

Projet d'initiation a la recherche: Segmentation automatique des ventricules cerebraux a partir d'images IRM

Mathieu Chalvidal
Etudiant CentraleSupélec
Msc Mathématiques appliqués, Msc Intelligence Artificielle
`mathieu.chalvidal@student.ecp.fr`

April 3, 2019



Sujet

Dans le cadre de sa veille technologique et de sa diversification industrielle, Dassault Systemes developpe un portefeuille de technologies logicielles autour de problematiques biomédicales et en s'appuyant sur son expertise en modélisation 3D.

Dans ce contexte et dans le but, par exemple, de guider un robot neuro-chirurgical, il peut être utile d'automatiser le processus qui, sur une image IRM, identifie des caractéristiques morphologiques du crane, notamment les ventricules cerebraux (en jaune ci-dessous).

On peut imaginer des algorithmes basés sur, par exemple, des approches de modélisation géométrique de type snake, de la templatisation 3D de la forme du ventricule, de matching itératif...

1 Présentation du probleme

La segmentation d'image est une opération de traitement d'images qui a pour but de rassembler des pixels entre eux suivant des critères pré-définis. Les pixels sont ainsi regroupés en régions, qui constituent un pavage ou une partition de l'image.

Dans le contexte de l'imagerie médicale, l'étape de segmentation définit la délimitation des tissus observés, voire la dénomination des structures anatomiques qu'elles composent, on parle alors de segmentation *sémantique*.

La segmentation des IRM cérébrales est une étape importante pour diverses applications médicales, cliniques et neuroscientifiques. Elle est complexe pour différentes raisons: faible résolution des images, présence d'artefacts lors de la reconstruction des signaux magnétiques, grandes variations entre individus.etc.

Ce travail a pour but de dégager des méthodes de segmentation autonomes, indépendantes d'un a priori médical, par des techniques d'apprentissage ou de traitement statistique du signal. En particulier, notre travail se portera essentiellement sur la segmentation des ventricules cérébraux humains. Pour ce faire, nous explorerons diverses formulations.

1.1 Fonctionnement IRM

Nous commençons par rappeler le principe de l'imagerie par résonance magnétique. Notons, qu'un excellent résumé des techniques de pré-traitement des images obtenues par résonance magnétique ainsi que certaines méthodes de segmentation est disponible en [1].

1.1.1 La résonance magnétique

L'imagerie par résonance magnétique (IRM) est une technique d'imagerie médicale non-invasive. Elle a de plus l'avantage de ne pas irradier le sujet, c'est pourquoi c'est une méthode de choix dans le domaine médical. Le principe de l'IRM utilise les propriétés de résonance magnétique nucléaire (RMN) des noyaux d'atomes contenus dans le corps du patient. En appliquant plusieurs champs magnétiques non-uniformes, on peut observer des différences dans les mouvements de précession de ces noyaux et en déduire une image en deux ou trois dimensions décrivant la composition des tissus dans le corps étudié. Plus spécifiquement, à chaque pixel (en deux dimensions) ou voxel (en trois dimensions) est associée une valeur entre 0 et 255, correspondant à une intensité en niveau de gris, correspondant à la moyenne de la résonance magnétique des atomes contenus à cet endroit. Typiquement, on obtient une résolution d'environ 2mm avec un scanner standard. Chaque acquisition prend une durée d'environ vingt minutes pour un cerveau d'adulte.

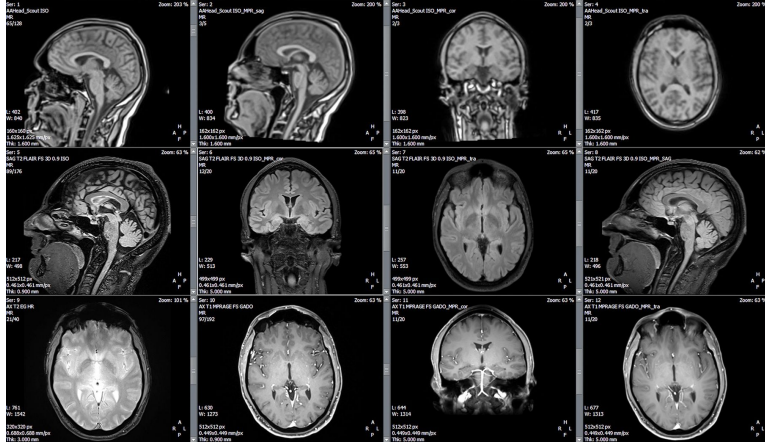


Figure 2: Planche d'une image cérébrale 3D acquise par résonance magnétique

1.2 Formulation du problème

Nous définissons ici le problème de segmentation d'une image cérébrale. Une image cérébrale obtenue par IRM est un ensemble discret de valeurs réelles proportionnelles au temps de relaxation du tissu présent dans un sous-volume unitaire appelé voxel. Nous assimilons donc une image à un élément de l'espace hilbertien $x \in \mathbb{R}^{l \times L \times h}$ où l représente la longueur, L la largeur et h la hauteur de l'image dans les 3 dimensions spatiales.

La segmentation d'une telle image correspond à la labélisation de chacun des voxels qui composent cette image. Nous assimilons donc cette opération à une application de $\mathbf{R}^{l \times L \times h}$ ayant pour image $\mathbf{s} \in \mathbf{A}^{l \times L \times h}$ où \mathbf{A} représente l'ensemble des labels possibles pour un voxel. On considère généralement 2 niveaux de description lorsque l'on parle de segmentation:

- une segmentation dite de **bas niveau**, qui correspond à la description de la matière contenue dans chaque voxel. Dans le cadre d'IRM cérébrales, cela correspond généralement à la délimitation des 3 matières corticales: matières blanche, matière grise, liquide céphalo-rachidien.
- une segmentation dite de **haut niveau**, qui correspond à une description structurelle et fonctionnelle, où la segmentation porte sur des régions anatomiques connues de type thalamus, ventricules, etc...

Dans notre étude, nous nous intéresserons essentiellement à la segmentation des ventricules cérébraux, sous-structure centrale du cortex qui présente une symétrie axiale dans le plan coronal. Les deux ventricules présentent chez le sujet adulte une forme allongée. Ces ventricules correspondent à une structure physiologiquement très bien délimitée puisqu'il s'agit de cavités remplies de

liquide céphalo-rachidiens à l'intérieur de la masse de matière blanche du cortex. Cela explique dès lors, le choix d'utiliser ces structures dans le domaine du recalage d'imagerie cérébrales. Leur géométrie est simple et très bien délimitée bien qu'elles présentent une relative variabilité inter-individuelle.

1.2.1 Notations mathématiques

Dans la suite, une image obtenue par résonnance magnétique est assimilée à un élément $x \in R^{l \times L \times h}$ et sa segmentation est $s \in A^{l \times L \times h}$ avec A est de cardinal fini, soit binaire: appartenance, non-appartenance), soit ternaire (matière grise, blanche ou liquide cr.)

Nous introduisons les notations usuelles pour manipuler la théorie des Probabilités. \mathbf{X} est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ Pour les techniques d'apprentissage, nous noterons $\theta \in \Theta$ les paramètres appris par le modèle.

1.2.2 Hypothèses et choix de méthodologie

Afin d'explorer un maximum de techniques de segmentation qui constituent le coeur de problème proposé ici, nous faisons l'hypothèse sur la qualité des données à disposition (exemple: prétraitement des images IRM déjà effectué, champs magnétique uniforme de l'IRM, encodage des IRM au format normé nifty, absence d'artefacts dans l'acquisition des images, etc...).

Par ailleurs, les techniques de segmentation apparaissent en première lecture, très peu robustes aux pathologies dégénératives qui influent sur la structure corticale (Type maladie d'Alzheimer) nous ajoutons donc l'hypothèse d'image acquise sur un sujet sain.

Concernant les données utilisées au cours de cette étude, nous utiliserons les ressources du NITRC : Neuro-Imaging Tools Ressources Collaboratory. Cette institution américaine regroupe des datasets public et privés. Nous utiliserons en particulier les images du Dataset IXI [2] issue de l'acquisition de 600 images de sujets sains, recueillies par le *Biomedical Image Analysis Group* à Imperial College de Londres. Le protocole d'acquisition consiste en l'acquisition pour chaque sujet de:

- Temps de relaxation T1
- Temps de relaxation T2
- PD-weighted images

Nous utiliserons par souci d'homogénéité des images utilisées, les acquisitions de l'Hopital Hammersmith à Londres utilisant un Modèle d'imageur Philips de puissance 3 Teslas.

2 Bibliographie des techniques existantes

Cette section liste et discute diverses formulations permettant d'effectuer une segmentation structurelle sans introduire d'a priori médical. Nous nous intéresserons spécifiquement à leur performance temporelle, la précision et les choix des hypothèses.

La littérature distingue essentiellement deux familles de modèles selon l'information structurelle qu'ils essaient de capturer.[3],[4] et [5]

- Les **méthodes de segmentation par contours** forment une première famille de modèles qui visent à délimiter les objets présents dans l'image selon leurs contours. Ces approches se basent essentiellement sur l'analyse des variations d'intensité dans l'image. Par définition, les contours sont les frontières où cette intensité est la plus forte.
- Les **méthodes de segmentation par régions** forment une seconde famille de modèles qui visent à délimiter les objets présents dans l'image selon des régions présentant des similitudes pour des critères donnés (texture, intensité...).

Dans les 2 cas, ces méthodes reposent souvent sur l'optimisation d'une fonction de coût formulée pour adhérer au mieux à l'instance présentée tout en gardant des propriétés de régularité ou d'optimalité intéressantes. Souvent la fonction de coût est formulée dans une approche dite compétitive: c'est à dire que plusieurs termes contraires sont en compétition dans une fonction d'aggrégation pour obtenir une solution. Le plus souvent, un terme de fidélité aux données (distance euclidienne des données et estimation) est associé à des contraintes de régularité (norme l_2 , norme de variation totale $\nabla_f \dots$) ou de sparsité (norme l_1). De façon intéressante, cette méthodologie est à retrouver dans des formulations très différentes du problème (modèles graphiques, estimation bayésienne, Deep learning... etc)

2.1 Méthodes de segmentation par région

2.1.1 Méthodes de clustering

Ces méthodes font l'hypothèse d'une homogénéité de l'intensité lumineuse des régions à segmenter ensemble. Cette hypothèse peut être critiquable à la fois du fait de artefacts liés à l'acquisition de données (biais de champ pour l'IRM) et du fait de la structure même de certaines régions. Toutefois, l'homogénéité que présentent les ventricules justifient d'essayer ces approches. Une autre limite évidente de ces modèles est la non-utilisation des données spatiales de l'image. En effet, pour ces algorithmes, l'image est traitée comme un nuage de point de \mathbf{R} . Logiquement, ces méthodes ne pourront dégager de solutions précises, notamment au niveau du contour des ventricules, néanmoins leur rapidité de calcul en font des outils de pré-traitement utiles, notamment pour une localisation plus précise dans l'IRM.

K-Moyennes

L'idée est d'utiliser l'algorithme des K-Moyennes [6] pour définir des centroïdes d'intensité lumineuse dans l'image. L'IRM est ensuite "seuillé" pour attribuer une appartenance des points à l'image. Dans notre cas, la spécification du nombre de classe répond à la présence des trois grandes classes de tissu dans l'IRM cérébral. (MB,MG,LCR)

De manière générale, l'algorithme des K-Moyennes vise à minimiser l'inertie d'un nuage de point vis-à-vis de la plus proche des moyennes de chaque point. Le critère optimisé est donc:

$$\mathbb{E}(\mu) = \frac{1}{2} \sum_{i=1}^N \min_{\mu_j \in C} (\|x_i - \mu_j\|^2) \quad (1)$$

Chaque μ_j est mis-à-jour à chaque itération de l'algorithme. La mise à jour des centroïdes peut être interprété comme en [7] comme un algorithme caché de descente du gradient ou l'itération, selon une séquence $(\lambda_t)_{t \in \mathbb{N}}$ t.q $\lambda_{t+1} < \lambda_t$, $\sum_{\mathbb{N}} \lambda_t = \infty$ et $\sum_{\mathbb{N}} \lambda_t^2 < \infty$ est :

$$\lambda_t \nabla \mu = \lambda_t \left(\sum_{i=1}^N (x_i - \mu_j) \right)_{j \in C} \quad (2)$$

Notons que la fonction \mathbb{E} est non-convexe mais différentiable par réécriture, ainsi le résultat de convergence pour l'opérateur α -averaged de la mise à jour du gradient conduit à la création d'une séquence de centroïdes convergeants vers un minimum local. Dans la pratique, il s'agira d'initialiser de différentes façons l'algorithme pour explorer ces différents minimum possibles. Dans le cas unidimensionnel de notre problème, la longueur et la forme des clusters d'intensité n'est pas essentiellement pertinente. Ce qui explique que pour un IRM de bonne qualité (pour le contraste et l'absence de biais de champ) les clusters retrouvés sont visuellement satisfaisants.

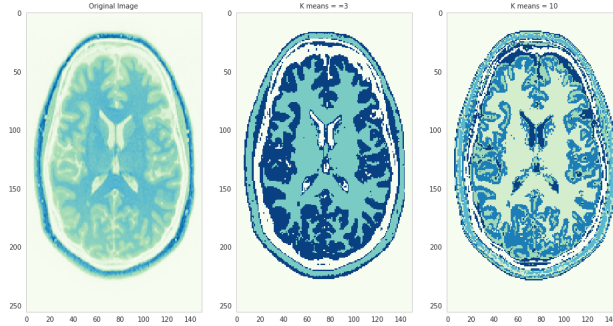


Figure 3: Coupe d'un cerveau et labélisation par K-Moyennes pour K=3 et 10

Mixture de Gaussiennes

Le clustering par mélange de gaussienne [8] introduit une formulation probabiliste pour la segmentation des intensités lumineuses de l'IRM. Cette formulation algorithmique fait l'hypothèse que chaque cluster d'intensité suit une loi normale (ici univariée de moyenne et variance inconnue). La formulation cherche à trouver la distribution des lois de chaque cluster qui rend compte au mieux de l'observation suivant un critère donné. L'avantage immédiat de cette formulation sur les K-moyennes est sa flexibilité vis-à-vis de la variance intra-cluster. En effet, pour des grandes zones tissulaires comme la matière grise encerclant les ventricules et qui couvre une importante part de l'image, la variance peut-être plus grande du fait de changement locaux du champ magnétique, de changements de densité des tissus, etc... Le mélange de gaussienne capture cette variabilité avec plus de succès que les K-Moyennes et présente donc une première amélioration. En revanche, la spatialité des points n'est pas prise en compte dans ce second modèle.

D'un point de vue algorithmique, le modèle de mélange de gaussiennes est un cas particulier de l'algorithme *Expectation-Maximization* (EM) [9]. Cet algorithme itératif permet, suivant certaines conditions d'obtenir un maximum local de la vraisemblance des données. La description mathématique de l'algorithme est donnée ci-après:

Formulation mathématique

Considérons un échantillon $\mathbf{x} = (x_1, \dots, x_n)$ suivant une loi de probabilité $f(x, \theta)$ paramétrée par $\theta \in \Theta$. L'algorithme EM cherche à déterminer le

paramètre θ maximisant la log-vraisemblance $\mathbf{L}(\mathbf{x}, \theta) = \sum_{i=1}^n \log(f(x_i, \theta))$. La loi de distribution directe étant complexe, il est souvent difficile d'optimiser cette fonction de vraisemblance. Il est alors utile de d'introduire une loi a priori $\mathbf{z} = (z_1, \dots, z_n)$ de loi connu et de densité de probabilité conditionnelle $f(z|x, \theta)$. On peut alors définir la log-vraisemblance de la loi jointe de (\mathbf{x}, \mathbf{z}) comme:

$$\mathbf{L}((\mathbf{x}, \mathbf{z}), \theta) = \sum_{i=1}^n \log(f(z_i|x_i, \theta)) + \sum_{i=1}^n \log(f(x_i, \theta)) \quad (3)$$

ou encore

$$\mathbf{L}(\mathbf{x}, \theta) = \mathbf{L}((\mathbf{x}, \mathbf{z}), \theta) - \sum_{i=1}^n \log(f(z_i|x_i, \theta)) \quad (4)$$

L'algorithme EM est basé sur une procédure itérative qui met à jour le paramètre courant noté $\theta^{(c)}$. On peut alors écrire l'espérance conditionnelle suivante:

$$\mathbb{E}_z[\mathbf{L}(\mathbf{x}, \theta)|\theta^{(c)}] = \mathbb{E}_z[\mathbf{L}((\mathbf{x}, \mathbf{z}), \theta)|\theta^{(c)}] - \mathbb{E}_z[\sum_{i=1}^n \log(f(z_i|x_i, \theta))|\theta^{(c)}] \quad (5)$$

ou encore, comme $\mathbf{L}((\mathbf{x}, \mathbf{z}), \theta)$ ne depend pas de \mathbf{z} :

$$\mathbf{L}((\mathbf{x}, \mathbf{z}), \theta) = \mathbb{Q}(\theta, \theta^{(c)}) + \mathbb{H}(\theta, \theta^{(c)}) \quad (6)$$

- $\mathbb{Q}(\theta, \theta^{(c)}) = \mathbb{E}_z[\sum_{i=1}^n \log(f(z_i|x_i, \theta))|\theta^{(c)}]$
- $\mathbb{H}(\theta, \theta^{(c)}) = \mathbb{E}_z[\sum_{i=1}^n \log(f(x_i, \theta))|\theta^{(c)}]$

On montre alors que la suite $\theta^{(c+1)} = \arg \max_{\theta} \mathbb{Q}(\theta, \theta^{(c)})$ converge vers un maximum local de la vraisemblance car $\arg \max_{\theta} \mathbb{Q}(\theta, \theta^{(c)}) \iff \arg \max_{\theta} \mathbb{E}_z[\mathbf{L}(\mathbf{x}, \theta)|\theta^{(c)}]$

L'algorithme consiste alors, jusqu'à convergence selon un critère donné des 2 étapes:

- Evaluation de l'espérance (par tirage aléatoire selon la loi de \mathbf{z}) $\mathbb{H}(\theta, \theta^{(c)})$
- Estimation par maximum de vraisemblance de $\theta^{(c+1)} = \arg \max_{\theta} \mathbb{Q}(\theta, \theta^{(c)})$

En pratique, on pourra faire tourner plusieurs fois l'algorithme avec des initialisations différentes de manière à obtenir de plus grandes chances d'obtenir un maximum global.

L'algorithme est initialisé avec les résultats de l'algorithme des K-moyennes précédant pour accélérer la convergence. Nous obtenons les résultats présentés dans la figure suivante.

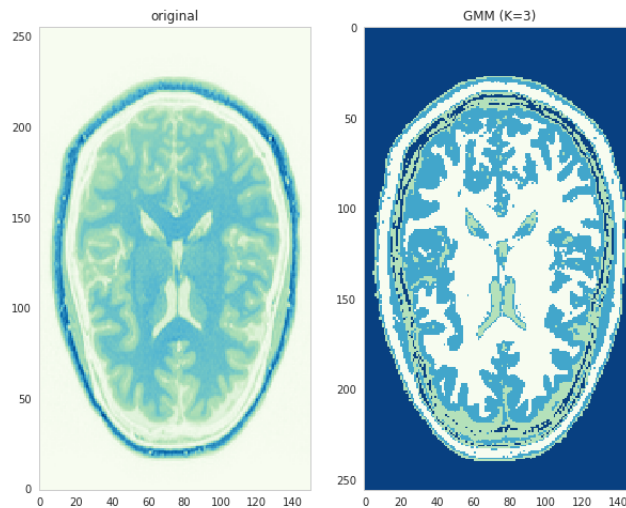


Figure 4: Coupe d'un cerveau et labélisation par GMM pour $K=3$

2.1.2 Modèles graphiques

L'utilisation des techniques de clustering ont permis de construire des premières approximations intéressants pour localiser d'un point de vue macroscopique les ventricules dans l'IRM. L'idée suivante est d'utiliser cette information pour construire un modèle plus sophistiqué dans la zone d'intérêt estimée.

L'objectif étant de délimiter précisément les contours des ventricules et leur forme dans l'espace, le nouveau modèle proposé doit prendre en compte la structure spatiale de l'image.

Les modèles graphiques sont une famille de modèles probabilistes largement utilisées dans les problèmes d'apprentissage et notamment en vision automatique pour leur capacité à capturer les relations spatiales à l'intérieur d'une image.[3] Ces modèles s'appuient sur une formulation graphique de l'image où chaque noeud est une variable aléatoire et les arrêtes du graphe constituent des relations de dépendance entre ces variables. Une sous-catégorie de ces modèles regroupe les champs aléatoires markoviens dont la principale propriété est de définir un graphe non-orienté acyclique. L'utilisation de ces modèles dans le contexte d'apprentissage va consister en l'optimisation d'une fonction de coût (aussi appelée énergie) définie comme la vraisemblance paramétrée de ce graphe vis-à-vis de l'image. [10] [11].

On note \mathbf{X}_i une variable aléatoire associée au pixel i représentant le label assigné au pixel i . Cette variable aléatoire prend ses valeurs dans un ensemble discret fini de valeurs $\{L_i\}_{i \in I}$. Nous notons \mathbf{X}_1 à \mathbf{X}_N les variable aléatoires ou N est le nombre de pixels dans l'image. Soit un graphe (V, E) où $V = \{\mathbf{X}_1, \dots, \mathbf{X}_N\}$ et une observation globale I . la paire (I, \mathbf{X}) peut être modélisée comme une champ aléatoire conditionnel par une distribution de Gibbs de la forme $\mathbb{P}(X = x|I) = \frac{\exp(E(x|I))}{Z(I)}$. $E(x|I)$ est appelée l'énergie de Gibbs de la configuration x et $Z(I)$ est la fonction de partition $\prod_i E(x_i|I)$. Dans la suite nous pouvons ignorer le conditionnement en I par souci de simplicité de lecture.

Nous définissons alors un graphe spécifique au cadre de la segmentation binaire des ventricules cérébraux. Soit le graphe (V, E) suivant,

- $V = \{\mathbf{X}_1, \dots, \mathbf{X}_N\}$
- $E = \{(\mathbf{X}_i, \mathbf{X}_j) \text{ t.q } i, j \text{ ont un relation de voisinage dans l'image.}\}$ Il est possible de définir plusieurs relations de voisinage pour une image tridimensionnelle comme la notre. Voir figure

Nous définissons ensuite l'énergie du graphe comme:

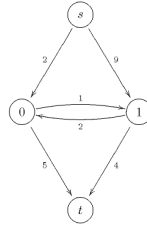
$$\mathfrak{E}(\mathbf{X}) = \sum_{X_i \in v} \Phi_u(X_i|I) + \lambda \left(\sum_{(i,j) \in E} \Psi_p(X_i, X_j) \right) \quad (7)$$

Cette fonctionnelle apparaît à nouveau comme un somme de termes compétitifs pour estimer V . Nous explicitons ci-dessous les termes de cette formule.

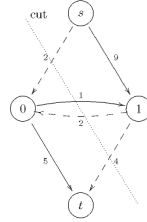
- $\Phi_u(X_i|I)$ représente le potentiel unitaire de la variable X_i prenant la valeur x_i . Il s'agit de l'opposée de la log-vraisemblance $-\mathbb{L}(X_i|I)$. Ce terme encourage ainsi une labelisation des pixels dans la classe la plus vraisemblable selon l'image I observée, afin d'obtenir la log-vraisemblance $-\mathbb{L}(X_i|I)$. Il s'agira d'estimer la densité de probabilité conditionnelle $\mathbb{P}(X_i = x_i|I), \forall i \in \text{Label}$. Cette densité est estimée par la méthode des estimateurs à noyaux présenté dans la section prétraitement pour chacune des zones $\{\text{arrière-plan, avant-plan}\}$ de l'image.
- $\Psi_p(X_i, X_j)$ représente les potentiels énergétiques des paires (X_i, X_j) . Ce potentiel peut être vu comme le coût d'associer simultanément (X_i, X_j) à (x_i, x_j) . Ce dernier pénalise des assignation de labels différents entre deux pixels voisins et favorise donc une estimation lissée de V .
- λ est un terme de régularisation qui permet de réguler la prépondérance de chacun des 2 termes dans l'énergie globale. Ce terme peut être choisi de manière optimale pour un critère de reconstruction donné.

Le problème final est donc le suivant:

$$\mathbf{X}_{MRF} = \underset{\mathbf{X} \in V}{\operatorname{argmin}} \mathfrak{E}(\mathbf{X}) \quad (8)$$



(a) Graphe orienté source-sink



(b) s-t cut du graphe

Figure 5: Graph orienté et s-t cut

Dans une configuration générale, ce problème est un problème de complexité NP-hard. Toutefois, la formulation idiomatique du graphe (problème de segmentation binaire) permet de reformuler le problème comme un problème de flot maximal dans le graphe (V,E) augmenté de deux noeuds représentant une appartenance à l'arrière-plan ou à l'avant-plan de l'image à segmenter. Cette reformulation tient au théorème du flot maximal:

Theoreme du flot maximal[12]

Pour tout graphe orienté G , tout couple (s,t) de sommets, et pour tout vecteur de capacités positives, la valeur maximale du flot de s à t est égale à la capacité d'une coupe minimale séparant s de t .

Ce Théorème est également interprétable comme un résultat de forte dualité pour le problème convexe défini plus haut. [13] Il existe dans la littérature un certain nombre de solveurs pour la recherche d'un coupe de capacité minimale dans un graphe que nous mettrons en oeuvre par l'intermédiaire de librairie **Pymaxflow**. [14]

Puisque notre image est tridimensionnelle, il est possible de proposer 2 formulations différentes:

- Formuler le problème par "strates" dans l'axe longitudinal de l'IRM. Cette formulation simplifie la complexité du problème puisque qu'une seule tranche de l'image est considérée. Toutefois, la simplification fait perdre l'information spatiale entre les couches et peut mener à des solutions aberr-

rantes le long de l'axe longitudinal. De plus l'estimation est bruitée par l'ensemble des zones présentant du liquide céphalo-rachidien à l'intérieur de l'image réduite utilisée pour la segmentation.

- Formuler le problème par dans sa globalité en augmentant la profondeur du graphe pour chaque tranche de l'image IRM. Ce problème est plus complexe à résoudre mais permet de garantir une solution physiquement réaliste le long des trois axes de l'image et présentant une meilleure robustesse aux zones extérieures aux ventricules de faible densité tissulaire.

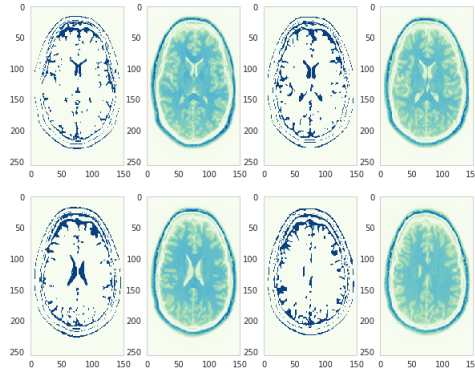
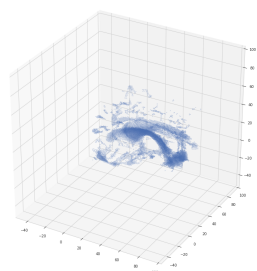
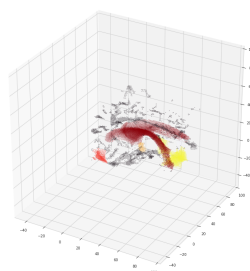


Figure 6: Coupe transversales aux altitudes $z=125, 128, 131$ et 134 et segmentations associées

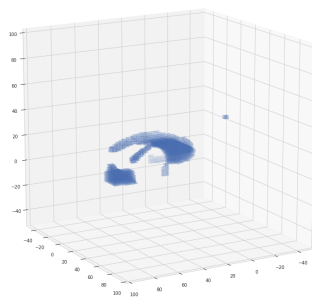
Les clusters de points encodant la position des ventricules semblent assez homogènes et denses pour opérer une étape de post-traitement par clustering afin de les isoler du reste. Pour cela, considérant la forme hautement anisotrope des clusters et la densité raisonnable des points qui le composent, l'utilisation d'un algorithme de type hiérarchique comme **hdbscan** [15] permet l'extraction des clusters d'intérêt.



(a) Visualisation 3D de la segmentation par couche de CRF. Notons la présence de nombreuses zones d'erreur à l'extérieur des ventricules.



(b) Labelisation des clusters issus du CRF par Hdbscan



(c) Segmentation obtenue par champs aléatoire conditionnel profond

2.2 Méthodes de segmentation par contours

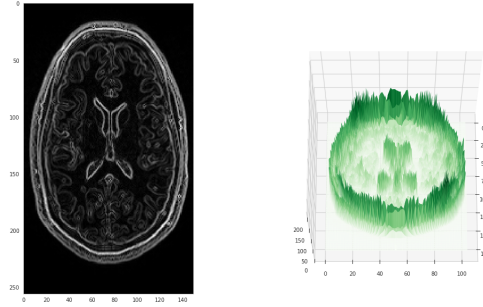
Nous présentons ici deux méthodes construisant une segmentation par contours. Ces méthodes se basent sur les gradients d'intensité lumineuse de l'image et sont donc plus robustes aux variations internes des zones à segmenter. Certaines méthodes nécessitent l'application de filtres différentiels afin de caractériser des points, lignes, ou surfaces singulières. Des opérateurs différentiels plus complexes permettent aussi de calculer localement des lignes de crête sur des surfaces d'iso-intensité. Ces approches n'intègrent aucune connaissance sémantique ni topologique et ne sont pas adaptées en présence de bruit ou de contraste faible. Toutefois, elles interviennent comme prétraitement dans de nombreuses approches.

2.2.1 Watershed - Ligne de partage des eaux

La segmentation par ligne de partage des eaux (*watershed algorithm* en Anglais) [16] désigne une famille de méthodes de segmentation d'image issues de la morphologie mathématique qui considèrent une image à niveaux de gris comme un relief topographique, dont on simule l'inondation.

Les algorithmes de calcul de la ligne de partage des eaux les plus efficaces ont une complexité qui dépend linéairement du nombre de pixels de l'image, ce qui les classe parmi les méthodes de segmentation les plus rapides. Ceci les rend particulièrement attrayants dans des applications traitant des images de grande taille, telles que les images à trois dimensions.

L'application de cette algorithmes n'est pas directement réalisé sur l'image brut mais sur une image prétraitée, souvent l'image par un opérateur de filtrage qui permet de mettre en valeur les contours (filtre de Sobel= approximation discrete de la norme du gradient, seuillage par projection sur une base d'ondelettes, etc...) voire figure 8a.



(a) Filtre de Sobel appliqué à une tranche de l'IRM (b) Analogie de l'image à un relief topographique

2.2.2 Marching cubes

L'algorithme des marching cubes [17] est un algorithme relativement ancien en computer vision permettant la reconstruction d'une forme dans un champ à partir des valeurs d'isosurface de ce champ. L'idée de l'algorithme est de parcourir le champ scalaire, prenant huit points à la fois (définissant ainsi un cube imaginaire), et de déterminer les polygones à créer pour représenter une partie de l'isosurface contenue dans ce cube.

Ceci fonctionne en créant un index dans un tableau précalculé des 256 configurations de polygones possibles ($2^8 = 256$) dans un cube, en traitant chacune des 8 valeurs scalaires comme un bit dans un nombre entier de 8 bits. Si la valeur scalaire est supérieure à la valeur de l'isosurface (i.e., est à l'intérieur de la surface), alors le bit correspondant est mis à 1, sinon il est mis à 0. La valeur finale après le test des 8 points est l'index de la bonne configuration polygonale dans le tableau précalculé.

Finalement, chaque sommet des polygones créés est placé à sa position finale le long de l'arête du cube, en interpolant linéairement les deux valeurs scalaires connectés par cette arête.

Bien que facile d'implémentation et très rapide, l'hypothèse d'application de cet algorithme est relativement forte puisqu'il s'agit de reconstruire des surfaces d'iso-valeurs dans l'image. Or rien n'indique que les contours des ventricules présentent une telle propriété. Toutefois, l'application de la norme du gradient point à point va permettre en pratique une reconstruction relativement bonne des ventricules. Un second réglage implique le choix de l'iso-valeur recherché dans l'image. Ce choix peut être semi-automatisé par une normalisation des valeurs de l'opérateur de gradient, toutefois cette valeur est intrinsèquement

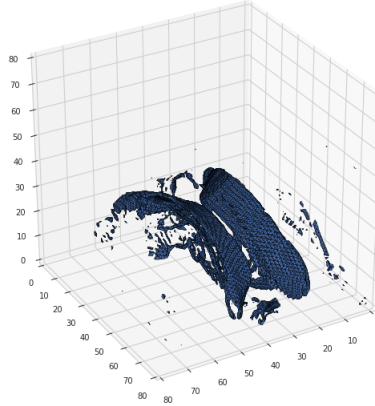


Figure 9: Segmentation obtenue par marching cubes sur la norme du gradient

liée aux propriétés de contraste qui peuvent varier d'un imageur IRM à l'autre.

Par ailleurs, le contour réalisé par l'algorithme nécessite une étape de post-traitement pour retirer les éventuelles zones extérieures aux ventricules ayant été segmentée comme l'indique la figure 9.

2.3 Méthodes d'apprentissage auto-supervisé

Une autre famille de méthode à fort potentiel est constituée par les méthodes de réseaux neuronaux de type convolutionnels. Ces méthodes ont récemment connu un fort développement du fait de leur performance de haut niveau dans des tâches de computer vision classique (détection de mouvement, recherche de similarité, segmentation sémantique). Ces méthodes sont toutefois fortement circonscrites par leur besoin massif de données d'entraînements, ce qui constitue un frein dans le contexte de non-supervision de l'apprentissage que nous avons retenu.

Self supervised-learning [18],[19],[20] et [21]

Pour contourner ce problème, nous reformulons le problème de segmentation classique par réseaux convolutionnels en un problème auto-supervisé. Le processus d'entraînement consiste en une première tâche classique de d'apprentissage supervisé où il est facile de générer un jeu de données exploitables. Cette tâche

dîte *prétexte* sera ici l'estimation d'un champ de déformation par un difféomorphisme $\mathbb{D} : R^{l \times L \times h} \longrightarrow R^{l \times L \times h}$ par l'intermédiaire d'un réseaux en sablier de neurones de type U-Net. Ce champ de déformation est assimilable dans le cas d'une vidéo au flot optique entre deux images. De fait, un réseaux correctement entraîné sera capable d'estimer le champ de déformation entre une image fixe (choisie comme étalon pour cette tâche) et une autre image IRM. Ce recalage est assimilable à une segmentation puisque l'estimation du champ de déformation donnera directement les coordonnées de points d'intérêt dans l'image IRM étudiée, comme par exemple la position des ventricules cérébraux.

Synthèse des données

Le jeu de données nécessaire est ici constructible synthétiquement. Il s'agit d'un ensemble $\mathfrak{D} = \{(f, \mathbb{D}_i(f))\}_{i \in I}$ où I est le cardinal du jeu de données et \mathbb{D}_i est un champ de déplacement difféomorphique $\mathbb{D} : R^3 \longrightarrow R^3$ de support fixe appliqué à l'image f . Il existe plusieurs manière caractériser \mathbb{D}_i . Nous pouvons par exemple définir $\mathbb{D} = I_d + \mathbf{u}$ où $\mathbf{u} : p = \mathbf{u}(p)$ représente un champ de déplacement. Les hypothèses portant sur cette transformation permettent de garantir l'homogénéité et la réalité physique des images IRM construites. Il est alors possible, afin de construire un dataset réaliste, de simuler des transformations qui représentent la variabilité des caractéristiques anatomiques observés chez le sujet humain. (volume de la boîte crânienne, dimensions des structures corticales...)

Formulation du problème

La tâche de recalage dite de *deformable registration* est ici encore formulable comme un problème d'optimisation:

$$\hat{\Phi} = \underset{\mathbb{D}}{\operatorname{argmin}} \mathcal{L}(f, \Phi \circ (\mathbb{D} \circ f)) \quad (9)$$

Ou encore, dans le cas d'un modèle paramétré comme c'est le cas pour un réseaux convolutionnel:

$$\hat{D}_\theta = \underset{\Phi}{\operatorname{argmin}} \mathcal{L}(f, \Phi_\theta \circ (D \circ f)) \quad (10)$$

Afin de garantir des propriétés de continuité et reconstruire une image réaliste, la fonction de coût \mathcal{L} peut être reformulé comme la somme d'un terme qui mesure la similarité d'une image à l'autre \mathcal{L}_{simu} et un terme de régularité \mathcal{L}_{smooth} pondéré par une paramètre de régularisation λ .

$$\hat{D}_\theta = \underset{\theta \in \Theta}{\operatorname{argmin}} \mathcal{L}_{simu}(f, \Phi_\theta \circ (D \circ f)) + \lambda \mathcal{L}_{smooth}(\Phi_\theta) \quad (11)$$

Pour notre expérimentation, nous retenons respectivement:

- la mean-squared error portant sur l'intensité lumineuses des deux images pour le terme de fidélité aux données $\mathcal{L}_{simu}(f, \Phi_\theta \circ (D \circ f)) =$

3 Conclusion

Nous présentons dans ce dossier plusieurs familles d’algorithmes pour la segmentation d’image cérébrales en trois dimensions de type IRM. Volontairement, nous avons choisi de traiter ce problème dans un cadre non-supervisé et en écartant l’introduction d’une expertise médicale de type atlas de forme etc... Ce choix est justifié par une volonté de proposer des méthodes robustes et neutres d’un point de vue de l’expertise médicale. Plusieurs méthodes nous ont conduit à des résultats de segmentation en 3D permettant une reconstruction visuellement satisfaisante des ventricules à partir d’une analyse de l’image brute seule. Ces méthodes sont relativement sobres du point de vue des ressources computationnelles qu’elles consomment et peuvent s’intégrer sans problème dans un programme d’imagerie médicale plus complexe. Néanmoins afin de mieux qualifier ces méthodes, il est nécessaire d’établir une liste de critères ainsi qu’un score de segmentation. Plusieurs méthodes ont été proposées pour établir ce score. Nous pouvons citer par exemple de l’indice de DICE, la distance de Jaccard entre deux ensembles, la norme de variation totale pour qualifier la régularité de la segmentation proposées. Par ailleurs, les méthodes proposées ici ne permettent pas de quantifier directement l’incertitude liée à l’estimation, ce qui représente un frein à leur application thérapeutique. Les formulations bayésiennes et d’apprentissage auto-supervisé apportent une réelle valeur de ce point de vue.

References

- [1] Ivana Despotovic, Bart Goossens, and Wilfried Philips. MRI segmentation of the human brain: challenges, methods, and applications. *Computational and mathematical methods in medicine*, page 23, 2015.
- [2] IXI – Information eXtraction from Images. <https://brain-development.org/ixi-dataset/>. Biomedical Image Analysis Group, Imperial College London.
- [3] Benoît Scherrer. *Segmentation des tissus et structures sur les IRM cérébrales: agents markoviens locaux coopératifs et formulation bayésienne*. PhD thesis, Institut National Polytechnique de Grenoble - INPG, 2008.
- [4] Rafael C. Gonzalez and Richard E. Woods. *Image segmentation*, chapter 10. Prentice Hall, 2008.
- [5] Herve Delingette. Segmentation d’images médicales, 2008.
- [6] Stuart P. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28:129–137, 1982.
- [7] J. Bengio and L. Bottou. Convergence properties of the kmeans. *NIPS*, 2007.

- [8] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
- [9] N.M. Laird A.P. Dempster and Donald Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.
- [10] Shuai Zheng, Sadeep Jayasumana, Bernardino Romera-Paredes, Vibhav Vineet, Zhizhong Su, Dalong Du, Chang Huang, and Philip H. S. Torr. Conditional random fields as recurrent neural networks. *IEEE International Conference on Computer Vision (ICCV)*, 2015.
- [11] Florence Forbes. Modelling structured data with probabilistic graphical models. *Statistics for Astrophysics- Classification and Clustering*, 77:195–219, 2016.
- [12] Lester R. Ford Jr. and Delbert R. Fulkerson. Maximal flow through a network, 1956.
- [13] Nikos Komodakis and Jean-Christophe Pesquet. Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems. *IEEE Signal Processing Magazine*, 32(6), 2015.
- [14] Yuri Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 2004.
- [15] Leland McInnes, John Healy, and Steve Astels. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11):205, 2017.
- [16] Jos Roerdink and A Meijster. The watershed transform: Definitions, algorithms and parallelization strategies. *FUNDINF: Fundamenta Informatica*, 41, 2000.
- [17] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3d surface construction algorithm. *ACM Siggraph Computer Graphics*, 21(4):163–169, 1987.
- [18] Longlong Jing and Yingli Tian. Self-supervised visual feature learning with deep neural networks: A survey. Fellow, IEEE.
- [19] Mert R. Sabuncu John Guttag Adrian V. Dalca Guha Balakrishnan, Amy Zhao. Voxelmorph: A learning framework for deformable medical image registration. *IEEE TMI: Transactions on Medical Imaging*, 2019.
- [20] Siyuan Shan, Xiaoqing Guo, Wen Yan, Eric Chang, Yubo Fan, and Yan Xu. Unsupervised end-to-end learning for deformable medical image registration. 2017.

- [21] Christodoulidis Stergios, Sahasrabudhe Mihir, Vakalopoulou Maria, Chasagnon Guillaume, Revel Marie-Pierre, Mougiakakou Stavroula, and Paragios Nikos. Linear and deformable image registration with 3d convolutional neural networks. In *Image Analysis for Moving Organ, Breast, and Thoracic Images*, pages 13–22. Springer International Publishing, 2018.
- [22] P. Fischer O. Ronneberger and T. Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.