vector valued expansion coefficient, $M^\ell \in \mathbb{C}^{(2\ell+1) \times N}$ the unitary transformation matrix and $f(x) = (f(x, n_1), \ldots, f(x, n_N))^T$ are just the $N$ measurements at voxel position $x$ corresponding to $N$ different directions. When omitting $x$ and shortly writing $a^\ell$ we consider $a^\ell : \mathbb{R}^3 \rightarrow \mathbb{C}^{2\ell+1}$ 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 $a^\ell$. Similar to Cartesian Fourier analysis where the power spectrum is translation invariant, the power spectra of the coefficients $a^\ell$ are rotation invariant (see e.g. [3]). In [5] 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 $|a^\ell(x)|$ describing the isotropy of the distribution, $|a^\ell(x)|$ representing the similarity to parallel fiber bundle (figure 2a) and $|a^\ell(x)|$ representing crossings (figure 2b).

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 $g(\sigma) := e^{-r^2/\sigma^2}$ is the 3D Gaussian function. What we obtain are expansion coefficients $b^\ell_i(x, \sigma) := (a^\ell * g(\sigma))(x)$ having the same properties as the expansion coefficients $a^\ell$ itself but also encoding the neighborhood. In addition to that we make use of so-called spherical tensor down-derivatives which we denote by $\nabla_1$ (For further details we recommend [5]). When applying this operator voxel-wise up to $\ell$ times to expansion coefficients $b^\ell_i(x, \sigma) \in \mathbb{C}^{2\ell+1}$ we successively gain new expansion coefficients $b^\ell_{i\ell}(0, \ldots, \ell) \in \mathbb{C}^{2(\ell-\ell)+1}$, with

$$\nabla_1(\cdots \nabla_1(a^\ell * g(\sigma))) ,$$

each describing certain characteristics like the local curvature of the smoothed fiber distribution (similar to an ordinary Taylor expansion). Due to the properties of $\nabla_1$ [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 $c : \mathbb{R}^3 \rightarrow \mathbb{R}^d$ with

$$c := \{ ||a^0||, ||a^1||, \ldots, ||b^0_0(\sigma_1)||, \ldots, ||b^0_2(\sigma_2)||, \ldots \} ,$$

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

3 Experiments

For our first experiments we use a database of the same size and being recorded using the same measurement parameters (resolution $2\text{mm} \times 2\text{mm} \times 2\text{mm}$ and $b$-value $1000 \text{mm}^2/\text{s}$) as used in [3] 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

H. Skibbe et al., in Proc. of the MVA, Nara, Japan, 2011
The RF classifier is based on a forest of decision trees, generalization due to an over-fitting to the training data. That is why 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 lose the ability of generalization due to an over-fitting to the training data. The RF classifier is based on a forest of decision trees, 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 5(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 the different regions of human brains from each dataset. No ground truth was available how ever, figure 6 shows very promising results.

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

<table>
<thead>
<tr>
<th>size</th>
<th>directions</th>
</tr>
</thead>
<tbody>
<tr>
<td>data 1</td>
<td>104 × 104 × 76</td>
</tr>
<tr>
<td>data 2</td>
<td>104 × 104 × 76</td>
</tr>
<tr>
<td>data 3</td>
<td>104 × 104 × 76</td>
</tr>
<tr>
<td>data 4</td>
<td>112 × 112 × 51</td>
</tr>
<tr>
<td>data 5</td>
<td>124 × 124 × 61</td>
</tr>
<tr>
<td>data 6</td>
<td>104 × 104 × 81</td>
</tr>
<tr>
<td>defect 1</td>
<td>104 × 104 × 69</td>
</tr>
<tr>
<td>defect 2</td>
<td>104 × 104 × 69</td>
</tr>
<tr>
<td>defect 3</td>
<td>104 × 104 × 69</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>GND</th>
<th>SHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF WM</td>
<td>19%</td>
<td>25%</td>
</tr>
<tr>
<td>GM</td>
<td>21%</td>
<td>26%</td>
</tr>
<tr>
<td>SVM WM</td>
<td>24%</td>
<td>26%</td>
</tr>
<tr>
<td>GM</td>
<td>34%</td>
<td>57%</td>
</tr>
</tbody>
</table>

4 Conclusion

We presented a new application generating probability maps for the different regions of human brains from...
Recall performance evaluated for each the GND using the Platt-approaches significantly.

(a) PR-graph comparing the (b) Detecting WM: Performance of the SHD and mance evaluated for each the GND using a RF classifier.

c) Detecting GM: Performance evaluated for each the GND using both SVM and RF classifiers.

d) PR-graph comparing the (c) Detecting GM: Performance evaluated for each the GND using both SVM and RF classifiers.

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

Figure 7. Comparing the performance of the SHD and the GND using the Platt-approaches significantly.

HARIDI 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


H. Skibbe et al., in Proc. of the MVA, Nara, Japan, 2011