

# Représentation et segmentation des fibres de matière blanche basées sur les zéros de la transformée en ondelettes et sur l'alignement de séquences

Côté Marc-Alexandre, Girard Gabriel, Wang Shengrui et Descoteaux Maxime

Département d'Informatique  
Université de Sherbrooke  
Sherbrooke (Qc), Canada, J1K 2R1  
{Marc-Alexandre.Cote, Gabriel.P.Girard, Shengrui.Wang, Maxime.Descoteaux}@usherbrooke.ca

## - Rapport de recherche #32 -

### Résumé.

L'imagerie par résonance magnétique de diffusion (IRMd) permet l'acquisition d'informations structurelles des tissus biologiques du cerveau. La matière blanche est composée d'un réseau de fibres responsables de la transmission des signaux électriques entre différentes zones cognitives de la matière grise. L'IRMd permet la reconstruction de ce réseau par un procédé appelé tractographie. Les algorithmes de tractographie réussissent à extraire des dizaines de milliers de fibres dans un cerveau. À ce jour, l'interaction et la sélection de faisceaux de fibres se font à l'aide d'outil de visualisation; le processus est très peu automatisé et long. Certains faisceaux sont très difficiles à sélectionner avec ces outils. Ce rapport a pour but de présenter les résultats préliminaires obtenus par une méthode innovatrice de sélection de fibres basée sur leurs caractéristiques géométriques. Chacune des fibres est soumise à une transformée en ondelettes afin d'en extraire une signature géométrique. Celle-ci sera par la suite comparée aux autres fibres par un algorithme d'alignement de séquences. La méthode calcule ensuite une distance entre les fibres trouvant ainsi les faisceaux de fibres ayant une grande similarité. Les résultats obtenus suite à des expérimentations avec des fibres réelles sont encourageants. Le traitement *a posteriori* des fibres fournit un nouvel outil pour l'étude de la connectivité cérébrale, notamment pour l'étude des maladies neurodégénératives et pour la compréhension du développement cérébral.

**Mot-clés:** fibres de matière blanche, imagerie par résonance magnétique de diffusion, ondelette, segmentation, visualisation.

# Table des matières

<b>Liste des figures</b>	<b>3</b>
<b>1 Introduction</b>	<b>4</b>
1.1 Matière blanche . . . . .	4
1.1.1 Interaction et sélection . . . . .	4
1.2 Problématique . . . . .	6
<b>2 Théories</b>	<b>9</b>
2.1 Transformée en ondelettes . . . . .	9
2.1.1 Zero-Crossing . . . . .	9
2.2 Alignement de chaînes . . . . .	10
<b>3 Méthodologie</b>	<b>12</b>
3.1 Notre approche . . . . .	13
3.1.1 Représentation alternative . . . . .	13
3.1.2 Représentation compacte . . . . .	14
3.1.3 Calcul de distance . . . . .	14
3.1.4 Segmentation . . . . .	15
3.2 Implémentation . . . . .	17
3.2.1 Translation et orientation invariant . . . . .	17
3.2.2 Représentation alternative . . . . .	17
3.2.3 Ré-échantillonnage . . . . .	19
3.2.4 Taille invariant . . . . .	19
3.2.5 Transformée en ondelettes . . . . .	19
3.2.6 Seuillage . . . . .	20
3.2.7 Zero-Crossing . . . . .	21
3.2.8 Alignement . . . . .	21
3.2.9 Segmentation . . . . .	21
<b>4 Résultats</b>	<b>23</b>
4.1 Améliorations . . . . .	24
<b>5 Conclusion</b>	<b>26</b>
5.1 Ouverture . . . . .	26
<b>Bibliographie</b>	<b>27</b>

## Liste des figures

1.1	Exemple de plusieurs faisceaux de matière blanche du cerveau humain avec une vue sagittale. Les faisceaux principaux de l'image sont numérotés de 1 à 7. Image de [Williams-et al97]. . . . .	5
1.2	Exemple de regroupement de fibres numérotées de 1 à 7 (a) et positionnement dans le cerveau humain (b). Images de [Williams-et al97]. . . . .	5
1.3	Image de la reconstruction des fibres de matière blanche par un algorithme de tractographie. Image obtenue par le logiciel <i>FiberNavigator</i> . . . . .	6
1.4	Interaction avec les fibres. Sélection du faisceau de la couronne rayonnante de l'hémisphère droit. Toutes les fibres traversant la boîte bleue sont affichées. De plus, les fibres intersectant une boîte rouge, ne sont pas affichées. Image obtenue par le logiciel <i>FiberNavigator</i> . . . . .	7
1.5	Interaction avec les fibres. Les volumes rouges sont des zones d'activation du cerveau obtenues par IRM fonctionnelle. Toutes les fibres traversant les volumes sont affichées. Image obtenue par le logiciel <i>FiberNavigator</i> . . . . .	7
2.1	Représentation <i>Zero-Crossing</i> d'un signal 1D. Image de Mallat [4]. . . . .	10
2.2	Exemple de la transformée en ondelettes (b) d'un signal (a) ainsi que sa représentation <i>Zero-Crossing</i> (c). Images de Mallat [4]. . . . .	10
2.3	Exemple d'alignement global utilisant l'algorithme de Needleman-Wunsch [1]. . . . .	11
2.4	Exemple d'alignement local utilisant l'algorithme de Smith-Waterman [1]. . . . .	11
3.1	Exemple de représentation alternative d'une fibre. . . . .	18
3.2	Exemple d'échantillonnage d'une fibre utilisant une interpolation spline cubique. . . . .	19
3.3	Transformées en ondelettes d'une fibre. . . . .	20
3.4	Seuillage <i>hard</i> d'une fibre. Seuil utilisé est de 0.01 . . . . .	20
3.5	Représentation <i>Zero-Crossing</i> d'une fibre à l'échelle $2^3$ . . . . .	21
3.6	Quatre fibres identifiées par leur couleur ( <b>B</b> leue, <b>R</b> ouge, <b>M</b> auve et <b>V</b> erte). $d(B, V) = 1.0698$ , $d(B, M) = 0.5554$ et $d(B, R) = 1.8788$ . . . . .	22
4.1	Dendrogramme représentant la segmentation hiérarchique <i>Complete Linkage</i> de 200 fibres choisies aléatoirement. . . . .	23
4.2	Groupe de fibres qui possèdent des caractéristiques semblables. . . . .	23
4.3	Groupes de fibres qui n'ont aucune caractéristiques communes. . . . .	24
4.4	Réseau de fibres artificielles. On y remarque la présence de faisceaux. . . . .	25

# 1 Introduction

Ce rapport présente l'avancement des travaux d'un projet de recherche portant sur la segmentation des fibres de matière blanche du cerveau humain. Ce projet est supervisé conjointement par les professeurs Maxime Descoteaux et Shengrui Wang dans le cadre des cours IMN767 - *Outils mathématiques du traitement du signal* et BIN701 - *Forage de données* respectivement. La section 1 du document introduit les méthodes génériques présentées dans l'un ou l'autre des deux cours et appliquées au projet. La section 2 présente la problématique, alors que la section 3 présente les outils mathématiques et informatiques proposés pour résoudre le problème posé. La section 4 est dédiée à l'implémentation et aux résultats préliminaires de la solution proposée. Les sections 5 et 6 terminent respectivement avec une conclusion et une ouverture sur d'autres applications de nos recherches.

## 1.1 Matière blanche

La matière blanche est avec la matière grise l'une des deux catégories de tissus du système nerveux. La matière blanche est composée de fibres nerveuses, les axones myélinisés et non myélinisés, regroupés en faisceaux (voir figures 1.1 et 1.2), connectant entre elles différentes parties du cerveau et transmettant des impulsions nerveuses. La compréhension du fonctionnement de notre cerveau fait partie des nouvelles frontières de la recherche biologique. Des milliers de chercheurs tentent de comprendre ce fascinant organe, mais la tâche est rude. Même s'il ne représente qu'environ 1,5 kilogramme de matière, il contient quelque 100 milliards de cellules nerveuses. L'étude de la connectivité cérébrale est un élément clé de la compréhension des maladies neuro-dégénératives telles la sclérose en plaques et l'Alzheimer ainsi que du développement du cerveau. Depuis 30 ans, l'imagerie par résonance magnétique (IRM) s'est imposée comme la méthode d'imagerie de référence pour le diagnostic de plusieurs maladies.

L'imagerie de diffusion par résonance magnétique (IRM de diffusion) est une technique d'imagerie médicale qui utilise la résonance magnétique pour observer la diffusion des molécules d'eau dans un objet. On peut notamment observer l'orientation des fibres des tissus humains, c'est par conséquent une technique de choix dans l'étude des fibres de matière blanche. Elle permet d'obtenir l'orientation des fibres dans l'espace avec une résolution millimétrique. Des techniques de traitement d'image permettant de reconstituer le trajet des fibres ont été développées. Une carte de la connectivité cérébrale peut alors être étudiée (voir figure 1.3).

### 1.1.1 Interaction et sélection

L'algorithme de tractographie produit un ensemble de courbes 3D positionnée dans l'espace. Chaque courbe représente une forte probabilité de présence de lien neuronal à cet endroit. La tractographie, à partir d'images en IRM de diffusion, génère quelques dizaines de milliers de fibres, voir quelques millions, selon les paramètres utilisés et la résolution des images de l'IRM. La figure 1.3 montre l'ensemble

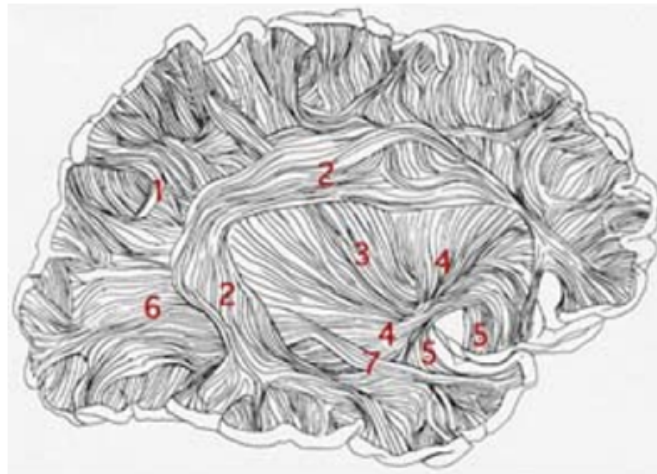
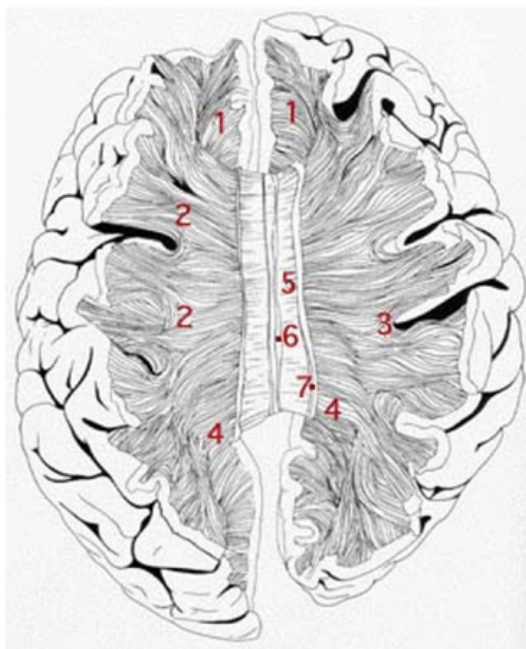
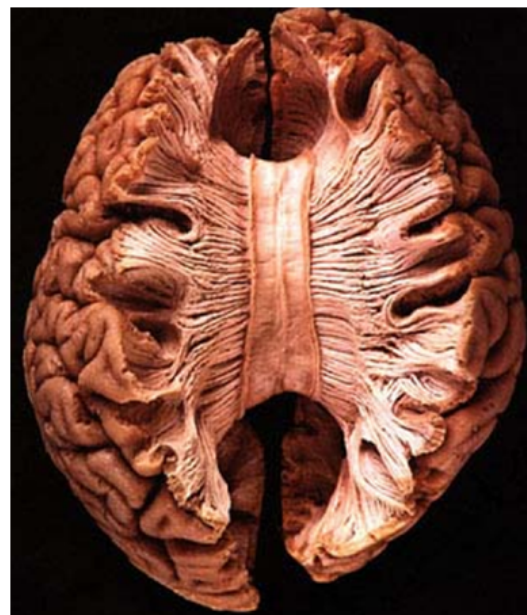


Fig. 1.1: Exemple de plusieurs faisceaux de matière blanche du cerveau humain avec une vue sagittale. Les faisceaux principaux de l'image sont numérotés de 1 à 7. Image de [Williams-et al97].



(a) Faisceaux principaux



(b) Cerveau humain, vue axiale

Fig. 1.2: Exemple de regroupement de fibres numérotées de 1 à 7 (a) et positionnement dans le cerveau humain (b). Images de [Williams-et al97].

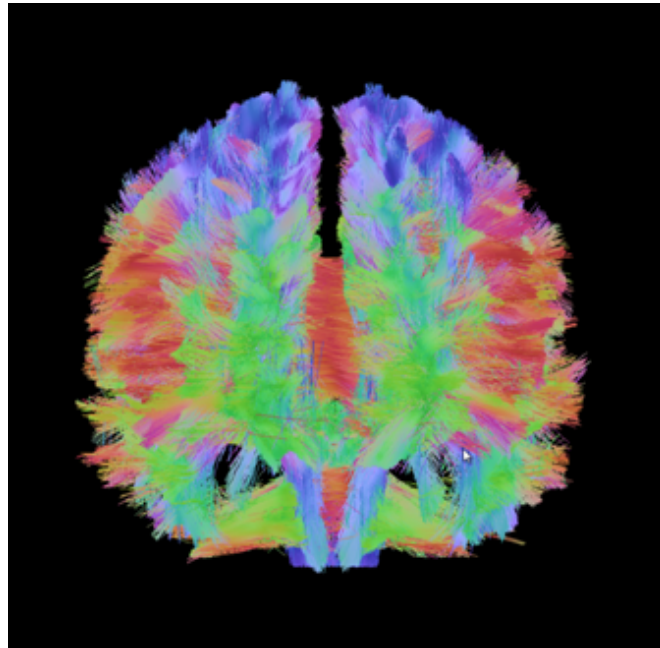


Fig. 1.3: Image de la reconstruction des fibres de matière blanche par un algorithme de tractographie. Image obtenue par le logiciel *FiberNavigator*.

des fibres obtenues.

Afin de pouvoir visionner des sous-ensembles de fibres, on utilise des volumes d'intérêt pour faire la sélection. On positionne des volumes dans l'espace et on conserve des fibres intersectant certains volumes et n'intersectant pas d'autres volumes. La figure 1.4 présente l'utilisation de boîtes et la figure 1.5 montre l'utilisation de volumes d'intérêt générés automatiquement à partir d'images d'IRM fonctionnelle. Il est également possible de combiner les différents volumes de sélection de fibres (*e.g.* volumes d'intérêt et boîtes de sélection). Il est aussi possible de sous-échantillonner les fibres avec des filtres. Le logiciel *FiberNavigator* propose deux types de filtre : filtrage basé sur la longueur des fibres ou bien un sous-échantillonnage aléatoire des fibres.

## 1.2 Problématique

La sélection de sous-ensembles de fibres est très importante dans l'étude de la connectivité cérébrale. L'utilisation de volumes de sélection est une bonne approche en ce sens où l'on s'intéresse souvent aux fibres étant physiquement à un endroit. En positionnant un volume à cet endroit, on obtient directement l'information recherchée. Cependant, la tâche de bien sélectionner un faisceau peut parfois être longue (*e.g.* lorsque plusieurs faisceaux se croisent en un même point). De plus, certains faisceaux comprennent moins de fibres ou des fibres plus courtes, il est alors plus difficile de bien sélectionner ce faisceau.

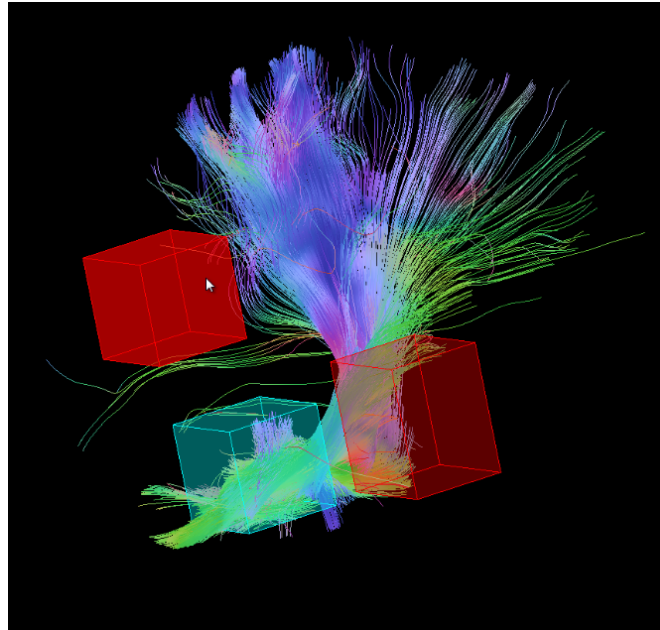


Fig. 1.4: Interaction avec les fibres. Sélection du faisceau de la couronne rayonnante de l'hémisphère droit. Toutes les fibres traversant la boîte bleue sont affichées. De plus, les fibres intersectant une boîte rouge, ne sont pas affichées. Image obtenue par le logiciel *FiberNavigator*.

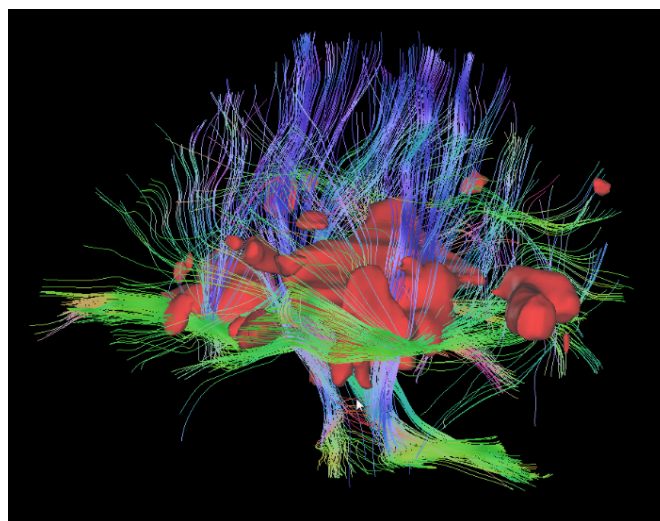


Fig. 1.5: Interaction avec les fibres. Les volumes rouges sont des zones d'activation du cerveau obtenues par IRM fonctionnelle. Toutes les fibres traversant les volumes sont affichées. Image obtenue par le logiciel *FiberNavigator*.

Pour certains faisceaux, l'utilisation de volumes n'est pratiquement pas applicable (*e.g.* les faisceaux liant les convolutions extérieures du cerveau). Ce rapport de recherche propose une méthode pour segmenter les fibres de matière blanche du cerveau basée sur la transformée en ondelettes et l'alignement de séquences. Dans le reste du document, la Section 2 explique la théorie des techniques employées (*i.e.* transformée en ondelettes et l'alignement de séquences), la Section 3 traite de la méthode proposée et des expérimentations effectuées, la Section 4 présente les résultats, et finalement la Section 5 conclut ce rapport de recherche.



## 2 Théories

### 2.1 Transformée en ondelettes

La transformée en ondelettes découle des limites de l'analyse de Fourier. Celle-ci ne permet pas de réaliser l'étude des composantes transitoires de durées différentes d'un signal, car elle repose sur des fonctions sinusoïdales, périodiques et à support infini. Au contraire, les ondelettes permettent de décomposer un signal en fréquences, comme la transformée de Fourier, mais également en positions dans le temps. De plus, l'analyse par ondelettes est basée sur un concept quelque peu différent du concept de fréquence: le concept d'échelle. Les ondelettes sont des copies les unes des autres, copies presque conformes puisqu'elles sont de forme constante et ne diffèrent que par leur taille. Un signal s'écrit alors sous la forme d'une superposition de telles ondelettes décalées et dilatées. Donc, la transformée en ondelettes est une fonction de deux variables: le temps et l'échelle (ou dilatation). Cela permet souvent une meilleure analyse des signaux présentant des discontinuités.

Intuitivement, une ondelette est une petite onde oscillant, généralement de moyenne nulle, avec une amplitude de départ et de fin à zéro. L'ondelette est choisie selon les caractéristiques du signal et les traitements requis. La transformée en ondelettes est la décomposition d'un signal en une somme pondérée de l'ondelette dilatée et translatée. La transformée en ondelettes dyadiques permet de réaliser une analyse des structures locales d'un signal avec un *zoom* qui dépend de l'échelle considérée.

Très souvent l'essentiel de l'information d'un signal se situe dans ses singularités et ses structures irrégulières. En traitement d'image, par exemple, les contours sont caractérisés par la discontinuité de l'intensité.

#### 2.1.1 Zero-Crossing

Mallat propose une représentation alternative de la transformée en ondelettes appelée *Zero-Crossing* [4]. Simplement, il s'agit de construire une courbe en onde carrée dont les changements de signe (*Zero-Crossing*) sont les mêmes que pour la courbe originale. De plus, l'aire sous la courbe entre deux zéros doit être la même pour les deux représentations. La figure 2.1 présente un exemple du respect de ces conditions. La figure 2.2 présente le traitement appliqué à un signal 1D jusqu'à la représentation *Zero-Crossing*.

Malgré sa forme, qui semble plus simple, la représentation *Zero-Crossing* pourrait être une représentation complète du signal original. Aucune preuve permet d'affirmer que cette représentation est complète, cependant aucun contre-exemple n'existe à ce jour [4]. Autrement dit, il n'existerait qu'un seul signal pour une représentation *Zero-Crossing* et une ondelette données, mais la preuve reste toujours à faire. Il est possible de retrouver une représentation en ondelettes à partir d'une représentation *Zero-Crossing* en  $O(n \log^2(n))$ . Pour ce travail, nous supposons que cette représentation est suffisamment complète pour des fins de comparaison.

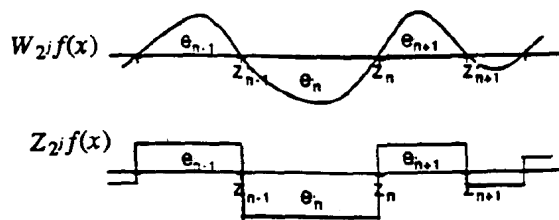


Fig. 2. Function  $Z_{2j}f(x)$  has the same zero-crossings and integral values as  $W_{2j}f(x)$  and is constant between two consecutive zero-crossings.

Fig. 2.1: Représentation *Zero-Crossing* d'un signal 1D. Image de Mallat [4].

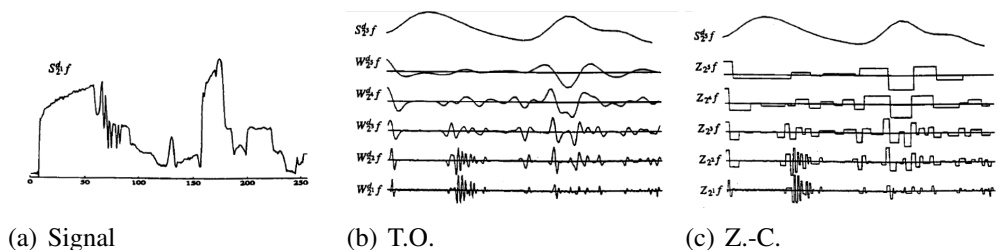


Fig. 2.2: Exemple de la transformée en ondelettes (b) d'un signal (a) ainsi que sa représentation *Zero-Crossing* (c). Images de Mallat [4].

## 2.2 Alignement de chaînes

L'alignement de chaînes (de séquences) est une manière de disposer les caractères de façon à identifier des zones de concordances (similarités). Lors de l'alignement, il existe trois opérations pour modifier une chaîne, soit: l'insertion, la suppression et la substitution. Une mesure de similarité est associée à chaque action. Par conséquent, le but est de maximiser la similarité totale de l'alignement entre deux chaînes.

Il existe deux grandes familles d'alignement de chaînes. La première est l'alignement global et consiste à aligner deux séquences sur toute leur longueur. Un exemple est présenté à la figure 2.3. La deuxième met l'accent sur l'alignement d'une sous-chaîne commune aux deux chaînes. À la figure 2.4 un exemple est donnée.

Généralement, une matrice de similarité est utilisée lors du processus, elle contient les valeurs de similarités pour chaque modification qu'on peut effectuer sur une chaîne. Il arrive parfois d'utiliser la notion de dissimilitude au lieu de similarité, dans ce cas une matrice de distance est plus appropriée. Il est également possible d'avoir recours à des fonctions de distance au lieu de la matrice de distance. Par conséquent, une fonction est nécessaire pour chaque opération de modification.

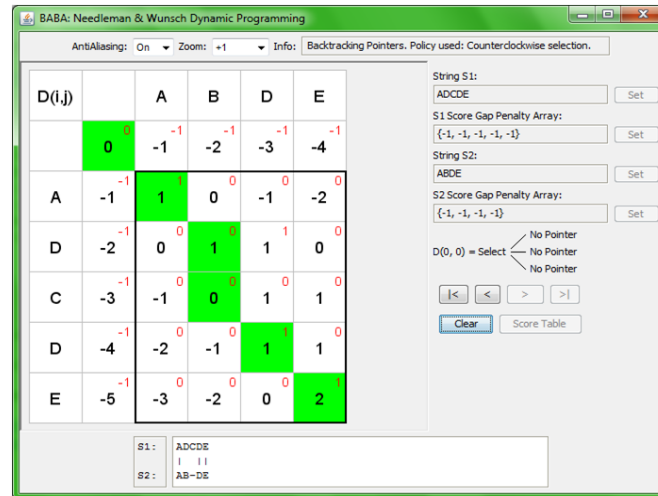


Fig. 2.3: Exemple d'alignement global utilisant l'algorithme de Needleman-Wunsch [1].

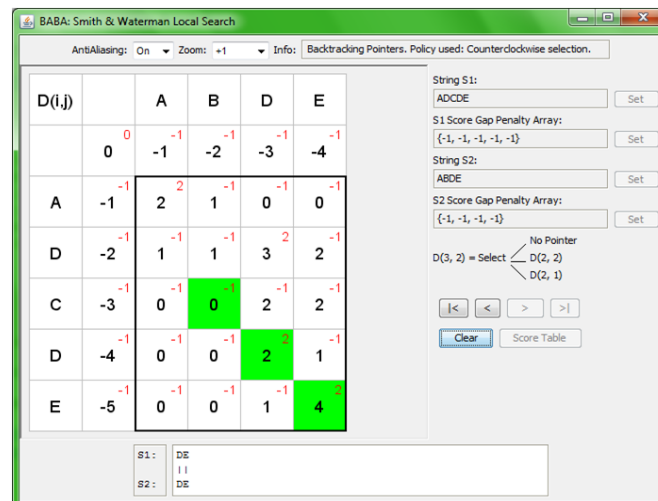


Fig. 2.4: Exemple d'alignement local utilisant l'algorithme de Smith-Waterman [1].

### 3 Méthodologie

L'observation de la figure 1.4 met en évidence une caractéristique des faisceaux de fibres que l'on pourrait exploiter pour faciliter la sélection. En effet, il n'est pas difficile de remarquer que la géométrie des fibres composant le faisceau sont similaires, du moins en plusieurs parties. On pourrait alors exploiter cette similarité pour extraire les faisceaux de fibres. Un outil comparant les fibres entre elles permettrait de concevoir une méthode de sélection basée sur les caractéristiques géométriques d'une fibre. L'utilisateur pourrait alors choisir une des fibres comprises dans le faisceau qui l'intéresse, et l'outil pourrait sélectionner les fibres dont la similarité est supérieure à un certain seuil choisi.

L'objectif de ce projet de recherche est de concevoir une méthode de comparaison de courbes 3D adaptée au problème de sélection mentionné ci-haut. Nous proposons une méthode pouvant respecter ou non les propositions suivantes sur les fibres: translation invariant, orientation invariant et taille invariant. L'approche présentée est fortement inspirée des travaux de Tieng *et al.* [6] et Mallat [4]. Tieng *et al* proposent une technique en vision numérique pour classifier des contours d'objets présents dans une image à partir d'une banque de contours d'objets connus. Les contours mesurés dans l'image sont bruités et déformés par la caméra. Ils traitent les contours comme des courbes 3D fermées et appliquent d'abord une transformation géométrique, puis une transformée en ondelettes et la représentation *Zero-Crossing*. Un algorithme d'alignement de séquences pour retrouver la classe d'une courbe donnée sera ensuite appliqué. Mallat, dans son article [4], présente le fondement et les propriétés de la représentation *Zero-Crossing*. Il introduit notamment l'idée que les zéros de la transformée en ondelettes, avec une ondelette bien choisie, correspondent aux positions des variations abruptes du signal. Ces deux références sont à la base de notre implémentation d'une méthode de segmentation des fibres de matière blanche.

### 3.1 Notre approche

Les sous sections suivantes décrivent les différentes techniques choisies pour traiter et comparer les fibres. La première étape sera de transformer les coordonnées cartésiennes 3D des fibres en une représentation simplifiée deux fois 1D et facilitant le respect des propositions mentionnées ci-haut. La seconde étape sera d'appliquer une transformée en ondelettes dyadiques pour extraire les détails des fibres. On présente par la suite une représentation alternative à la transformée en ondelettes simplifiant le calcul de la mesure de similarité des fibres. Finalement, le calcul de similarité entre deux fibres sera présenté.

#### 3.1.1 Représentation alternative

Les fibres sont composées d'un vecteur de points positionné dans un volume 3D. Afin de faciliter la comparaison, on propose d'appliquer une transformation de la courbe 3D en deux courbes 1D [6]. La première étape est de trouver le barycentre (centre de masse) de la courbe. Puisque les points représentant les fibres sont tous équidistants et qu'ils ont une masse égale, le barycentre est par conséquent le point moyen en X, Y et Z. On applique ensuite une translation afin de déplacer le barycentre de la fibre à l'origine. La fibre est, après cette étape, translation invariant.

La seconde étape consiste à calculer l'orientation relative des axes d'inerties principaux passant par le barycentre de la fibre. Le but est d'aligner les axes principaux de la fibre avec les axes en coordonnées monde. Ensuite, les axes d'inerties les plus importants sont respectivement alignés sur les axes Z, Y et X en coordonnée monde. Cette opération rendra la comparaison des fibres orientation invariant.

La fibre est par la suite projetée sur les axes X et Y. Ensuite, la distance de chaque point par rapport à l'origine (0,0,0) est calculée. Ce vecteur de mesure constitue la première représentation 1D,  $R_1$ . On fait de même pour  $R_2$ , mais cette fois on projette la fibre seulement sur l'axe Z.  $R_1$  et  $R_2$  à eux seuls ne sont pas suffisant pour reconstruire la fibre originale. Nous avons également besoin du vecteur  $\alpha$ , qui correspond aux angles respectifs de chaque point par rapport au centre. Ce vecteur sera mis de côté par la suite, quoique essentiel à la reconstruction de la fibre, il contient peu d'information sur la structure géométrique de la fibre [6].

$$\begin{aligned} R_1(t) &= \sqrt{x(t)^2 + y(t)^2} \\ R_2(t) &= z(t) \\ \alpha(t) &= \arctan\left(\frac{y(t)}{x(t)}\right) \end{aligned}$$

où  $t$  est l'indice des points de la fibre. Il est possible dès cette étape de gérer la taille invariant des fibres. Il suffit de normaliser  $R_1$  et  $R_2$  avant d'effectuer la transformée en ondelettes.

### 3.1.2 Représentation compacte

La transformée en ondelettes dyadiques fournie 512 valeurs sur neuf échelles. Ce vecteur de valeurs est long à comparer. Tieng *et al* [6] proposent une représentation numérique alternative et équivalente à la représentation *Zero-Crossing* présentée précédemment. Il s'agit d'utiliser une séquence de binômes où chaque élément de la séquence caractérise un zéro. Le premier terme du binôme  $\mu$  est l'amplitude de la courbe entre deux zéros, alors que le second terme  $\rho$  est la distance entre deux zéros de la courbe. Les binômes sont numériquement sauvegardés dans des nombres complexes. Les nombres complexes ont l'avantage de simplifier le calcul de distance entre deux éléments. Tieng *et al* [6] proposent de lier les extrémités de la courbe en ondes carrés. Contrairement à eux, nous traitons des courbes ouvertes, il est alors plus logique de maintenir les extrémités indépendantes.

$$\begin{aligned}\xi_j^{xy} &= \mu_j^{xy}(r) + \imath \cdot \rho_j^{xy}(r) & r = 1, \dots, R_j^{xy} \\ \xi_j^z &= \mu_j^z(r) + \imath \cdot \rho_j^z(r) & r = 1, \dots, R_j^z\end{aligned}$$

où  $R_j^{xy}$  et  $R_j^z$  représentent le nombre de changement de signe à l'échelle  $2^j$  en ce qui concerne la représentation  $R_1$  et  $R_2$  respectivement.

### 3.1.3 Calcul de distance

Pour comparer deux fibres entre elles, nous utilisons des algorithmes d'alignement de chaînes combinés à la représentation *Zero-Crossing* compacte de chaque fibre. Comme mentionné plus haut, cette représentation peut être vu comme une séquence de nombres complexes. Puisque les composantes des séquences sont des nombres plutôt que des symboles, on peut en tirer avantage et s'en servir pour créer des fonctions de distance qui utilisent ces informations numériques supplémentaires.

Afin d'être en mesure de comprendre les formules de cette sous-section, expliquons d'abord certaines notations utilisées.  $[R_j^{xy}]_f$  et  $[R_j^{xy}]_g$  représentent le nombre de zéros de la représentation *Zero-Crossing* de  $R_1$  à l'échelle  $2^j$  en ce qui concerne la fibre  $f$  et  $g$  respectivement. Également,  $\xi_j^{xy}(r)$  est la valeur complexe, partie réelle identifiant l'amplitude alors que celle imaginaire représente la distance entre deux zéros, à l'indice  $r$  de la représentation compacte (i.e. *Zero-Crossing*) à l'échelle  $2^j$ . Enfin, le facteur d'échelle, utilisé pour compenser la variance de la taille que peut avoir la représentation compacte, est  $\Gamma = \frac{\sum_{n=1}^N R_1 f(n)}{\sum_{n=1}^N R_1 g(n)}$ , où  $N$  est le nombre de points.

### Insertion

La valeur calculée pour l'insertion représente la norme euclidienne de la composante à insérer divisée par la composante moyenne de la chaîne d'où provient cette composante. Autrement dit, c'est la

longueur du vecteur associée à la composante qu'on désire ajouter, le tout normalisé par rapport aux autres composantes de la chaîne.

$$\gamma(\lambda \rightarrow [\xi_j^{xy}(r+m)]_g) = \frac{\|[\xi_j^{xy}(r+m)]_g\|}{\frac{1}{[R_j^{xy}]_g} \sum_{q=1}^{[R_j^{xy}]_g} \|[\xi_j^{xy}(q)]_g\|}$$

### Suppression

La valeur calculée pour la suppression est sensiblement identique à celle pour l'insertion, sauf que utilise la composante à retirer au lieu de celle à insérer.

$$\gamma([\xi_j^{xy}(r)]_f \rightarrow \lambda) = \frac{\|[\xi_j^{xy}(r)]_f\|}{\frac{1}{[R_j^{xy}]_f} \sum_{q=1}^{[R_j^{xy}]_f} \|[\xi_j^{xy}(q)]_f\|}$$

### Substitution

La distance entre deux composantes est représentée par la norme euclidienne de la différence de ces composantes fois un facteur d'échelle, le tout divisé par un facteur de normalisation. En d'autre mots, c'est la longueur Euclidienne pondérée du vecteur associée à la différence entre les composantes à substituer, le tout normalisé.

$$\gamma([\xi_j^{xy}(r)]_f \rightarrow [\xi_j^{xy}(r+m)]_g) = \frac{\Gamma \|[\xi_j^{xy}(r)]_f - [\xi_j^{xy}(r+m)]_g\|}{\sqrt{\Gamma \|[\xi_j^{xy}(r)]_f\| \|[\xi_j^{xy}(r+m)]_g\|}}$$

#### 3.1.4 Segmentation

Dans leur article, Tieng *et al* [6] se concentrent sur la classification plutôt que la segmentation. Par conséquent, ils comparent chaque courbe à la courbe modèle représentant chaque classe. En ce qui nous concerne, il n'y a pas de fibre modèle, on cherchera donc à regrouper les fibres en segments (*clusters*). Ainsi, les fibres les plus similaires vont être dans un même groupe, tandis que les fibres les plus éloignées seront dans deux groupes différents.

Le regroupement est effectué à l'aide de l'algorithme de segmentation hiérarchique, plus particulièrement le critère pour fusionner deux groupes est le *Complete Linkage*, c'est-à-dire que la distance

entre deux segments est représentée par la distance entre les deux fibres, provenant de groupes différents, les plus éloignées.



## 3.2 Implémentation

Cette section décrit l'avancement du projet. Le projet est inspiré des travaux de Tieng *et al* [6] et de Mallat [4]. Tieng *et al* proposent une technique en vision numérique pour classifier des contours d'objets présents dans une image à partir d'une banque de contours d'objets connus. Les contours mesurés dans l'image sont bruités et déformés par la caméra. Ils traitent les contours comme des courbes 3D fermées et appliquent d'abord une transformation géométrique, puis une transformée en ondelettes et la représentation *Zero-Crossing*. Un algorithme d'alignement de séquences pour retrouver la classe d'une courbe donnée sera ensuite appliqué. Mallat, dans son article [4], présente le fondement et les propriétés de la représentation *Zero-Crossing*. Il introduit notamment l'idée que les zéros de la transformée en ondelettes, avec une ondelette bien choisie, correspondent aux positions des variations abruptes du signal. Ces deux références sont à la base de notre implémentation d'une méthode de segmentation des fibres de matière blanche.

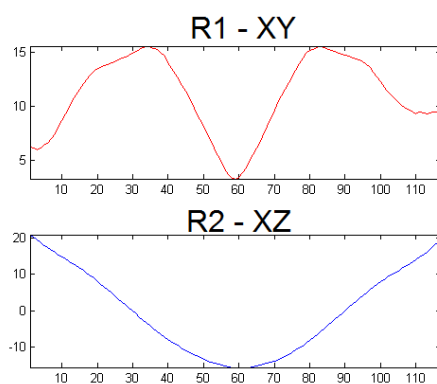
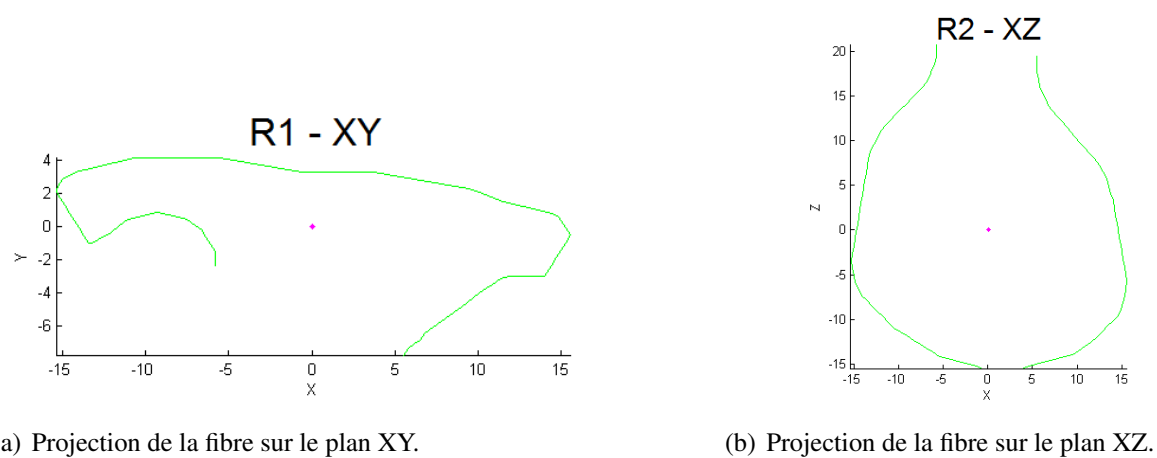
Les fibres utilisées pour tester notre approche proviennent d'un cerveau anonyme qui a passé une IMRd. Les données recueillies ont alors été traitées afin d'obtenir une représentation des fibres en utilisant des algorithmes de suivi de fibres (tractographie).

### 3.2.1 Translation et orientation invariant

Tout d'abord, chaque fibre doit subir une transformation afin de la rendre invariante à la translation et à l'orientation. Cela permettra le regroupement de fibres similaires, mais qui ne sont pas positionnées au même endroit dans le cerveau, ni orientées de la même façon. Nous avons également pensé que la modification de la fibre, pour la rendre invariante à l'orientation, devrait être effectuée au besoin. En effet, on peut penser, par exemple, à segmenter seulement une région des fibres du corps calleux qui se retrouvent à plusieurs endroits dans le cerveau mais avec des orientations différentes. Il en va de même pour l'invariance à la translation.

### 3.2.2 Représentation alternative

Ensuite, les représentations  $R_1$  et  $R_2$  sont calculées pour chaque fibre. Comme mentionné précédemment, la représentation  $\alpha$  est mise de côté, puisque peu d'information autre que l'angle où se situe chaque point de la fibre  $y$  est contenu. Pour notre implémentation, nous avons opté pour ces deux représentations proposées par Tieng *et al* [6], mais d'autres pourraient être envisagées. La figure 3.1 illustre les représentations  $R_1$  et  $R_2$  d'une fibre quelconque.



(c) Représentation R1 et R2 d'une fibre.

Fig. 3.1: Exemple de représentation alternative d'une fibre.

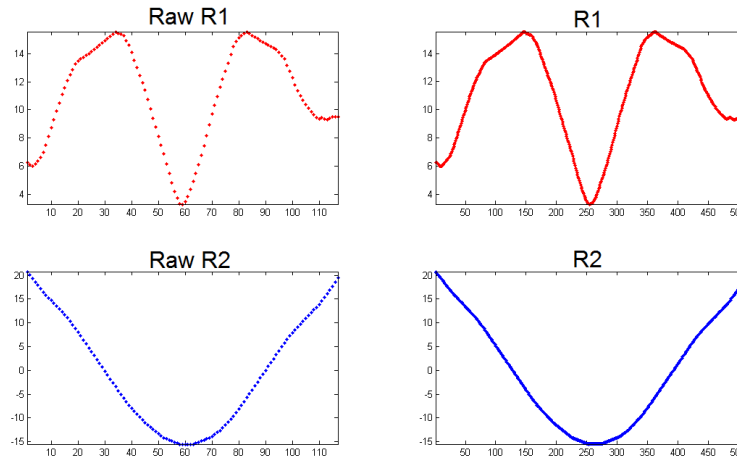


Fig. 3.2: Exemple d'échantillonnage d'une fibre utilisant une interpolation spline cubique.

### 3.2.3 Ré-échantillonnage

Puisque les fibres n'ont pas toutes le même nombre de points, nous effectuons un ré-échantillonnage, à l'aide d'une interpolation spline cubique, afin d'obtenir 512 points. Cela permet également d'utiliser une transformée ondelettes dyadiques pour faire ressortir les caractéristiques d'une fibre. D'autres méthodes d'interpolation peuvent être envisagées, par exemple l'interpolation linéaire. Voici un exemple tiré de notre implémentation à la figure 3.2.

### 3.2.4 Taille invariant

Si désiré, il est possible d'effectuer un changement d'échelle, permettant ainsi de comparer des fibres de tailles différentes.

### 3.2.5 Transformée en ondelettes

Par la suite, les transformées en ondelettes des deux représentations  $R_1$  et  $R_2$  sont calculées dans le but de faire ressortir les caractéristiques, en lien avec les discontinuités, de la fibre. Dans notre implémentation, l'ondelette utilisée est celle de la spline cubique. D'autres existent et pourraient être plus adaptées à ce problème, nommons seulement la *mexican hat* suggérée par Mallat. Pour une visualisation de la transformée en ondelettes d'une fibre, voir la figure 3.3.

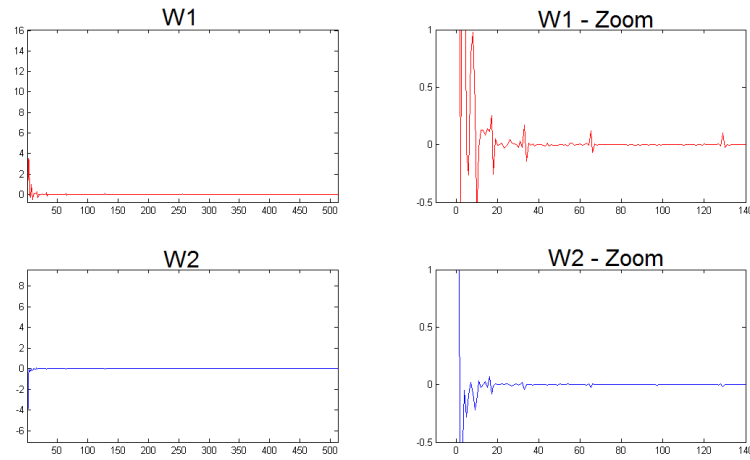


Fig. 3.3: Transformées en ondelettes d'une fibre.

### 3.2.6 Seuillage

Suite à cela, un seuillage dur est appliqué afin d'atténuer les faibles variations que peut comporter les transformées en ondelettes. Pour y parvenir, un seuil empirique de 0.01 a été choisi. Bien évidemment, cette valeur critique a un grand impact sur la représentation *Zero-Crossing*. Voir la figure 3.4.

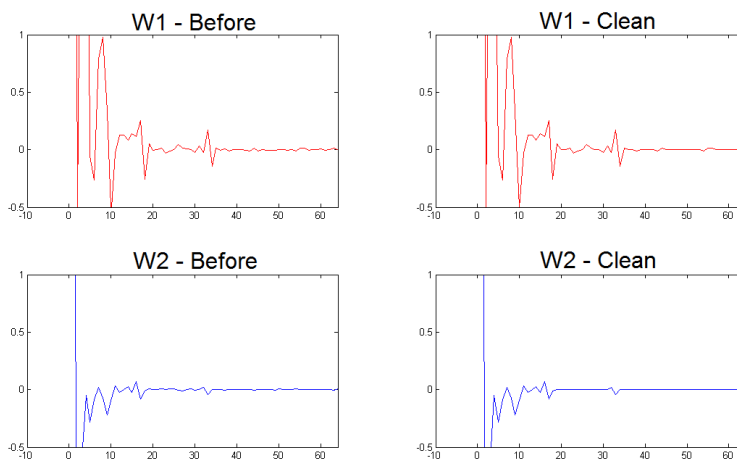


Fig. 3.4: Seuillage *hard* d'une fibre. Seuil utilisé est de 0.01

### 3.2.7 Zero-Crossing

Dernière étape du pré-traitement des fibres, on calcule la représentation *Zero-Crossing*, expliquée précédemment. Cela permettra de représenter une fibre en une série de séquences. Cette représentation est un aspect important de l'article de Tieng *et al* [6] dont on s'est fortement inspiré. Grâce à cette suite de séquences, il est possible de comparer deux fibres entre elles et par le fait même d'en ressortir une mesure de distance. Cette représentation est illustrée à la figure 3.5.

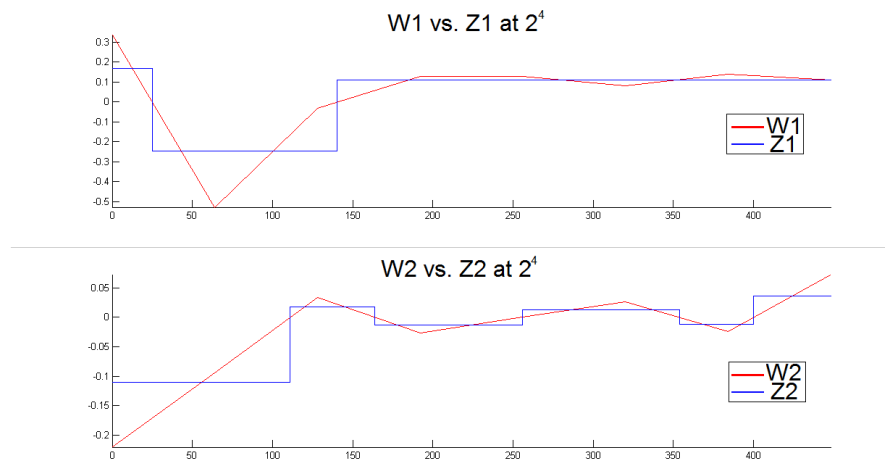


Fig. 3.5: Représentation *Zero-Crossing* d'une fibre à l'échelle  $2^3$ .

### 3.2.8 Alignement

La première étape pour la segmentation est l'alignement des fibres deux à deux. L'algorithme utilisé est celui de l'alignement global basé sur la programmation dynamique. De plus, trois fonctions de distance ont été adaptées pour ce type de comparaison tel que mentionné dans l'article de Tieng *et al* [6]. Nous avons décidé de les adopter puisque notre problème est similaire, en certain point, au leur. Des fonctions de distance convenant à notre problème doit être envisagées ainsi que le choix de l'algorithme d'alignement. La figure 3.6 montre des distances obtenues entre des fibres.

### 3.2.9 Segmentation

Finalement, les fibres similaires sont regroupées à l'aide de la segmentation hiérarchique, tel que mentionné précédemment. Pour ce faire, cet algorithme nécessite une table de distances qui est calculée en effectuant l'alignement deux à deux de toutes les fibres. On termine par choisir le nombre de groupes désirés. Cette dernière façon de procéder n'assure pas d'obtenir les meilleurs segments, mais offre un

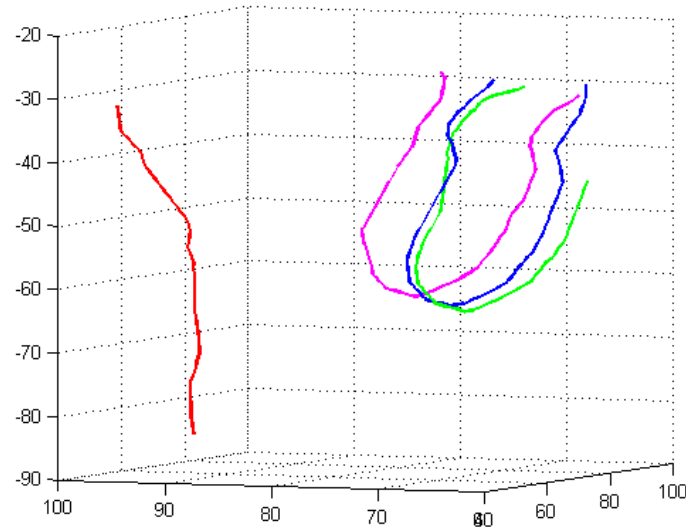


Fig. 3.6: Quatre fibres identifiées par leur couleur (**B**leue, **R**ouge, **M**auve et **V**erte).  $d(B, V) = 1.0698$ ,  $d(B, M) = 0.5554$  et  $d(B, R) = 1.8788$

premier aperçu. Cependant, il existe diverses méthodes pour calculer la compacité d'un groupe et la *sparsité* de l'ensemble des groupes.

## 4 Résultats

Les résultats de notre implémentation sont encourageants. En regardant la figure 4.1, on peut y voir plusieurs groupes distincts. Parmi les groupes les plus "purs", on voit clairement que les fibres y appartenant possèdent plusieurs caractéristiques semblables. Par exemple, à la figure ?? on peut voir un tel segment. On remarque en effet que la majorité des fibres sont en formes de U. Quelque unes, par contre, n'ont pas cet aspect. Cela est dû au fait qu'aucune mesure de pureté n'a été utilisée. Elle permettrait de déterminer si un groupe pourrait devenir plus homogène en le subdivisant ou en le joignant avec un autre. Donc la segmentation pourrait éventuellement être améliorée.

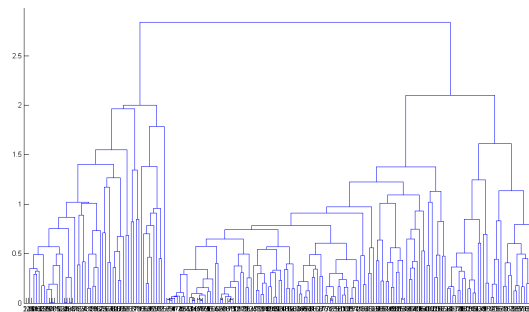


Fig. 4.1: Dendrogramme représentant la segmentation hiérarchique *Complete Linkage* de 200 fibres choisies aléatoirement.

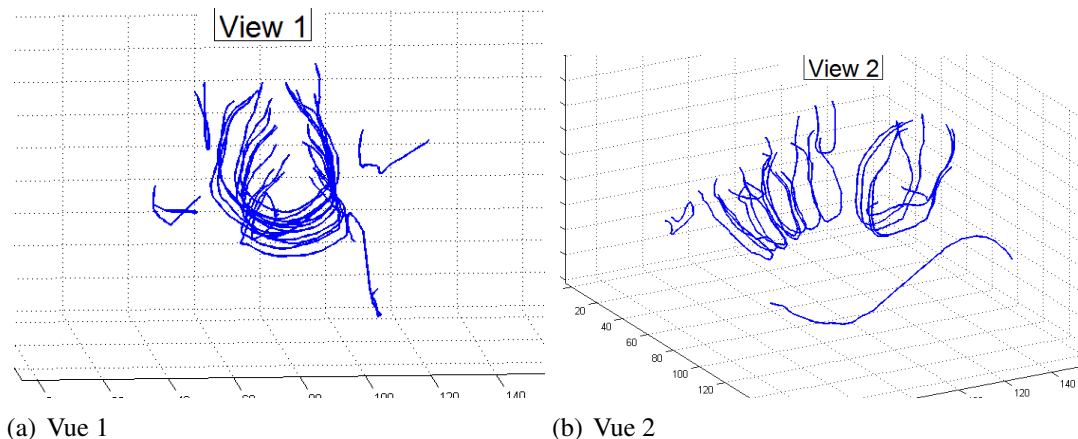


Fig. 4.2: Groupe de fibres qui possèdent des caractéristiques semblables.

Évidemment, il y a des groupes dont les fibres n'ont aucun point commun. La figure 4.3 illustre bien ce fait. Encore une fois, une bonne mesure de pureté pourrait améliorer la séparation et la fusion de groupes mal formés ou incomplets.

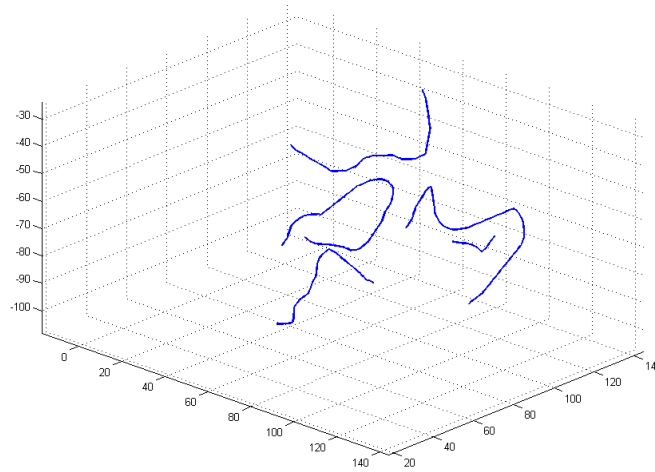


Fig. 4.3: Groupes de fibres qui n'ont aucune caractéristiques communes.

Bien entendu, plusieurs paramètres du processus de segmentation de fibres ont besoin d'être déterminés empiriquement, ce qui complexifie les choses. Le choix de la représentation alternative, l'ondelette, et les fonctions de distance sont quelques exemples qui affectent grandement les résultats. Ces paramètres restent à être ajustés à notre problématique.

## 4.1 Améliorations

Suite à notre implémentation, quelques améliorations (corrections) nous paraissent importants pour l'avancement du projet.

Tout d'abord, la transformation nécessaire pour rendre les fibres orientation invariant n'est pas au point. Des techniques doivent être explorées pour arriver au résultat escompté.

En second lieu, la représentation alternative, proposée par Tieng *et al* [6], semble convenir à notre problème, mais nous aimerions en essayer une autre. Plutôt que d'utiliser la distance avec le barycentre de chaque point dans le plan XY, nous nous servirions de la projection de la fibre sur trois plans, soit : XY, XZ et YZ. De la redondance d'informations sera ajoutée par les trois représentations ( $R_1$ ,  $R_2$  et  $R_3$ ). Nous pensons que cette redondance pourra faciliter la comparaison.

Troisièmement, nous savons que la distance entre deux points consécutifs d'une fibre est constant, donc il serait intéressant d'utiliser cette connaissance pour améliorer la représentation alternative ainsi que le ré-échantillonnage. De plus, concernant l'interpolation que nous effectuons présentement, nous avons quelque doute sur le fait que la taille des fibres est altérée par cette modification. Par conséquent, une étude approfondie de cette partie sera nécessaire.

Ensuite, en ce qui concerne l'ondelette à utiliser, nous n'avons pas encore décidé laquelle prendre.



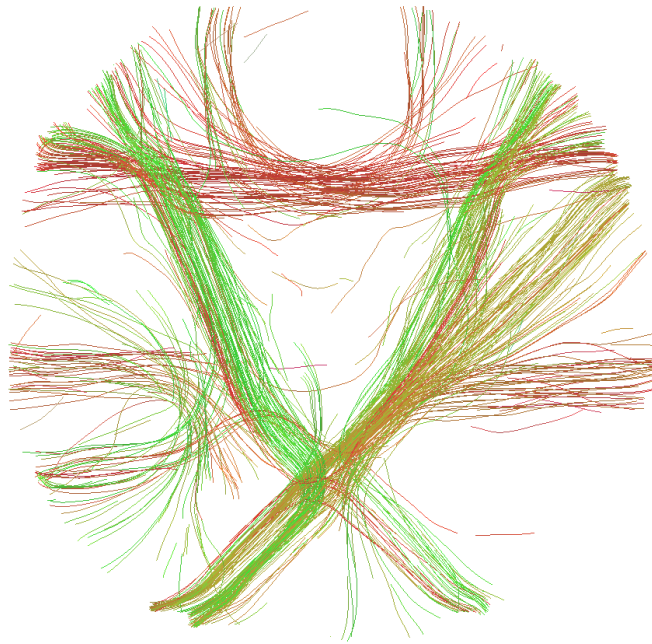


Fig. 4.4: Réseau de fibres artificielles. On y remarque la présence de faisceaux.

Puisque qu'on s'intéresse à l'allure générale de la fibre, les ondelettes possédant un grand support seront favorisées. Par contre, il faut tout de même être en mesure de bien déterminer les changements brusques des fibres et donc le support ne devra pas être trop grand. Il va de soi, que nous essayerons celle proposée par Tieng *et al* [6] et celle de Mallat [4], soit : la dérivée d'une spline cubique et la *mexican hat*. Également, il serait intéressant d'en regarder d'autres, notamment celles qui ne sont pas dyadiques, dans le but de trouver celle qui caractérisera le mieux une fibre de matière blanche.

Par la suite, au sujet de la comparaison des fibres, plus particulièrement les fonctions de distance, nous devons les adapter pour qu'elles reflètent mieux la similitude que possèdent les fibres appartenant à un même faisceaux. Pour y arriver, nous avons l'intention de nous servir d'un cerveau artificiel, voir la figure 4.4 pour un exemple. Ce *phantom* a été obtenu en faisant passer une IRM à des objets représentant le comportement des fibres dans le cerveau [5]. De plus, l'utilisation de cette vérité terrain nous aidera à bien déterminer la valeur des paramètres empiriques.

Également, toujours en ce qui a trait à la segmentation des fibres, la nouvelle approche d'alignement de séquences développée par Kelil *et al* [3], nommée *Aligner*, est définitivement une option à envisager. L'algorithme est une méthode hybride entre l'alignement global et local. De plus, il gère l'alignement multiple qui peut s'avérer utile pour regrouper les fibres d'un même faisceau.

Finalement, comme mentionné plus haut, une mesure de pureté adéquate pour faire de l'auto-segmentation hiérarchique de fibres devra être ajoutée afin de faciliter le choix des groupes.

## 5 Conclusion

En conclusion, l'approche présentée dans ce rapport a été conçue pour la classification de contours bruités en trois dimensions. De notre côté, pour la segmentation de fibres de matières blanches, nous avons repris les concepts généraux en ce qui concerne le traitement de courbes 3D destiné à la comparaison deux à deux. Les résultats préliminaires montrent que cette méthode peut être utile à notre problème. Par conséquent, nous améliorerons cette méthode pour qu'elle s'adapte mieux à la segmentation de fibres.

Cet automne, nous examinerons en détail les points énoncés dans la Section 4.1.

### 5.1 Ouverture

L'étude des fibres de matière blanche *in vivo* est un domaine de recherche récent. L'imagerie par résonance magnétique de diffusion repousse sans cesse les limites de la visualisation du cerveau et de ses connectivités. Ces données sont peu utilisées en clinique, mais renferment pourtant une quantité d'informations à ne pas négliger sur l'état de santé du cerveau d'un patient.

Dans ce projet de recherche, nous tentons une technique de caractérisation des fibres de matière blanche pour les comparer entre elles. La caractérisation des fibres est un élément clé dans l'étude approfondie des fibres pour un cerveau donné, mais également pour l'étude d'une population de cerveaux. Il est trivial d'imaginer de nombreuses techniques permettant d'extraire des informations additionnelles d'un cerveau à partir de la structure de sa matière blanche. Il serait possible notamment de faire des statistiques sur les proportions de fibres appartenant aux faisceaux principaux et de signaler toute anomalie.

De plus, il serait possible de détecter de nouveaux faisceaux suite à une coupure ou un déplacement de la matière blanche (e.g. causé par une tumeur). Les algorithmes de tractographie calculant la structure de matière blanche sont encore rudimentaire. On pourrait imaginer combiner un algorithme de tractographie à celui de l'analyse d'une population de cerveaux afin d'améliorer le suivi de fibres en éliminant les fibres impossibles (*outliers*).

Bref, la caractérisation des fibres de matière blanche permet d'informatiser le traitement *a posteriori* des fibres, ouvrant ainsi la porte à une utilisation plus grande de cette connaissance.

## References

- [1] Norman Casagrande. Basic-algorithms-of-bioinformatics applet, May 2010. <http://baba.sourceforge.net>.
- [2] N. C. Jones and P. A. Pevzner. *An Introduction to Bioinformatics Algorithms (Computational Molecular Biology)*. Massachusetts Institute of Technology.
- [3] A. Kelil, S. Wang, and R. Brzezinski. Aligner, August 2010. <http://prospectus.usherbrooke.ca/aligner>.
- [4] S. Mallat. Zero-crossings of a wavelet transform. *IEEE Trans. Inform. Theory*, 37(14):1019–1033, 1991.
- [5] C. Poupon, B. Rieul, I. Kezele, M. Perrin, F. Poupon, and J. Mangin. New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (hardi) models. *Magn. Reson. Med.*, 60(6):1276–1283, December 2008.
- [6] Q.M. Tieng and W.W. Boles. Space curve representation and recognition based on wavelet transform zero-crossings. *Journal of Mathematical Imaging and Vision*, 13(1):5–16, August 2000.
- [7] D. Wassermann, M. Descoteaux, and R. Deriche. Diffusion Maps Clustering for Magnetic Resonance Q-Ball Imaging Segmentation. *International Journal on Biomedical Imaging, special issue on Recent Advances in NeuroImaging Methodology*, 2008, 2008.