

2.1 Introduction

L'imagerie médicale joue un rôle très important dans l'observation chez l'humain. Elle permet en effet d'observer la structure et le comportement d'organes grâce à un ensemble important de techniques d'acquisition : imagerie par résonance magnétique (IRM), scanographie par rayon X, échographie, etc.

L'objectif de notre travail est la segmentation d'images IRM de cerveau. Nous allons présenter dans ce chapitre la problématique de la segmentation elle-même et enfin différentes méthodes de segmentation s'appliquant à cette problématique.

2.2 La segmentation

Classiquement, on peut définir la segmentation comme étant une partition de l'image I en un nombre N de sous-ensembles R_i appelés régions, homogènes dans le sens d'un ou plusieurs critères, comme le niveau de gris, tels que :

- La segmentation doit être complète (c'est-à-dire, chaque pixel doit être affecté à une classe).
- Les pixels appartenant à la même région doivent être connectés.
- Les régions doivent être disjointes.

En termes mathématiques :

$$\begin{cases} U_{i=1}^n = R_i = 1 \\ R_i \neq \emptyset \forall i = 1 \dots n \\ R_i \cap R_j = \emptyset \forall i, j \text{ avec } i \neq j \end{cases} \quad (2.1)$$

Où R_i est l'ensemble des régions formant l'image I .

Dans un contexte médical, les régions R_i correspondent aux différentes structures anatomiques constituant les différentes régions d'intérêt. La détermination automatique du nombre de régions N est une vraie problématique [19]. Souvent, la valeur N est supposée être connue comme étant une information à priori sur les structures anatomiques en investigation.

Il est difficile de définir d'une manière absolue, une bonne méthode de segmentation. Le choix d'une technique est lié à :

- La nature de l'image (éclairage, contours, texture, etc.)
- Aux opérations en aval de la segmentation (compression, reconnaissance des formes, mesures, etc.)
- Aux primitives à extraire (droites, régions, textures, etc.)
- Aux contraintes d'exploitation (temps réel, espace mémoire, etc.)

2.3 Segmentations d'images cérébrales

Dans cette partie nous introduisons le but de la segmentation d'images cérébrales et l'apport qui porte la segmentation automatique dans le domaine médical ainsi que les difficultés trouvées dans les images cérébrales.

Segmenter, ou délimiter les structures cérébrales est une étape fondamentale pour l'analyse d'images cérébrales. Elle permet de faire une séparation des différents tissus cérébraux (matière grise, matière blanche, liquide céphalorachidien, etc..) ainsi que d'éventuelles pathologies cérébrales. Une bonne segmentation permet d'aider le médecin à prendre une décision finale, avant son geste chirurgical.

De nombreuses études des tissus du cerveau dans les images d'IRM ont été effectuées et rapportées dans la littérature, telles que des techniques basées sur les réseaux neuronaux, la logique floue, des méthodes statistiques, coupe de graph [20] et bien d'autres, nous présenterons les principales méthodes de segmentation utilisées dans la suite du chapitre.

2.3.1 La segmentation automatique des IRM cérébrales

La segmentation manuelle des images IRM est extrêmement coûteuse en temps humain (des chercheurs dans une étude sur la schizophrénie ont segmenté 50 images par an et par chercheur travaillant à temps plein [21]) en plus l'intervention humaine reste une source potentielle d'erreurs : au cours de longues séances d'interaction, l'opérateur humain change son comportement, la fatigue le fera commettre des erreurs et le résultat manque de fiabilité et de robustesse.

En comparaison avec la segmentation manuelle qui est par essence une approche de contour, la segmentation semi-automatique, peut être une approche contour ou région. Sur le plan informatique, l'approche région est souvent plus facile que l'approche contour. En effet, il peut être difficile de fermer le contour et donc de quantifier la surface segmentée. Alors que la segmentation entièrement automatique est l'objectif extrême. Elle a de nombreux

avantages, en plus d'être indépendante de l'opérateur elle est rapide et reproductible. Comme dans le cas de la segmentation semi-automatique, les approches contours où région peuvent être envisagées pour la segmentation entièrement automatique.

Donc l'automatisation de la segmentation est nécessaire pour deux raisons :

- Le temps du tracé manuel par un expert sur un tissu spécifique est long,
- L'expertise humaine peut produire des erreurs difficilement contrôlables et reproductibles.

La segmentation des IRM cérébrales présente des particularités par rapport aux d'autres domaines d'applications de la segmentation comme la segmentation des cartes routières, ou la segmentation des visages, ceci est dû principalement aux raisons suivantes :

- L'objet à segmenter (l'anatomie du cerveau).
- Le processus d'acquisition IRM qui génère des artefacts (défauts).

2.4 Les différentes approches

La segmentation est un vaste sujet d'étude et fait partie des grands thèmes de l'imagerie numérique. Comment préférer l'une ou l'autre est un débat ouvert qui fait rage dans bien des laboratoires.

Nous allons donc présenter dans cette section diverses techniques connues de segmentation des images cérébrales en les organisant selon l'approche qui les régit. Ainsi, nous avons retenu deux grandes approches, la segmentation par contours, et par régions. Cette classification est représentée sur la figure 2.1.

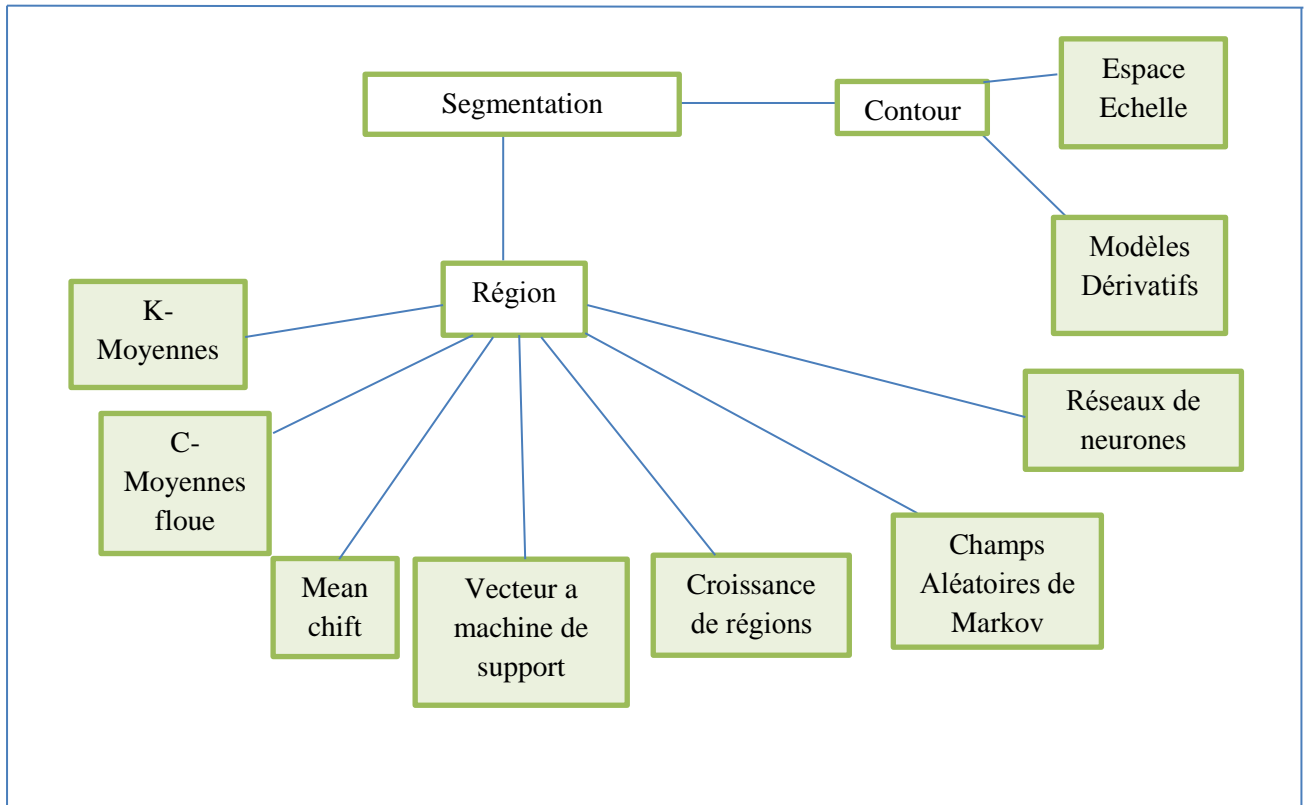


Figure 2.1 Classification des différentes méthodes

2.4.1 Région

L'approche région de la segmentation utilise des techniques d'identification et de localisation d'ensembles connexes de pixels. Ces méthodes ont pour but de partitionner l'image en plusieurs classes.

Par la suite, nous nous intéressons plus particulièrement aux méthodes à base de réseaux de neurones, K-means, C-means, puis aux méthodes par croissance de régions et bien d'autres.

2.4.1.1. Classification par K-Means (K-Moyens)

L'algorithme des k-means est utilisé pour classer les objets à partir de leurs propriétés en k parties (ou clusters) en supposant que les attributs forment un espace vectoriel. L'objectif est de minimiser la variance intra-cluster V :

$$V = \sum_{i=1}^K \sum_{x_j \in S_i} |x_j - \mu_i|^2 \quad (2.2)$$

Où S_i , $i = 1, 2, \dots, k$ sont les k clusters et μ_i est le point moyen des points $x_j \in S_i$

L'algorithme de K-means se résume dans les étapes suivantes :

1. Partitionnement des données en k sous-ensembles (non vides)
2. Calcul des centres des groupes de la partition courante
3. Les données sont affectées au groupe dont leur centre est plus proche
4. Retour à l'étape 2
5. Arrêt lorsque les groupes sont constants.

Dans la littérature Lachman et coll. [22] ont proposé d'utiliser les statistiques de concurrence dans l'algorithme des k-means pour segmenter les tissus cérébraux. Une version adaptative de cet algorithme a été utilisée par Yan et coll. [23] pour segmenter le cerveau avec des résultats plus au moins satisfaisants.

En 2002 Bustillo [24] la variation de volume des noyaux caudés chez des patients atteints de schizophrénie.

Cet algorithme est parmi les algorithmes de segmentations les plus populaires parce qu'il est extrêmement rapide en pratique. En termes de performance, cet algorithme ne garantit pas un optimum global. La qualité de la solution dépend grandement des ensembles initiaux et peut, en pratique, être bien en deçà de l'optimum global. Comme l'algorithme est très rapide, une méthode courante est de le lancer plusieurs fois et de retourner la meilleure partition [20] en plus l'intervention d'un expert est généralement nécessaire pour fixer le nombre K de classes.

2.4.1.2. C-moyennes floues

Dunn [25] a introduit l'algorithme c-moyennes floues (fuzzy C-means ou FCM) qui, est un algorithme qui repose sur le même principe que le K-Means mais qui a recours à la logique des ensembles flous. Bezdek s'y est intéressé [26] et a développé cet algorithme. Dans le cas des c-moyennes floues, la fonctionnelle L à minimiser est :

$$L = J_w(x, v; y) = \sum_{j=1}^n \sum_{i=1}^k (x_{ij})^w \|y_j - v_i\|_A^2 \quad (2.3)$$

Où n est le nombre de points à traiter, k le nombre de classes désirées, $w \in [1; +\infty[$ est le poids de fuzzyfication, $V = (V_1; ..V_k)$ est le vecteur des centres de classes, $\|.\|_A$ est un produit scalaire où A est une matrice définie positive et $x = [x_{ij}] \in R^{k \times n}$.

L'algorithme suivant montre les étapes de FCM :

1. initialiser les centres.
2. fixer le coefficient flou.
3. calcul la partition floue initiale (La matrice d'appartenance)

4. répéter :
 - Calcul des nouveaux centres
 - Calcul de la nouvelle partition floue

Tant que le critère d'arrêt n'est pas vérifié.

En [27] Bensaid et al. Proposent un algorithme semi-automatique permettant d'introduire la notion d'ensemble d'appartenance. L'apprentissage peut être de taille variable et son influence sur la classification est plus ou moins importante.

Les algorithmes FCM souffrent du manque de prise en compte d'informations contextuelles et fournissent des classifications moins efficaces en présence de bruit. Pour corriger ce défaut, une segmentation floue Markovienne est proposée dans [28] et appliquée dans [29] à la segmentation des principaux tissus et d'éventuelles tumeurs sur des volumes pondérés en T1.

Dans [30], Zouaoui a appliqué une architecture de coopération entre le FCM et le PCM (C-means possibilistes) pour la segmentation d'images IRM cérébrales.

Alors que Dernièrement Yang et coll. [31] ont proposé des FCM plus robustes (moins sensible au bruit) pour la segmentation d'IRM cérébrales.

Cependant, le grand avantage de ces méthodes par rapport aux classifications pures est de graduer l'appartenance des points aux différentes classes.

2.4.1.3 Mean Shift

L'algorithme du Mean Shift, introduit par Fukunaga [32] puis remis au got du jour par Comaniciu [33], recherche le " mode " ou point de plus haute densité d'une distribution de données. Une méthode élégante pour localiser les maxima locaux d'une fonction de densité est la recherche des zéros de son gradient.

Le gradient de l'estimation non paramétrique par le noyau K_H est :

$$\hat{\nabla} f(x) = \nabla \hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \nabla K_H(x - x_i)$$

2.4.1.4 Machine à Vecteurs de Support

Le principe des machines à vecteurs de support (en anglais support vector machine ou svm) [34] est simple : nous allons déplacer un problème complexe, souvent non-linéaire, dans un espace où le problème est linéaire ou, tout du moins, plus simple. Ainsi, nous avons projeté les données par une transformation ϕ dans un espace de dimension supérieure et calculer un séparateur dans cet espace de Hilbert séparable.

Ce séparateur est appelé hyperplan et les points de caractéristiques les plus proches de celui-ci définissent des plans appelés vecteur de support. Pour obtenir une segmentation robuste, il faut maximiser la marge, i.e. la distance entre l'hyperplan et les vecteurs de support.

Ce principe de séparation de l'espace de caractéristiques s'applique très bien à la segmentation d'image [35]. Les machines à vecteurs de support sont utilisées pour segmenter les lésions de la matière blanche [36,37] mais aussi pour la classification des cerveaux [38]. Lee et Col. Proposent de les utiliser conjointement avec les champs de Markov pour créer les Support Vector Random Fields qui leur permettent de segmenter des tumeurs cérébrales [39].

2.4.1.5 Croissance de régions

Les algorithmes de croissance de régions sont fondés sur l'agrégation progressive des pixels à partir d'un point initial appelé « germe », selon un ou plusieurs critères. Les points voisins du germe sont successivement examinés ; s'ils vérifient les critères d'agrégation, ils sont inclus à la région et deviennent à leur tour un germe ; dans le cas contraire, ils sont rejetés.

Dans le cas de la segmentation d'images cérébrales, Les méthodes fondées sur la croissance de régions permettent aussi bien la segmentation de structures saines [40] que la segmentation de structures pathologiques [41].

Warfield et al. Présentent une méthode d'identification des tumeurs de la matière blanche. Cette méthode suppose la segmentation préalable du cortex, de la matière grise et de la matière blanche. En particulier, le cortex est segmenté par croissance de régions, guidée par des connaissances anatomiques a priori sur la forme de la structure, ainsi que par des contraintes d'intensité et de gradient [42]. Dans [43], Thiran et al. proposent une méthode de croissance de régions pour la segmentation du cerveau sur des images IRM 3D. L'algorithme

contient une procédure de sélection automatique des germes. La croissance des régions est guidée, non pas par une information de type gradient, mais par les niveaux radiométriques de l'image. La segmentation obtenue permet la labellisation automatique de sulci et gyri.

Le ventricule latéral et le troisième ventricule sont segmentés par croissance de régions et la morphologie mathématique dans [40] par Schnack, et Hulshoff.

Dernièrement et Dans [41], la croissance de régions est directement utilisée pour la segmentation des lésions cérébrales. Le critère de croissance est basé sur une mesure de contraste. Le taux de classification a été d'environ 90%.

La segmentation par croissance de régions est simple à mettre en œuvre. L'avantage principal de cette technique se réside dans la prise en compte des informations contextuelles par le mécanisme de croissance. Alors que la disposition des germes initiaux représente un inconvénient. En effet, celle-ci nécessite quasi-systématiquement l'intervention d'un opérateur, en particulier lorsqu'il s'agit de détecter des tumeurs, qui, par définition, n'ont pas de localisation a priori dans le cerveau, et sur lesquelles nous ne disposons d'aucune connaissance anatomique [42].

2.4.1.6 Champs Aléatoires de Markov

La nature des images médicales, notamment l'aspect de texture aléatoire de celle-ci nous pousse à supposer qu'elles sont une réalisation d'un processus stochastique gouverné par des paramètres qui se calculent à partir de l'image en question [44]. C'est la règle des méthodes à base des champs aléatoires de Markov (MRF : « Markov Random Fields » en anglais), cet outil s'avère très efficace pour la segmentation d'image du fait qu'il permet d'exprimer l'interaction spatiale présentée par les pixels (voxels) (un pixel a une forte probabilité d'appartenir à la même classe de ses voisins). Dans la littérature Les premières techniques [45,46] ont utilisé le MRF sur des IRM de cerveaux avec des résultats prometteurs.

En 1997 Salzenstein [28] a introduit la théorie des ensembles flous dans les champs aléatoires de Markov et cette idée a été reprise avec quelque modification dans [47] pour améliorer la segmentation d'IRM cérébrales.

Le formalisme des champs de Markov permet aisément d'exprimer des contraintes de voisinage entre les pixels et ainsi d'améliorer la classification obtenue. Cette technique a aussi comme rôle d'utiliser les résultats d'autres algorithmes de segmentations tel que Le K-means pour corriger certaines imperfections essentiellement liées au bruit engendré par le matériel

d'acquisition et à la nature texturée des organes échantillonnés. C'est pour cela qu'on parle de régularisation Markovienne.

Les effets de volumes partiels des IRM sont problématiques pour les champs aléatoires de Markov puisqu'ils peuvent mener à une mauvaise classification des voxels.

2.4.1.7 Classification par les Réseaux de neurones

Un réseau de neurone est une succession de couche de neurones dont chacune prend ses entrées à partir des sorties des couches précédentes. Chaque couche i est composée de N_i neurones.

Les entrées de chaque neurone sont pondérées par des poids dits synaptiques de sorte que les contributions des neurones des couches précédentes sont chacune multipliées par ces poids synaptique et additionnées pour donner une valeur globale qui sera affectée à un neurone donné.

La fonction de sortie dit fonction d'activation qui introduit une non-linéarité à chaque étape. Ceci montre l'importance du choix judicieux d'une bonne fonction de sortie : un réseau de neurones dont les sorties seraient linéaires, n'aurait aucun intérêt [20]. Il est possible de donner au réseau de neurones des boucles en liant les sorties d'une couche avec les entrées de couches précédentes. Cela donnerait au réseau de neurones d'autres propriétés intéressantes mais augmentera sa complexité.

La figure 2.2 montre la structure d'un neurone artificiel. Le calcul de la sortie est obtenu à travers la sommation pondérée des entrées et puis l'application d'une fonction d'activation.

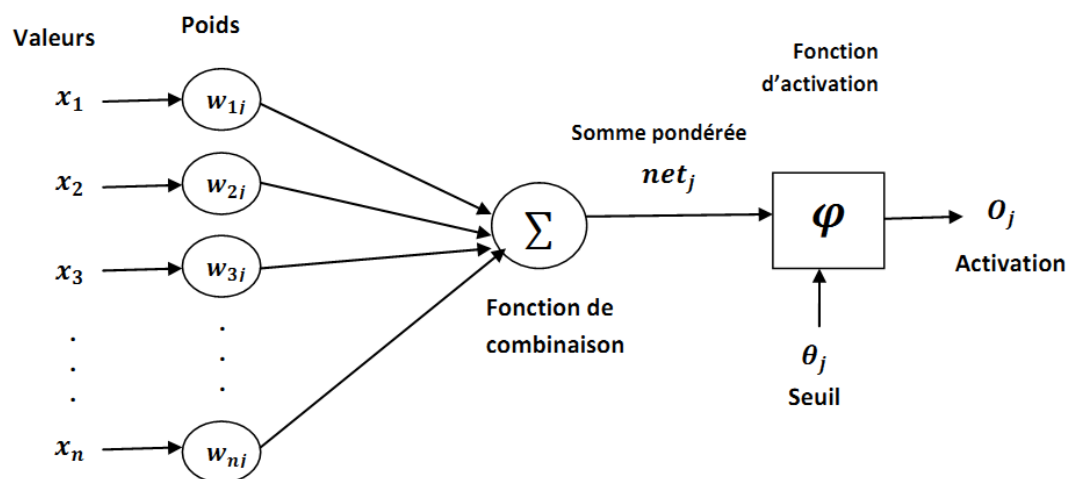


Figure 2.2 Structure d'un neurone artificiel.

Les réseaux de neurones sont utilisés comme classifieurs en segmentation d'images médicales.

L'apprentissage du réseau de neurones consiste à ajuster convenablement les valeurs des poids des synapses. Dans le domaine de l'imagerie médicale, cet apprentissage est basé sur une base d'images médicales dont le résultat de la segmentation est connu à l'avance (apprentissage supervisé).

Le réseau de neurone accepte en entrée un vecteur de caractéristique tiré à partir des IRMs et donne en sortie la classe attribué à chaque image. Il est possible de passer plus d'information en entrée par exemple le volume de l'image, afin de donner plus de robustesse à la classification.

Dans la littérature Sammouda et coll. [48] ont choisi d'utiliser les réseaux de neurones de Hopeld (réseaux à deux couches) pour la segmentation des pathologies du cerveau (tumeurs et sclérose en plaques - SEP -).

Blonda [49,50] est le premier qui a introduit la théorie des ensembles flous dans les réseaux de neurones pour la segmentation des tissus cérébraux et des lésions de SEP. Un réseau de neurones à deux couches appelé Learning Vector Quantization (LVQ) a été proposé par Alirezaie et coll [51], ils ont prouvé la supériorité du LVQ par rapport aux habituels réseaux de quatre couches à rétropropagation, grâce notamment à son insensibilité aux variations de niveaux de gris entre les différentes coupes des IRM. Après Une automatisation de la segmentation des sillons corticaux a été faite par Rivière et coll. [52] par l'utilisation d'une association de perceptrons multi-couche(PMC).

En 2003 une autre combinaison des réseaux de neurones cette fois ci avec la morphologie mathématique était faite par Hult et coll. [53] pour la segmentation multimodale de l'hippocampe.

Dernièrement Yeh et coll. [22] ont créé une nouvelle méthode basée sur l'algorithmique génétique(AG), la logique floue (LF) et les LVQ pour la segmentation des IRM cérébrales.

Parmi les inconvénients des RNNs (réseaux de neurones) on cite L'intervention manuelle lors de l'étape d'apprentissage et le manque d'interopérabilité des résultats de segmentation obtenue (boite noire) pour corriger ce dernier les réseaux de neurones sont généralement combinée avec d'autres techniques telles que la logique floue, les AGs etc., pour donner plus d'efficacité et de transparence.

2.4.2 Contour

Contrairement aux approches régions, qui cherchent à former des zones homogènes, les approches contours (ou frontières), il s'agit de reconnaître les zones de transition et de localiser au mieux la frontière entre les régions. Parmi les principales techniques qui appartiennent à cette approche. On distingue les modèles dérivatifs et les modèles d'espace-échelle.

2.4.2.1 Espace – échelle

L'espace-échelle (scale-space) est une théorie formelle pour la manipulation des structures d'images à des échelles différentes de telle sorte que les structures soient successivement supprimées avec un paramètre « t » d'échelle qui est associé à chaque niveau dans la représentation de l'espace d'échelle.

La représentation en espace-échelle d'une image $f(x,y)$ est une famille de signaux dérivés $L(x, y, t)$ qui est défini la convolution de $f(x, y)$ avec le noyau gaussien $g(x,y,t)$

tel que

$$g(x, y, t) = \frac{1}{2\pi t} e^{-(x^2+y^2)/2t} \quad (2.4)$$

$$L(x, y, t) = g(x, y, t) * f(x, y)$$

$t = \sigma^2$ est la variance de la gaussienne et avec une condition initiale $L(x, y, 0) = f(x, y)$.

Dans [54] Lindeberg a essayé de lier les minimums locaux et les points selles à travers les échelles. Il a proposé une représentation de l'image intitulée « scale-space primal sketch » qui explicite les relations entre structures à différentes échelles.

Lachmann [55] a appliqué des techniques différentes du scale-space sur l'imagerie médicale en 3 dimensions. Henkel [56] à son tour a revu les différentes techniques de segmentation scale-space.

Enfin et dans [57] Germond a utilisé L'espace-échelle pour la segmentation du cerveau (matière blanche et matière grise).

2.4.2.2 Méthodes dérivatifs

Les méthodes dérivatives permettent de trouver les contours dans une image d'une manière simple. On peut détecter le point de contour par la détermination du maximum de la norme du gradient ou bien en étudiant le passage par zéro du Laplacien.

Nom	Mx	My
Roberts	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$
Perwitt	$1/3 \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$1/3 \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$
Sobel	$1/4 \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$1/4 \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$
Kirsch	$1/15 \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}$	$1/15 \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}$
Laplacien	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 0 \\ 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Tableau 2.1 - Principaux masques de calcul de gradient.

Dans la littérature Canny [58,59] représente l'un des piliers des techniques d'optimisation, il démontre qu'on ne peut pas obtenir à la fois une bonne détection et une bonne localisation du contour. Pour cela, Canny a ajouté un critère en plus pour définir un détecteur optimal : la non-multiplicité des maximums locaux. Il développe un filtre à réponse impulsionnelle finie.

Dans [60] Monga et coll. ont utilisé le paradigme des dérivées partielles dans le cadre de La segmentation d'images biomédicales, notamment appliqué aux IRM. Ils ont adapté les filtres de Deriche à des images en 3 dimensions via une implémentation récursive.

Les méthodes dérivatives, employées pour la segmentation des IRM cérébrales, sont relativement peu nombreuses malgré qu'ils soient rapides précises et ne nécessitant pas d'apport d'information a priori parce que elles ne permettent pas, en pratique, d'obtenir des contours fermés. Elles sont en effet très sensibles au bruit, les images bruitées doivent être lissées au préalable par le filtre de gauss par exemple. En pratique le lissage et la dérivation sont réunie dans un seul filtre [61]. Cette combinaison minimise la probabilité de détecter des

faux contours. Enfin, ces méthodes fournissent fréquemment de la sous- ou de la sur-segmentation.

2.5 L'algorithme EM

L'algorithme d'Espérance-Maximisation et une méthode classique non supervisée [62]. C'est un algorithme proposé par Dempster en 1977 et a été appliqué dans nombreux domaines, notamment en segmentation d'image. L'algorithme a été présenté par Wells puis développé par Van Leemput [63].

2.5.1 L'idée de l'algorithme EM

Si le nombre des composantes dans un mélange est connu, l'algorithme EM (Expectation Maximisation) proposé par Dempster [64] permet d'approcher une fonction à partir d'une initialisation.

Le problème de l'estimation des paramètres des composantes dans les mélanges se base sur le principe de maximisation de la vraisemblance (maximum likelihood) : c'est-à-dire que les paramètres des composantes sont choisis de manière à maximiser la probabilité des exemples d'appartenir au mélange.

Connaissant les paramètres du mélange Φ , la probabilité de l'exemple i de la base d'apprentissage est donnée par :

$$P_k (X_i | \Phi) = C ((X_i | \Phi) = \sum_{k=1}^K P_k \cdot \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2}(\frac{X_i - \mu_k}{\sigma_k})^2} \quad (2.5)$$

Si les probabilités sont considérées indépendantes pour les N exemples, la probabilité globale de la base est donnée par le produit des probabilités de tous les exemples :

$$P_k (\vec{X} | \Phi) = \prod_{i=1}^N (\sum_{k=1}^K P_k \cdot \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2}(\frac{X_i - \mu_k}{\sigma_k})^2}) \quad (2.6)$$

L'algorithme EM fonctionne en décomposant le problème en deux étapes nommées étape Estimation et étape Maximisation. Dans l'étape E, il calcule une valeur λ_{ki} qui exprime pour chaque exemple i de la base, avec quel degré il correspond à chaque composante k du mélange. Dans l'étape M, les λ_{ki} sont considérés constants et les paramètres de chaque composante peuvent être calculés séparément pondérant chaque élément i de la base par son degré d'appartenance λ_{ki} à la composante k . Pour l'estimation des paramètres d'une composante, il reste donc à maximiser :

$$P_k (\vec{X} | \Phi, \vec{\lambda}_k) = \prod_{i=1}^N (\frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{\lambda_{ki}}{2 \sum_{i=1}^N \lambda_{ki}} (\frac{X_i - \mu_k}{\sigma_k})^2}) \quad (2.7)$$

Maximiser $P_k(\vec{X}|\Phi, \vec{\lambda}_k)$ est équivalent à maximiser son logarithme $\ln P_k(\vec{X}|\Phi, \vec{\lambda}_k)$:

$$P_k(\vec{X}|\Phi, \vec{\lambda}_k) = \sum_{i=1}^N \ln\left(\frac{1}{\sqrt{2\pi}\sigma_k}\right) - \frac{\sum_{i=1}^N \lambda_{ki}(X_i - \mu_k)^2}{2\sigma_k^2 \sum_{i=1}^N \lambda_{ki}} \quad (2.8)$$

Le maximum relativement à μ_k est en une position de dérivée partielle en μ_k nulle :

$$\frac{\partial}{\partial \mu_k} \ln(P_k(\vec{X}|\Phi, \vec{\lambda}_k)) = 0 \quad (2.9)$$

$$\Rightarrow -\sum_{i=1}^N \frac{\lambda_{ki} \cdot (X_i - \mu_k)}{\sum_{i=1}^N \lambda_{ki} \cdot \sigma_k^2} = 0 \quad (2.10)$$

$$\Rightarrow \sum_{i=1}^N \lambda_{ki} \cdot X_i = \sum_{i=1}^N \lambda_{ki} \cdot \mu_k \quad (2.11)$$

$$\Rightarrow \mu_k = \frac{\sum_{i=1}^N \lambda_{ki} \cdot X_i}{\sum_{i=1}^N \lambda_{ki}} \quad (2.12)$$

De la même manière, le maximum relativement à σ_k est en une position de dérivée partielle en σ_k nulle :

$$\frac{\partial}{\partial \mu_k} \ln(P_k(\vec{X}|\Phi, \vec{\lambda}_k)) = 0 \quad (2.13)$$

$$\Rightarrow -\frac{1}{\sigma_k} + \sum_{i=1}^N \frac{\lambda_{ki} \cdot (X_i - \mu_k)^2}{\sum_{i=1}^N \lambda_{ki} \cdot \sigma_k^3} = 0 \quad (2.14)$$

$$\Rightarrow \sigma_k^2 = \frac{\sum_{i=1}^N \lambda_{ki} \cdot (X_i - \mu_k)^2}{\sum_{i=1}^N \lambda_{ki}} \quad (2.15)$$

$$\Rightarrow \sigma_k = \sqrt{\frac{\sum_{i=1}^N \lambda_{ki} \cdot (X_i - \mu_k)^2}{\sum_{i=1}^N \lambda_{ki}}} \quad (2.16)$$

La valeur P_k est simplement calculée à partir des λ_{ki} :

$$P_k = \sum_{i=1}^N \lambda_{ki}$$

Les formules retrouvées ici sont celles d'une estimation classique du poids p , de la valeur centrale μ et de l'écart-type σ d'une gaussienne à partir d'une série d'exemples avec la seule différence que les exemples sont pondérés selon leur valeur d'appartenance λ_{ki} .

2.5.2 Description de l'algorithme EM

L'algorithme EM s'effectue selon les deux étapes décrites ci-dessous :

Étape E

On effectue la mise à jour des λ_{ki} lors de l'itération (m). Lors de l'initialisation de l'algorithme, on fixe les $p_k; \mu_k; \sigma_k$ pour calculer les valeurs initiales pour des λ_{ki} ;

$$\lambda_{ki}^{(m)} = \frac{P_k^{(m-1)} \cdot g(x_i, \mu_k^{(m-1)}, \sigma_k^{(m-1)})}{\sum_{j=1}^K P_k^{(m-1)} \cdot g(x_i, \mu_j^{(m-1)}, \sigma_j^{(m-1)})} \quad (2.17)$$

Ensuite les mélanges eux-mêmes sont mis à jour à partir des exemples pondérés par les λ_{ki} .

Étape M

(mise à jour des $p_k; \mu_k; \sigma_k$)

$$\begin{aligned} P_k^{(m)} &= \sum_{i=0}^N \lambda_{ki}^{(m)} \\ \mu_k^{(m)} &= \frac{\sum_{i=0}^N \lambda_{ki}^{(m)} x_i}{P_k^{(m)}} \\ \sigma_k^{(m)} &= \sqrt{\frac{\sum_{i=0}^N \lambda_{ki}^{(m)} (x_i - \mu_k^{(m)})^2}{P_k^{(m)}}} \end{aligned} \quad (2.18)$$

Avec :

x_i exemples de la base d'apprentissage

λ_{ki} appartenance de l'exemple i à la gaussienne k

2.6 Conclusion

Dans cet état de l'art, nous avons présenté un certain nombre de méthodes dédiées à la classification et à la segmentation des structures cérébrales saines ou pathologiques classées selon leurs approches. La revue qui est faite ici n'est pas exhaustive mais nous avons essayé de mettre en évidence la diversité des techniques abordées par les différents auteurs.

Cependant chacune d'entre elles à ses qualités et ses défauts, et c'est sur cette base que doit se faire le choix de l'emploi de l'une ou l'autre de ces techniques. La diversité des techniques est représentative de la complexité du problème de segmentation d'images IRM cérébrales, Ces difficultés sont liées, d'une part, aux objets traités et, d'autre part, au particularisme des images IRM. Et nous avons présenté notre méthode d'application ce qui « Estimation-Maximisation » son idée et son description.