

POLITECNICO DI TORINO

Master degree course in Biomedical Engineering

Master Degree Thesis

Segmentation of White Matter Tractograms Using Fuzzy Spatial Relations



Supervisors:

prof. Filippo MOLINARI
prof. Isabelle BLOCH

Co-Supervisor:

prof. Pietro GORI

Candidate

Alessandro DELMONTE

March 2018

This work is subject to the Creative Commons Licence

Summary

The aim of this master thesis *Segmentation of White Matter Tractograms Using Fuzzy Spatial Relations* is the creation and validation of an innovative segmentation tool, called Fuzzy Tracts, for whole-brain tractograms applications, with the goal of extracting reproducible white matter fiber bundles using prior knowledge combined with a spatial fuzzy sets approach and an innovative clustering metric.

The segmentation of anatomically significant fiber bundles is the main strategy to analyze tractograms, which are often composed by hundreds of thousand of fibers. Different segmentation methodologies have been proposed, each suffering from a multitude of problems, the most common ones being the consistency of the results, the time needed for the procedure and strong assumption on inter-subjects spatial correspondence.

This thesis, after introducing the diffusion imaging methods commonly used, testing state of the art segmentation methods and highlighting the problems encountered, proceeds detailing the algorithm proposed, emphasizing the spatial fuzzy sets approach and the innovative clustering metric implemented.

Finally, the results are validated on the Uncinate Fasciculus use-case, focusing on inter-subject reproducibility and anatomical coherence, illustrating the improvements achieved through the use of comparisons over other techniques.

Acknowledgements

I would first like to thank my thesis advisors. Prof. Filippo Molinari at the Polytechnic University of Turin, for his immense help and for the passion he transmitted to me in regards of the medical imaging world. prof. Isabelle Bloch at the Télécom ParisTech, for all the opportunities she gave me, her unconditional support, trust and mentoring changed completely my perspective on the professor-student relationship.

I would also like to thank prof. Pietro Gori for being the best chief and supervisor every student could wish for. Not only his door were always open but his constant participation, support and advice were key to my contribute to the project.

I would also like to acknowledge all of the people from the BIOMED and IMAGES teams at Télécom ParisTech for their valuable suggestions and the welcoming environment they helped to create.

Finally, I must express my very profound gratitude to my grandmother, parents, sister and other family members and friends for providing me with unflinching support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Alessandro Delmonte

Contents

Summary	III
Acknowledgements	IV
I Introduction	1
1 Diffusion MRI (dMRI)	3
1.1 Diffusion Tensor Imaging (DTI)	5
1.2 Constrained Spherical Deconvolution (CSD)	7
1.3 Human Connectome Project (HCP)	10
1.3.1 Data	11
1.3.2 Preprocessing	12
2 Tractography	15
2.1 Tracking Algorithms	16
2.2 Open Challenges	19
2.3 MRTrix3	21
3 Use Case: Uncinate Fasciculus (UF)	23
II State Of The Art	27
4 Segmentation Methods	29
4.1 Manual Segmentation	30
4.2 Clustering Methods	32
4.3 Different Approaches	34

5	White Matter Query Language (WMQL)	35
5.1	Queries	36
5.2	Limitations	38
III	Fuzzy Tracts	39
6	Concept	41
7	Spatial Fuzzy Sets	45
8	Functional Varifolds (fVar)	49
9	Implementation	53
9.1	Fuzzy Sets computation	55
9.2	Initial Segmentation	60
9.2.1	WMQL	60
9.2.2	Sets Intersections	62
9.3	FA Extraction	64
9.4	Clustering	66
9.4.1	Metric	67
9.4.2	Clustering Algorithms	69
9.5	Final Segmentation	71
10	Other Features	73
10.1	File Format Support	73
10.2	Visualization	74
10.3	Performance	75
11	User Guide	77
IV	Results	81
12	Comparisons	83
12.1	Fuzzy Tracts and Manual Segmentation	83
12.2	Fuzzy Tracts and WMQL	86
12.3	Inter-Subject Consistency	87
13	Conclusions	89
14	Future Work	93

V	Appendices	95
A	CoBCoM	97
B	OHBM	99

Part I

Introduction

Chapter 1

Diffusion MRI (dMRI)

Diffusion MRI is an imaging technique that estimates the water particles diffusion in a tissue (Figure 1.1). Diffusion, also called Brownian motion [13], is the microscopic movement of the particles due to thermal collisions. This thermally-driven motion, that flows from high density area to low density area, is always present for temperatures different to the absolute zero and its description, first postulated by Einstein [24], can be summarize as: $x_{rms} = \sqrt{2Dt}$.

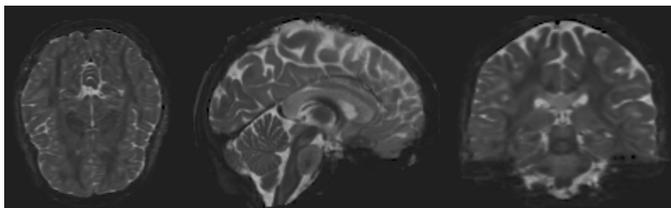


Figure 1.1. Diffusion Weighted Image from the Human Connectome Project dataset.

D is called diffusion coefficient and it is a function of the temperature T , the viscosity of the tissue η and the radius of the particle: $D = \frac{k*T}{6*\pi*r*\eta}$.

In the body there is a proportion between the water contained in the cells and in the extracellular matrix, kept in balance from the Na/K pumps. The extracellular part usually has more freedom to move in any direction since the water contained in a cell has a high probability to collide with the cell walls and other molecules. The displacement of the water within a cell will thus follow the path designed by the structural disposition of the walls, creating an high level of diffusion anisotropy

for directional architectures like in the cases of white matter and muscles fibers.

For these types of movements the diffusion coefficient can not be described by a single values, as it varies depending on the direction, but it assumes the shape of a tensor (more information at Chapter 1.1).

It is important to point out that what is being seen in the diffusion images is not the real water diffusion map, since a lot of other body processes are active in the background (microcirculation, respiratory movement, ...). The measure is better defined as the Apparent Diffusion Coefficient (ADC) [10], the coefficient that keeps into account all the contributes.

To obtain a DWI (Diffusion Weighted Image) the imaging technique exploited is the MRI, with particular configuration of the gradients [86]. Two gradients, the first one at 90° and the second one at 180° , are applied in sequence (Figure 1.2). In this way, the spins of stationary molecules are not affected, since the second gradient reset it to their natural position, meanwhile the change in the spin of the water stays noticeable due to its movement to a different position between the two impulses.

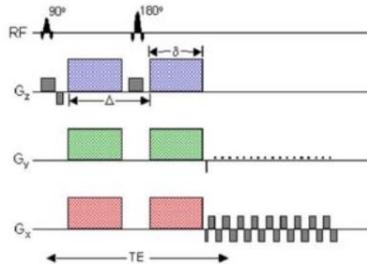


Figure 1.2. Diffusion EPI sequence. In grey the imaging gradients, in color the diffusion gradients for three directions.

The gradient parameters influence the output images, highlighting or less the diffusion events. It is possible to express the diffusion based on the gradient energy and the base intensity (the values of the T2 image) as: $S = S_0 * e^{-bD}$.

Where b , called bvalue, is a measure of the gradient power and timing. Its unit, s/mm^2 , represents a time over an area, the opposite of D . Increasing the bvalue accentuates the weighting towards the diffusion, making the water flow more noticeable, but at the same increasing the noise level. Typical diffusion images are obtained with bvalues not higher than $3000 s/mm^2$.

The direction of the gradient is instead described by the b-vectors. Each b-value has a b-vector associated, the gradient orientation at that instant. The combination of b-values and b-vectors allows to know the gradient configuration, in all their part, for each diffusion image obtained.

Diffusion images are mostly used for brain applications, usually combined with tractography algorithms (Chapter 2), but are not limited to them. In the recent years the diffusion imaging expanded to pelvis applications [75] and whole-body imaging [47].

Since the diffusion associated to a gradient configuration express only the diffusion along a particular direction, in order to have a better understanding of the diffusion phenomena more gradients must be applied, following a predefined scheme.

1.1 Diffusion Tensor Imaging (DTI)

The Diffusion Tensor Imaging [51] utilizes a gradient scheme to induce the diffusion in multiple direction and create a diffusion matrix that characterizes the behavior of the water particles in a 3D structure.

Since what is sensed is the ADC intensity S , the values D of diffusion can be extracted applying an inverse formula [86]: $D = \frac{-\ln(\frac{S_2}{S_1})}{b_2 - b_1}$ using a baseline image S_1 (for example the T_2).

Filling the smallest tensor configuration with diffusion measures means that three gradient must be applicable in a cartesian plane :

$$D = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}$$

The three main eigenvectors are positioned on the diagonal of the tensor and are equals to each other in the situation of isotropic diffusion. The tensor is symmetrical in respect to the main diagonal ($D_{xy} = D_{yx}$, $D_{xz} = D_{zx}$, $D_{yz} = D_{zy}$).

While at least three gradient directions must be implemented, increasing the number of gradient orientations improves the accuracy of the measures. Modern DTI application often use 6, 21 or 31 directions [31] to better the accuracy of the

results and reduce the noise sensitivity.

From the tensor, various indicative diffusion measures can be extracted to more appropriately define the water behavior and, consequently, the anatomical structure that originated the coefficients (Figure 1.3).

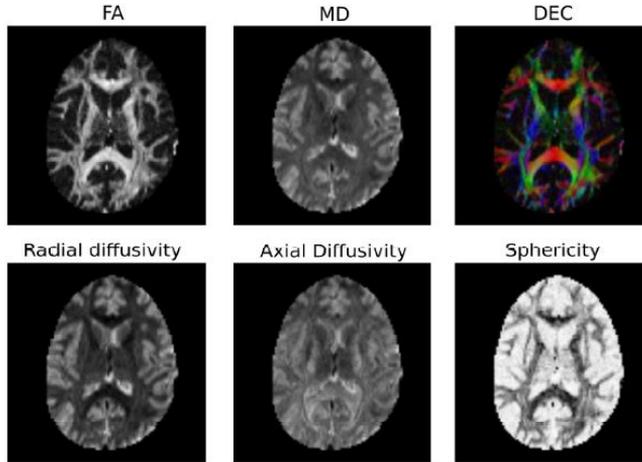


Figure 1.3. Metric images extracted from the diffusion tensor (MD = mean diffusivity, FA = fractional anisotropy, DEC = directional encoded color).

The three principal eigenvectors are used to determine the directions of the water diffusion [66]. In particular the one with the highest value gives an information about the main direction of diffusion, moving along its orientation means following the trail that the majority of the water molecules pursue.

The other two eigenvector instead detail how the water move along the traversal planes. Their main, also called radial diffusivity, could be confronted to the main diffusion to verify the displacement variance.

With a three directions gradient scheme an equal number of diffusion image are produced. Compressing the different results into one image allows to have a compact and more comprehensible view [61], the so called trace image express the final diffusion coefficient as the average of the tensor eigenvalues:

$$S_{DWI} = \sqrt[3]{S_x S_y S_z} = S_0 e^{-b \left(\frac{D_{xx} + D_{yy} + D_{zz}}{3} \right)}.$$

The trace image offers some advantages to the singles DWI since, being the

mean of multiple sources it is more robust to noise. The observed homogeneity and low inter-subject variability [74] constitute a good option for the aid in the detection of abnormal diffusivity phenomena.

Fractional Anisotropy

One of the most indicative diffusion metric extracted from the DTI is the fractional anisotropy [66]. It is used to estimate the intra-voxel degree of water flow directionality. The FA is computed with the equation:

$$FA = \sqrt{\frac{3}{2} \frac{\sqrt{(D_{xx}-M)^2+(D_{yy}-M)^2+(D_{zz}-M)^2}}{D_{xx}+D_{yy}+D_{zz}}}.$$

Where M is the local trace value.

The fractional anisotropy indicates if the local diffusivity follow a preferred direction or if it has a more isotropic behaviour. High values means a more accentuate anisotropy, expressing a more constrained water flow. The immediate association between large FA values and myelinated axons presence [46] is emblematic of the axon population that lies beneath.

Extensive studies have been carried out trying to find a correlation between the FA values with age [79], sex [81], lesions [107] [9] and other types of markers.

Additionally, the fractional anisotropy tends to be a characteristic parameter of white matter fiber bundles, making it a useful parameter for the formalization of their properties.

1.2 Constrained Spherical Deconvolution (CSD)

Main drawback of the diffusion tensor approach, that confines its applicability condition to simple analysis, is the innate limitation to solve the crossing fibers problem.

Having a typical DWI resolution floating between 1.5 mm and 3 mm, poorly comparable with the radius of white matter axons (not larger than 20 μm), the diffusion imaging techniques provide just an estimation on the behavior of the fibers group that pass through a voxel [69]. Furthermore, if the diffusion tensor is being used for more advanced tractography tasks, like for segmentation purposes (Chapter 2), the risk of an error propagation is particularly high [65].

Various approaches have been proposed to overcome these limitations, trying to extend the tensor method to support more complex imaging strategies. The most innovative data acquisitions protocol usually exploits the HARDI (High Angular Resolution Diffusion Imaging) scheme [92]. In this sampling configuration a large quantity of gradients (at least 28) are applied (Figure 1.4), changing the directions while keeping a constant bvalue. The sampling scheme, build around a sphere around the brain, also called shell, aims to maximize the angular contrast. Furthermore, using multiple shells with different gradient intensities it is even possible to reach a high resolutions without a loss in the useful information provided by the use of multiple bvalue, for what is called multi-shell multi-value approach.

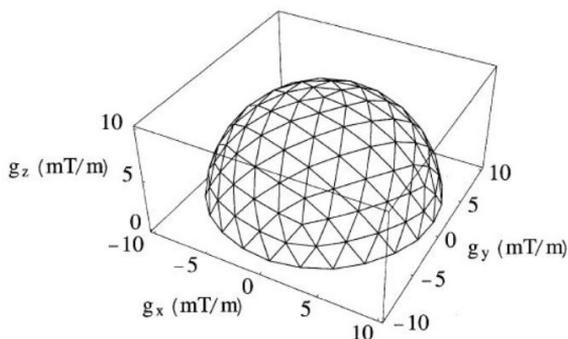


Figure 1.4. HARDI sampling scheme formed by 126 gradient direction sampled on a tessellated sphere.

The Constrained Spherical Deconvolution (CSD) [91] exploits and extends the diffusion tensor approach with a key difference. While the DTI use a discrete number of directions, the CSD assumes a continuous function, assimilating the HARDI sampling scheme to an infinite number of gradient orientation.

In the Spherical Deconvolution a fiber response R , produces looking at the DWI attenuation of a group of coherently-oriented fiber (identified observing where the fractional anisotropy tends to zero), is computed in function of the elevation angle θ .

With the assumptions that the response function is the sum of all the response functions generated by the underlying axon population and that the response function of every axon of the brain would be the same if oriented in the same direction, finding the local orientation means deconvolving the local fiber response with the function $R(\theta)$ (Figure 1.5).

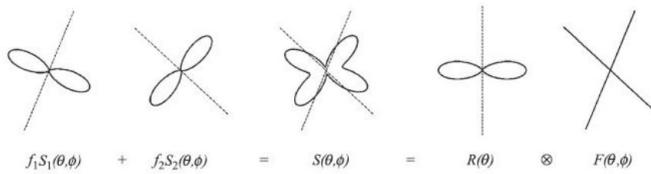


Figure 1.5. A voxel containing two fibers with different orientation θ_1 e θ_2 is represented by the local response function $S(\theta)$. Through the deconvolution with $R(\theta)$ it is possible to estimated the fiber population distribution $F(\theta)$.

The DWI can be expressed as a convolution between the response function and the FOD (fiber orientation distribution), represented as lobes, using spherical coordinates. In order to add a consistency to the mathematical procedure the introduction of a low-pass filter is necessary to remove the high sensitivity to the noise that caused spurious, not anatomically relevant, peaks in the responses [90]. This constrain allows to have just one visible peaks oriented in the direction of the highly probable fiber direction and to obtain values next to zero in every other part of the support.

CSD methods can estimate the FOD and recognize intravoxel fiber crossing wherever the angle between the two streamlines is at least 40%, making it one the most used method for tractograms estimation (Figure 1.6).

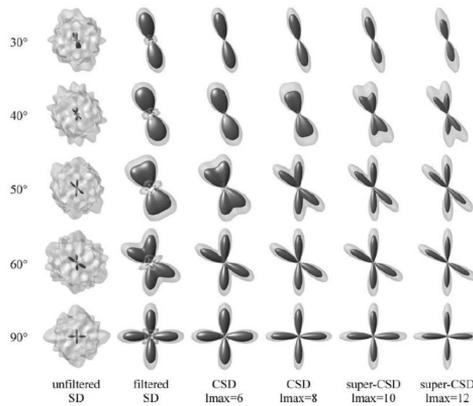


Figure 1.6. Fiber orientation distribution obtained with the spherical deconvolution method for different separation angles and harmonic degrees (l_{max})

1.3 Human Connectome Project (HCP)

The Human Connectome Project is an ongoing research program that aims to significantly improve the understanding of the human brain and to define and map its connectivity network [97].

The primary objective is the creation of an extensive and accurate database, making it available to the scientific community. With more than 19 active research studies, spanning from brain mapping, diseases characterization and aging evaluations the HCP consortium provides the largest font of data for brain functionality related researches.

To achieve the imaging precision searched important customization procedures have been applied to commercial MRI scanner and specific MRI sequences have been prepared [96]. The special 3T scanner prepared is finally able to reach gradient intensities of 300mT/m with a signal-to-noise ratio of around 70%. Additionally, a 64-channel 7T scanner has been used for imaging procedures on around 200 subjects (Figure 1.7).

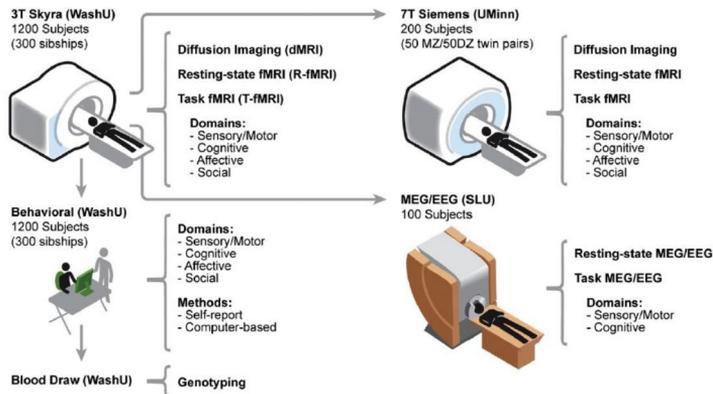


Figure 1.7. Summary of the HCP dataset and scanners prepared for the image acquisitions.

The diffusion pulse sequence, traditionally an EPI, has also been subject of heavy modifications in order reach high imaging resolutions while reducing the scanning time. The solution adopted capitalize on a multiplexed EPI train approach, allowing a whole brain scansion to be carried out in under a second reaching a resolution of 2 mm.

The HCP coverage does not stop to diffusion data but also expands to functional MRI (both resting state and task related), structural MRI, MEG and EEG.

1.3.1 Data

The validation of Fuzzy Tracts is based on diffusion HCP data due to their high resolution, high SNR and number of subjects involved. The dataset used, that can be downloaded from <https://db.humanconnectome.org/>, is formed scanning 1200 young healthy subjects.

The subject population is composed by siblings and not siblings with an age under of 35. From them the images relative to 100 unrelated subject have been extracted and used as the final benchmark database. The relative structural and diffusion are the base of the results presented in this thesis (Figure 1.8).

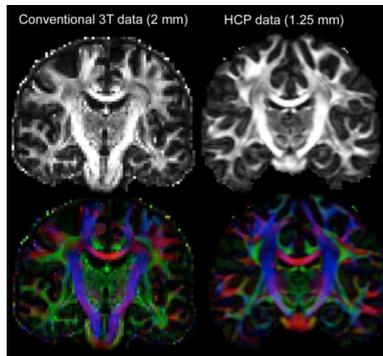


Figure 1.8. Confront between the diffusion images produced by a conventional scanner (left) and by the customized HCP machine (right).

The diffusion data present a resolution of 1.25mm isotropic and are obtained using a multishell approach with bvalues ranging from 50 s/mm^2 (used as the b_0 image to avoid vascular artifacts) to 3.000 s/mm^2 . The full specific can be found in table 1.1

The diffusion acquisition protocol consists of six scans (with three gradient tables applied first in a right-left direction and after in the left-right encoding) for a total of circa sixty minutes. Three shell, with bvalues 1000, 2000 and 3000 s/mm^2 , have been sampled, alternating with six b_0 scans. All the diffusion data have associated two structural images: a T1 weighted and a T2 weighted.

Parameter	Value
<i>Sequence</i>	<i>Spin – Echo EPI</i>
<i>TR</i>	5520 <i>ms</i>
<i>TE</i>	89.5 <i>ms</i>
<i>Flip angle</i>	78°
<i>Resolution</i>	1,25 <i>mm isotropic</i>
<i>Echo spacing</i>	0.78 <i>ms</i>
<i>Shells</i>	1000, 2000, 3000 <i>s/mm²</i>

Table 1.1. HCP diffusion data acquisition parameters.

The final files folder, with a size of 4 Gb, contains the image files, reconstructed in NIFTI format with an-hoc prepared algorithm [93], the gradient table divided in bvectors (for the directions) and bvalues (for the intensities) and a brain mask, obtained using Freesurfer .

The data are available both raw and processed with a specific HCP artifact-removing pipeline.

1.3.2 Preprocessing

The extensive HCP processing protocol [32] aims to reduce the artifact effects that afflict the dataset. These artifacts slightly differs from the standard MRI problems due to the very unique characteristic of the HCP diffusion images. The usage of preprocessed data over their raw counterpart is highly suggested since they have been cleaned by difficult to analyze issues, like gradient non-linearities. Furthermore, the use of a common space for the processed images allows a more accurate inter-subject analysis for relevant statistical studies.

The pipeline covers both structural data and diffusion data. For the case of T1w and T2w image the process is split in 3 part: pre, with and post Freesurfer. A first analysis corrects the spatial distortion, aligns and registers the images to the MNI space [17]. A segmentation operation is carried out to obtain one cortical parcellation and a whole-brain labelling. At last, the images are adapted to the NIFTI file format and a brain mask is produced.

The structural scans, acquired with an isotropic resolution of 0.7 mm, necessitate a downsampling to 1 mm in order to be fit for the Freesurfer recon-all pipeline [20]. The final parcellation outputs, with resolution of 256x256x256 per 1 mm isotropic, will contain the automatic segmentation of white matter other than the cortical and pial surfaces.

The diffusion pipeline (Figure 1.9) tries instead to eliminate the artifacts belonging to the special protocol used. First the intensities are aligned exploiting the b_0 images information, capturing different T_2 images at evenly spread time intervals during the diffusion acquisition it is possible to estimate the intensity drift [100].

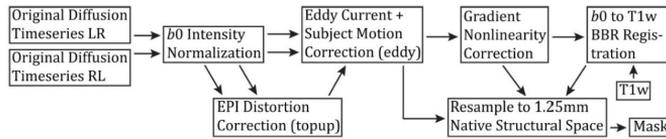


Figure 1.9. Diffusion preprocessing pipeline applied to the HCP raw data.

The eddy currents and gradient non-linearity problems are tackled with the use of FSL [84] and a Gaussian estimator. After a final registration between the b_0 and the T1w is achieved, the data are masked with the structural data brain mask to reduce the file size.

Chapter 2

Tractography

The restraints in the diffusion of water molecules due to different types of obstacles, such as cells and membrane cause a change in in the diffusion coefficient perceived in DWI. In the portions where water has a free diffusion, for example in cerebrospinal fluid, an high diffusion coefficient will be present and, moreover, the diffusion tensor associated will carry isotropic proprieties [3].

In the white matter the molecules diffusion along the ax of a fiber will be less restricted than the in the perpendicular directions. This observation can be exploited to determine the fibers pathway. Following the trace of the diffusion images changes, streamlines could be tracked down, starting from a seed point, in all their lengths due to their relationship with the underlying fiber population [43].

Tractography is the medical imaging operation that reconstructs 3D fibers starting from diffusion data and proprieties (Figure 2.1). Locating the pathway of fiber systems enable to delineate the connectivity of an anatomical regions with a certain level of confidence [2], before exclusive to invasive methods.

In the last years this imaging technique saw a boost in importance thanks to scientific advances that augmented the applicability on routine medical tasks, with the main utilization being the pre-operative planning for brain procedures [85], estimation of recovery in functionality after traumatic incidents [39] and brain connectivity mapping.

Tractography not only offers a new point of view to interpreter pathological white matter abnormalities, but can also all-around improve the way the anatomy of the brain is defined [4]. Not limited to the nervous system, it can be complementary to the standard magnetic resonance exams, especially for type of lesions

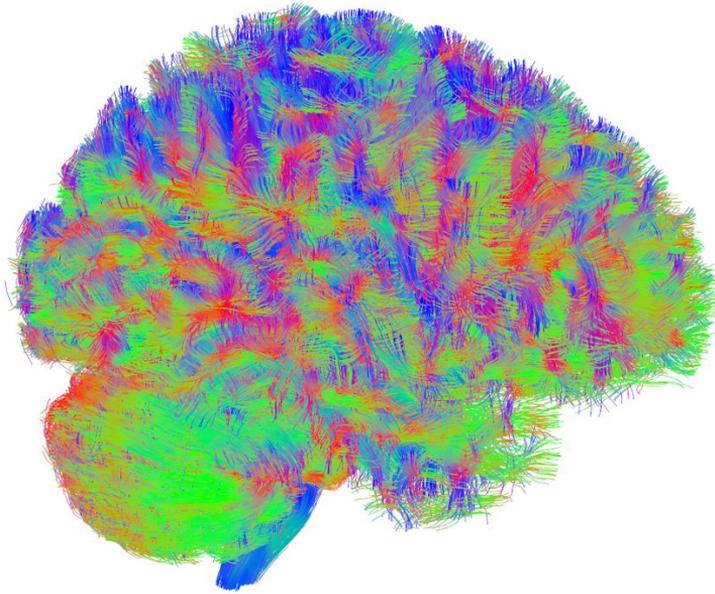


Figure 2.1. Whole brain tractogram computed with the deterministic algorithm SDSTREAM on MRTrix3, starting from HCP data.

particularly difficult to find with standard imaging procedure [98].

Initially based only on the DTI approach, during the years different methods have been developed with the goal to improve the accuracy of the results with more and more sophisticated fiber tracking algorithms been presented to the scientific community.

Tractography at the moment is at its highest peak with a great number of contributor all around the world ready to push even further this innovative imaging technique to reach even higher level of confidence and to offer clinicians useful new decision tools.

2.1 Tracking Algorithms

Fiber tracking algorithm have the goal to trace back all the axons trajectories using the diffusion imaging data. These algorithms can be classified in two main groups: deterministic methods and probabilistic methods.

A deterministic approach will always result the same output for the same input data. The most common algorithms use propagation techniques: fibers are reconstructed connecting adjacent voxels following a selected criteria[65]. This type of tracking was born with the FACT algorithm (Section 2.1) and it is still used nowadays.

Deterministic algorithms follow a three step process: initialization, propagation, termination. The initialization is based on the choice of seed points, from which fibers will be tracked down. The propagation consist in applying a method to connect the voxels along the streamlines length. The termination criteria is necessary to define when a fiber should end.

Most software suites allows to decide where the seed points should be placed, using random initialization or ROI selection [42]. The decision on the reconstruction algorithm should instead be made with in mind the result wanted and the specific application the tractogram will be used for. At last, the termination criteria could be based on diffusivity information (like fractional anisotropy or FOD peaks distributions), geometries (e.g. considering the curvature of the fiber) or through the use of masks [82].

Associating just one fiber to each seed point considered, deterministic methods are unable to find branches and resolve fiber crossing problems. Furthermore, the lack of a confidence score to a fiber pathway implies that local noise-related diffusion errors could considerably deviate the pathways, making them not coherent with the anatomical distribution [54]. Finally, the track termination procedure contemplate the application of thresholds, not always intuitive to set[108].

Probabilistic tracking algorithms, on the other end, try to tackle these issues considering the possibilities of splitting the fiber pathway during its reconstruction [6]. Multiple pathway could be found using an estimation of the uncertainty of the direction [7].

The clinical limitation of the probabilistic algorithms are due to the huge quantity of fiber represented in the tractograms (that makes them really difficult to interpreter), the slowness of the tracking (that limits the real-time application) and, even more importantly, the nature of the output [8]. In fact, the output tractograms does not picture the real axonal distribution but it symbolize a volume of connections probabilities. The inevitable creation of anatomical non-senses induce the necessity of an accurate analysis, based on prior knowledge, to extract the relevant information.

FACT

The first tracking algorithm conceived was the FACT (Fiber Assignment by Continuous Tracking) [64]. This approach showed how it is possible to reconstruct the path of an axons using the diffusion tensor imaging, the study focused on rat brains and set the stepping stone for all the following tractography tracking algorithm implementation.

FACT is a deterministic algorithm that can trace down axons in a 3D grid using the eigenvectors of the diffusion tensor. Due the impossibility of using the fractional anisotropy information the tracking is only based on the largest diffusivity direction. While a more immediate approach would suggest to link a voxel to the most adjacent one that faces the same direction of the most prominent eigenvector, errors and huge discrepancies between the tracking and the ground truth could be observed since the diffusion direction are limited to eight.

The algorithm introduces a discrete approach using an edge-to-edge propagation instead of a center-to-center technique (Figure 2.2). Using just the principal direction it is possible to obtain, with limited resolution, a good approximation of the trajectory of a white matter fiber. The tracking is concluded when a factor R , that weights the directions of the neighbor voxels, becomes lower than a threshold.

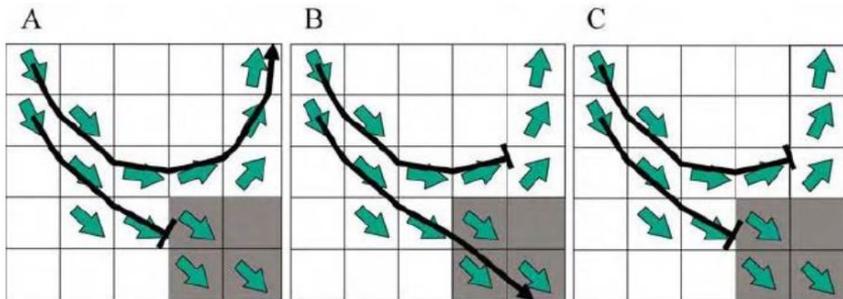


Figure 2.2. FACT tracking procedure and termination conditions (A: fibers entering in low FA zone. B: sharp turns. C: combined conditions). The green arrows indicate the voxel main diffusion directions.

SD_STREAM

The SDSTREAM [89] is a deterministic tracking algorithm implemented in the MRTrix3 software suite that utilizes the FOD computed with the Constrained Spherical Deconvolution to determine the fiber paths.

The local orientation of the fiber is considered as from peak of the FOD, computed using a gradient-descend algorithm. The major problem appears when there is a presence of small shoulders in the distribution (Figure 2.3). If the tracking, for just a local point, deviates to a shoulder instead of following the real maximum peak, the results could lose any anatomical validity. The termination strategy stops the tracking when a fiber exceed a curvature constrain. This constrain is kept low in order to allows tracking in the situation of small shoulder in the FOD.

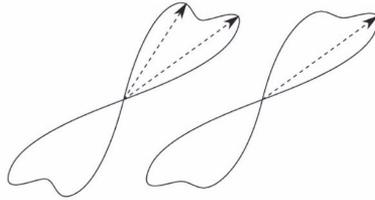


Figure 2.3. Local maximum, represented by bumps in the fiber orientation distribution, can severely alter the tracking path.

The validation of Fuzzy Tracts has been carried on tractograms obtained with this algorithm.

2.2 Open Challenges

Ten years have passed since the birth of tractography as we know it and a good amount of obstacles are still unsolved [40].

Diffusion tensor tractography algorithm tends to use the main direction of diffusion to reconstruct the axons. While this could be considered correct from a diffusivity point-of-view there is no guarantee that the fiber path delineated has a meaningful information in the axonal scale. Even more, it is nearly impossible to distinguish between fibers that merge, or approach, into each other (Figure 2.4). More sophisticated algorithms based on the Constrained Spherical Deconvolution (CSD) are able to differentiate them but locally, in the voxel space, the behavior is not determined [27].

Tractography can not distinguish between axons polarity. Whether a fiber is afferent or efferent does not modify the diffusion information, making any assumption on the direction a completely blind guess. Furthermore, given that the current spatial resolution of the diffusion imaging is not compatible with the axonal dimension [78] a streamline is the representation of a certain number of real axons,

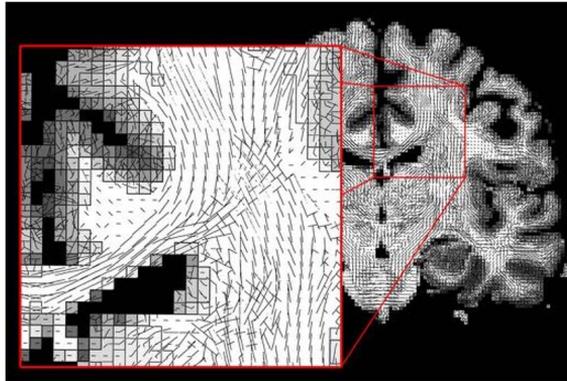


Figure 2.4. In particular points of the diffusion images up to five fiber bundles can merge in the space occupied by only two voxels, making them indistinguishable.

thus making assumptions on the polarity of a group of fibers intrinsically wrong there might be differences in the axons composing it.

False positives, especially in large tractograms, are all over the place [88]. This is the combination of different, not easy to tackle, problematic. Additionally, it is not even simple to define what a false positive is, since these type of fibers could be more correctly described as the probability of a mis-estimation from the tractography algorithm. The most problematic causes are the over representation of large bundles due to seeding mechanisms, inaccuracy in the white matter / grey matter interface and resolution issues. Bottlenecks are frequent and often gathers fibers with totally different paths in the space of a couple of voxels [56], making a thrust worthy reconstruction nearly impossible.

Although some of this obstacles seem insurmountable due to technological limitations, huge steps have already been made towards more faithful and consistent tractographies. The most prominent works are relative to the post filtering of tractograms in order to tackle the false positive problems [15] [83]. These approaches include the incorporation of microstructural proprieties, removing streamlines that does not represent the diffusion data. In particular since a streamline is the visual interpretation of a group of axons it must have a volume, it is not independent by the others of the same functional bundle and the density must be conserved [19].

In conclusion, while at the present time different issues still afflict the tractography based analysis, a lot of methods are currently being developed to push forward

the accuracy of this image modality, the most promising being the inclusion of microstructures information. Additionally, the correlation to functional MRI data, the advance in post-processing filters and the continue progress in the understanding of the brain anatomy make the tractography a research worthy topic with a lot of breakthrough expected in the near future.

2.3 MRTrix3

With so many different diffusion imaging techniques, fiber orientation estimations options, tractography algorithms, data formats and processing steps introduced in the scientific literature the necessity of having a software suite that manages all these aspects is crucial. MRTrix3 [89] proposes to perform these and other operations wrapping different commands and selections in an intuitive environment.

The suite, downloadable from the creators site <http://mrtrix.readthedocs.io/>, is a command line based tools that implements all the above functionalities while being constantly kept up-to-date with newly introduced algorithms, by a team fully dedicated to its maintenance in regards to tractography and diffusion imaging innovations.

In this work MRTrix3 has been used to perform the fiber tracking on the Human Connectome Project data (Chapter 1.3). A list of commands based on the specific needs has been prepared for a serial execution.

Due to the accuracy searched a simple diffusion tensor fitting would not satisfy the requirements, for this reasons the constrained spherical deconvolution was chosen for the computation of the fiber orientation distribution. The SD_STREAM (Chapter 2.1) tractography algorithm allows to obtain more consistent inter-subject results, useful for validation purposes, over other probabilistic approaches. Finally, a filtering operation was omitted as not considered important for the validation of Fuzzy Tracts.

It is worth noting that the software works with specific file formats, not supported by other neuroimaging application. While not problematic for visualization purposes, since a visualization tool is already implemented in the suite, it forces a format conversion operation during the development process.

Starting from the DWI, the values associated to the gradients (bvalues and bvectors files) and a brain mask the process is divided in three parts:

- Response function computation: `dwi2response`;

- Fiber orientation distribution extraction: `dwi2fod`;
- Tracking: `tckgen`.

The `dwi2response` function computes the CSD fiber response for a group of streamlines with the same orientation. It is possible to use more sophisticated algorithm to demarcate the contribution of the WM, GM and CSF to the response [21].

From the response and diffusion data the fiber orientation distribution is computed voxel-wise. With multi-shell data it is possible to obtain a more robust estimation in white matter / grey matter boundaries and in brain parts formed by the cerebrospinal fluid [41].

Finally, the tracking can be carried out with a selected algorithm (e.g. `SD-STREAM`) after selecting the seed sampling method and the fiber-specific parameters (length, number, ...).

Chapter 3

Use Case: Uncinate Fasciculus (UF)

The presence and organization of white matter fiber bundles, constituted by multiple axons having similar geometry and function, has been known for year, with the first accurate, and today still relevant, atlas dating the early 50' [45].

The advances in the diffusion MRI introduce new ways of defining white matter structures, for the first time in a non-invasive fashion. Due to resolution problems, far from being solved, it is still impossible to define the connectivity of the brain on an axonal scale but, at the same time, the macroscopic organization is clearer than ever.

Different atlases [1] [68] report the major fiber bundles organization basing the definitions on diffusion data analysis (Figure 3.1) and the validation on prior knowledge, obtained from already tested classical method, like ex-vivo dissection.

The new opportunity of having an in-vivo representation of the connectivity allows an even more specific characterization of the group of fiber composing the brain, thus making a bundle-based investigation more compelling then a whole-brain study (Chapter 4). Furthermore, the validation of a tool on large tractograms would not be a simple and immediate task.

Of the multitude of anatomically known and literature defined bundles the choice for the validation of Fuzzy Tracts feel onto the Uncinate Fasciculus for two major reasons: the quantity of articles that it figures in and the rather smooth and packed spatial geometry (Figure 3.2).

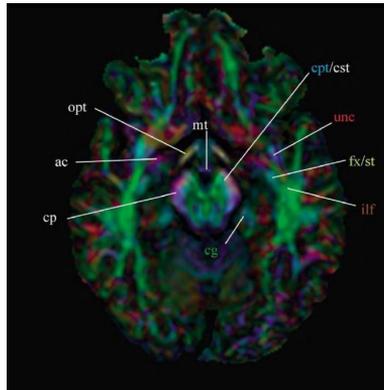


Figure 3.1. Section of the Mori atlas based on the colored fractional anisotropy map.

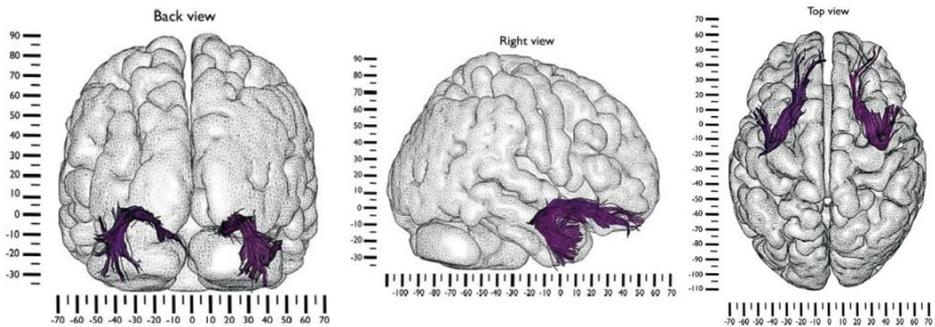


Figure 3.2. Uncinate Fasciculus obtained through manual ROI positioning (Catani technique).

The neurological role of this bundle is still debated with most studies reporting a link between a pathological reduced activity of the Uncinate Fasciculus and memory issues [99], lesions especially creating issues during the retrieval of memories [58] and names. It has also been proposed the involvement of the bundle in the language regulation, with deficit observed after the removal of the fibers.

The retrieval of the bundles definitions can follow two main strategies: ex-vivo dissection and in-vivo imaging. The dissection-based methods allow to distinguish fiber bundles that do not have well recognizable contours in MRI scans [44].

The Uncinate Fasciculus takes its name from its particular shape [23] resembling an hook. This peculiar characteristic has been validated and considered accurate by nearly everyone in the neuroimaging community [76] [71].

While the definition about the shape is vastly accepted there are huge discrepancies between interpretations regarding the more detailed anatomy of the bundle. In particular, dissection techniques produced different results, often reporting conflicting definitions about the relative position to the insula [60] [71] and in the anatomy of the portion that lies in the temporal lobe [59] [73].

The most important contribution from a radiological point of view are given in form of atlases [16] [102] and refers to the Uncinate Fasciculus as a pack of fiber that links the temporal pole to the frontal pole arching at the level of the external capsule.

A clustering of the bundle showed five fiber groups [37] that connects slightly different regions. These regions have been evaluated from dissections studies and the exact termination of each part mapped.

An added difficulty in the examination of the Uncinate Fasciculus is its proximity to other vast fiber bundles. In particular, the Inferior Fronto-Occipital Fasciculus (IFOF) intersects the UF in the external capsule [57] and both of them join the claustrorocortical fibers under the inferior limiting insular sulcus [76] (Figure 3.3).

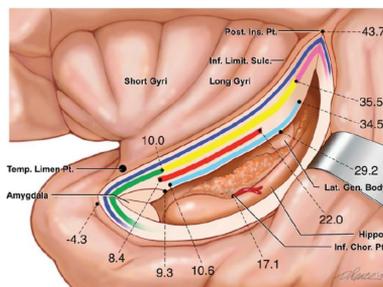


Figure 3.3. Fibers with different origins and functions all run close to each others in the temporal stem. In green the fibers belonging to the UF. In yellow the fibers belonging to the IFOF.

The characteristic fractional anisotropy is a difficult information to define since a lot of variations have been observed for different age, sex and pathologic conditions. While there is no evidence of a consistence asymmetry in the FA values for the two hemispheres [101] [94] it has been noted a significant increase of the FA

values moving towards the frontal lobe (Figure 3.4).

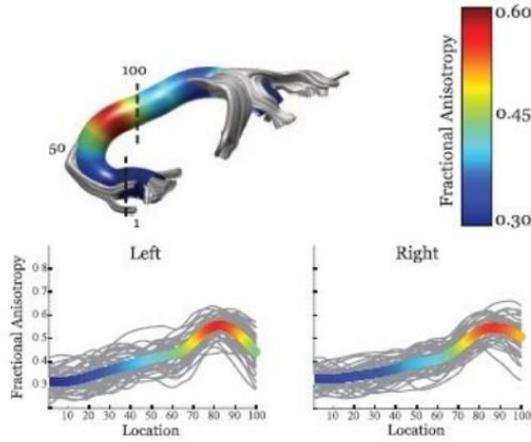


Figure 3.4. Typical FA profile in the UF obtained meaning the data from 48 subjects.

The asymmetry has also been evaluated in terms of volume occupied by the fibers. In this parameter most of the studies give concordant results, a noticeable amount of asymmetry has been observed with the right hemisphere expressing one third more fiber than its counterpart. The results stands even for different sex and in the presence of pathologies [38].

Part II

State Of The Art

Chapter 4

Segmentation Methods

Working on whole-brain tractograms is not an easy task due to the huge number of streamlines that compose them. Not only they are very difficult to interpret visually but also from a computational point of view operating with millions of streamlines could quickly chock up the machine virtual memory and CPU.

In order to provide clinicians useful information about the brain behaviour there is a necessity to divide tractograms into anatomically significant parts. Various segmentation methods have been developed during the years with the goal to extract group of fiber with similar proprieties and biological functions.

For functionality recovery estimation and surgical planning the tractogram analysis usually focuses on the specific type of white matter fibers involved in the biological processes correlated to the pathology. A segmentation algorithm not only should provide accurate results but it must be able to adapt to the specific operation in regards to both functionalities and applicability.

The diversity of the streamline species (commisurals, associations and projection fibers are all present in a tractogram) adds a supplementary level of complexity to the problem, especially for larger fiber bundles composed by a variagated population.

Furthermore, the consistency and precision of the results are not parameters immediate to evaluate since universally accepted bundle definitions have not been found yet. The analysis of the results can only be done in qualitative way and its a prerogative of every segmentation algorithm to allow the clinician to perform a critical interpretation of the output, for an eventual modification or acceptance.

The segmentation techniques vary greatly between each other for concept, implementation and applicability but three main categories can be identified: manual segmentation, clustering algorithms and mixed approaches.

4.1 Manual Segmentation

For this type of segmentation methods the fibers selection is carried out directly operating on diffusion images, manually selecting the zones (set of voxels) where the fiber bundle is thought to pass by. This parts, called Region of Interest (ROI), could be additionally be classified in region of inclusion (highlighting the fiber path) and region of exclusion (brain parts not crossed by the bundle).

Since the ROI are directly drawn onto the DWI or on the, often more indicative, FA map, an extensive knowledge of the human brain anatomy must be possessed by the user as in both cases the streamlines are not directly visible.

While different ROI positioning strategies have already been defined for the most common fiber bundles [1] [101] (Figure 4.1) the final decision is in the hand of the operator, making the clinician experience a valuable added factor.

The manual drawing of the ROI, often carried out on external software on PCs or tablets, consists in the creation of different masks with the use of a pointing device (mouse or stylus). Due to the human factor introduced this method does not allow to obtain a consistency in the results even for multiple segmentation of the same bundle, in the same subject, by the same operator.

There are mainly three types of operation used to exploit the region of interests approach [102]: AND, CUT, NOT. An AND between two ROI will select all the fibers that pass through both of them, the CUT, in a similar fashion, is an AND operation where the streamlines are truncated in the position they cross the ROI, the NOT operation transforms the region of interest in a region of exclusion (Figure 4.2).

Instead of a real-time segmentation operation the drawing of the region could be carried out on a standard brain space in order to form a mask image to be used for subsequent segmentation [110]. While this approach can speed up the manual processing the validity of the results is not guaranteed due to potential lack of correspondence between the subject brain structure and the reference space.

The intrinsic variability tied to the human aspect involved makes the manual segmentation a low reproducible and of difficult validation technique. Furthermore,

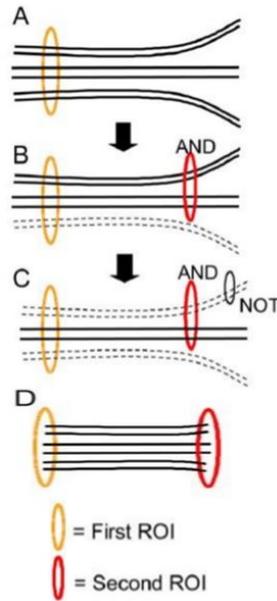


Figure 4.1. ROI position strategy edited by Catani.

it is not possible to adapt an ROI-based approach to statistical studies since the time-consuming segmentation operation scales really poorly with the number of subjects.

Ultimately, the pros of using a full manual method (customization and control over the segmentation process) are counterbalanced by a vast quantity of cons (time of the procedure and result consistency over all) limiting severely the field of application of this approach.

Figure 4.2. Different region of interest position strategies. a) first ROI b) the AND eliminates the fibers that do not pass through it c) the NOT eliminates the fibers that pass through it d) the CUT combines the AND operation with a truncation.



4.2 Clustering Methods

While manual segmentation methods attempt to divide a tractogram starting from the delineation of the regions traversed by the fibers, using prior anatomical knowledge and user experience, clustering algorithms exploit the mathematical properties of the streamlines to aggregate the most similar ones.

The clustering methods usually offer very little interaction with the users, separating tractograms into fiber groups (clusters) only based on objectives and quantitative information. Each cluster should contain streamlines very similar between each others (minimizing the intra-cluster variability) and well separated from the other clusters (maximizing the inter-cluster distance). The exact definition of the concept of similarity is central to every clustering algorithm and must be carefully designed to fit the algorithm characteristics to the application they need to perform.

One of the most immediate, and commonly used, distance metric is the MDF function (Figure 4.3). A point-wise distance function (like an Euclidean or an Hausdorff equation) computes and means the distance contribute for each pair of point composing two streamlines. The final distance will be considered as the minimum between the result obtained from the two fibers as they are and the result in the situation one of them is flipped, in order to avoid issues related to orientation

mismatches. This metric requires a downsampling of all the fibers to the same number of points to have a correspondence along their entire lengths. This basic metric has been implemented in simple clustering algorithms [30] that aggregates the fibers when the distance is under a certain threshold while creating a new cluster otherwise.

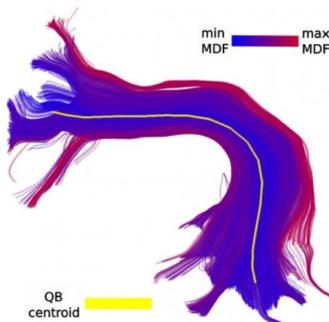


Figure 4.3. UF fibers colored with their MDF values associated, calculated as the distance from the bundle centroid.

Similar techniques, also using pairwise distance metrics, can compute the distance between each fiber of the tractogram and reference fibers obtained through a manual segmentation with ROI placement [110] or being used in combination with atlas based approaches [36].

The main drawback distinctive of quick-to-implement metrics is the very little quantity of fiber information (diffusion and structural proprieties in articular) taken into consideration, the necessity of a downsampling operation (reducing the precision to maintain a pairwise correspondence) and the adaptability to large dataset, making them a valid choice only when the goal is to perform a qualitative segmentation of small tractograms.

The diffusion information, usually not taken into consideration by classical clustering algorithm, could characterize and well distinguish fiber bundles. The behavior of the fractional anisotropy for example has been shown to be consistent between different subjects [109].

Clustering techniques differentiate from each other for the use of different metrics, for example estimating the curvature and torsion of the streamlines [18] or their connectivity [14], and different cluster-creation methods [63], often based on

splitting techniques already introduced in other fields, like the k-means and c-means algorithms.

The choice on what metric and algorithm fit better to the implementation required, completely in the hand of the user, could create a lot of confusion and pitfalls, especially when the decision has to be expressed by someone without an engineering background. Finally, the impossibility to develop a clustering technique able to easily adapt to every tractography application introduce the need to institute a more universal segmentation approach.

4.3 Different Approaches

Due to the intrinsic limitation of both fully manual and fully automated approaches, hybrid methods are more and more relevant and developed. Trying to combine multiple sources of information could bring the solidity, so hard sought by the segmentation algorithm.

The ROI-based segmentation idea could be modified to create an automatized method, more fit to multiple subject studies, using an atlas of the brain to guide the ROI positioning [50]. The use of an atlas allows to reduce of the time-consuming operation of drawing the inclusion and exclusion regions to a one-time operation. A deep parcellation of the brain assigns to every white-matter portion a label number, that can be sequentially decoded to a literature defined anatomical area.

In these atlas-based methods the modus operandi tends to be very close to the one of a fully manual segmentation with the atlas generation being fully hand-made by an expert. Each subject scans will then be registered to the atlas and the fibers that pass through the chosen label number will be segmented.

Voxel-based approaches try to assign each voxel to the most probable fiber bundle they belong to, using models to express the local diffusion proprieties and tracts connectivity [5] or directly looking at the fiber orientation distributions [104].

Finally, fully hybrid approaches can combine clustering algorithms with atlases and tissue segmentation methods [53] or use newly defined strategies, like spatial relations between structures (Chapter 5).

The vastness of segmentation options today available, greatly differing in concept, is an indicator of the importance of the topic and shows the need for a solid and complete ultimate tool. Fuzzy Tracts tries to fill this aspect with a complete hybrid approach as described in Chapter 6.

Chapter 5

White Matter Query Language (WMQL)

The White Matter Query Language (WMQL) [106] is an atlas-based segmentation method that introduces a new approach: an human readable segmentation language. The users have the ability to define a white matter fiber bundle through the use of a specific syntax, built using neuroanatomical terms (Figure 5.1).

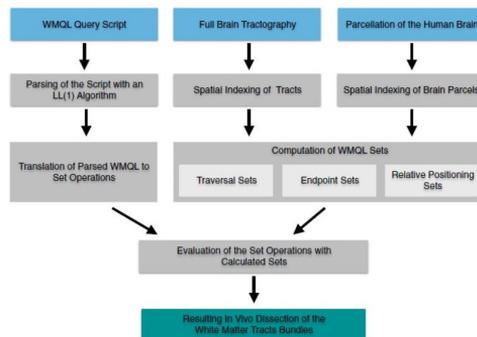


Figure 5.1. WMQL pipeline.

With the language proposed a bundle can be identified through the use of different types of position based definitions [105]. A fiber can be specified as a line that start from an anatomical area and terminate in another one or it can be described as a line that pass through a particular zone. Finally, relative spatial relation could be used to frame the bundle in a bounding box selection.

The feature that stands out is the concept of the near-to-English segmentation language proposed. Having a textual reference could impact the way clinicians defines fiber bundles, standardizing the approach. It also fits very well when the tool must be used by people without a engineering/mathematical background, like neuroscientist. Furthermore, this approach allows to easily modify a bundle definition, formalizing the description to keyword and names the users only concern is to respect the introduced notation.

The combination of all the spatial position based definition forms the exclusion/inclusion rules, that applied element-wise to all the tractogram fibers produce the final segmentation.

The application of the algorithm requires three preliminary steps:

- Complete brain labelling (amygdala, putamen, thalamus, ...) in order to define the reference (atlas-based approach);
- Collection of information about the relative position of the fiber bundle in respect to the labels (query);
- Fine tuning of the algorithm parameters.

With a query, an atlas and a tractogram, WMQL can be run from the command line to extract a segmentation in the TrackVis file format [103].

5.1 Queries

The list of definitions, also called query, has to be compiled directly by the user and it offers the maximum liberty in regards of the customization possibility (Figure 5.2).

A parser has the goal of read the queries and transform the descriptive content in labels and operators more suited for computational purposes. The queries interprets the use of some keywords and label names. The fiber connectivity is expressed defining where its starting and ending point lie (e.g. *endpoints_in(temporal.side)*), it is important to highlight that this definition does not distinguish between starts and ends, distinguishing the two requires the use of the operator *and*. For the crossing area option no keywords are needed (e.g. *temporal.side*) but it is possible to specify the percentage of the fiber points that must pass trough the region. These keywords and those expressing relative positions are reported in table 5.1.

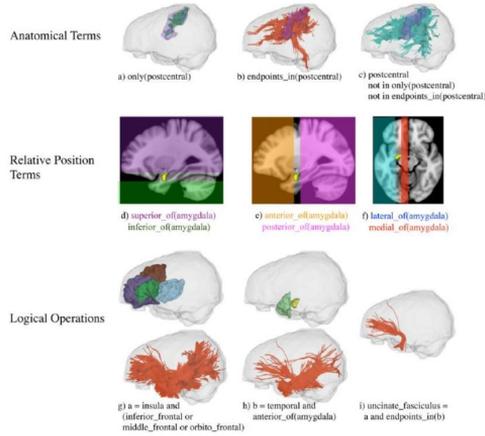


Figure 5.2. Terms available for defining WMQL queries and example of the UF segmentation.

Keywords		
<i>and</i>	<i>endpoints_in</i>	<i>lateral_of</i>
<i>or</i>	<i>both_endpoints_in</i>	<i>inferior_of</i>
<i>not : in</i>	<i>anterior_of</i>	<i>superior_of</i>
<i>not</i>	<i>posterior_of</i>	
<i>only</i>	<i>medial_of</i>	

Table 5.1. WMQL supported relations for queries definition.

In addition, a last parameter, given directly through the command line, allows to define the minimum length, in millimeters, the fibers should have to belong to the bundle.

The difficulty about defining a query is the lack of gold standard definitions. In fact, as explained in Chapter 3, even for a relatively compact and reproducible bundle like the Uncinate Fasciculus every author slightly modify the phrasing and it's not rare to find discordant definitions.

In the end the delineation of the queries must be modified for each image/atlas pair to best fit the the results to the application searched. The final query used in Fuzzy Tracts combines the UF definition from the papers of Catani [16], Wakana [101] and Leng [52].

5.2 Limitations

Like every other atlas-based approach (Chapter 4.3) WMQL is very sensitive to the correspondence between the parcellation file and the image in analysis. The combination of user defined queries and atlas-image differences diminish greatly, not only the inter-subject reproducibility, also the level of confidence in subject-specific studies.

Noticeable discrepancies can be observed between the results obtained with literature based queries and experimental queries (Figure 5.3, an operator could doubt the correctness of the definitions and be disoriented by the differences between the segmentation produced and the expected output. Finally, the quality of the tractograms can affect the introduction of other errors, the presence of false positives that have a path similar to the true positives could produce misleading results.

While the last issue can be fixed adding new definitions a posteriori of an initial trial segmentation, the atlas and queries induced variability are obstacle not easy to overcome making the necessity to extend the framework is particularly evident in situations where the confidence is a key factor, like a pre-surgical planning.

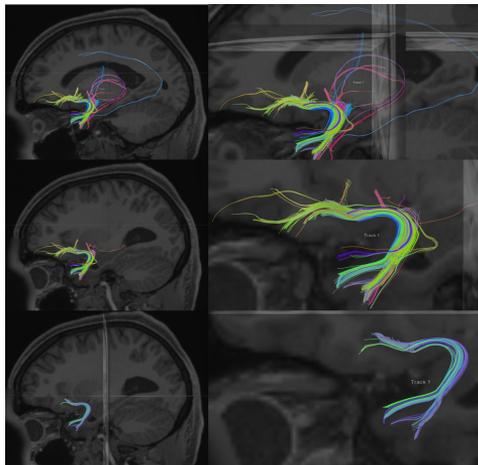


Figure 5.3. The inter-subject variability does not allow to have consistent bundle definitions. The necessity to add custom information to remove the outliers is evident in the UF segmentation. Top: WMQL default query. Middle: Combination of Catani and Wakana definition. Bottom: Experimental query.

Part III

Fuzzy Tracts

Chapter 6

Concept

Fuzzy Tracts, the tractograms segmentation algorithm proposed, has the goal of accurately segment the majors white matter fiber bundles in an accurate and reproducible way, keeping an user-centered environment.

The segmentation algorithms proposed in the literature all fail short on producing optimal performances and applicability due to the different problems that affect them.

The ROI approach, being completely user dependant, could output results overall not rrepresentative of the bundles and highly discordant between operators since there is no gold standard region of interest/exclusion position strategy. In addition the method has a really low scalability, the time-consuming operation of manually delineate subject specific region-of-interest makes it not feasible for vast dataset applications and statistical analysis [50].

The clustering algorithms, on the other, hand offer little to no interaction with the users and are often too dependant on the metric selected or the classification algorithm chosen. Furthermore the metrics are, most of the time, based on difficult to define parameters and the clustering methods might require additional, not known a priori and difficult to validate, assumptions [63].

Finally, the atlas based approaches are not able to extract meaningful information when the inter subject correspondence hypotheses is void. This methods offers an high scalability as the strategy can be defined just one time and applied on each subject of the dataset, making them very useful in statistical analysis, but less relevant for subject specific studies.

Excluding the manual ROI segmentation, the other methods could be combined together in order to find the optimal compromise between results quality, feasibility, number and coherence of the hypothesis.

Fuzzy Tracts, exploiting the spatial relation definition concept introduced by WMQL, tries to improve the reproducibility and to produce more consistent outputs adding an interactive confidence level thresholding to the process and integrating and extending the already existent framework.

With a new view, an uncertainty factor is modeled defining spatial fuzzy sets. Assigning a degree of uncertainty to the output segmentation allows to perform a critical analysis of the results, thus enabling an on-the-fly validation.

The absence of standard definitions, in respect of positions, geometries and diffusion information, limits the possibility of creating a fully automatized tool. The user must be able to use his experience and prior knowledge to customize and optimize the segmentation result in order to best fit the output to the particular needs.

In conjunction with the fuzzy approach an innovative metric for fiber clustering is here implemented and tuned. This allows to improve the feeling of control of the process adding ways to manage the part of the algorithm usually in the hand of the machine.

Since using just one feature is not adequate to obtain anatomically meaningful clusters, the metric should include both structural (position, geometry, ...) and diffusion (anisotropy, diffusivity, ...) information. The choice for the metric fell on the functional varifold, adding the FA data to the fiber proprieties.

Furthermore, the tool needs to be feasible for the usage in the most common tractography applications and environments. This implies that it must be practical and convenient to use also by people with different knowledge background, not limited to computer science. The computational times must be coherent with the surgeon request, in case of pre-surgery planning, or with the data scientist needs, for statistical analysis applications.

The great number of options supported by the tool does not interfere with the operator experience, the process is all around kept user-friendly reducing the final decision to the choice of a threshold value. More expert operators will find the opportunity to use supplementary fine tuning parameters to improve the segmentation, managing every step of the process.

Finally, Fuzzy Tracts inherits by the WMQL the human readable language idea.

The queries already defined can still be used in the situation the user wants to use the WMQL framework in an initial step but if, on the contrary, the segmentation strategy wants to deploy all the new native functionality, the same queries can be intuitively adapted to the new configuration file.

All the other tuning options (metrics parameters, thresholding, ...) can also be easily adjusted in the same settings file, thus avoiding the dispersion of information in multiple location of the code, keeping a more clean layout.

In this part of the thesis the processes behind the tool will be presented in their entirety, covering all the aspects from the input to the final segmentation, detailing the theoretical basis and the practical implementation, and finally performing a critical analysis of the results obtained highlighting the pro and cons of the software.

Chapter 7

Spatial Fuzzy Sets

The idea behind a fuzzy set approach is to complete the problem delineation with information expressed in a natural way and modeled to reflect the human thinking. Fuzzy logic tries to implement a conceptualization of the thought processes the experts create, often composed by vague and ambiguous terms relevant only with a knowledge of the context, when solving a specific task. Fuzzy logic model the uncertainty admitting continuous degrees of membership to different sets. While the crispy logic decisions follows a binary approach, either a condition is true or it is false, a fuzzy classes makes use of membership function, defined in an assigned interval, to assign to each element an association score.

The fuzzy approach is divided into two different part: the definition of a rule and the application to the result. While for the crispy logic the state of the output (true or false) is equal to the result of the evaluation of the rule in a fuzzy set there is not a concept of state but more of degree of membership.

The interaction between fuzzy rules are extensions of logical operation AND (set intersection, also called t-norm) and OR (set union, also called t-conorm), their conceptual definitions are: how much an element belong to both sets? how much an element is in wither sets?

A last step of the process is the defuzzification. To obtain meaningful results the uncertainty must be resolved in order to reach a crispy decision (e.g. fiber is in bundle or fiber is outlier). The defuzzification aggregates the results in output from the application of the rules to the dataset. The definition of how to aggregate these results is key to the modellization of the final output.

The fuzzy sets applications are not limited to the biomedical field, it covers

topics like management systems, geography, ... [48] and more generally wherever the crispy approach is not able to solve situation naturally answered by the common human experience [70].

In image processing the spatial fuzzy sets are the natural extension of the described uncertainty framework, for a direct application on 2D and 3D structures. Maintaining the same process already defined the fuzzy rules are computed directly on a object and evaluated in all its points.

The class of the spatial fuzzy set methods used in Fuzzy Tracts tries to assign a degree of uncertainty to the definition of directional relative position, concept often not well delineated. Fuzzy approaches provide a way to interpret these definitions and compute their consistency. While topological positions are deeply investigated not much is done for directional relations due to the difficulty to quantify definitions as 'right of' and 'left of', especially in situations like Figure 7.1.

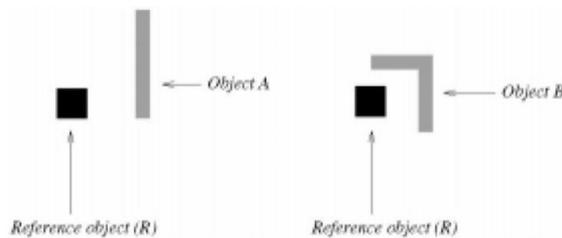


Figure 7.1. While for the case on the left is possible to define the object in respect to the reference, for the image on the right it is not possible to decide exactly if the object B is to the right or up to R.

One first approach, introduced in [62], consists in computing the angle θ formed by the segment that joins two points and the reference ax (equal to one of the four main direction of a 2D image). In this situation the membership functions are dependant on θ and defined in $[-\pi, \pi]$. Initially accepting only on the four principal direction, the extension on multiple direction and 3D cases adds a versatility of use.

A method to increase the directions number from four to eight combines the precedent direction definitions to form the oblique directions, still remaining limited in the precision. The extension to a 3D image is possible using a different way to define the directions, expressing them as the combination of two angles α_1 (defined in $[0, 2\pi]$) and α_2 (defined in $[-\pi/2, \pi/2]$).

In order to obtain a more precise fuzzy representations the relative position

could be computed with the fuzzy pattern matching approach [22], composed by two steps: a fuzzy landscape is prepared based on the reference object points, the target is directly mapped on the fuzzy set.

Let's consider two object R (reference) and A. The membership function, defining the degree of direction of A in respect to R, $\mu_A(R)$, can be represented on the image S in the interval [0,1], where high values means high probability and low values picture a low probability. The uncertainty of how much A, in all its points, satisfies the spatial relation in respect to R is then based in function of $\mu_A(R)(x)$ and $\mu_A(x)$.

The fuzzy pattern matching define probability as:

$$\prod_{\alpha_1, \alpha_2}^R(A) = \sup_{x \in S} t[\mu_{\alpha}(R)(x), \mu_A(x)]$$

The definition of $\mu_A(R)$ is thus central part of the whole expression. The membership function definition, being user dependant and specific to the results needed, must be carefully crafted.

The membership function used by Fuzzy Tracts considers the visibility space saw by a reference object in the direction specified [11].

For each point in S, the point of R which form the smallest angle β between the segment that joins them and the direction chosen, is identified. The landscape in a point P will then be defined as: $\mu_A(R)(P) = f(\beta_{min}(P))$ where f is a function defined in $[0, \pi]$ that can assume values in the range [0,1], for example the following:

$$\mu_A(R)(P) = \max(0, 1 - \frac{2 * \beta_{min}(P)}{\pi}).$$

Furthermore, seeing the problem from another perspective it is possible to indicate $\mu_A(R)$ as a fuzzy dilation of a structural element v (defined on the support).

This brings the notation used by Fuzzy Tracts to:

$$v(P) = \max[0, 1 - \frac{2}{\pi} - \arccos(\frac{\vec{OP} * \vec{\mu}_{\alpha}}{\|\vec{OP}\|})].$$

Where P is a point of the support and O the origin, in this tools the origin has been assumed has the center of the image (Figure 7.2). The final dilation algorithm will then consist on applying:

$$D_v(\mu)(P) = \max_{Q \in S} t[\mu(Q), v(P - Q)].$$

Due to computation time issues this approach has been chosen even if it brings an approximation of the results. Another more precise implementation, using a propagation algorithm [12] is available but not directly wrapped into Fuzzy Tracts.

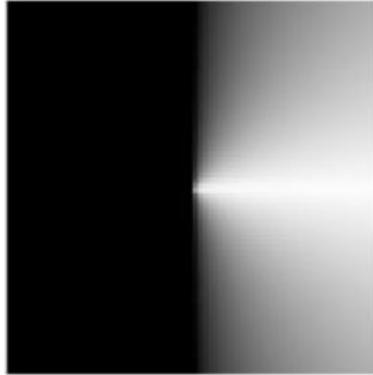


Figure 7.2. Structuring element centered in the middle of the image for the definition 'to the right to'.

Chapter 8

Functional Varifolds (fVar)

Most of the streamlines clustering algorithms, as reported in Chapter 4, fails short in some key aspect and, even more importantly, do not combine diffusion data with streamlines geometries. This loss of information originates a low level of specificity that limits greatly the quality of the results.

The functional varifold, an extension of the weighted currents framework [34], combines tracts morphology and microstructure information, using the fractional anisotropy in the similarity measure formulation.

In this approach every fiber is approximated to a line composed by a certain number of segments and described by their centers and tangents. With this view a fiber X formed by P segments can be rewritten as:

$$V_{X,f}(\omega) \approx \sum_{p=1}^P \omega(x_p, \beta_p, f_p) c_p.$$

It is immediate to define the inner product between the streamline X and another one Y (composed by Q segments) using this notation:

$$\langle V_{(X,f)}, V_{(Y,g)} \rangle = \sum_{p=1}^P \sum_{q=1}^Q k_f(f_p, g_q) k_x(x_p, y_q) k_\beta(\beta_p, \gamma_q) c_p d_q.$$

Where f and g are a specific function computed along the fibers points (the fractional anisotropy in our case). The three kernels that appear in the expression regulates the sensitivity and the response to each factor.

In the implementation used by Fuzzy Tracts k_f and k_x are assumed as gaussian kernel and k_β is a Cauchy-Binet kernel, in this way the final expression can

expressed as:

$$\langle V_{(X,f)}, V_{(Y,g)} \rangle = \sum_{p=1}^P \sum_{q=1}^Q e^{-\frac{\|f_p - g_q\|^2}{\lambda_m^2}} * e^{-\frac{\|x_p - y_q\|^2}{\lambda_w^2}} * \left(\frac{\beta_p^T * \gamma_q}{c_p * d_q} \right)^2 * c_p * d_q.$$

This formulation does not require a downsampling of the fibers like most others point-to-point correspondence based clustering algorithms do (to have every streamlines with the same number of points). Furthermore, a reorientation of the fiber is also not necessary since the metric weights the differences in segments features keeping in consideration the distances between them.

The functions f and g are exploited to model the diffusion information given by the fractional anisotropy map while the other terms consider the differences in the geometry of the streamline (Figure 8.1). Wanting to add even more structural information with a new approach the distances between start-points and end-points are inherited from the weighted currents. This modification forces a preprocessing consisting on orienting the fibers in the same direction since, if not, it would not be possible to define what are the points considered as start and those considered as ends (Chapter 9.4.1).

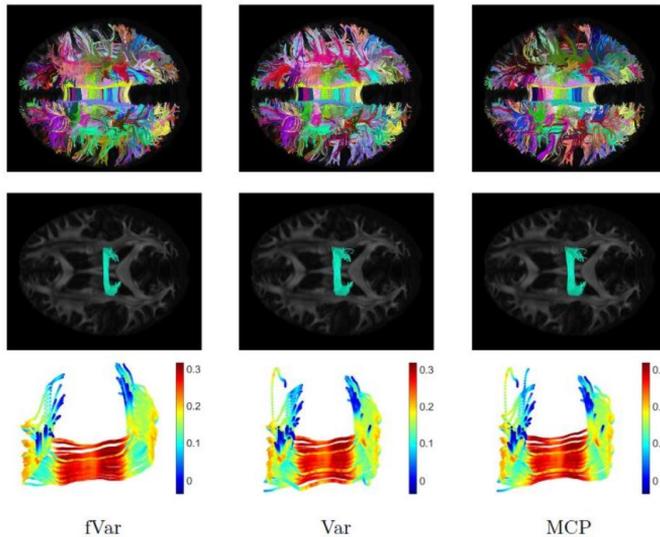


Figure 8.1. Whole brain clustering using the functional varifolds (left), the varifold without considering the diffusion information (center) and the MCP distance (right).

In the weighted currents framework [35] the new information about the connectivity can be added to the already defined distance metric. Assuming two fibers X and Y it is possible to specify the inner product as:

$$\langle C_X, C_Y \rangle = k_a(f_a, t_a) * k_b(f_b, t_b) * \sum_i^N \sum_j^M [\alpha_i^T k_g(x_i, y_i) \beta_j].$$

Combining the weighted currents to the varifolds and applying the kernel as previously described, the final inner product metric implemented in Fuzzy Tracts between two streamlines X (composed by N points) and Y (composed by M points) can be expressed as:

$$\langle C_X, C_Y \rangle = e^{-\frac{\|f_a - t_a\|^2}{\lambda_a^2}} * e^{-\frac{\|f_b - t_b\|^2}{\lambda_b^2}} * \sum_i^N \sum_j^M [e^{-\frac{\|x_i - y_j\|^2}{\lambda_w^2}} * e^{-\frac{\|FA_i - FA_j\|^2}{\lambda_m^2}} * (\frac{\alpha_i^T * \beta_j}{|\alpha_i| * |\beta_j|})^2 * |\alpha_i| * |\beta_j|].$$

Where f_a and f_b are the extremities of the streamline X and t_a and t_b are the extremities of the streamline Y . In the parenthesis instead there are the terms that regulate the distance based on the geometry, x and y are the coordinates of the center of the segments composing the respective streamlines. The difference between the FA values are computed in these centers and the last term confront the difference between the tangent between those segments.

The widths of four kernels must be imposed by the user in function of the application searched:

- λ_a : regulates start-points distances contribute (mm);
- λ_b : regulates end-points distances contribute (mm);
- λ_w : regulates segment distances contribute (mm);
- λ_m : regulates FA differences contribute ([0,1]).

The kernels should be chosen with in my mind the anatomical characteristics of the bundle to segment, since they bring a 'real-world' information. When a contribute exceed the kernel width by three times the lines will be considered orthogonal for that feature.

Applying the equation to every pair of fibers composing the tractogram produces an inner product matrix, with dimension $K \times K$ (where K is the number of streamlines), that can be exploited to compute a distance matrix or other measures, like the cosine similarity.

The matrix, being symmetric in respect to the main diagonal, can be obtained looping only on the upper half in order to save computational time and hardware

resources.

The kernel widths searched for Fuzzy Tracts are bundle based, values that allows to produce similar results for different subjects are sought. To improve the consistency it is suggested to impose more relaxed parameter, in particular for the Uncinate Fasciculus use case the four kernel were tuned to: $\lambda_a = 5 \text{ mm}$, $\lambda_b = 6 \text{ mm}$, $\lambda_w = 3 \text{ mm}$, $\lambda_m = 0.05$.

Chapter 9

Implementation

As noted in the tool introduction (Chapter 6), the whole process to obtain the segmented output fiber bundle is composed by different segmentation methodologies, an atlas based approach is used for the definition and the computation of the spatial fuzzy sets, meanwhile a clustering technique, that exploits the functional varifold metric, anticipates the final thresholding (Figure 9.1).

The procedure can be seen as a succession of different steps, each performed in series to the previous.

In total five major section of the algorithm can be identified:

1. Fuzzy Sets computation;
2. Initial Segmentation;
3. Fractional Anisotropy extraction;
4. Clustering;
5. Final Segmentation.

The data needed are the tractogram and either the folder containing the diffusion data of the subject or a precomputed FA map. This structure gives the user the ability to obtain a tractography with the modalities he prefers and furthermore, it is possible to use suites of softwares or stand-alone package that would be difficult, if not impossible, to wrap in the main code. Finally, it speeds up the segmentation as the modern tractography algorithms usually require a powerful machine and a low-level coded algorithm.

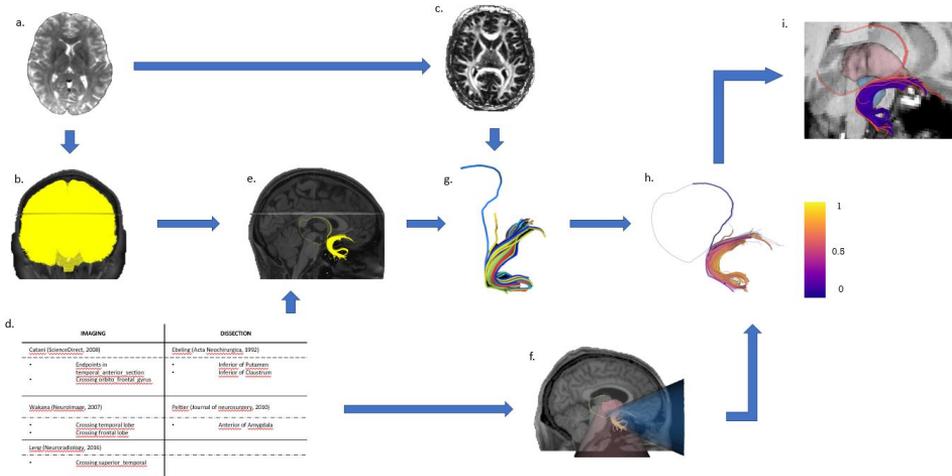


Figure 9.1. Fuzzy Tracts Pipeline. From the DWI (a) both a whole brain tractogram (b) and an FA map (c) are computed. Through the use of spatial relations (d) an initial segmentation (e) is performed with WMQL. The same relations are applied to produce the spatial fuzzy sets (f), in the figure just the amygdala and putamen are showed. Combining diffusion, connectivity and geometric information a set of clusters is generated. The fuzzy values are then mapped onto the prototypes (g) of each cluster and propagated to the other fibers (h). Using a threshold the user is able to modify the final segmentation (i)

Different options and configuration files are available, for more information it is best to check the User Guide (Chapter 11).

Before running the code it is mandatory to configure the tool. This is done editing one configuration file (or two, if the WMQL option is chosen). The file follows the JSON format [25]. This format has been preferred over others markup languages, like XML, for its user friendliness, and over other configuration types of files, like INI, for its more powerful features.

The configuration files must specify: the name of the fiber bundle (it will compose the name of the output file), the spatial relation in respect to the six principal directions (anterior, posterior, left, right, up, down) using the Freesurfer notation [28] and the width of the kernel used by the function varifold.

After all the parameters are set, the spatial fuzzy sets are computed for each bundle defined. There are no limitation on the number of bundles to segment at the same time or the number of sets defined, as long as they are coherent with the

Freesurfer definitions and the labels used are present in the atlas.

An initial segmentation of the whole brain tractogram is performed either using the White Matter Query Language [106] or using the sets. The diffusion tensor is then fitted to the diffusion image and the fractional anisotropy is extracted and given as input to the functional varifolds. After the distance matrix is computed, using the fVar, a clustering algorithm (density based or spectral clustering) divides the initial segmentation into small groups of fibers. The fuzzy sets are finally mapped on the centroids of the clusters and through the use of a threshold the resulting output can be fine tuned to best fit the user needs.

9.1 Fuzzy Sets computation

The fuzzy sets definitions and computing are central to the whole process. They are created once for each fiber bundle at the beginning of the pipeline and then used in every other part of the code.

The input parameter needed for this steps are the atlas, from which a mask will be extracted, the Freesurfer label legend and the configuration file information about the relative positions. Being the feature that makes this segmentation algorithm stands out from the others, it has been decided that the definition of at least one spatial relation is mandatory, since, if not, the whole spatial fuzzy sets approach would be voided.

For each direction is possible to define one or more spatial relations, although it is better to avoid replicating definitions as that would not add any additional information and, at the same time, increase the computational time. It is not mandatory to set the spatial relations in regard to every of the six directions implemented, the user has the option of leaving the field, corresponding to the relation not defined, to null (special character of the JSON data format).

Initially the directions are transformed from human-readable keywords to sets of values, indicating the two angles (α_1 and α_2), that will be used to define the 3D baseline for the landscape set computation (as explained in Chapter 7). Since the concept of anterior/posterior, up/down and left/right is not well-defined without a reference space and since the modern MRI scanners can pick up the data from any position in respect to the magnet, the angles definition needs to take into account the orientation of the image they are referred to, in this case the atlas.

The principal space convention are the RAS (Right-Anterior-Superior) space

convention, also called neurological, and the LAS (Left-Anterior-Superior), called radiological [87]. In the neurological convention in particular, the orientation could be associated with the view of the patient from the top of the head, having the right side of the image showing the right side of the patient.

While most of the softwares use the RAS (right-anterior-superior) space convention, due being the most common one in neurological analysis, there is the need to obtain the versatility of working with different images. Fuzzy Tracts automatically manages the angles values based on the atlas orientation, allowing every space convention possible. The angles, in function of the orientation, are listed in 9.1.

	L	R	A	P	I	S
<i>right_of</i>	$\alpha_1 = 0,$ $\alpha_2 = 0$	$\alpha_1 = \pi,$ $\alpha_2 = 0$				
<i>left_of</i>	$\alpha_1 = \pi,$ $\alpha_2 = 0$	$\alpha_1 = 0,$ $\alpha_2 = 0$				
<i>anterior_of</i>			$\alpha_1 = -\pi/2,$ $\alpha_2 = 0$	$\alpha_1 = \pi/2,$ $\alpha_2 = 0$		
<i>posterior_of</i>			$\alpha_1 = \pi/2,$ $\alpha_2 = 0$	$\alpha_1 = -\pi/2,$ $\alpha_2 = 0$		
<i>superior_of</i>					$\alpha_1 = 0,$ $\alpha_2 = \pi/2$	$\alpha_1 = 0,$ $\alpha_2 = -\pi/2$
<i>inferior_of</i>					$\alpha_1 = 0,$ $\alpha_2 = -\pi/2$	$\alpha_1 = 0,$ $\alpha_2 = \pi/2$

Table 9.1. Direction angles in function of definitions and orientations

Once the angles are defined it is possible to obtain the direction of the vector specifying the direction as:

$$\begin{aligned} dir_X &= \cos(\alpha_2) * \cos(\alpha_1) \\ dir_Y &= \cos(\alpha_2) * \sin(\alpha_1) \\ dir_Z &= \sin(\alpha_2) \end{aligned}$$

Using the three vectors computed a first fuzzy set is generated, the fuzzy landscape associated to a direction is computed by looking at the angle from which each voxel is seen from the center of the image and which is the closest from that specific direction [11].

Centering the landscape to the middle point of the image allows to have an initial reference set that can be moved around for the processing of the final, anatomically significant, spatial set of the region configured, following a direct approach.

The initial fuzzy set is an image, defined in all its domain, having high values in the voxels right in the line specified by the vector, that decreases moving away from it, until reaching the exact opposite direction, where the voxels will have the value 0.

Normalizing the fuzzy value in the interval $[0,255]$ makes it easy to notice that all the voxels with high values (around 255) are the most certain to be in the direction chosen for the object, and that this probability decreases with an equal speed in circles centered around the vectors defined line, forming what could be pictured as cones of fuzzy values (Figure 9.2).

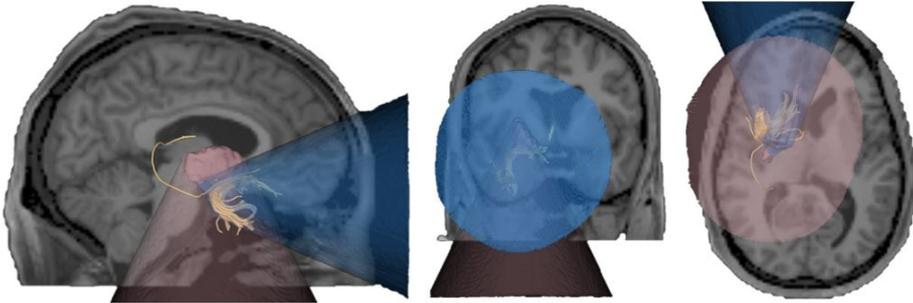


Figure 9.2. Fuzzy sets obtained in respect to the Amygdala and the Putamen projected on the UF. Their support widths can create figures resembling cones of fuzzy values.

Once the fuzzy layout is defined, the anatomical information are introduced. For each direction, the labels of the regions indicated are extracted. Since the regions are written in the configuration file in an human readable way they need to be translated in their Freesurfer respective label number. To do so a legend of the Freesurfer parcellation files has been created and it is looped over to find the correspondence.

After extracting all the labels a binary mask is prepared.

The initial fuzzy landscape is projected onto all the mask object points, moving the apex of the fuzzy cone (in the layout set fixed on the center of the image) to every object point.

Additionally, varying the support width (which means regulating the speed

the set decreases the fuzzy probability moving away from the line) the focus of the fuzzy projection will be increased, reducing of the fuzzy cone apertures. The widths can be modified with an additional parameter k . For example, assuming $k = \pi$ the support is constrained to the bounding box that starts on the farthest point of the mask from the direction indicated.

To save up computational power and consequently time, the projection is not moved on every object point of the mask but instead the operation is reduced to the boundary points. This does not introduce any error as the fuzzy cones centered on the internal points of the object would be anyway enclosed by the ones starting from the boundary points, thus not affecting the quality of the final result.

Finding the boundary points of the mask is a procedure composed by a basic morphological operation and a binary logical instruction. A dilation [95] of the mask, previously created from the labels, using the smallest 3D structuring element possible, a 3x3x3 cube, is performed. This operations sets to one the 26-neighbourhood of every object point, increasing their total number.

Finally, an XOR operation between the original mask and the dilated one finalize the extraction of the boundaries. The element-wise XOR keeps only the voxels that swapped their value (Table 9.2) which implies, since these correspond to the neighbourhood of the original object, the boundary of the region.

A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

Table 9.2. XOR Truth Table.

Once every boundary point has the reference fuzzy landscape centered on him, this multitude of sets is joined into one, forming the final spatial uncertainty mask relative to the label. The joining operation consists of the element-wise maximum of all the sets, equivalent to the fuzzy OR operation.

The shift of the fuzzy landscape is a computational demanding operation and highly costly in terms of processing time. This aspect depends on the size of the atlas and on the type of the data contained in the fuzzy sets.

The fastest approach, an array slicing with a memory allocation, has been identified and extended for the 3D case. In addition the newly created shift function allows to specify a fill value for the points that would move out of boundary. For this application the filler has been imposed to 0, as it can be seen as the lowest fuzzy membership value accepted.

The final shift algorithm is able to move the initial fuzzy landscape to a boundary point each 0.025 s.

To further improve the computational time the whole process has been parallelized. Handling the multi-core approach adds the problem of memory leaking: since the boundary points could be in the order of tens of thousands and each time a fuzzy set is moved a new array is created, the memory of the machine could very quickly start to suffer and return an error.

To manage the memory the solution implemented is the batch computing. Instead of obtaining all the fuzzy sets and then merge them together, the data are processed in sections (batch), at the end of each batch cycle the result is used by the next one and it continues until all the boundary points have a landscape associated.

The size of every batch should be the maximum possible to make the most out of the amount of memory available. To do so first some data are extracted:

- Number of processing cores;
- Quantity of free virtual memory;
- Dimensions of a single fuzzy set $\frac{\text{array_size} * \text{machine_architecture}}{\text{fuzzy_value_size}}$.

Where `array_size` is the number of element of a fuzzy set, the `machine_architecture` can be 64 or 32 and the `fuzzy_value_size` depends of the format of the data contained (e.g. float64).

The max number of array possible to be computed at the same time is then:

$$N = \frac{\text{available_memory}}{\text{set_dimension}} * \text{security_factor}.$$

Where `security_factor`, if set just under one, gives a breathing room to the system, not allowing the complete filling the memory.

The computation is finally spread over all the processing cores found.

9.2 Initial Segmentation

The real segmentation process is split into two parts. A first operation is implemented to greatly reduce the number of fibers to work with, opposed to a second part that fine tunes the output.

The initial segmentation has the sole purpose of reducing the number of fibers adopted for the fVar computation and the fuzzy sets mapping. This is crucial since the input tractograms are often composed by a very large quantity of streamlines, reaching even the order of millions, thus making the computational time of the whole process not fit for clinical applications and generating potential memory leaks.

In particular, the memory leaks could occur during the construction of the similarity matrix based on the functional varifold. Since this structure has dimension $N \times N$ (where N is the number of fibers) and contains float values (that occupies 8 bytes), working with just 20000 fibers can result on a 20 Gb array. This means that if the input is formed by one million lines and the code has to run on a typical office PC configuration (that may have up to 32 Gb of RAMs) the initial segmentation task must eliminate more than the 95% of the tractography streamlines.

This step of the pipeline should greatly reduce the fibers number while being careful of not eliminating true positives. The accuracy of the segmentation is not a concern in this phase as that will be tackled in following steps.

The tool provides two options, one slower but more accurate (using WMQL) and one less demanding, using the already created fuzzy sets, that performs a sloppier fiber eliminations.

9.2.1 WMQL

The White Matter Query Language, which working methods were explained in chapter 5, can not be considered accurate enough to be used as a stand-alone segmentation tool. In particular, the major problems are the low fidelity of the outputs presented, that often contain easy-to-identify outliers, and consistency issues between subjects, due to the fact that the correspondence paradigm is not always a valid assumption [72].

Although the noticeable limitations, WMQL is an all-around useful tool. Its outputs, while not being perfects, are more often than not of easy interpretation by clinicians.

Even if a good quantity of outliers is noticeable, the tool makes an overall good job to keep all the true positives (or at least what could be considered so), as long as the queries are not too strictly defined. This aspects forms for sure an optimal starting point for the pipeline and a candidate as an initial segmentation operator.

Using queries with few definitions the results are less anatomically accurate and do not well represent the bundle, but at the same time, the inter subject variability is reduced (Figure 9.3). This can be exploited adapting the query definition to the new goal of reducing the quantity of streamlines as much as possible without losing the necessary anatomical information.

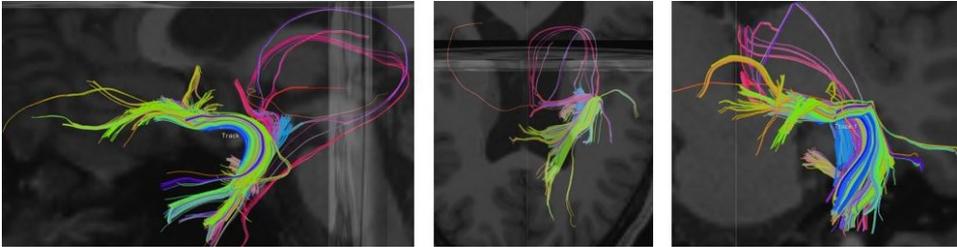


Figure 9.3. Initial segmentation using WMQL with few definitions. While the true fibers are segmented a great quantity of outliers is still visible. From left to right: side view, top view, frontal view.

Due to slight differences between the geometrical position of the same bundle in different subjects, it is not possible to be certain of not eliminating true positives using just one very strict query. On the other hand, looser queries might not be sufficient to exclude a meaningful quantity of fibers.

A compromise between the accuracy and the security of maintaining all the true positives has to be reached. Being able to define this type of queries involves have a good understanding about the importance of every single annotation found in literature.

For example it is possible to statistically choose the definition searching the ones that are found the most in the scientific literature. This assures that the representations are well known to the medical community but it could happen that particular relations are too strict and produce a deletion of true positives.

An other approach could involve using only the definition found in the most cited papers, as a guarantee of their correctness, and have an expert validate the best fitting ones. This is the suggested modus operandi and the one applied in the

use cases presented.

For the Uncinate Fasciculus instance (Chapter 3) this is done combining definitions appearing in three different highly cited papers, all coming from a radiological point of view, for a total of six characterizations, listed in 9.3.

Catani [16]	<ul style="list-style-type: none"> • endpoints in (temporal_anterior_section.side) • crossing orbito_frontal_gyrus.side
Wakana [101]	<ul style="list-style-type: none"> • crossing temporal.side • crossing frontal.side
Leng [52]	<ul style="list-style-type: none"> • crossing superior_temporal.side • crossing orbito_frontal_gyrus.side

Table 9.3. UF definition for WMQL initial segmentation.

Once the query is prepared the same atlas used for the computation of the fuzzy sets is be used as input for the WMQL tool. The algorithm will manage the creation of the command pipe to call the instruction and will proceed to create the temporary working folder and perform the garbage collection at the end.

It is important to point out that the White Matter Query Language is distributed independently and it is the user job to configure it. Without a correct installation this initial segmentation method will not be used.

9.2.2 Sets Intersections

There are situation when using WMQL might not be the optimal choice. Using an external tool to integrate a functionality means that the dependencies must be satisfied for both WMQL and Fuzzy Tracts, often increasing the number of libraries needed.

Furthermore the WMQL could have high computational times for large dataset and not always being clear enough in errors management situations.

As the configuration of the dependencies is not always handy and it is preferable to have full control on the tool, with a different view, a native initial segmentation option has been added.

The included way of performing the segmentation provides a method that exploits the previously computed fuzzy sets. In this case, the fibers kept are the one that pass through the intersection (the fuzzy AND operation) of every fuzzy set, defined for each bundle.

This is done in two steps: first a mask is created performing a logical *AND* operation between the different sets (binarized in order to have a logical true wherever a probability is expressed), every fiber is then mapped on the mask image. All the streamlines that have no points at all positioned in the defined region, found checking the maximum value of the mapping, are eliminated.

The high aperture of the fuzzy cones support, defined during their computation (by default π), assures that it is always possible to find an intersection between sets as long as all the spatial definitions indicated in the configuration file result coherent with each others (Figure 9.4).

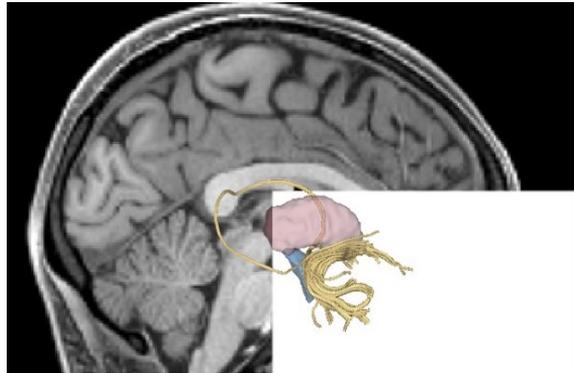


Figure 9.4. The fuzzy sets intersection could be seen as the bounding box (white) that incorporates all the anatomical label defined. The fibers that have at least one point inside of it are maintained.

It has been decided to keep every fiber with at least one point in the region of intersection for two reasons. Firstly the loose threshold ensures that every true positives are kept and, more importantly, even if it would be possible to perform a more strict selection, for example imposing the minimum percentage of valid points of a fiber, that would require the user to decide a new, not easy to be tuned, parameter.

While this approach scales well for large tractograms in terms of speed, its inaccuracy put some limitations on its application. Having a initial segmentation

that does not always reduce the number of streamlines down to a few thousands, likely to happen with a vast number of initial streamlines, might cause memory leaks in the following steps.

This methodology is then appropriate in case of medium sized tractograms and when it is possible to define a lot of spatial relations during the bundle definition phase.

9.3 FA Extraction

Once the initial segmentation is obtained, the next big step is dividing it into sub-clusters.

In order to produce more precise fiber clusters it is necessary to utilize different streamlines characteristics in the metric. One of these is the fractional anisotropy.

To assign a FA value to each fiber of the tractogram (Figure 9.5) is first necessary to compute the fractional anisotropy map from diffusion weighted images. Since the FA is a metric distinctive of the diffusion tensor (Chapter 1.1), that must be computed in advance.

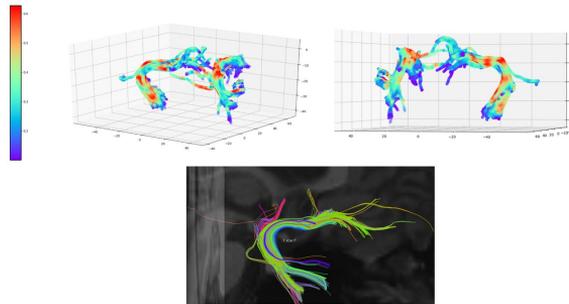


Figure 9.5. Mapping of the fractional anisotropy image on the fibers in output from the initial segmentation algorithm. Each streamline has a number of FA values associated equal to its number of points.

It is possible both to generate the fractional anisotropy map on-the-fly each time a segmentation is launched or to pre-compute it. This option is particular useful during the debug phase or to tune the parameters in a faster way, since some computational time is saved. In addition, it allows to utilize different tensor fitting methods: if a more modern and precise tensor fitting operation is introduced to

the scientific community the associated FA map can be easily used in Fuzzy Tracts.

The tool can use two different methods for the on-the-fly tensor fitting and FA extraction: using the MRTrix3 interface [89] or the DiPy interface [29].

The first implementation wraps the MRTrix3 (Chapter 2.3) commands associated to these tasks while the second one works directly in the tool environment.

If MRTrix3 is installed on the PC and the path is configured in the bashrc file in order to have the instructions available in the command prompt they will be used. Since, as mentioned before, MRTrix3 splits the whole DWI to FA process in two phases (DWI to DTI and DTI to FA) it is necessary to define a pipeline to concatenate the instructions (`dwi2tensor` and `tensor2metric`). In addition, the commands should be wrapped in the Fuzzy Tracts programming language independently of MRTrix3.

This is done using Nipype [33]. Not only it allows to create complex pipelines and to manage them in each of their parts, it also includes the wrapping for the most used neuro-imaging tools.

A pipeline, also called work-flow, is defined as a sequence of nodes. Each node wraps a single instruction and it is formed by different fields, with the mandatory ones being the input file and output file. The concatenation of the nodes consists of assigning the output file of one node to the input of the next. If necessary, it is also possible to branch the workflow to create more complex pipelines that are composed by commands from different softwares.

The nodes that wrap the MRTrix3 commands into the Nipype interface are:

- `FitTensor`: wraps the instruction `dwi2tensor`;
- `TensorMetrics`: wraps the instruction `tensor2metric`.

`FitTensor` requires in input, apart from the DWI image, the files containing the information about the bvalues and bvecs and optionally a brain mask to reduce the computation field. It produces in output the DTI file.

`TensorMetric`, on the other hand, allows to perform various operations, such as the extraction of mean diffusivity and fractional anisotropy. Specifying in input the DTI file previously created and as a option the string `FA`, it outputs the fractional anisotropy map.

Starting from the nodes a workflow is built. The process takes place in the temporary folder of the system, that will be removed after the loading in memory

of the final array, in order to avoid the waste of computational resources by useless files (garbage collection operation).

The suggestion is to use this approach whenever possible, as the MRTrix3 interface produces solid results, the pipeline is well optimized and the computation parallelized.

The second approach exploits the DiPy library. It is very similar to the previous method but has one huge drawback, it works directly in an high-level programming language, thus being particularly slow in the computation.

In this case, first the bvalues and bvecs are loaded and joined together to form a gradient table, this is used to create the tensor model. The tensor model is then fit to the diffusion image and the fractional anisotropy extracted.

9.4 Clustering

Having both an initial segmentation of the whole-brain tractogram and the diffusion information, a set of clusters is computed in order to reduce the final step to work directly with sub-bundle and not with the full quantity of streamlines.

The goal is to reduce the fibers to few prototypes of the clusters, for this aspect it is very important that the clusters are composed by fibers as similar as possible from each others (Figure 9.6). In this step the anatomical coherence of the clusters is not important. It doesn't matter if a cluster corresponds to a set of fibers with specific anatomical functions as long as the intra-cluster variability is kept low.

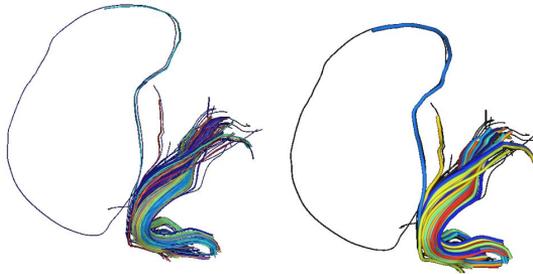


Figure 9.6. Left: the initial segmentation of the UF is clusterized using the fVar. Right: each cluster is pictured with a prototype.

Compact representations of fiber bundles are widely used in order to ease the

load on the machine that has to perform streamline-based operations, like segmentation and transformations [111].

Usually the clusters are approximated to prototypes, fibers that better portray the characteristics of the set. To obtain them it is possible to execute basic operations, like selecting the most central line [30], or have a specific metric [67].

9.4.1 Metric

Having a solid metric is necessary to generate consistent cluster. The metric chosen, the functional varifold [35], exploits geometrical information as well as diffusion information, as discussed in Chapter 8.

In particular the Fuzzy Tracts implementation of the fVar weights four parameters for the similarity matrix computation:

- Distances between fiber start points;
- Distances between fiber end points;
- Distances between fiber segments;
- Distances between fiber anisotropies.

For all of these measures it is necessary to specify a kernel width in order to tune the sensitivity of the metric towards each of them. The kernels bring an information about the geometry and features of the bundles, thus they should be tuned with in mind an a prior anatomical knowledge.

Since often it is not possible to know all the information about the specific bundle the user has the chance of leaving the width of the kernel null. This avoids to have a random initialization and tuning, not forcing the user to guess values difficult to decide.

Being the concept of start points and end points not well defined for the fibers of a tractograms (the tractography algorithms often random initialize the seed points used to start the tracking), all the streamlines must be reoriented in the same direction.

For the reorientation, first a random fiber is taken as reference, all the fibers will be rotated in accordance with this reference direction. To decide if a fiber needs to be changed, the distance between the fiber and the reference is computed, the streamline is then flipped and the distance recomputed another time. If the

difference *normal fiber / reference* is higher than the difference *flipped fiber / reference* then the streamline will be re-orientated.

The distances are computed as the sum of the norm between the first point of the fiber to the first point of the reference and the norm between the last point of the fiber to the last point of the reference.

In addition, it is available for the use (and sometimes suggested) the option to downsample the fibers. Since the functional varifold, when computing the distance matrix, compares every point of a streamline to every point of another one, reducing the number of points could result in huge improvements of the computational time (Table 9.4).

Fiber Points	Computational Time (s)
50	7.015
100	13.464
200	38.377
500	211.497

Table 9.4. Computational time needed for the fVar matrix computation for different downsampling values. The lower the better. Tested on WMQL segmentation of the UF (circa 300 streamlines).

There are two ways to downsample a streamline, either setting every fibers with a selected number of point or reducing the points by a percentage, treating each line independently. Since the functional varifold does not need a pairwise correspondence between points both methods are applicable.

In the end the choice fell on using a fixed number of point. This allows a possible future developer to change the metric without having to touch anything else in the code. If the option is selected without a specific value, all the fibers will be downsampled to the number of points of the shortest one.

During the application of the fVar two matrix are computed: one containing the inner products and one containing the distances. Each of them has a dimension $N \times N$, where N is the number of streamlines.

The distance matrix can be easily obtained from the inner products, as defined in the fVar approach:

$$C_X - C_Y = \sqrt{\langle C_X, C_X \rangle + \langle C_Y, C_Y \rangle - 2 * \langle C_X, C_Y \rangle}$$

Computing the inner product matrix just one time, in the beginning, allows to save a lot computational time, since all it remains is to extract the data and perform basic mathematical operations.

Once the clustering algorithm is applied (more details on the next subsection 9.4.2) for every fiber group a prototype is extracted. Again, this operation can be done working directly with the fVar notation.

In particular, the fiber considered as a prototype is the closest one to the mean of the cluster, computed following this equation:

$$M_{cl} = \frac{\sum_{i \in cl}^N S_i}{N} = \frac{\sum_{i \in cl}^N \sqrt{\langle C_i, C_i \rangle}}{N}$$

Where cl is the cluster analyzed, N is the number of streamlines belonging to that cluster and S_i is the i^{th} streamline of cl.

After having associated a prototype to each cluster it is possible to proceed with the final segmentation.

9.4.2 Clustering Algorithms

Deciding which cluster algorithm to use is not an easy task, it is necessary to define exactly what are the wanted proprieties of the clusters.

For Fuzzy Tracts the clusters are important for the choice of the prototypes. The prototype extraction is needed to reduce the number of streamlines on which the final segmentation will be carried on.

The priority is more on obtaining a low intra-cluster variability than an anatomical coherence. In addition the number of cluster is not a decisive factor, as well as its population size. Having clusters composed by just two fibers will improve the final segmentation speed of 50% .

Three different approaches have been tested:

- Hierarchical clustering (Dendrogram);
- Spectral Clustering;
- Density based clustering (DBSCAN).

Since the Dendrogram [77] requires an user action, the decision on where to cut the tree, it has been excluded by the selection after the first testing phases.

Fuzzy Tracts enables to perform the clustering operation with either the spectral clustering or the density based DBSCAN algorithm.

Spectral Clustering

The Spectral Clustering [55] works directly on the inner products matrix, thus avoiding the extra (even if relatively fast) computation of the distance matrix.

While it requires in input the definition regarding the number of cluster searched, the algorithm divides the set automatically, without necessitating the user to interpret the ordered tree. It could be considered a more feasible implementation, for Fuzzy Tracts, of the hierarchical clustering methods, in respect to the Dendogram.

The Spectral Clustering initially creates a graph from the similarity matrix, it then computes the unnormalized laplacian L of the adjacency matrix and extracts the first k eigenvectors (where k is the number of cluster imposed). The eigenvectors are sorted in column order and, using a standard k -means algorithm, the row vectors are classified into k clusters.

The assumption on the number of cluster determines its limitation and in particular the poor scalability in situations of multiple fiber bundles segmentation.

DBSCAN

The DBSCAN [26] is the option selected by default. It allows to divide the streamline set without specifying the number of clusters, very important feature that makes it more universal, since it can be applied to every bundle without a specific knowledge about its composition.

This algorithm requires two parameters: ϵ and n . One more of the Spectral Clustering, but easier to impose. In fact, since the clusters should just be indicative of a small group of fibers there is no need to modify them based on the wanted bundle.

Another important factor is the introduction of the concept of noise. Fibers too distant to the others, in function of the ϵ , could be 'non classified'. This could be exploited to instantly detect the outliers. It must be noted that the outliers, found as noise, are not necessarily anatomical outliers. This happens because packs of not-meaningful streamlines could be clustered together, if they are sufficiently compact, while coherent streamlines could be found as noise if they are alone in their trajectory.

DBSCAN defines three types of points: cores (they have n other points in the ϵ neighbourhood), reachables (they reside in a core point neighbourhood but are not cores by themselves) and noise (not reachable by any other point). The clusters are formed by the union of every core and reachable point that are linked together (Figure 9.7).

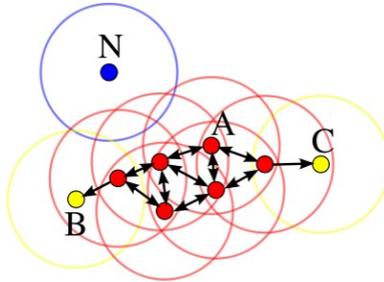


Figure 9.7. DBSCAN point definitions and cluster formation. Red: object point, Yellow: reachable point, Blue: noise.

This algorithm offers a good robustness to outliers, does not require an assumption on the number of clusters and it offers a good consistency, being insensitive to the point positions (it can classify non linearly separable clusters). It is especially a good choice for packed data, like in the case of streamlines all belonging to a bundle, as the density is all around similar.

9.5 Final Segmentation

The final segmentation is composed by two actions: the fuzzy mapping and the thresholding.

The fuzzy mapping is done by joining the information of the fuzzy cones to the centroids extracted after the clustering. Since the streamlines are composed by a multitude of points there is the need for the delineation of a method to assign to each fiber just one fuzzy value.

Initially each centroid is mapped on the first fuzzy cone defined during the initial phase of the pipeline. The coordinates of the points are converted from real space (mm) to the image space (voxels) through an inverse affine transformation. The value of the voxel with the same coordinate as the point will be assigned to it, this for all the points of the streamline.

To reduce the fuzzy information to one value per streamlines, the mean of the fuzziness of all the point belonging to that line is computed.

The procedure is repeated for every other fuzzy cone. In the end, the centroids will have a number of fuzzy values associated equal to the number of cones. The final fuzziness will be taken as their minimum.

The fuzziness is then normalized in the range $[0,1]$ and finally propagated from a centroid to every other streamlines in the cluster. The outliers, defined as noise by the DBSCAN algorithm, are associated with the fuzzy value 0.

Having every fiber with a fuzzy value associated allows to perform a final operation of global thresholding. This is the user final decision and must be based on the experience and knowledge about the bundle examined.

The thresholding will eliminate from the segmentation all the fiber with a lower fuzzy score, considered less certain to belong to the bundle (Figure 9.8).

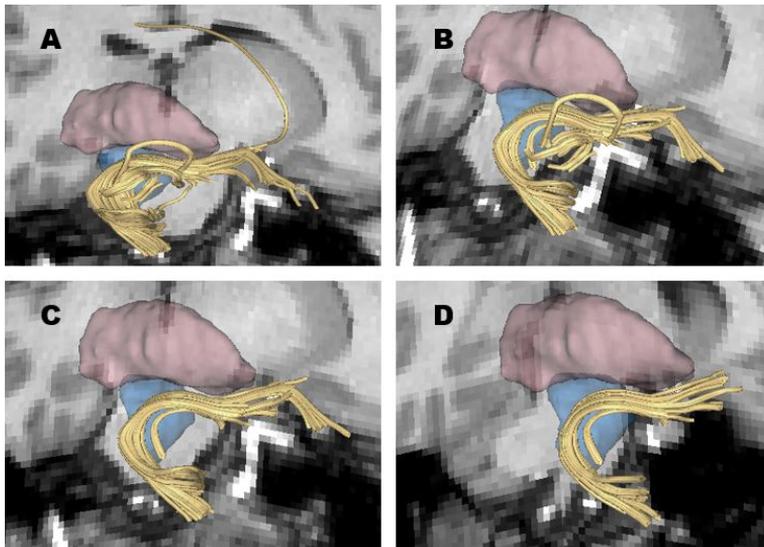


Figure 9.8. The final thresholding fine tunes the segmentation obtained initially with WMQL or the sets intersection acting on the uncertainty associated to each fiber. The parameter must be chosen in order to eliminate the outliers without losing true positives. Threshold values: A=0, B=0.25, C=0.5, D=0.75.

Chapter 10

Other Features

The core of Fuzzy Tracts is the innovative method introduced, the fuzziness computation and consequent clustering technique take most part of the tool code. A well-made software has also the necessity to fit in the context it was developed for and interact with the users in meaningful ways.

The tool should be easy to use even for someone who lacks a background in computer science, offering an intuitive interface. For this reason there are no parameters hard coded and everything can be configured by a simple text file that follows the JSON format.

While more expert users could tune every single input parameter, for novices and untrained staff, the final decision on the segmentation is made modifying one single value, the threshold, intuitively positioned directly in the command line and not in a file.

10.1 File Format Support

A white matter segmentation tool has the necessity to support every major file format used by the neuroscientific community. The support must cover both the tractograms and the images (DWI and atlas).

There are currently three main format for tractogram files: TRK [103], TCK (MRTrix3 file format) and VTK [80]. The first two are associated to particular softwares, `.trk` being the file format of TrackVis and `.tck` belonging to MRTrix3. These two formats inherit the popularity of the suite they belong to, with time they became two of the most used in the neuroscience.

VTK, on the other hand, is a widespread file format not limited to biomedical applications. It is able to define not only streamlines but also meshes, objects... . Its popularity derives from its versatility.

It is important to support every of this files to give the user a smooth experience. In addition the segmentations should contain the newly introduced information about the tracts, like the fuzzy score.

In order to keep a consistency between the input files provided by the user and the segmentation produced, the output tractogram will be saved in the same format.

For the input images the support is limited to the NIfTI format [49]. This comes from the necessity to work with the Human Connectome Project dataset [97] and to have a well detailed header.

The header of the a NIfTI file brings a lot of information. It procures the orientation of the image, useful when creating the fuzzy sets direction, an affine matrix for real space to voxels space streamlines transformation and other image proprieties such as voxel size and resolution.

Finally it is not a difficult task to obtain NIfTI files starting from DICOM images or img+hdr files, with plenty of converter already available.

10.2 Visualization

The visualization options should serve the user as an on-the-fly validation tool. It is important to print to video the results between steps for evaluating if the configuration is done correctly or not.

There are different visualization windows, one after each pipeline node (Figure 10.1). First the initial segmentation is shown, this allows to decide whether or not all the true positives of the bundle are kept or, in the case of the fuzzy cones initial segmentation approach, if the spatial relations are well focused on the bundle stem.

After the clustering a specific color is assigned to every group and they are all printed together, giving the ability to judge the result of both the clustering algorithm and the metric. If the results are not considered satisfactory it is necessary a more precise tuning of the kernels widths. In addition, in parallel, a figure highlighting the centroids of the clusters with every other fiber in opacity gives a

better perspective of the quality of the clustering.

In a final figure is ultimately shown the final segmentation output after the fuzzy-based thresholding. The fiber that have been segmented are highlighted and the other streamlines, belonging to the initial segmentation, are put in the background with a different color. The color scheme maximize the contrast using opposite colors, with the fiber eliminated in red for an added intuitiveness.

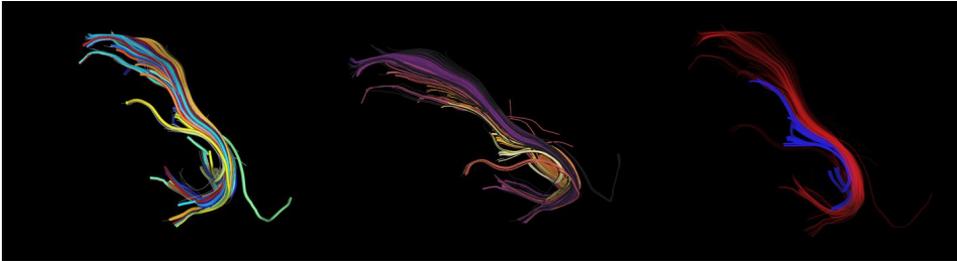


Figure 10.1. Visualization window offered by Fuzzy Tracts. From left to right: cluster prototypes with fiber in background, fuzzy set mapping, final thresholding with outliers in background.

10.3 Performance

The performances and the relatives optimization of the tool, especially in computational time and hardware required, are a huge part of the potential success it might have.

The main applications of diffusion imaging, and tractography especially, are pre-surgical planning, statistical analysis and recovery estimation (Chapter 2), it is then vital to achieve computational time compatible with these tasks. In particular, the surgical planning requires a fast analysis, with expert desiring to have a segmentation ready in under ten minutes.

To improve the run-time two options are available: using more powerful hardware or exploiting the hardware supplied at its maximum capacity. Since the first solution is not always viable due to budget limits or impossibility to host high-performances machines it is better to make the most out the accessible resources.

In this thesis all the tests have been carried on a tower station with the following characteristics:

CPU Intel® Xeon® E5-1620 0 @ 3.60GHz
RAM 16 Gb
GPU NVIDIA® Quadro 600

Fuzzy Tracts will greatly benefit from an high-end processor and a large quantity of virtual memory.

The most intense tasks are, in weighted order: the fuzzy sets computation, the functional varifold matrix preparation and the clustering algorithm. In addition, for the fuzzy sets and the *fVar* the memory leaks could cause issues since these operations work on huge arrays in memory.

The solution adopted is the parallel computing. Each cycle is performed in parallel, at the same time, between all the processor cores available (four with the cited equipment).

The table 10.1 reports the difference in speed with and without the parallel computing approach. Both the set definition and the clustering exploit all the cores accessible through the whole algorithm, meanwhile the *fVar* has just some functions parallelized.

Operations	Parallelized (s)	Non Parallelized (s)
<i>FuzzySets</i>	138.598	496.64
<i>fVar</i>	35.187	38.475
<i>Clustering</i>	0.002	0.003

Table 10.1. Differences in computational time for three major operations between the standard implementation and the parallel computing (two fuzzy set definitions, 165 fibers, downsampling to 200 points)

Furthermore, to avoid memory leaks during the fuzzy sets computation an approach called batch computing has been adopted (Chapter 9.1).

Chapter 11

User Guide

This concise user guide serves as an introduction to aid first time users and operators that wants to explore all the possibilities Fuzzy Tracts offers.

Fuzzy Tracts is written in Python 2.7 and can be download from the GitHub page of the author: <https://github.com/aledelmo>.

Dependencies

A couple of dependencies are needed to exploit all the functionalities:

- DiPy <http://dipy.herokuapp.com/>;
 - Nibabel <http://nipy.org/nibabel/index.html>.
 - * NumPy;
 - * SciPy.
- NiPype <http://nipype.readthedocs.io/en/latest/index.html>;
- VTK <https://www.vtk.org/>;
- Sklearn <http://scikit-learn.org/stable/install.html>.

For the installation of the single dependency it is better to refers directly to the specific library guide. The suggestion is to use an Anaconda environment or install the package through the use of pip.

Configuration Files

There are two configuration file at the disposal of the user, one mandatory and one optional.

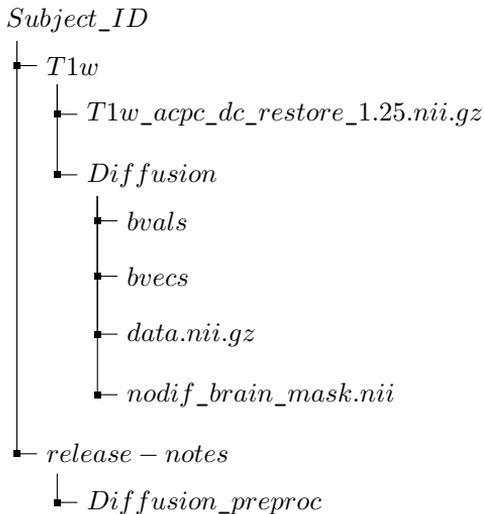
In the main folder of Fuzzy Tracts the *config.json* file has to be used to define the spatial relations for the computation of the fuzzy sets and to specify the kernel width for the functional varifold. The definition must be based on the Freesurfer parcellation definitions (found in the file *./scripts/data/FreeSurfer.txt*).

In *./scripts/WMQL* it is possible to find the optional configuration file relative to the White Matter Query Language (*WMQL_query.qry*). This become mandatory when the option *-w* is selected and it must follow the WMQL queries definition framework.

Data

Fuzzy Tracts is built around the HCP dataset. The folder that contains the diffusion data used for the segmentation needs to be organized in the same way. It is possible to download subject specific data from: <https://db.humanconnectome.org/>.

The folder structure follows the pattern:



Usage

Launch Fuzzy Tracts as a standard Python script, after moving the to its master folder:

```
cd fuzzy_tracts
```

\$ `python fuzzy_tracts.py input_file t <options>`

Where `input_file` is the whole-brain tractogram in tck, trk or vtk and `t` is the threshold in the interval [0,1].

Options

The option available are listed in table 11.1.

At least one options between `-fa` and `-f` is mandatory, using a precomputed FA map could be really useful during the debugging and parameters tuning as it avoids to perform the same operation over and over. Both of this options require the insertion of a string pointing to the location of the file or folder chosen.

<code>-w</code>	<code>-WMQL</code>	Initialization using WMQL
<code>-f</code>	<code>-File_Folder</code>	Name of the files folder
<code>-FA</code>	<code>-FA</code>	Precomputed FA map
<code>-r</code>	<code>-resample</code>	Downsampling streamlines
<code>-s</code>	<code>-show</code>	Use the visualization tools.

Table 11.1. Fuzzy Tracts option table.

Resample expects in input a positive integer number. If the wanted number of points is not know a priori the option can accept the parameter 'max' to resample all the streamlines to number points of the shortest fiber.

Show enables the on-the-fly visualization tools.

Output

The filename of the segmentation produced will report the name ID of the bundle, as defined in the configuration files, the threshold value and FuzzyTracts. The output file will assume the same format as the input tractogram.

Part IV

Results

Chapter 12

Comparisons

In this part of the thesis the segmentation results obtained with Fuzzy Tracts will be presented in regards to three subjects from the Human Connectome Project dataset. The fuzzy sets generated during the process were computed using the more accurate propagation technique, with two definitions formulated (anterior of Amygdala, inferior of Putamen) and both the kernel width and clustering parameter were left untouched for a consistent analysis.

The segmentation methods are qualitatively confronted through the visual validation of their outputs, focusing on the Uncinate Fasciculus segmentation. The benefits introduced by Fuzzy Tracts are highlighted for both accuracy and consistency.

12.1 Fuzzy Tracts and Manual Segmentation

One of the original goal of the algorithm was to obtain the same fidelity as a manual segmentation method without being as tedious and time-consuming.

The process of drawing the region of interest on the FA map or on the diffusion image, although being extremely long, allows to have an high level of customization to the output, making it very user dependent and application focused.

Using the atlas defined by Catani [16] a segmentation has been produced with the definition of two regions of inclusion, one in the temporal lobe and one in the frontal lobe (Figure 12.1). The presence of evident outliers in regions not considered be part of the UF by any reference made necessary the use of a region of exclusion positioned on the coronal plane, centered in the middle of the brain.

While this region of interest has not been explicitly formulated in any papers, the anatomical correctness can be confirmed through a visual analysis of the results obtained by different authors. While this added ROE solves the problems of the evident outliers passing over the Corpus Callosum, all the streamlines that deviates the expected path at the level of the curvature are not eliminated.

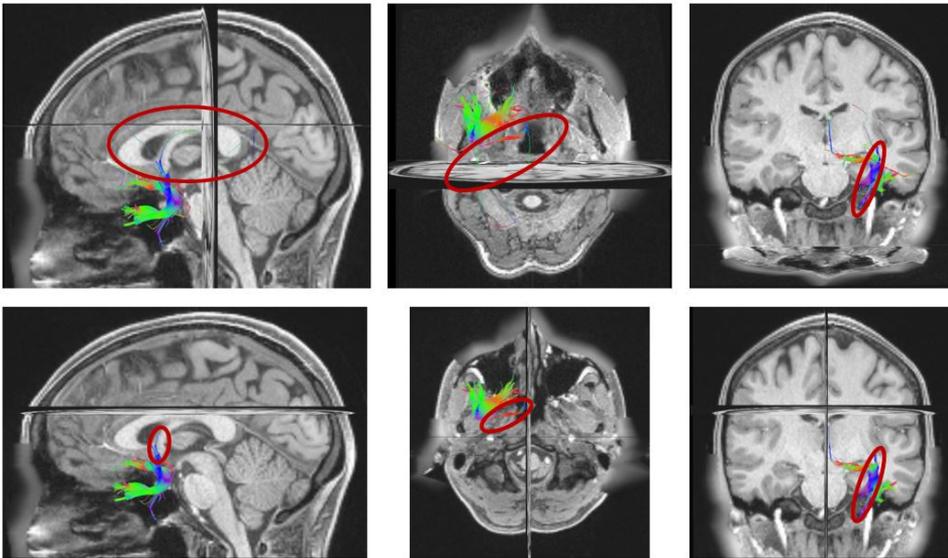


Figure 12.1. Top row: segmentation with the placement of two ROI as defined by Catani. Bottom row: segmentation with an additional ROE on the coronal plane positioned in the middle of the Corpus Callosum. Circled in red the suspected outliers.

In the Fuzzy Tract implementation, in particular with the initial segmentation technique that exploits WMQL, the expression of the false positives is initially (prior to the threshold application) very similar to the manual segmentation case.

Following the fuzzy cones mapping over the streamlines initially identified (Figure 12.2), the outlier fibers, that in the previous segmentation required the delineation of a region of exclusion for their removal, are here easily identified. Since the vast majority of the points composing them is situated behind the Amygdala and they are not completely inferior of the Putamen, the resulting uncertainty score associated to these relation assumes a value next to zero (identified in blue in the picture).



Figure 12.2. Uncertainty score mapped onto the Uncinate Fasciculus fibers.

The fine tuning of the segmentation, requiring the imposition of a threshold value, is carefully performed with the goal of not removing the true positives present. The positions of the anatomical area (completely posterior and slightly superior to the bundle), defined by the relative spatial relations, creates a smooth decrease in the uncertainty values, moving from the back of the complex to its inner portion.

A threshold of 0.5 has been imposed after a consultation with a neurosurgeon. In the achieved segmentation (Figure 12.3) the paths of all the fibers can be identified in the scientific literature, with no documented outlier visible. The frontal / temporal lobe connection is kept intact as well as the 'hook' shape, the lateral-medial bending and the characteristic fan shaped temporal termination.

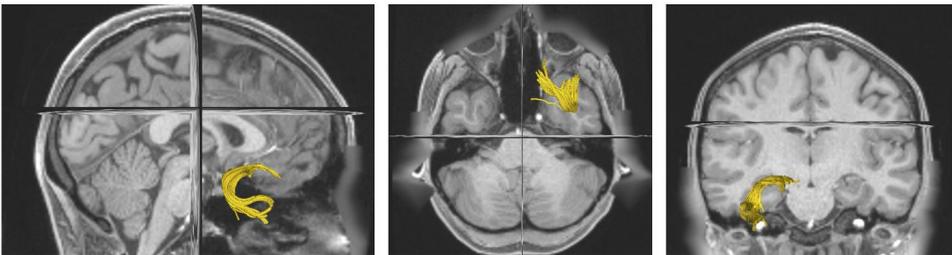


Figure 12.3. Final segmentation after the threshold application.

12.2 Fuzzy Tracts and WMQL

In respect to the White Matter Query Language, the tool proposed offers a more dynamic and precise segmentation, extending the core concepts, like the near-to-English language and the relative relation approach, with the spatial fuzzy sets integration. The fine tuning thresholding operation allows to noticeably improve a WMQL initial segmentation in an operator-centered manner.

Main problem of WMQL is the correspondence paradigm, the assumption every atlas approach is founded on. As already anticipated, differences between the positions of brain structures are easy to find even in mono-zygotic twins, making the spatial relation formulations based on atlases (usually created meaning different brain parcellation or through registration methods) weak and sensible to conformation changes and pathological deformations.

Furthermore, more complex tractogram reconstruction algorithms tend to produce a number of false positives higher in respect to simple diffusion tensor based methods. These false positives could assume very twisted paths, often crossing multiple brain region not biologically associated. The combination of the atlas/subject differences and the tractogram initial composition could create situations particularly hard for WMQL (Figure 12.4) to segment.

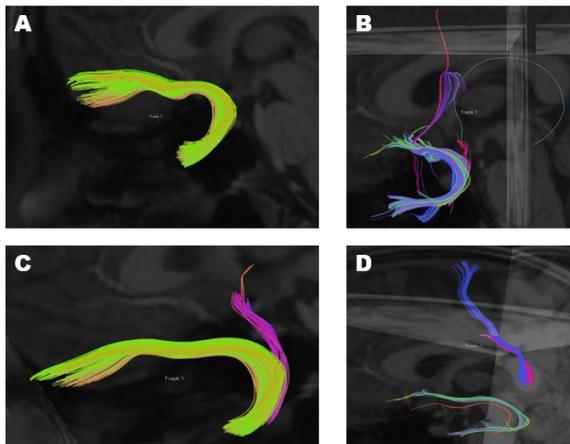


Figure 12.4. Segmentation obtained with WMQL for different tracking algorithms and atlases. A: FACT and MNI152, B: SDSTREAM and HCP parcellation, C: FACT and HCP parcellation, D: SDSTREAM and MNI152.

In this optic Fuzzy Tracts completely compensate the lack of structural correspondence adding a level of confidence to the output. While huge errors in the atlas could still invalidate the results, an interpretation of the relative position fuzziness allows to quickly identify the problems and proceed with a new attempt based on more accurate data.

When the correspondence paradigm is not completely respected but the discrepancies are constrained to small local conformation changes, the ability to manually intervene with a confidence score completely compensate them, making the tools less sensible to this type of artifacts.

12.3 Inter-Subject Consistency

In order to evaluate the consistency of the results produce by Fuzzy Tracts it necessary to distinguish between the inter-subject reproducibility and the inter-operator reproducibility.

Starting from the latter, the tool shares a lot of similarity in respect to the manual segmentation methods. Being the output an ad-hoc created segmentation, with the definition of different parameters all along the algorithm, the analogies between the results from are limited to the differences between the imposed clustering kernels and threshold value.

On the other hand, using the same options on the same image the result obtained will absolutely not change, making the segmentation deterministic in this regard. Furthermore, to evaluate the inter-subject variability, indicative of the solidity of the clustering algorithm and the thresholding operation, the algorithm has been applied on different HCP subject using the same specifics (Figure 12.5).

For the three subject showed here and in Chapter 13 (Figure 13.1 and 13.2) the segmentation were computed for three different threshold values (0, 0.5 and 0.75) without changing the fVar kernel widths. In two cases out of three the confirmed 'best fit' threshold was identified as 0.5, while in the last case as 0.75. It is then safely to assume that the solidity of the clustering algorithm is high enough to support different segmentations without incurring in gross discrepancies from the attended results.

Additionally, the definition of the fuzzy sets support angle as π , that could be considered very large, not only allows to have a reliable initial segmentation, but also increases the reproducibility of the thresholding effect. While tighter projections could improve security level in regards of removing all the outliers present

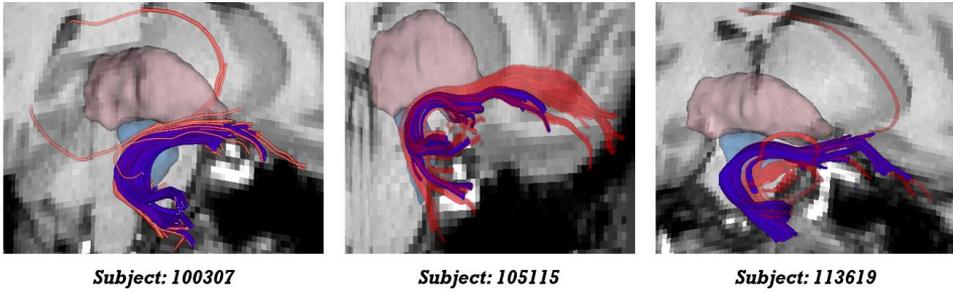


Figure 12.5. Fuzzy Tracts segmentation for three HCP subjects using the same parameters in every phase of the algorithm.

(Figure 12.6), a looser set enhances the uncertainty resolution, making the fibers with medium fuzzy values more distinguishable from each others.

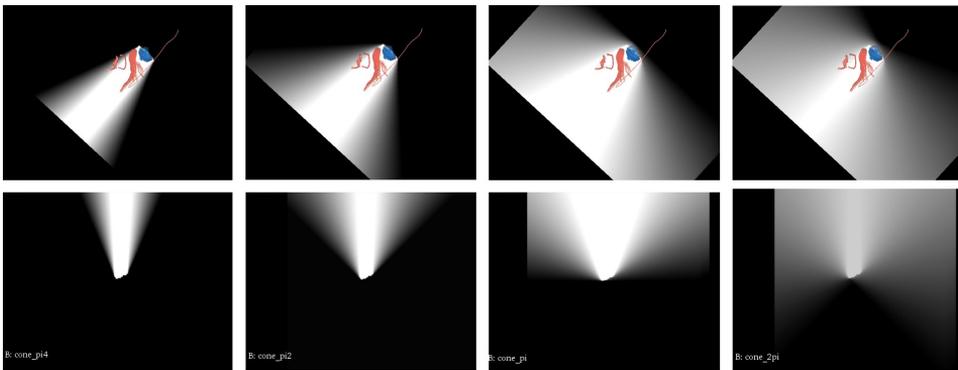


Figure 12.6. Fuzzy set support for different apertures, in respect to the spatial relation 'anterior of Amygdala'. From left to right: $\frac{\pi}{4}$, $\frac{\pi}{2}$, π , 2π

Chapter 13

Conclusions

Fuzzy Tracts, an innovative tractogram segmentation tool, proposes a new approach to the white matter fiber bundles identification, shifting to the user the control on the output quality while remaining easy to set up and feasible for the use in surgical contexts in both speed and results consistency.

Inheriting from the state of art a near-to-English language format for the spatial definitions of a fiber bundle, Fuzzy Tracts extends the concept of an anatomically informed segmentation introducing a spatial fuzzy set approach for a dynamic outliers identification.

A clustering approach, that exploits both structural and diffusion proprieties, combined with an atlas-base segmentation allows to reach high performance reducing the computational resources demand in a significant way without losing the coherency to the bundle definition.

An user-centered approach allows to perform a customized analysis while eliminating the need for tedious manual segmentation operations or the applications of not controllable clustering algorithms.

The accuracy of the results is dependent on the ability of the user to well define the fiber bundle in interest via clustering kernel parameters, relative positions and a final thresholding effect. Having different parameters does not interfere with the consistency of the results or the difficulty to operate the tool since the weights of the effects they produce are very different, with the fuzzy-based thresholding being the most crucial to impose (Figure 13.1). In fact, taking the use-case of the Uncinate Fasciculus, the tool was tested on a multitude of HCP subjects without the necessity to change any kernel widths or anatomical position specification.

The anatomical coherence is based foremost on the relative position definition. Adding spatial information increases the quality of the initial segmentation, making the following fine tuning operations easier and potentially more accurate. When these relations are not available or not validated by multiple sources the tool is able to work with gross indications, with as little as two definition being sufficient to obtain the desired results (during the tests carried on the UF only the relation to the Amygdala and the Putamen were used).

The consistency of the results are one of the focal point of the concept behind Fuzzy Tracts. The level of customization applicable by the users creates an environment where no wrong segmentation can be produced. Additionally, different clinicians confirmed the consistency and quality of the results, for different HCP subjects, obtained through the use of the same parameters in all the steps of the algorithm.

Fuzzy Tracts is expected to be used for both statistical studies and subject specific investigation, with a programmed development that should expand its applicability through the implementation of a multi-bundle support based on a dynamic user interface.

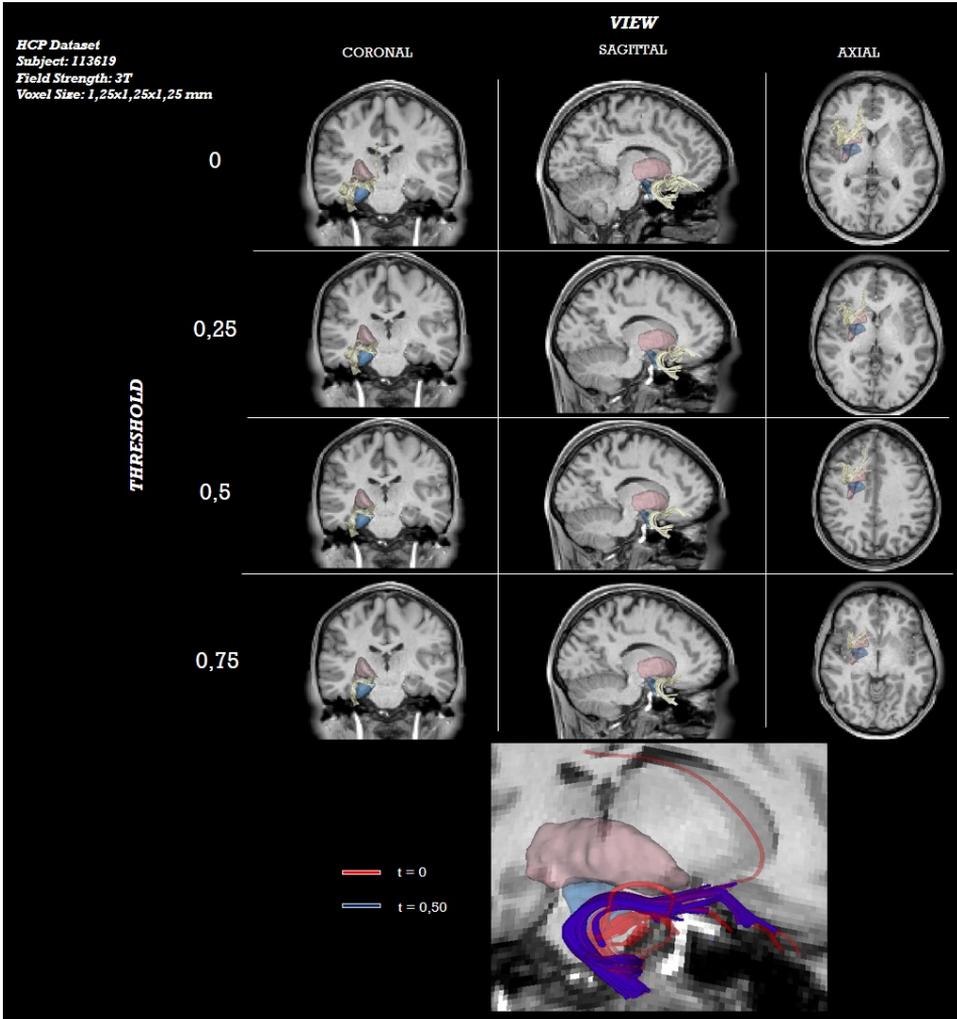


Figure 13.1. Uncinate Fasciculus segmentation obtained with Fuzzy Tracts for HCP subject 113619.

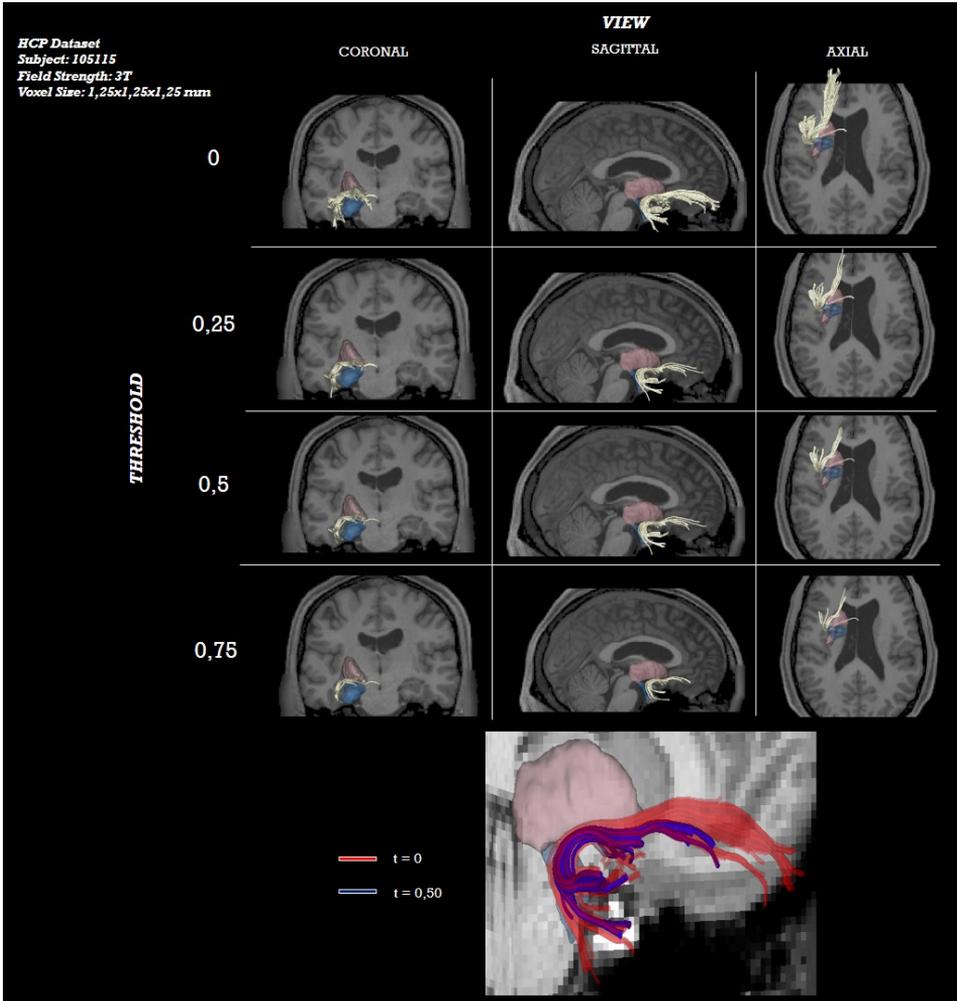


Figure 13.2. Uncinate Fasciculus segmentation obtained with Fuzzy Tracts for HCP subject 105115.

Chapter 14

Future Work

The preparation of the Fuzzy Tracts segmentation tool is part of the *Meta Tracts: Parsimonious multi-resolution representations for modeling, visualizing and statistically analyzing brain tractograms* project, led by prof. Pietro Gori and spread across three different research labs.

The project, spanning on a three years period, aims to change the tractogram visualization procedures using a multi-resolution structure, performing the validation on fiber bundles extracted with automatic segmentation methods, for the final goal of producing statistical analysis of both healthy and pathological subjects.

In this context Fuzzy Tracts serves as the base from which the final segmentations will be obtained. The further development of the tool is programmed to tackle the current limits observed and to push forward the multi-bundle segmentation aspects.

The spatial fuzzy sets work will include the wrapping of the more precise propagation method that could replace the current landscape based direct method. Additionally, the support for the clustering of large tractograms will be introduced in order to completely get rid of the WMQL initial segmentation approach. Finally, small performance improvement tricks could be applied to reach an higher portability of the tool on low-specs machines.

The validation on more white matter fiber bundles will be a central topic, especially focusing on the Inferior Fronto-Occipital Fasciculus. An extensive bibliographic research will be carried out to better define the IFOF and, due to its spatial proximity to the already tested UF, the 'exchange' of fibers in the overlapping regions will be observed in function of the threshold selected.

Part V

Appendices

Appendix A

CoBCoM

In the early developing phases the state of the art segmentation procedures were deeply investigated. In particular, the research focused on the White Matter Query Language introduction and applicability.

The occasion to discuss with the WMQL creator, prof. D. Wassermann, coincided with the 2017 CoBCoM (Computational Brain Connectivity Mapping) event, held in Juan-les-Pins (France) on November.

During this five days conference various cutting-edge brain analysis methods were presented, covering tractography, connectomes, functional MRI and electrophysiological imaging techniques (here attached the conference program).

With over 28 talks this conference offered the chance to discuss with one of founder of the Human Connectome Project (whose data where used for the validation of Fuzzy Tracts), DiPy developers (library extensively used during the project) and to other high-regarded researcher in the diffusion imaging field.

The slides of the speakers can be found at the following link:
<https://project.inria.fr/cobcom2017/abstracts/>

From the multitude of speeches the most interesting ones were found to be:

- M. Descoteaux: Open challenges in tractography: Addressing tractography biases and tackling the false-positive problem;
- C. Lenglet: Advances in diffusion MRI from the human connectome project;
- A. Leemans: Quality assessment and correction methods for diffusion MRI data;

- R. Verma: Multimodal patho-connectomics;
- D. Wassermann: Computational neuroanatomy of the human brain white matter and beyond.

An extensive discussion with prof. Wassermann was very elucidating about the characteristic of the WMQL, its limitations and possible improvements. This encounter helped to build the foundation of what successively became Fuzzy Tracts. It was also the occasion to share the advances in tractogram segmentation methods with other researcher working on the topic, with a poster session fully dedicated to the diffusion MRI.

MONDAY 20		TUESDAY 21		WEDNESDAY 22	
8:45	RACHID DERICHE – Welcome address	9:00	CHRISTOPHE LENGLET – Advances in diffusion MRI from the human connectome project	9:00	ALEXANDER LEENANS – Quality assessment and correction methods for diffusion MRI data ExploreDTI demo
9:00	MARCUS KATSER – A Tutorial in Connectome Analysis I: Topological and Spatial Features of Brain Networks	10:30	COFFEE BREAK	10:30	COFFEE BREAK
10:30	COFFEE BREAK	11:00	FRANCO PESTILLI – Multidimensional encoding of brain connectomes	11:00	AUROBRATA GHOSH – Image quality transfer
11:00	MARCUS KATSER – A tutorial in connectome analysis II: Topological and spatial features of brain networks	12:30	LUNCH BREAK AND POSTER SESSION	11:45	SUDHIA K. PATNAK – Hollow nanotube textile based phantoms for ground validation of truth measurements of diffusion MR images: Testing compartmental models and fiber tractography
12:30	LUNCH BREAK	14:00	OLIVIER COULON – Seeding connectivity: Cortical parcellation	12:30	LUNCH BREAK AND POSTER SESSION
14:00	GIORGIO INDOCENTI – The development of neural connections obeys to algorithmic principles	15:30	COFFEE BREAK	14:00	DIERIK VAN DE VELLE – Functional connectivity models: from blobs to (dynamic) networks
15:00	MAXIME DISCOTEAUX – Open challenges in tractography: Addressing tractography biases and tackling the false-positive problem	15:45	MAUREN CLERC – Forward models for multimodal functional imaging	15:30	COFFEE BREAK
16:00	COFFEE BREAK	16:30	ALEXANDRE GRABHOFF – Methods for functional brain imaging with MEG and EEG: Better preprocessing, better inverse methods and better group studies ... with better software tools	16:00	ALFRED ANWARANDE – From links between functional and structural connectivity to quantitative plasticity of brain connectivity
16:30	MAXIME CHAMBERLAND – Advances in structural and functional connectivity visualization using the Fibernavigator	18:00	END OF DAY	17:30	SAMUEL DELABARRES-GAUTIER – Inferring information flow in the white matter of the brain: New information provided by the fusion of dMRI and M/EEG
17:00	EMMANUEL CARUYER – Validating tractography pipelines with the help of simulated phantoms				
17:30	GABRIEL GIRARD – Dipy: From Diffusion Data to Bundle Analysis				
18:00 WELCOME COCKTAIL					
THURSDAY 23		FRIDAY 24			
9:00	EVREN ÖZARSLAN – Recent advances in microstructure determination: From confinement tensors to the orientationally-averaged signal	9:00	GAIL VAROQUAUX – Functional connectomics: extracting and quantifying functional connectivity in fMRI Practical work: using scikit-learn and nilearn for resting-state fMRI biomarkers		
10:30	COFFEE BREAK	10:30	COFFEE BREAK		
11:00	CARL-FREDRIK WESTIN – Multidimensional diffusion MRI	11:00	MOD K. CHUNG – Persistent homological brain network analysis		
12:00	GLORIA MENEGAZ – A generalized SMT-based framework for diffusion MRI microstructural model estimation	12:30	LUNCH BREAK		
12:30	LUNCH BREAK AND POSTER SESSION	14:00	RAGNI VERMA – Multimodal patho-connectomics		
14:00	LUC FLORACK – Riemannian and Finslerian geometry for diffusion weighted magnetic resonance imaging	14:45	MAXIME GUYE – Structure-function relationship in epilepsy: Insights from multimodal imaging and computational models		
15:30	COFFEE BREAK	15:30	DIERIK WASSERMANN – Computational neuroanatomy of the human brain white matter and beyond		
16:00	TREDDORE PAPADOPOULOU – Invariants and shape: Application to diffusion MRI	16:15	RACHID DERICHE – Concluding remarks		
16:45	XAVIER PENNEC – Geometric statistics for computational anatomy				
18:00	END OF DAY				

Appendix B

OHBM

In order to contribute, either with a poster or an oral presentation, to the Organization for Human Brain Mapping (OHBM) 2018 annual meeting held in Singapore on June, 17-21 an abstract submission has been prepared. The abstract has been accepted on date March 1st, 2018

The institution, leader in the field of brain mapping studies, focus its work in gathering researcher from all around the world to improve the understanding of the functional and anatomical organization of the brain. The program of the annual meeting will span from functional and diffusion MRI to electrophysiological monitoring method, nuclear medicine and newly introduced imaging techniques.

For more information please follow the link to the main page of the organization: <https://www.humanbrainmapping.org/>

The full abstract can be found attached here.

Segmentation of White Matter Tractograms Using Fuzzy Spatial Relations

Alessandro Delmonte^{1,*}, Isabelle Bloch¹, Dominique Hasboun², Corentin Mercier¹, Johan Pallud³, and Pietro Gori¹

¹Télécom ParisTech, Paris, France

²AP-HP, Groupe Hospitalier Pitié-Salpêtrière, Paris, France

³Department of Neurosurgery, Hôpital Sainte-Anne, Paris, France,

*alessandro.delmonte@telecom-paristech.fr

Keywords: Diffusion MRI, Tractography, Segmentation, White Matter Anatomy, Fiber Pathways.

Introduction

Clinicians or researchers often want to isolate precise white matter tracts in order to test hypotheses related to a pathology. Segmenting white matter into reproducible tracts is difficult due to the huge amount of fibers and their vague anatomical definitions. The most common method, selecting fibers passing through manually delineated ROIs, is time-consuming and poorly reproducible for tracts with convoluted trajectories [4]. Differently, manually segmented ROIs can be transferred from training images to test subjects via non-linear deformations [4]. The resulting segmentation might not be accurate when training and test images do not share the same topology (e.g. due to a tumor). Other methods rely on unsupervised clustering and use a labeled atlas [5]. With a different perspective, a near-to-English query language (WMQL) allows the user to interactively define divisional terms, relative clauses and logical operations to segment anatomical tracts [1]. However, the shape of the resulting tracts can vary among subjects since WMQL is based on simple binary relations and bounding boxes. We propose to include qualitative anatomical definitions, modeled as fuzzy sets [6], into a semi-automatic segmentation algorithm, and to automatically compute an “anatomical coherence score” for every fiber of the tractogram.

Methods

Whole brain tracts were computed for 10 subjects from the HCP Project [7] using the SDSTREAM deterministic tracking algorithm of MRTrix3 [3], and the fractional anisotropy (FA) was extracted from the diffusion tensor. An initial segmentation, based on an exhaustive list of spatial relations (Fig.1) is performed with WMQL [1]. All anatomical labels are computed using FreeSurfer. To reduce the computational load and ease the visualization, a clustering step is implemented. Every cluster is then approximated with a prototype, computed as the fiber closest to the cluster centroid. We propose a new metric, called weighted functional varifolds, which takes into consideration the geometry, connectivity [4] and FA [8] of a fiber. The parameters were manually tuned based on experiments and anatomical considerations. Two fibers, with possibly different numbers of points, are similar only when they follow adjacent paths, their endpoints are close to each other and their FAs, mapped point-wise, are similar. An “anatomical coherence score” is then assigned to every prototype, based on the satisfaction of spatial relations, with respect to anatomical structures, defined as fuzzy sets in the image domain, thus coping with their intrinsic imprecision [6]. First, for every fuzzy set, we compute the mean value along the points of the prototypes. Then, we assign to the prototype the minimum score value, which is propagated to the fibers of the cluster. By thresholding it, clinicians can segment a tract in a fast and easy way and, notably, adapt the segmentation to a specific task.

Results

The evaluation focused on the Uncinate Fasciculus (Fig.1). The whole-brain tractogram (b) is initially segmented using WMQL (e) based on an exhaustive list of spatial relations (d). Then, we use the DBSCAN algorithm to separate fibers into clusters. Two other techniques (hierarchical and spectral) gave comparable results (not shown here). In (g) the prototypes are displayed. The fuzzy sets (f) are computed from the previous list and mapped onto the fibers (h). The coherence score is normalized between 0 and 1 (h). The final segmentation output is obtained setting a threshold of 0.5 (i). Fig. 2 shows the effects of the thresholding for a single subject (a), and the results for three subjects (b) where the threshold of 0.5 has been identified as the most anatomically accurate and reproducible by a neurosurgeon.

Conclusions

We presented a new semi-automatic tool for segmentation of white matter tractograms combining shape, connectivity, FA and anatomical information. The segmentation can be adapted in an intuitive way tuning a single parameter.

Bibliography

[1]

D. Wassermann et al (2013), 'On Describing Human White Matter Anatomy: The White Matter Query Language, *Med Image Comput Comput Assist Interv*, vol. 16, pp. 647–654.

[2]

P. Gori et al. (2016), 'Parsimonious Approximation of Streamline Trajectories in White Matter Fiber Bundles', *IEEE Transactions on Medical Imaging*, vol. 35.12, pp. 2609–2619.

[3]

J.-D. Tournier, F. Calamante, A. Connelly (2012), 'MRtrix: Diffusion tractography in crossing fiber regions', *International Journal of Imaging Systems and Technology*, vol. 22.1, pp. 53–66.

[4]

Y. Zhang et al. (2010), 'Atlas-guided tract reconstruction for automated and comprehensive examination of the white matter anatomy', *NeuroImage*, vol. 52.4, pp. 1289–1301.

[5]

P. Guevara et al. (2012), 'Automatic fiber bundle segmentation in massive tractography datasets using a multi-subject bundle atlas', *NeuroImage*, vol. 61.4, pp. 1083–1099.

[6]

I. Bloch (2005), 'Fuzzy spatial relationships for image processing and interpretation: a review', *Image and Vision Computing*, vol. 23.2, pp. 89–110.

[7]

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., et al. (2013). 'The WU-Minn human connectome project: an overview', *Neuroimage*, vol. 80, pp. 62–79.

[8]

K. Kumar, P. Gori, B. Charlier, S. Durrleman, O. Colliot, C. Desrosiers (2017), 'White matter fiber segmentation using functional varifolds', *MICCAI Workshop, Springer (LNCS)*.

Additional information

The author(s) declare no competing financial interests.

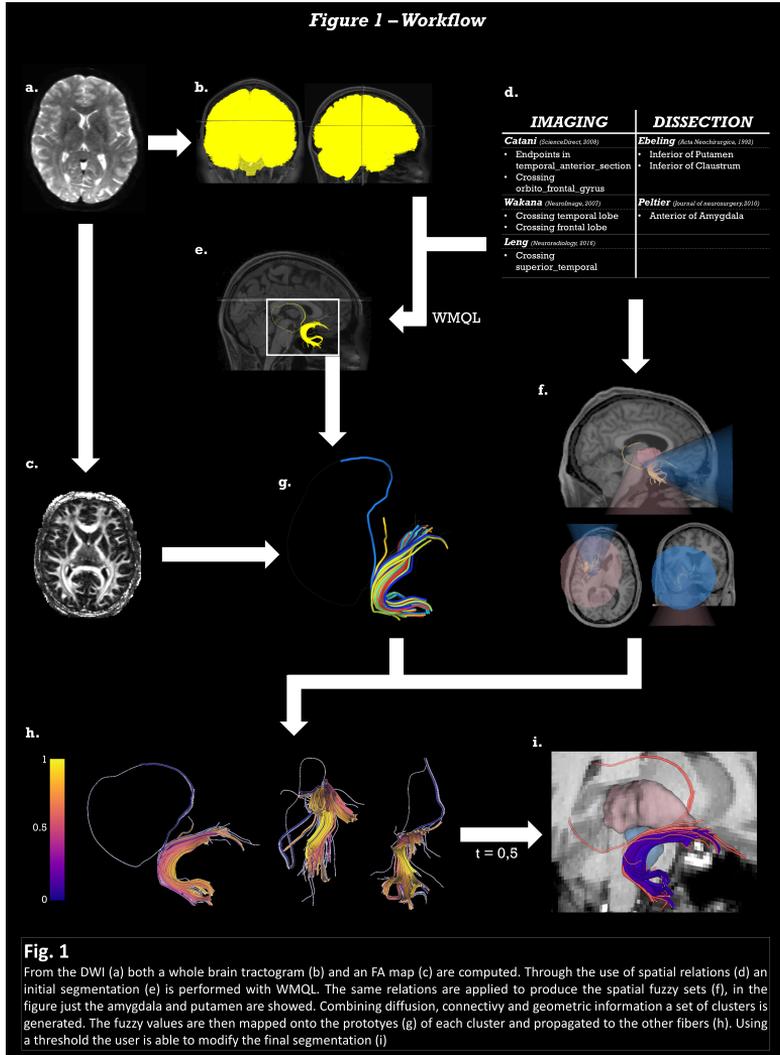


Figure 2 – Thresholding Effects Comparison and Reproducibility

HCP Dataset
Subject: 100307
Field Strength: 3T
Voxel Size: 1,25x1,25x1,25 mm

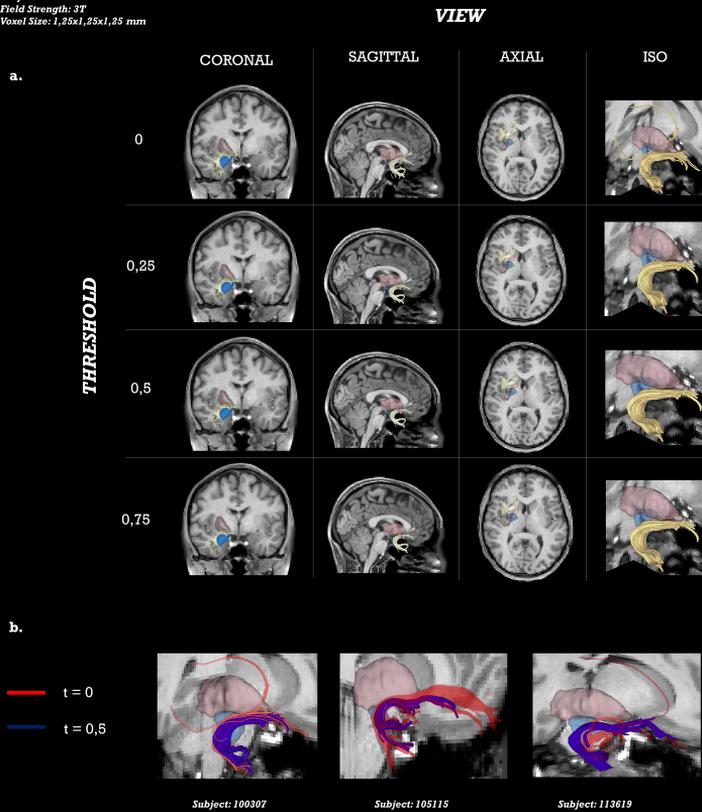


Fig. 2

a. Effects of different level of thresholding on a single subject (t=0 keeps all fibers, t=1 removes all fibers).
b. Inter subjects reproducibility. The threshold value 0,5 has been identified by a neurosurgeon as the most anatomically coherent and reproducible. The results have been validated and evaluated as anatomically coherent.

Bibliography

- [1] MRI atlas of human white matter. OCLC: 728068603.
- [2] Y. Assaf and O. Pasternak. Diffusion tensor imaging (DTI)-based white matter mapping in brain research: A review. 34(1):51–61.
- [3] P. J. Basser, J. Mattiello, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. 66(1):259–267.
- [4] P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. 44(4):625–632.
- [5] P.-L. Bazin, J. Bogovic, D. Reich, J. L. Prince, and D. L. Pham. Belief propagation based segmentation of white matter tracts in DTI. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 943–950. Springer.
- [6] T. Behrens, M. Woolrich, M. Jenkinson, H. Johansen-Berg, R. Nunes, S. Clare, P. Matthews, J. Brady, and S. Smith. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. 50(5):1077–1088.
- [7] T. E. J. Behrens, H. J. Berg, S. Jbabdi, M. F. S. Rushworth, and M. W. Woolrich. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? 34(1):144–155.
- [8] T. E. J. Behrens, H. Johansen-Berg, M. W. Woolrich, S. M. Smith, C. a. M. Wheeler-Kingshott, P. A. Boulby, G. J. Barker, E. L. Sillery, K. Sheehan, O. Ciccarelli, A. J. Thompson, J. M. Brady, and P. M. Matthews. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. 6(7):750–757.
- [9] T. Beppu, T. Inoue, Y. Shibata, N. Yamada, A. Kurose, K. Ogawara, A. Ogawa, and H. Kabasawa. Fractional anisotropy value by diffusion tensor magnetic resonance imaging as a predictor of cell density and proliferation activity of glioblastomas. 63(1):56–61.
- [10] D. L. Bihan. The 'wet mind': water and functional neuroimaging. 52(7):R57–R90.
- [11] I. Bloch. Fuzzy relative position between objects in image processing: a morphological approach. 21(7):657–664.
- [12] G. Borgefors. Distance transformations in arbitrary dimensions. 27(3):321–345.

- [13] R. Brown. *A Brief Account of Microscopical Observations Made ... on the Particles Contained in the Pollen of Plants, and on the General Existence of Active Molecules in Organic and Inorganic Bodies.*
- [14] A. Brun, H.-j. Park, H. Knutsson, C.-f. Westin, H. M. School, and M. Informatics. *Coloring of DT-MRI Fiber Traces Using Laplacian Eigenmaps.*
- [15] C. F. Caiafa and F. Pestilli. Multidimensional encoding of brain connectomes. 7(1).
- [16] M. Catani and M. Thiebaut de Schotten. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. 44(8):1105–1132.
- [17] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans. Automatic 3d intersubject registration of MR volumetric data in standardized talairach space. 18(2):192–205.
- [18] I. Corouge, S. Gouttard, and G. Gerig. Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)*, pages 344–347 Vol. 1.
- [19] A. Daducci, A. D. PalÁz, A. Lemkaddem, and J. P. Thiran. COMMIT: Convex optimization modeling for microstructure informed tractography. 34(1):246–257.
- [20] A. M. Dale, B. Fischl, and M. I. Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. 9(2):179–194.
- [21] T. Dhollander, D. Raffelt, and A. Connely. Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion MR data without a co-registered t1 image. In *Proc ISMRM Workshop on Breaking the Barriers of Diffusion MRI*, volume 5.
- [22] D. Dubois, H. Prade, and C. Testemale. Weighted fuzzy pattern matching. 28(3):313–331.
- [23] U. Ebeling and D. von Cramon. Topography of the uncinate fascicle and adjacent temporal fiber tracts. 115(3):143–148.
- [24] A. Einstein. Uber die von der molekularkinetischen theorie der warme geforderte bewegung von in ruhenden flussigkeiten suspendierten teilchen.
- [25] B. Erfianto, A. K. Mahmood, and A. S. A. Rahman. Modeling context and exchange format for context-aware computing. In *2007 5th Student Conference on Research and Development*, pages 1–5.
- [26] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise.

- [27] P. Fillard, M. Descoteaux, A. Goh, S. Gouttard, B. Jeurissen, J. Malcolm, A. Ramirez-Manzanares, M. Reisert, K. Sakaie, F. Tensaouti, T. Yo, J.-F. Mangin, and C. Poupon. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. 56(1):220–234.
- [28] B. Fischl. *FreeSurfer*. 62(2):774–781.
- [29] E. Garyfallidis, M. Brett, B. Amirbekian, A. Rokem, S. Van Der Walt, M. Descoteaux, and I. Nimmo-Smith. *Dipy*, a library for the analysis of diffusion MRI data. 8.
- [30] E. Garyfallidis, M. Brett, M. M. Correia, G. B. Williams, and I. Nimmo-Smith. *QuickBundles*, a method for tractography simplification. 6.
- [31] M. Giannelli, M. Cosottini, M. C. Michelassi, G. Lazzarotti, G. Belmonte, C. Bartolozzi, and M. Lazzeri. Dependence of brain DTI maps of fractional anisotropy and mean diffusivity on the number of diffusion weighting directions. 11(1):176–190.
- [32] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, and M. Jenkinson. The minimal preprocessing pipelines for the human connectome project. 80:105–124.
- [33] K. Gorgolewski, C. D. Burns, C. Madison, D. Clark, Y. O. Halchenko, M. L. Waskom, and S. S. Ghosh. *Nipype*: a flexible, lightweight and extensible neuroimaging data processing framework in python. 5:13.
- [34] P. Gori, O. Colliot, L. Marrakchi-Kacem, Y. Worbe, F. F. De Vico, M. Chavez, S. Lecomte, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman. A prototype representation to approximate white matter bundles with weighted currents. 17:289–296.
- [35] P. Gori, O. Colliot, L. Marrakchi-Kacem, Y. Worbe, F. D. V. Fallani, M. Chavez, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman. Parsimonious approximation of streamline trajectories in white matter fiber bundles. 35(12):2609–2619.
- [36] P. Guevara, D. Duclap, C. Poupon, L. Marrakchi-Kacem, J. Houenou, M. Leboyer, and J.-F. Mangin. Segmentation of short association bundles in massive tractography datasets using a multi-subject bundle atlas. In *Iberoamerican Congress on Pattern Recognition*, pages 701–708. Springer.
- [37] J. Hau, S. Sarubbo, G. Perchey, F. Crivello, L. Zago, E. Mellet, G. Jobard, M. Joliot, B. M. Mazoyer, N. Tzourio-Mazoyer, and L. Petit. Cortical terminations of the inferior fronto-occipital

- and uncinata fasciculi: Anatomical stem-based virtual dissection. 10:58.
- [38] J. R. Highley, M. A. Walker, M. M. Esiri, T. J. Crow, and P. J. Harrison. Asymmetry of the uncinata fasciculus: A post-mortem study of normal subjects and patients with schizophrenia. 12(11):1218–1224.
- [39] S. H. Jang, S. H. Kim, and H. D. Lee. Recovery of an injured prefronto-caudate tract in a patient with traumatic brain injury: A diffusion tensor tractography study. 31(11):1548–1551.
- [40] S. Jbabdi and H. Johansen-Berg. Tractography: Where do we go from here? 1(3):169–183.
- [41] B. Jeurissen, J.-D. Tournier, T. Dhollander, A. Connelly, and J. Sijbers. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. 103:411–426.
- [42] H. Jiang, P. C. M. van Zijl, J. Kim, G. D. Pearlson, and S. Mori. DtiStudio: resource program for diffusion tensor computation and fiber bundle tracking. 81(2):106–116.
- [43] D. K. Jones, A. Simmons, S. C. Williams, and M. A. Horsfield. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. 42(1):37–41.
- [44] E. L. Kier, L. H. Staib, L. M. Davis, and R. A. Bronen. Anatomic dissection tractography: a new method for precise MR localization of white matter tracts. 25(5):670–676.
- [45] L. Klingler. Atlas cerebri humani.
- [46] P. Kochunov, P. Thompson, J. Lancaster, G. Bartzokis, S. Smith, T. Coyle, D. Royall, A. Laird, and P. Fox. Relationship between white matter fractional anisotropy and other indices of cerebral health in normal aging: Tract-based spatial statistics study of aging. 35(2):478–487.
- [47] D.-M. Koh, M. Blackledge, A. R. Padhani, T. Takahara, T. C. Kwee, M. O. Leach, and D. J. Collins. Whole-body diffusion-weighted MRI: Tips, tricks, and pitfalls. 199(2):252–262.
- [48] R. Ladner, F. E. Petry, and M. A. Cobb. Fuzzy set approaches to spatial data mining of association rules. 7(1):123–138.
- [49] M. Larobina and L. Murino. Medical image file formats. 27(2):200–206.
- [50] I. N. C. Lawes, T. R. Barrick, V. Murugam, N. Spierings, D. R. Evans, M. Song, and C. A. Clark. Atlas-based segmentation of white matter tracts of the human brain using diffusion tensor tractography and comparison with classical dissection. 39(1):62–79.

- [51] D. Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: Concepts and applications. 13(4):534–546.
- [52] B. Leng, S. Han, Y. Bao, H. Zhang, Y. Wang, Y. Wu, and Y. Wang. The uncinate fasciculus as observed using diffusion spectrum imaging in the human brain. 58(6):595–606.
- [53] H. Li, Z. Xue, L. Guo, T. Liu, J. Hunter, and S. T. Wong. A hybrid approach to automatic clustering of white matter fibers. 49(2):1249–1258.
- [54] N. Lori, E. Akbudak, A. Snyder, J. Shimony, and T. Conturo. Diffusion tensor tracking of human neuronal fiber bundles: Simulation of effects of noise, voxel size and data interpolation. In *Proc. International Society for Magnetic Resonance in Medicine*, page 775.
- [55] U. V. Luxburg. *A Tutorial on Spectral Clustering*.
- [56] K. H. Maier-Hein, P. F. Neher, J.-C. Houde, M.-A. Cote, E. Garyfallidis, J. Zhong, M. Chamberland, F.-C. Yeh, Y.-C. Lin, Q. Ji, W. E. Reddick, J. O. Glass, D. Q. Chen, Y. Feng, C. Gao, Y. Wu, J. Ma, H. Renjie, Q. Li, C.-F. Westin, S. Deslauriers-Gauthier, J. O. O. Gonzalez, M. Paquette, S. St-Jean, G. Girard, F. Rheault, J. Sidhu, C. M. W. Tax, F. Guo, H. Y. Mesri, S. David, M. Froeling, A. M. Heemskerk, A. Lee-mans, A. Bore, B. Pinsard, C. Bedetti, M. Desrosiers, S. Brambati, J. Doyon, A. Sarica, R. Vasta, A. Cerasa, A. Quattrone, J. Yeatman, A. R. Khan, W. Hodges, S. Alexander, D. Romascano, M. Barakovic, A. Auria, O. Esteban, A. Lemkaddem, J.-P. Thiran, H. E. Cetingul, B. L. Odry, B. Mailhe, M. S. Nadar, F. Pizzagalli, G. Prasad, J. E. Villalon-Reina, J. Galvis, P. M. Thompson, F. D. S. Requejo, P. L. Laguna, L. M. Lacerda, R. Barrett, F. Dell Acqua, M. Catani, L. Petit, E. Caruyer, A. Daducci, T. B. Dyrby, T. Holland-Letz, C. C. Hilgetag, B. Stieltjes, and M. Descoteaux. The challenge of mapping the human connectome based on diffusion tractography. 8(1).
- [57] J. Martino, C. Brogna, S. G. Robles, F. Vergani, and H. Duffau. Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data. 46(5):691–699.
- [58] J. Martino and E. M. De Lucas. Subcortical anatomy of the lateral association fascicles of the brain: A review. 27(4):563–569.
- [59] J. Martino, P. C. De Witt Hamer, F. Vergani, C. Brogna, E. M. de Lucas, A. Vazquez-Barquero, J. A. Garcia-Porrero, and H. Duffau. Cortex-sparing fiber dissection: an improved method for the

- study of white matter anatomy in the human brain. 219(4):531–541.
- [60] J. Martino, F. Vergani, S. G. Robles, and H. Duffau. New insights into the anatomic dissection of the temporal stem with special emphasis on the inferior fronto-occipital fasciculus: implications in surgical approach to left mesiotemporal and temporoinsular structures. 66(3):4–12.
- [61] L. Minati and W. P. Weglarz. Physical foundations, models, and methods of diffusion magnetic resonance imaging of the brain: A review. 30A(5):278–307.
- [62] K. Miyajima and A. Ralescu. Spatial organization in 2d segmented images: representation and recognition of primitive spatial relations. 65(2):225–236.
- [63] B. Moberts, A. Vilanova, and J. J. Van Wijk. Evaluation of fiber clustering methods for diffusion tensor imaging. In *Visualization, 2005. VIS 05. IEEE*, pages 65–72. IEEE.
- [64] S. Mori, B. J. Crain, V. P. Chacko, and P. C. van Zijl. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. 45(2):265–269.
- [65] S. Mori and P. C. M. van Zijl. Fiber tracking: principles and strategies - a technical review. 15(7):468–480.
- [66] P. Mukherjee, J. Berman, S. Chung, C. Hess, and R. Henry. Diffusion tensor MR imaging and fiber tractography: Theoretic underpinnings. 29(4):632–641.
- [67] L. J. O’Donnell, C.-F. Westin, and A. J. Golby. Tract-based morphometry for white matter group analysis. 45(3):832–844.
- [68] K. Oishi, K. Zilles, K. Amunts, A. Faria, H. Jiang, X. Li, K. Akhter, K. Hua, R. Woods, A. W. Toga, G. B. Pike, P. Rosa-Neto, A. Evans, J. Zhang, H. Huang, M. I. Miller, P. C. van Zijl, J. Mazziotta, and S. Mori. Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter. 43(3):447–457.
- [69] E. Ozarslan and T. H. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. 50(5):955–965.
- [70] R. P. Haining, M. Goodchild, and S. Gopal. The accuracy of spatial databases. 16:243.
- [71] J. Peltier, S. Verclytte, C. Delmaire, J.-P. Pruvo, O. Godefroy, and D. Le Gars. Microsurgical anatomy of the temporal stem: clinical relevance and correlations with diffusion tensor imaging fiber tracking. 112(5):1033–1038.
- [72] J. S. Peper, R. M. Brouwer, D. I. Boomsma, R. S. Kahn, and H. E. Hulshoff Pol. Genetic influences

- on human brain structure: A review of brain imaging studies in twins. 28(6):464–473.
- [73] L. Pescatori, M. P. Tropeano, A. Manfreda, R. Delfini, and A. Santoro. Three-dimensional anatomy of the white matter fibers of the temporal lobe: Surgical implications. 100:144–158.
- [74] C. Pierpaoli, P. Jezzard, P. J. Basser, A. Barnett, and G. Di Chiro. Diffusion tensor MR imaging of the human brain. 201(3):637–648.
- [75] A. Qayyum. Diffusion-weighted imaging in the abdomen and pelvis: Concepts and applications. 29(6):1797–1810.
- [76] E. C. Ribas, K. Yagmurlu, H. T. Wen, and A. L. Rhoton. Microsurgical anatomy of the inferior limiting insular sulcus and the temporal stem. 122(6):1263–1273.
- [77] L. Rokach and O. Maimon. Clustering methods. In *Data Mining and Knowledge Discovery Handbook*, pages 321–352. Springer, Boston, MA. DOI: 10.1007/0-387-25465-X_15.
- [78] K. Schilling, Y. Gao, V. Janve, I. Stepniewska, B. A. Landman, and A. W. Anderson. Can increased spatial resolution solve the crossing fiber problem for diffusion MRI? 30(12).
- [79] V. J. Schmithorst, M. Wilke, B. J. Dardzinski, and S. K. Holland. Correlation of white matter diffusivity and anisotropy with age during childhood and adolescence: A cross-sectional diffusion-tensor MR imaging study. 222(1):212–218.
- [80] W. J. Schroeder, L. S. Avila, and W. Hoffman. Visualizing with VTK: A tutorial. 20(5):20–27.
- [81] Y.-W. Shin, D. J. Kim, T. Hyon, H.-J. Park, W.-J. Moon, E. C. Chung, J. M. Lee, I. Y. Kim, S. I. Kim, and J. S. Kwon. Sex differences in the human corpus callosum: diffusion tensor imaging study. 16(8):795–798.
- [82] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly. Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. 62(3):1924–1938.
- [83] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly. SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. 119:338–351.
- [84] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. J. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. M. Brady, and P. M. Matthews. Advances in

- functional and structural MR image analysis and implementation as FSL. 23 Suppl 1:S208–219.
- [85] N. Soni, A. Mehrotra, S. Behari, S. Kumar, and N. Gupta. Diffusion-tensor imaging and tractography application in pre-operative planning of intra-axial brain lesions. 9(10).
- [86] E. O. Stejskal and J. E. Tanner. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. 42(1):288–292.
- [87] S. C. Strother. Evaluating fMRI preprocessing pipelines. 25(2):27–41.
- [88] C. Thomas, F. Q. Ye, M. O. Irfanoglu, P. Modi, K. S. Saleem, D. A. Leopold, and C. Pierpaoli. Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. 111(46):16574–16579.
- [89] J.-D. Tournier, F. Calamante, and A. Connelly. MRtrix: Diffusion tractography in crossing fiber regions. 22(1):53–66.
- [90] J.-D. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. 35(4):1459–1472.
- [91] J.-D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. 23(3):1176–1185.
- [92] D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. 48(4):577–582.
- [93] K. Ugurbil, J. Xu, E. J. Auerbach, S. Moeller, A. T. Vu, J. M. Duarte-Carvajalino, C. Lenglet, X. Wu, S. Schmitter, P. F. Van de Moortele, J. Strupp, G. Sapiro, F. De Martino, D. Wang, N. Harel, M. Garwood, L. Chen, D. A. Feinberg, S. M. Smith, K. L. Miller, S. N. Sotiropoulos, S. Jbabdi, J. L. R. Andersson, T. E. J. Behrens, M. F. Glasser, D. C. Van Essen, and E. Yacoub. Pushing spatial and temporal resolution for functional and diffusion MRI in the human connectome project. 80:80–104.
- [94] N. V. Malykhin, L. Concha, P. Seres, C. Beaulieu, and N. J. Coupland. Diffusion tensor imaging tractography and reliability analysis for limbic and paralimbic white matter tracts. 164:132–42.
- [95] M. Van Droogenbroeck and H. Talbot. Fast computation of morphological operations with arbitrary structuring elements. 17(14):1451–1460.

- [96] D. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. Curtiss, S. Della Penna, D. Feinberg, M. Glasser, N. Harel, A. Heath, L. Larson-Prior, D. Marcus, G. Michalareas, S. Moeller, R. Oostenveld, S. Petersen, F. Prior, B. Schlaggar, S. Smith, A. Snyder, J. Xu, and E. Yacoub. The human connectome project: A data acquisition perspective. 62(4):2222–2231.
- [97] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. J. Behrens, E. Yacoub, K. Ugurbil, and WU-Minn HCP Consortium. The WU-minn human connectome project: an overview. 80:62–79.
- [98] M. I. Vargas, J. Delavelle, H. Jlassi, B. Rilliet, M. Viallon, C. D. Becker, and K.-O. Lovblad. Clinical applications of diffusion tensor tractography of the spinal cord. 50(1):25–29.
- [99] R. J. Von Der Heide, L. M. Skipper, E. Klobusicky, and I. R. Olson. Dissecting the uncinate fasciculus: disorders, controversies and a hypothesis. 136(6):1692–1707.
- [100] S. B. Vos, C. M. W. Tax, P. R. Luijten, S. Ourselin, A. Leemans, and M. Froeling. The importance of correcting for signal drift in diffusion MRI. 77(1):285–299.
- [101] S. Wakana, A. Caprihan, M. M. Panzenboeck, J. H. Fallon, M. Perry, R. L. Gollub, K. Hua, J. Zhang, H. Jiang, P. Dubey, A. Blitz, P. van Zijl, and S. Mori. Reproducibility of quantitative tractography methods applied to cerebral white matter. 36(3):630–644.
- [102] S. Wakana, H. Jiang, L. M. Nagae-Poetscher, P. C. M. van Zijl, and S. Mori. Fiber tract-based atlas of human white matter anatomy. 230(1):77–87.
- [103] R. Wang, T. Benner, A. G. Sorensen, and V. J. Wedeen. Diffusion toolkit: a software package for diffusion imaging data processing and tractography. In *Proc Intl Soc Mag Reson Med*, volume 15. Berlin.
- [104] D. Wassermann, M. Descoteaux, and R. Deriche. Diffusion maps clustering for magnetic resonance q-ball imaging segmentation. 2008:1–12.
- [105] D. Wassermann, N. Makris, Y. Rath, M. Shenton, R. Kikinis, M. Kubicki, and C.-F. Westin. On describing human white matter anatomy: The white matter query language. 16(0):647–654.
- [106] D. Wassermann, N. Makris, Y. Rath, M. Shenton, R. Kikinis, M. Kubicki, and C.-F. Westin. The white matter query language: a novel approach for describing human white matter anatomy. 221(9):4705–4721.
- [107] D. J. Werring, C. A. Clark, G. J. Barker, A. J. Thompson, and D. H. Miller. Diffusion tensor

- imaging of lesions and normal-appearing white matter in multiple sclerosis. 52(8):1626–1632.
- [108] K. Yamada, K. Sakai, K. Akazawa, S. Yuen, and T. Nishimura. MR tractography: A review of its clinical applications.
- [109] J. D. Yeatman, R. F. Dougherty, N. J. Myall, B. A. Wandell, and H. M. Feldman. Tract profiles of white matter properties: Automating fiber-tract quantification. 7(11):e49790.
- [110] Y. Zhang, J. Zhang, K. Oishi, A. V. Faria, H. Jiang, X. Li, K. Akhter, P. Rosa-Neto, G. B. Pike, A. Evans, A. W. Toga, R. Woods, J. C. Mazziotta, M. I. Miller, P. C. M. van Zijl, and S. Mori. Atlas-guided tract reconstruction for automated and comprehensive examination of the white matter anatomy. 52(4):1289–1301.
- [111] O. Zvitia, A. Mayer, and H. Greenspan. Adaptive mean-shift registration of white matter tractographies. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 692–695.