



Optimisation de séquences de segmentation combinant modèle structurel et focalisation de l'attention visuelle. Application à la reconnaissance de structures cérébrales dans des images 3D.

Geoffroy Fouquier

► To cite this version:

Geoffroy Fouquier. Optimisation de séquences de segmentation combinant modèle structurel et focalisation de l'attention visuelle. Application à la reconnaissance de structures cérébrales dans des images 3D.. domain_other. Télécom ParisTech, 2010. Français. NNT : . pastel-00006074

HAL Id: pastel-00006074

<https://pastel.hal.science/pastel-00006074>

Submitted on 19 May 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



École Doctorale
d'Informatique,
Télécommunications
et Électronique de Paris

Thèse

présentée pour obtenir le grade de docteur

de l'École Nationale Supérieure des Télécommunications

Spécialité : Signal et Images

Geoffroy Fouquier

**Optimisation de séquences de
segmentation combinant modèle structurel
et focalisation de l'attention visuelle.
Application à la reconnaissance de
structures cérébrales dans des images 3D.**

Soutenue le 22 février 2010 devant le jury composé de

**Serge Guillaume
Philippe Tarroux
Jamal Atif
Michel Desvignes
Isabelle Bloch**

Rapporteurs

Examineurs

Directeur de thèse

Résumé

Nos travaux portent sur l'interprétation d'une scène dont nous possédons un modèle, représentant l'agencement spatial des objets contenus dans cette scène. Dans le cadre d'une segmentation séquentielle permettant de reconnaître les objets les uns après les autres en fonction des étapes antérieures, nous utilisons la connaissance spatiale du modèle pour optimiser la séquence de segmentation à effectuer à partir d'un objet de référence vers un objectif à segmenter. Nous proposons pour cela d'optimiser un chemin dans un graphe représentant les objets de la scène (nœuds) et leurs relations spatiales (arcs). Deux approches sont proposées.

La première approche effectue une optimisation à partir de l'information spatiale du modèle uniquement, en évaluant un critère de pertinence de chaque chemin. L'évaluation est effectuée de manière indépendante sur chaque arc dans un premier temps, puis nous proposons une manière de représenter un chemin entier, permettant d'évaluer la pertinence du chemin à partir de cette représentation.

La deuxième approche s'intègre dans un processus de segmentation séquentielle, vu comme l'exploration progressive d'une image à partir d'un objet de référence. Nous utilisons une modélisation d'une technique pré-attentionnelle, une carte de saillance, afin de guider le processus de segmentation séquentielle, en intégrant à l'approche structurelle des informations de saillance extraites de l'image à interpréter.

Le domaine d'application de ces approches est la segmentation des structures sous-corticales du cerveau dans des images IRM 3D dont certaines présentent des pathologies.

Abstract

Sequential segmentation optimization using a structural model and focus of visual attention. Application to the recognition of internal brain structures in 3D magnetic resonance images (MRI).

We aim at recognizing a 3D scene described by a 3D image and a structural model, i.e., a model that describes the spatial arrangement of the objects. The sequential segmentation framework is considered. This allows us to segment and recognize objects in a sequential way, using at each step the previously recognized object to guide the segmentation of the next ones. We propose to use the spatial information included in the model to optimize the segmentation sequence from a reference object to a selected target. This sequence is viewed as a path in a graph where a node represents an object and an edge carries the spatial relation information between two objects.

We propose to use the spatial information included in the model to optimized the segmentation sequence from a reference object to a selected target. This sequence is view as a path in a graph where vertex represents objects and edges represents spatial relations.

Two approaches are proposed. The first one proposes to evaluate the relevance of a path according to the generic available knowledge. This estimation is realized either on each spatial relation independently or directly on a fuzzy subset that represents the whole path at once. The best path according to a criterion is then selected and the objects may be segmented.

The second approche proposes to integrate the segmentation sequence optimization directly into a sequential segmentation framework. The optimization uses a spatial model of the scene modeled as a graph and also a saliency map to guide the segmentation. The latter can be seen as an image exploration process.

Both approaches are used for segmentation and recognition of internal brain structures in 3D magnetic resonance images. We also propose an adaptation of these methods to cope with pathological cases (e.g., brain tumors).

Remerciements

Je voudrais remercier tout particulièrement Isabelle et Jamal pour la direction de mon stage puis de cette thèse. Pour m'avoir fait confiance tout d'abord avec un profil atypique, puis pour les conseils précieux, l'aide et le soutien. Merci à Isabelle d'être toujours présente et si prompt à relire notre prose. Merci à Jamal pour son amitié et son accueil en Guyane. J'ai eu parfois plus le sentiment d'une longue collaboration que d'une direction, tout en ayant beaucoup appris, scientifiquement mais aussi humainement, alors merci !

Je remercie Michel Desvignes pour avoir accepté de présider mon jury, Serge Guillaume et Philippe Tarroux pour avoir été les rapporteurs de ces travaux, ainsi que l'ensemble des membres du jury pour leur évaluations et leur précieux conseils.

Je remercie également Jérôme, Saïd, Sylvain, Ceyhun et Réda pour leur amitié, leur aide et leur soutien au long de ces années qui n'ont pas toujours été aisées. Avoir un avis, une confirmation ou simplement un oreille a été pour moi plus que important et nécessaire, ainsi que les moments de détente entre amis.

Je remercie tous les doctorants et postdoc que j'ai croisé à télécom avec lesquels j'ai pu passer de bons moments, parfois collaborer. En particulier Emi, Olivier, Jérémie, David, Vincent, Céline, Antonio, Racha, Carolina, Julien, Nicolas et tous les autres. Je remercie très chaleureusement Patricia pour s'occuper de tous les doctorants, pour ses attentions, le café du matin, les bonbons ou l'aspirine mais aussi pour sa compagnie. Je remercie également Catherine Vazza et Florence Besnard ainsi que l'ensemble du département TSI, en particulier Marc, Laurence et avec une pensée pour Francis. Je remercie également Sophie-Charlotte pour son impeccable gestion du réseau et sa réactivité.

Je remercie également mes parents pour m'avoir encouragé et soutenu tout au long de mes études jusqu'à cet aboutissement. Je n'aurais pas pu faire tout cela sans eux et je suis reconnaissant. Je remercie également Thibaud et sa famille et Florent pour toutes ces bonnes années et de leur soutien. Enfin, je remercie Ana, les plus belles découvertes ne sont pas les plus attendues. Merci d'avoir été et d'être toujours à mes côtés.

Table des matières

Résumé	3
Abstract	5
Remerciements	7
Introduction	13
1 Segmentation et reconnaissance de structures cérébrales : les approches par modèle	17
1.1 Interprétation d'images et vision cognitive	18
1.1.1 Interprétation et acquisition	18
1.1.2 Systèmes à base de connaissances	19
1.1.3 La vision cognitive	21
1.1.4 Conclusion sur les systèmes d'interprétation d'images	21
1.2 Segmentation et reconnaissance d'images cérébrales à l'aide d'un modèle	21
1.2.1 Représentations de l'anatomie cérébrale	23
1.2.1.1 Atlas et modèles de forme	25
1.2.1.2 Représentation structurelle de l'anatomie cérébrale	28
1.2.2 Reconnaissance avec un modèle structurel	30
1.2.2.1 Segmentation et reconnaissance par mise en correspondance du modèle	30
1.2.2.2 Approche itérative de la segmentation	32
1.2.2.3 Approche globale par contraintes	33
1.3 Conclusion	33
2 Les mécanismes de l'attention	35
2.1 Qu'est ce que l'attention ?	35
2.1.1 Définition de l'attention	36
2.1.2 L'unité attentionnelle	37
2.2 Le pré-attentionnel	39
2.2.1 Les différentes théories pré-attentionnelles	42
2.2.1.1 « Feature integration theory »	42
2.2.1.2 « Guided search theory »	43
2.2.1.3 « Texton theory »	43
2.2.1.4 « Similarity theory »	44
2.2.2 Conclusion sur les théories pré-attentionnelles	44
2.3 Les cartes de saillance	44
2.4 Les cartes de saillance adaptées aux images IRM	47
2.4.1 Pré-traitements	47

2.4.2	Filtrage par caractéristique	48
2.4.3	Génération des pyramides	49
2.4.4	Fusion des cartes de discontinuités	50
2.4.5	Cartes de saillance	53
2.4.6	Masquage des cartes de saillance	53
2.4.7	Résultats	54
2.5	Conclusion	56
3	Le modèle de connaissance	57
3.1	Graphe de relations spatiales	58
3.1.1	Les relations spatiales pour l'imagerie médicale	58
3.1.2	Graphe de relations spatiales	60
3.1.3	Notations	61
3.2	Sources de connaissances	62
3.2.1	Connaissance experte et textuelle	62
3.2.2	Connaissance extraite automatiquement	62
3.2.3	Connaissance extraite de manière semi-interactive	64
3.2.3.1	Les traces de l'utilisateur	64
3.2.3.2	Récupération des objets	65
3.2.3.3	Interprétation des traces	66
3.2.3.4	Création du modèle	67
3.2.4	Conclusion sur les sources de connaissances	68
3.3	Formalisme flou pour les relations spatiales	69
3.3.1	Représentation de la relation de distance	69
3.3.2	Représentation de la relation d'orientation	70
3.3.3	Représentation de l'adjacence	71
3.3.4	Autres relations	72
3.3.5	Notations des paysages flous	72
3.4	Base de données d'images cérébrales	72
3.5	Apprentissage des paramètres des intervalles flous	74
3.5.1	Cadre général pour l'apprentissage des intervalles flous	75
3.5.2	Un exemple d'apprentissage	76
3.5.3	Le cas de la distance	78
3.6	Conclusion	78
4	Optimisation avec représentation des structures	81
4.1	Raisonnement avec représentation de la forme des structures	82
4.1.1	Raisonnement dans le cas « sain »	83
4.1.2	Évaluation de la pertinence d'un arc	84
4.1.3	Sélection du « meilleur » chemin	86
4.1.4	Expériences	88
4.2	Raisonnement dans le cas pathologique	92
4.2.1	Degré de stabilité des relations spatiales	92
4.2.2	Adaptation de l'approche aux cas pathologiques	95
4.3	Optimisation globale de la pertinence d'un chemin	97
4.3.1	Fusion des connaissances spatiales	98
4.3.2	Évaluation du chemin par mesure de son entropie	101
4.3.3	Expériences	101

4.3.4	Adaptation aux cas pathologiques	102
4.3.5	Conclusion sur l'approche globale	102
4.4	Conclusion	102
5	Optimisation avec information visuelle	105
5.1	Utilisation d'une information visuelle	106
5.1.1	Attention visuelle et segmentation séquentielle	106
5.1.2	Saillance et difficulté de segmentation	108
5.1.3	Apprentissage de la saillance	109
5.1.4	Un critère reposant sur la saillance	112
5.1.4.1	Critère simple sans apprentissage	112
5.1.4.2	Critère utilisant une mesure EMD	112
5.1.5	La saillance des tumeurs cérébrales	115
5.2	Segmentation séquentielle avec un critère fondé sur la saillance	117
5.2.1	Exploration progressive de l'image	117
5.2.2	Graphe spatial	117
5.2.3	Filtrage du graphe	119
5.2.4	Domaine de recherche	121
5.2.5	Intégration de la saillance	124
5.2.6	Sélection du prochain objet	126
5.2.7	Le processus de segmentation	127
5.2.8	Mise à jour du graphe	131
5.2.8.1	Pas de segmentation	132
5.2.8.2	Il y a une segmentation	133
5.2.8.3	Échec de la segmentation d'une structure	136
5.2.8.4	Structure de contrôle	136
5.2.8.5	Mise à jour du graphe	136
5.3	Expériences	138
5.3.1	Déroulement du processus	138
5.3.2	Les séquences de segmentation	139
5.3.3	Les résultats de segmentation	143
5.3.4	Résultats dans les cas pathologiques	146
5.4	Conclusion	146
6	Conclusion et perspectives	151
6.1	Synthèse des contributions	151
6.1.1	Optimisation de chemin avec représentation des structures	151
6.1.2	Optimisation de chemin avec saillance	152
6.2	Perspectives	154
6.2.1	Optimisation avec représentation des structures	154
6.2.2	Optimisation avec information visuelle	154
Annexes		157
A Liste des publications		157

B	Cartes de saillance	159
B.1	Les cas sains	160
B.1.1	IBSR 01	160
B.1.2	Oasis 02	162
B.1.3	Les autres cas sains	163
B.2	Les cas pathologiques	178
B.2.1	Cas 1	178
B.2.2	Les autres cas pathologiques	179
C	Image segmentation as inexact graph matching using high-level attributes	189
C.1	Abstract	189
C.2	Introduction	189
C.3	Model, image and deformation graphs	191
C.4	Attributes and cost function	192
C.4.1	Vertex cost : intrinsic features for each class of the model	192
C.4.2	Edge cost : reflecting the structure	193
C.4.3	Connectivity	194
C.4.4	Cost function	195
C.5	Matching algorithm and optimization	195
C.6	Experiments	196
C.7	Conclusions	198
C.8	Acknowledgment	198
	Bibliographie	199

Introduction

L'interprétation des images est une tâche complexe, autant par la diversité des moyens de représenter une image et des approches associées permettant de réaliser son interprétation, que par la subjectivité du résultat attendu. L'objectif de l'interprétation est de pouvoir reconnaître les objets qui composent une scène et leurs relations.

L'utilisation d'un modèle de la connaissance se heurte au problème du saut sémantique, c'est-à-dire la différence entre la description d'un objet par une connaissance générique et exprimée en langage naturel d'une part, et sa représentation numérique d'autre part. Dans notre cas, il s'agit de la difficulté de faire le lien entre la connaissance générique et les parties de l'image qui lui correspondent. Cependant, décrire les objets qui composent une scène et leurs relations est une manière naturelle de décrire une scène et qui est cohérente avec la manière dont le système visuel explore une scène. Les modèles représentant une image comme un ensemble d'objets structurés sont donc bien adaptés à cette tâche. Parmi ces modèles, la théorie des graphes fournit un cadre permettant de représenter plusieurs niveaux de connaissance, objet ou région et connaissance structurelle.

Une manière naturelle de décrire les relations entre les différents objets qui composent une scène est de décrire leurs positions relatives, par exemple « *l'objet A est à droite de l'objet B* ». De plus les relations spatiales, grâce à leur imprécision intrinsèque, sont appropriées pour modéliser l'imprécision de ces relations. Il existe différentes manières de prendre en compte l'information spatiale, que ce soit pour la segmentation ou pour la reconnaissance des structures. Nos travaux se placent dans le cadre de l'interprétation d'une scène guidée par un modèle décrivant l'agencement spatial des objets composant la scène. Nous proposons d'exploiter au mieux la connaissance spatiale d'une scène à interpréter, mais aussi la connaissance extraite de l'image elle-même dès qu'elle est disponible. La problématique de ces travaux est principalement la suivante : comment explorer l'image de la manière la plus propice à son interprétation. Si l'exploration correspond à une séquence de segmentation, alors nous souhaitons connaître la meilleure séquence de segmentation possible d'une image en fonction de l'information disponible.

En fonction du type de connaissance disponible à propos d'une scène (experte, extraite automatiquement, ...), le modèle spatial généré va permettre un raisonnement spatial plus ou moins puissant. La constitution d'un modèle de l'agencement spatial d'une scène n'est pas l'objet de nos travaux, même si cette question est abordée lors de la présentation du modèle de la connaissance.

Le domaine d'application nous permettant d'illustrer nos contributions est celui de l'imagerie cérébrale. La segmentation et la reconnaissance des structures sous-corticales du cerveau représente une tâche complexe d'interprétation en raison de la radiométrie non discriminante des structures, de la forme complexe que peuvent prendre ces structures et de la grande variabilité inter-patients. Pour ces raisons, la segmentation des images cérébrales est le plus souvent guidée par un modèle. De plus, l'agencement spatial des structures cérébrales est stable (dans le cas sain). L'information spatiale est donc pertinente dans ce cas. Il existe de nombreuses représentations structurelles de l'anatomie cérébrale, l'ontologie de la FMA par exemple ([Rosse et Mejino \(2007\)](#)), ainsi que des méthodes de segmentation des structures sous-corticales utilisant ce type de

représentation. Nous proposons des approches dans le cadre de cette application, pour déduire de la représentation structurelle et de l'image à interpréter la séquence de segmentation.

Connaissance générique

Nous avons une connaissance qui provient de descriptions anatomiques : nous connaissons les différentes structures du cerveau, et nous connaissons les relations spatiales entre elles. Ces descriptions sont le plus souvent textuelles. Par exemple, le noyau caudé est « proche du ventricule latéral ». Une telle relation est intrinsèquement imprécise, ce qui permet de prendre en compte ses variations inter-patients. Nous utilisons donc un formalisme qui permet de conserver cette imprécision. Le formalisme flou est particulièrement adapté pour modéliser l'imprécision de ces relations (Bloch (2005)).

Il y a plusieurs manières de représenter une relation spatiale. Les représentations que nous utilisons répondent à cette question (Bloch (2005)) : « À partir d'un objet de référence A , quels sont les points de l'espace qui satisfont une relation R calculée à partir de A ». Par exemple, si nous avons une relation *à droite de* A , nous représentons cette relation dans l'espace de l'image, et à chaque point correspond un degré de satisfaction de la relation « à droite de A ». La représentation de la relation est donc directement dépendante de la forme de l'objet de référence. À partir de ces relations nous proposons une approche permettant de sélectionner un chemin de segmentation et répondant à cette question : à partir d'un objet donné, quelle est la meilleure séquence de segmentation permettant de segmenter un objet cible donné. Cette approche repose sur la connaissance spatiale, ainsi que sur des représentations des objets qui proviennent de la connaissance générique.

Connaissance extraite de l'image

En outre, nous souhaitons également utiliser les informations qui sont extraites de l'image. D'une part, nous souhaitons pouvoir prendre en compte l'information de l'image pour adapter son exploration et ainsi mieux prendre en compte les particularités d'une image. D'autre part, nous souhaitons tenir compte de l'exploration réalisée pour guider la suite du processus. Pour cela, nous considérons deux types d'information : une information extraite d'une manière globale de l'image, et les segmentations qui seront réalisées au cours du processus. Nous effectuons dans ces travaux un parallèle entre un processus de segmentation séquentielle et une modélisation, selon la théorie de l'intégration des caractéristiques, d'un processus d'attention visuelle. Dans cette modélisation, une étape pré-attentionnelle, calculée sur l'ensemble de l'image permet d'attirer le faisceau attentionnel sur des parties de l'image. À l'étape attentionnelle, cette petite partie de l'image sera analysée. Nous proposons une approche utilisant un mécanisme pré-attentionnel, les cartes de saillance, pour guider un processus de segmentation séquentielle. Le principe est de réaliser une exploration progressive de l'image, où le choix des segmentations successives sera effectué en utilisant l'information spatiale et selon un critère dérivé des cartes de saillance. Cette approche permet d'utiliser non seulement l'information globale de saillance, mais aussi l'information extraite de l'image après chaque segmentation.

Gestion des cas pathologiques

Des pathologies peuvent intervenir dans les images cérébrales, en particulier, nous nous intéressons au cas des tumeurs cérébrales. Il existe de nombreux types de tumeurs, avec des comportements spatiaux différents (Khotanlou (2008)). Parmi les comportements spatiaux classiques, les tumeurs peuvent déplacer, déformer, voire détruire des structures cérébrales. Les relations spatiales sont également affectées. Il est donc nécessaire d'adapter le raisonnement spatial pour être

capable de gérer ces cas pathologiques. Nous présentons, pour chacune des approches, comment les cas pathologiques peuvent être pris en compte.

Structure du document

Ce document est composé des chapitres suivants.

Le chapitre 1 présente une étude bibliographique non exhaustive sur les systèmes d'interprétation d'image, en particulier les systèmes à base de connaissances d'une part, puis les méthodes d'interprétation d'images cérébrales.

Dans le chapitre 2 nous présentons une étude bibliographique portant cette fois sur la notion d'attention visuelle, et des mécanismes qui la modélisent. Nous présentons plus en détail le mécanisme pré-attentionnel permettant la génération des cartes de saillance tel qu'il a été décrit dans la littérature, puis nous proposons des adaptations permettant de calculer des cartes de saillance adaptées à l'imagerie cérébrale.

Le modèle de la connaissance générique utilisé dans notre étude est présenté dans le chapitre 3. Nous discutons des sources de connaissances autres que la connaissance experte utilisée en imagerie cérébrale. Nous présentons également le formalisme de représentation des relations spatiales ainsi que la manière dont les paramètres de ces relations sont appris.

Le chapitre 4 présente une première approche qui vise à optimiser des chemins de segmentation, en utilisant la connaissance spatiale du modèle ainsi que des représentations des structures issues de la connaissance générique. Cette méthode est également adaptée pour prendre en compte les cas pathologiques qui peuvent se présenter en imagerie cérébrale.

Une seconde approche intégrant le mécanisme pré-attentionnel dans un processus de segmentation séquentielle est présentée dans le chapitre 5. Nous présentons en détail comment intégrer l'information de saillance pour guider la segmentation et comment cette information peut être utilisée après segmentation.

Le chapitre 6 récapitule les travaux développés dans les chapitres précédents et présente des perspectives de recherche envisageables.

La liste des publications en relation avec ces travaux se trouve dans l'annexe A. Nous présentons dans l'annexe B des résultats de génération de cartes de saillance sur toutes les images de notre base de données. Enfin l'annexe C présente une application pour la segmentation d'un modèle discuté dans le chapitre 3, utilisant une connaissance fournie par l'utilisateur.

Chapitre 1

Segmentation et reconnaissance de structures cérébrales : les approches par modèle

Nos travaux portent sur une tâche d'interprétation des images, avec une application particulière à la reconnaissance des structures cérébrales dans le cerveau humain. Cette tâche est effectuée en utilisant une connaissance a priori de la scène, sur les objets et sur leur structure, connaissance modélisée à l'aide d'un graphe. L'objectif de ce chapitre est de présenter les travaux existants se rapportant aux différents aspects de cette tâche.

L'interprétation des images correspond à l'analyse d'une image ou d'une scène permettant de décrire les objets composant la scène et leurs relations, c'est-à-dire extraire la sémantique de l'image, afin de la comprendre. Cette problématique est un problème de perception de l'environnement par des capteurs (« visual perception ») qui peut être divisé en trois catégories ([Trivedi et Rosenfeld \(1989\)](#)) :

La neurophysiologie ou l'étude des mécanismes biologiques de la vision. L'humain est capable d'interpréter une scène souvent sans difficultés et de manière automatique. De nombreux travaux cherchent à modéliser la vision humaine et les différents mécanismes permettant l'exploration d'une scène. Le chapitre [2](#) présente la notion d'attention visuelle, ainsi que les mécanismes bio-inspirés des phases attentionnelles et pré-attentionnelles de la vision.

La psychologie perceptive qui consiste à comprendre les aspects psychologiques de la perception. Certains aspects sont abordés dans le chapitre [2](#). En revanche, notre étude se limite à des systèmes bio-inspirés plutôt que psycho-réalistes.

La vision artificielle c'est-à-dire les mécanismes permettant de faire comprendre à une machine ce qu'elle « voit » au travers de capteurs. Nos travaux se situent dans cette dernière catégorie.

Nous commençons dans ce chapitre par introduire la problématique de la vision artificielle en présentant les différents types de systèmes d'interprétation d'images, ainsi que la notion de vision cognitive. Dans une deuxième partie, nous dresserons un panorama des méthodes de reconnaissance des structures sous-corticales du cerveau utilisant un modèle de l'anatomie. Cette problématique constitue le domaine d'application de nos travaux.

1.1 Interprétation d'images et vision cognitive

La première théorie de la vision numérique a été proposée dans [Marr \(1982\)](#) et propose une architecture en trois niveaux que tout système de traitement de l'information doit respecter pour demeurer cohérent. Ces travaux vont inspirer la plupart des systèmes de traitement de l'information par la suite. Les différents niveaux proposés par sa théorie sont les suivants :

- un niveau abstrait : le « quoi » et le « pourquoi » ([Marr \(1976\)](#)), c'est-à-dire que doit-on faire, la théorie, les données en entrée ;
- un niveau de la représentation : le comment, les structures de données, les algorithmes ;
- un niveau de réalisation : l'implantation des algorithmes, reliée au matériel.

Marr propose également un système de vision passif et ascendant (sans information a priori) permettant la représentation en trois dimensions et via la stéréoscopie d'images en deux dimensions, et qui repose sur la perception visuelle. Les trois niveaux de ce système sont :

- l'ébauche primitive : où des primitives sont extraites de l'image et regroupées selon des règles proches de la Gestalt ([Desolneux et al. \(2008\)](#)) ;
- l'ébauche 2,5D : qui effectue une carte de profondeur des objets présents dans la scène ;
- la représentation 3D de la scène.

Le système de Marr produit une représentation en trois dimensions d'une scène à partir de projection en deux dimensions. Cette représentation est forcément incomplète en l'absence d'information a priori (mis à part les règles de regroupement de l'ébauche primitive), qui empêche toute interprétation sémantique des objets ou de la scène.

1.1.1 Interprétation et acquisition

Le système de Marr est un système passif, qui ne tient pas compte des possibilités offertes à l'observateur d'interagir avec son environnement. De nombreuses approches exploitant ces possibilités ont été proposées.

« **Active Vision** » L'approche de vision active ([Aloimonos et al. \(1988\)](#)) tient compte de l'aspect séquentiel de l'attention du système visuel biologique, en considérant la perception comme une exploration. Le but est, en ajoutant des points de vue, de contraindre plus le problème, et ainsi de mieux poser un problème de vision, que les auteurs considèrent souvent comme mal posé ([Aloimonos et al. \(1988\)](#)). Dans cette approche, les mouvements de l'observateur (tête, « covert » ou « overt attention », c'est-à-dire déplacement de l'attention visuelle par le déplacement des yeux, ou sans mouvement des yeux) sont représentés par les mouvements des capteurs.

« **Active Perception** » Cette approche ([Bajcsy \(1988\)](#)) voit le problème de perception comme un problème de stratégie de contrôle des capteurs, une meilleure stratégie permettant d'obtenir plus d'information sur la scène et l'environnement. La notion de perception active est l'étude des différentes modélisations des stratégies de contrôle des capteurs. Ces travaux introduisent des raisonnements sur la connaissance de la scène.

« **Animate Vision** » La vision animée ([Ballard et Brown \(1992\)](#)) est issue de l'étude des mouvements dans le cadre d'une tâche visuelle. Le mode exploratoire d'une image a été illustré par [Yarbus \(1967\)](#) et sera présenté dans le chapitre suivant à la section 2.1. La vision animée a un point de vue assez proche de la vision active, en considérant la perception comme un problème mal posé qu'il est nécessaire de mieux contraindre, en ajoutant de l'information. Mais dans cette approche, [Ballard et Brown \(1992\)](#) modélisent le focus attentionnel, c'est-à-dire la limitation de

l'attention à une zone très restreinte de l'espace visuel, afin d'analyser uniquement les zones les plus significatives de l'espace, mais d'une manière plus attentive.

« **Purposive Vision** » La vision à dessein ([Aloimonos \(1990\)](#)) se focalise sur la vision guidée par une tâche et non pas par les données uniquement. Les différentes explorations d'une image en fonction de la tâche à accomplir ont également été illustrées par [Yarbus \(1967\)](#). L'attention est dirigée sur certaines parties de l'image en fonction de l'objectif, et d'autres parties peuvent être ignorées. De même, certaines parties prennent une grande importance en fonction du but poursuivi. Il s'agit ici d'extraire les représentations de l'image les plus adéquates en fonction de l'objectif, et les modules de traitement les plus adéquats pour traiter ces représentations. Le principe est de séparer le problème en sous-problèmes et de définir un gestionnaire permettant la recomposition, ce qui permet d'améliorer la tâche de perception et de reconnaissance ([Tsotsos \(1994\)](#)). Cette approche peut être reliée à la problématique de la recherche visuelle (« visual search »), qui propose d'adapter la notion de saillance en fonction de l'objectif poursuivi. Cette notion de saillance et l'approche de la recherche visuelle sont présentées dans la partie [2.2.1](#).

Vision « passive » Les systèmes d'interprétation d'images que nous avons présentés voient le problème de l'interprétation d'image comme une mise en correspondance entre des projections en deux dimensions d'une scène vue depuis un capteur et le modèle en trois dimensions de cette scène. Ces systèmes « actifs » utilisent les capteurs pour mieux contraindre le problème et obtenir de l'information supplémentaire.

Dans nos travaux, l'acquisition des images est séparée de leur traitement, ce qui empêche l'utilisation des approches actives. Nous sommes donc dans le cas d'une vision « passive » par opposition aux systèmes actifs présentés. Cela est particulièrement important dans notre domaine d'application, en imagerie médicale où les acquisitions ne peuvent pas être contrôlées. De plus, les acquisitions d'images médicales nous fournissent un volume en trois dimensions de la scène, et non pas une projection en deux dimensions de la scène. En imagerie cérébrale plus particulièrement, la scène complète est comprise dans le volume fourni.

1.1.2 Systèmes à base de connaissances

Il existe plusieurs approches méthodologiques de l'interprétation d'images. Nous utilisons dans nos travaux une représentation de l'agencement spatial de la scène. Il est donc nécessaire, pour avoir cette connaissance spatiale, d'avoir une représentation explicite de la connaissance utilisable dans le processus. Nous nous intéressons donc plus particulièrement aux systèmes à base de connaissances.

Ces systèmes (on pourra se référer à [Crevier et Lepage \(1997\)](#) et [Le Ber et al. \(2006\)](#) pour une revue de ces systèmes) modélisent la connaissance a priori sur une scène, ainsi que la connaissance nécessaire à son interprétation. On peut distinguer trois types de connaissances ([Matsuyama et Hwang \(1990\)](#)) :

- la connaissance a priori de la scène, des objets qui la composent et des relations entre ces objets. Le modèle de l'agencement structurel d'une scène que nous utilisons dans nos travaux entre dans cette catégorie ;
- la connaissance nécessaire à l'extraction d'information de l'image, c'est-à-dire une connaissance spécifique aux images utilisées permettant d'obtenir des primitives à partir de cette image ;
- la connaissance permettant de faire le lien entre la connaissance de la scène et la connaissance de l'image. Le problème du lien entre la connaissance a priori apportée par le modèle

et l'information extraite de l'image est connu comme le problème du *saut sémantique*.

En pratique, les systèmes tels que celui de Marr (Marr (1982)), SIGMA (Matsuyama et Hwang (1990)) ou eCognition¹ utilisent une architecture à deux niveaux :

- l'extraction d'informations de l'image, c'est-à-dire la couche de bas-niveau, permettant de fournir des primitives au système. Cela peut être une segmentation par exemple ;
- un niveau d'interprétation effectuant le lien entre l'information de l'image et le modèle.

La figure 1.1 présente ces deux niveaux et leurs relations dans un système à base de connaissances.

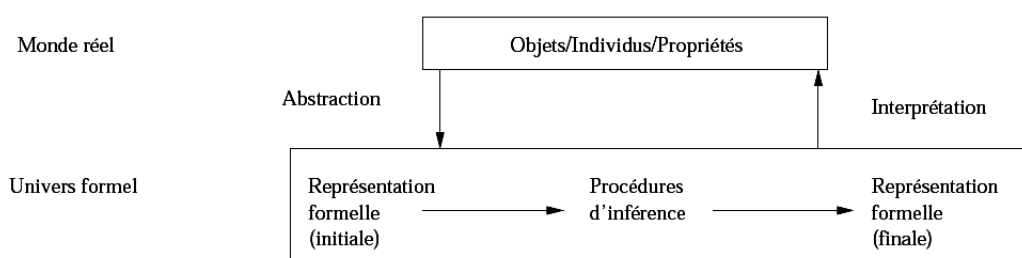


FIG. 1.1 – Représentation et manipulation des connaissances dans un système à base de connaissances [Figure extraite de Le Ber et al. (2006)].

L'observation d'images en deux dimensions pour modéliser une scène en trois dimensions pose des problèmes d'occultation, qui ne concernent pas ces travaux, la scène complète étant observée. L'extraction d'information de l'image est sujette aux problèmes classiques de la segmentation, la sur-segmentation par exemple. L'utilisation d'un modèle spatial apporte une information structurelle généralement stable qui est utilisée dans des systèmes à base de connaissances pour améliorer la reconnaissance.

Utilisation d'un modèle de la connaissance spatiale Les relations spatiales sont communément admises comme jouant un rôle important dans l'interprétation d'une scène. L'information spatiale peut être vue du point de vue sémantique comme un attribut d'objet du modèle (avec des relations topologiques par exemple). Cette information peut également être utilisée pour le raisonnement, en particulier si les caractéristiques des objets ne permettent pas de les discriminer.

L'approche de Le-Ber et Napoli (2002) utilise des relations spatiales topologiques (suivant le formalisme RCC-8) pour la classification de paysages agricoles. Les relations sont hiérarchisées sur un treillis de Galois. Dans cette approche, les relations sont représentées en tant que concepts, c'est-à-dire que les relations sont représentées par des objets propres qui renseignent sur les primitives intrinsèques des relations, mais aussi en tant que relations entre des concepts, pour faire un lien entre deux concepts (deux classes de terrains par exemple). Les relations topologiques sont fréquemment utilisées pour l'interprétation d'images satellitaires (Alboody et al. (2008) par exemple).

Des représentations floues des relations spatiales (Bloch (2005)) sont utilisées pour la reconnaissance. Les méthodes permettant la reconnaissance des structures cérébrales utilisant un modèle structurel de l'anatomie cérébrale sont des exemples d'utilisation de l'information spatiale pour l'interprétation des images. Nos travaux utilisent ce genre de représentations qui sont présentées dans le chapitre 3.

¹On pourra trouver une description à cette adresse : <http://earth.definiens.com>

1.1.3 La vision cognitive

La vision cognitive a été introduite sur un constat qui fait consensus sur les systèmes d'interprétation d'images : ils manquent de robustesse, ils sont souvent difficiles à adapter et souvent très dépendants d'un domaine d'application. La vision cognitive - qui regroupe différents domaines tels que la vision par ordinateur, la reconnaissance de formes, l'intelligence artificielle, la robotique, l'apprentissage ou encore les sciences cognitives - propose de construire des systèmes plus robustes en les dotant de capacités cognitives ([Vernon \(2008\)](#)).

Les caractéristiques d'un système de vision cognitive ont été définies comme les suivantes ([Vernon \(2006\)](#); [P. Auer \(2005\)](#); [Granlund \(1999\)](#)) :

- la capacité de suivre un objectif ;
- de s'adapter à des cas nouveaux ;
- d'anticiper les objets et les événements.

Un tel système doit donc être capable d'apprendre, de s'adapter, de faire des choix, et de développer de nouvelles stratégies.

La notion de capacité cognitive est une notion qui reste vague, et dont il existe plusieurs interprétations en fonction des modèles concernés. Il existe néanmoins deux familles d'approches : l'approche cognitive qui est fondée sur les systèmes de représentation et de traitement de l'information symbolique, et les systèmes émergents, regroupant les systèmes connexionnistes et les systèmes dynamiques (entre autres). Il existe en outre des systèmes hybrides.

On pourra se référer à [Vernon \(2008\)](#) pour une étude complète des différents aspects de la vision cognitive.

1.1.4 Conclusion sur les systèmes d'interprétation d'images

Parmi les systèmes d'interprétation d'images, nous nous plaçons parmi les systèmes à base de connaissances, et plus précisément, parmi les systèmes utilisant l'information spatiale pour raisonner.

Dans le cadre de notre domaine d'application, nous avons un problème « simplifié » de vision, au sens où nous avons en entrée du processus une image en trois dimensions, et pas une projection de cette scène sur une image en deux dimensions. De plus, nous avons la garantie d'avoir l'intégralité de la scène. De plus, le modèle du cerveau, que nous présenterons dans la partie suivante, nous permet de connaître les objets présents dans la scène.

Nous avons donc un système d'interprétation qui est dépendant de notre domaine d'application, un des problèmes pointés par l'approche de la vision cognitive. Nous allons à présent nous intéresser à notre domaine d'application, la segmentation des structures cérébrales, à l'aide d'un modèle structurel de l'anatomie.

1.2 Segmentation et reconnaissance d'images cérébrales à l'aide d'un modèle

Les IRM cérébrales sont des images qui présentent une faible résolution, en particulier par rapport à la taille des structures internes (sous-corticales) telles que celles que nous considérons dans nos travaux. De plus, la radiométrie des différentes structures cérébrales n'est pas suffisante pour discriminer les structures entre elles. Les structures peuvent présenter une différence d'intensité faible par rapport à la matière qui les entoure, mais certaines structures peuvent présenter une radiométrie similaire. Il n'est donc pas possible de segmenter les structures cérébrales en s'ap-

puyant sur cette information uniquement. La figure 1.2 présente un exemple d'IRM cérébrale avec quelques structures internes qui ont été pointées.

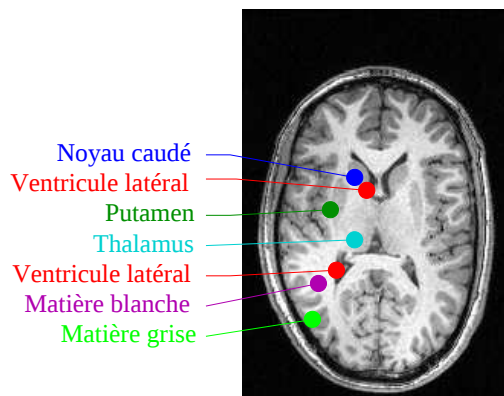


FIG. 1.2 – Une coupe d'image IRM du cerveau avec quelques structures internes étiquetées. Les structures sont présentes de manière symétrique dans les deux hémisphères. La matière blanche englobe les structures présentées. La matière grise est située plutôt sur le bord du cerveau.

Les IRM cérébrales présentent en outre une grande variabilité. D'une manière générale, les structures internes présentent des formes complexes et soumises à des variations. La figure 1.3 présente des coupes extraites d'IRM cérébrales de la base OASIS (Marcus *et al.* (2007)) et de la base IBSR². Nous pouvons clairement voir sur ces images les variations, en particulier sur la forme du cerveau en général, mais également sur les ventricules latéraux au centre de l'image, même si les coupes ne sont pas exactement les mêmes sur cette figure, les images n'étant pas recalées dans la base OASIS. Nous pouvons également observer les différences d'intensité entre ces images. Leur segmentation est donc un problème complexe, qui nécessite une connaissance a priori sur la scène. Cette connaissance peut concerner les caractéristiques des structures ou encore leur agencement spatial. La segmentation des structures cérébrales doit donc être guidée par un modèle de l'anatomie cérébrale.

L'apparition de pathologies, en particulier de tumeurs cérébrales, est un problème qu'il est nécessaire de prendre en compte dans le modèle utilisé. Pour une revue des différents types de pathologies, on pourra se référer à Khotanlou (2008); Khotanlou *et al.* (2007). Les tumeurs cérébrales peuvent avoir différents comportements spatiaux, selon qu'elles vont infiltrer les tissus (et donc modifier la radiométrie), ou s'insérer entre des structures (tumeur refoulante). Dans ce dernier cas en particulier, les structures cérébrales peuvent être déplacées, déformées voire détruites. L'aspect des tumeurs varie également selon qu'elles sont nécrotiques ou provoquent l'apparition d'un œdème. D'une manière générale, l'aspect, la localisation et le comportement spatial des tumeurs varient, ce qui rend difficile une modélisation des tumeurs.

Nous allons présenter les différents modes de représentation des images cérébrales, puis nous allons présenter deux grandes familles de méthodes pour la segmentation des structures cérébrales. La première famille correspond aux méthodes modélisant les structures cérébrales ou leurs caractéristiques. Dans ces méthodes, l'agencement spatial des structures est en général induit par le modèle, c'est-à-dire pas exprimé de manière directe. La deuxième famille de méthodes propose d'utiliser une représentation structurelle de l'anatomie et se focalise moins sur les caractéristiques

²Internet Brain Segmentation Repository. The MR brain data sets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at <http://www.cma.mgh.harvard.edu/ibsr/>

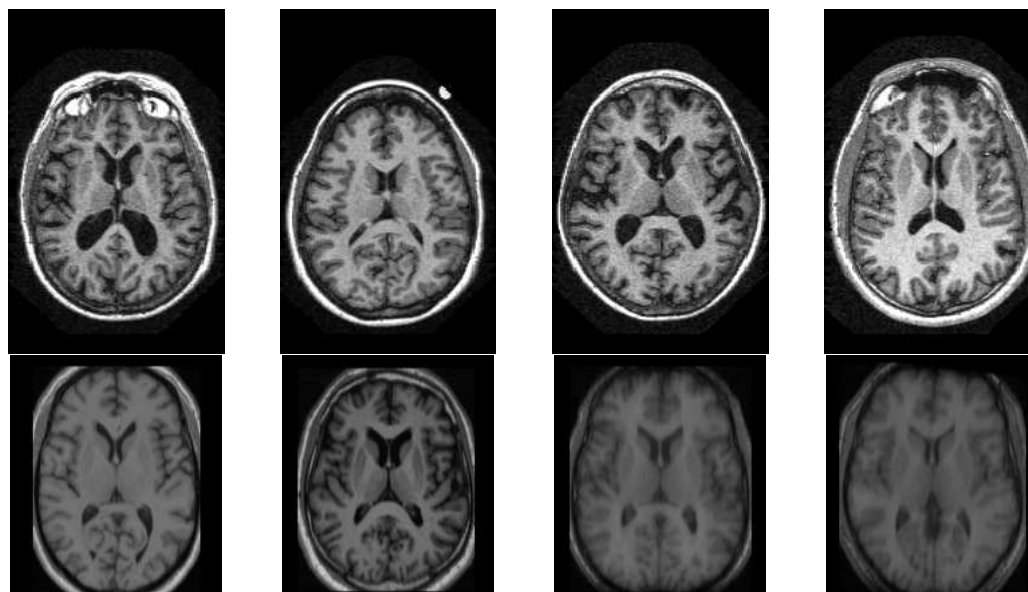


FIG. 1.3 – Coupes d’IRM cérébrales (T1) de la base OASIS (en haut, les coupes sont proches mais ne sont pas exactement les mêmes. [Marcus et al. \(2007\)](#)) et la base IBSR (en bas).

des structures. Les approches que nous proposons dans ces travaux se situent dans cette deuxième famille.

1.2.1 Représentations de l’anatomie cérébrale

L’anatomie cérébrale dispose de descriptions anatomiques ([Waxman \(2000\)](#); [Hasboun \(2005\)](#)) sous forme de nomenclature ou d’atlas morphologique et fonctionnel. Ces nomenclatures permettent l’identification de structures, et également de faire le lien entre les différents noms possibles d’une structure. Elles peuvent également contenir des informations sur les caractéristiques des structures.

Plusieurs descriptions anatomiques sous forme de hiérarchie ont été proposées dans la littérature, nous en présentons à présent quelques unes.

Neuronames :

Neuronames ([Bowden et Dubach \(2003, 2005\)](#)) propose une hiérarchie de l’anatomie cérébrale où le cerveau (humain et macaque) est décomposé en 550 structures dites « primaires » et près de 850 éléments au total. Des définitions, des synonymes et des traductions des noms de chaque structure sont proposés. Une présentation de la hiérarchie ainsi qu’un navigateur est proposé à cette adresse : <http://braininfo.rprc.washington.edu/>. Une capture du navigateur est présentée dans la figure 1.4.

MEsH (« Medical subject headings ») :

MEsH ([Lipscomb \(2002\)](#)) propose des descriptions médicales de plus de 25000 termes, et en particulier des structures cérébrales. Les descriptions sont accessibles par ordre alphabétique ou dans une hiérarchie sous forme d’arbre. Une définition pour chaque structure est fournie. Un navigateur, dont une capture est présentée par la figure 1.5, est disponible à cette adresse : <http://www.nlm.nih.gov/mesh/MBrowser.html>.

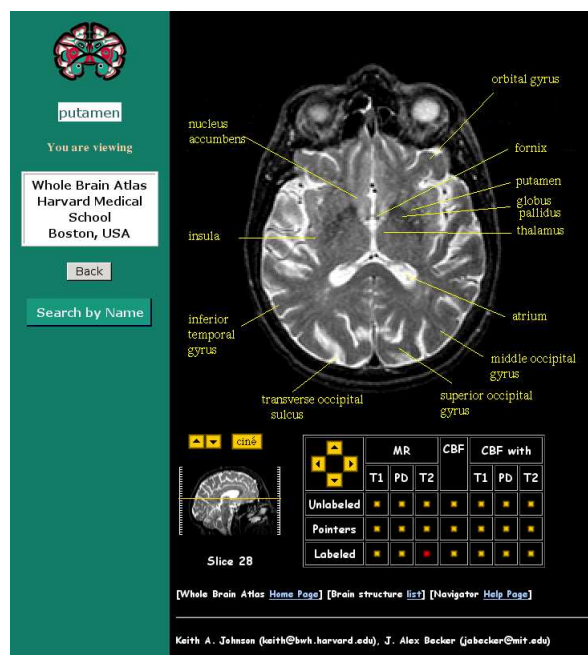


FIG. 1.4 – Une capture d'écran du navigateur dans la hiérarchie neuro-
names qui est hébergée par l'université de Washington à cette adresse :
<http://braininfo.rprc.washington.edu/>.

National Library of Medicine - Medical Subject Headings

2010 MeSH

MeSH Descriptor Data

[Return to Entry Page](#)

Standard View. [Go to Concept View](#). [Go to Expanded Concept View](#)

MeSH Heading	Caudate Nucleus
Tree Number	A08.186.211.730.885.287.249.487.550.184
Annotation	part of the neostriatum
Scope Note	Elongated gray mass of the neostriatum located adjacent to the lateral ventricle of the brain.
Allowable Qualifiers	AB AH BS CH CY DE EM EN GD IM IN ME MI PAPH PP PS R RE RI SE SU TR UL US VI
Date of Entry	19990101
Unique ID	D002421

MeSH Tree Structures

[Nervous System \[A08\]](#)

[Central Nervous System \[A08.186\]](#)

[Brain \[A08.186.211\]](#)

[Prosencephalon \[A08.186.211.730\]](#)

[Telencephalon \[A08.186.211.730.885\]](#)

[Cerebrum \[A08.186.211.730.885.287\]](#)

[Basal Ganglia \[A08.186.211.730.885.287.249\]](#)

[Corpus Striatum \[A08.186.211.730.885.287.249.487\]](#)

[Neostriatum \[A08.186.211.730.885.287.249.487.550\]](#)

[Caudate Nucleus \[A08.186.211.730.885.287.249.487.550.184\]](#)

[High Vocal Center \[A08.186.211.730.885.287.249.487.550.484\]](#)

[Putamen \[A08.186.211.730.885.287.249.487.550.784\]](#)

FIG. 1.5 – Une capture d'écran du navigateur de la base de la « Na-
tional Library of Medicine ». L'adresse du navigateur est la suivante :
<http://www.nlm.nih.gov/mesh/MBrowser.html>.

FMA (« Foundational Model of Anatomy ») :

L'ontologie de la FMA (Rosse et Mejino (2007)), disponible à cette adresse : <http://sig.biostr.washington.edu/projects/fm> vise à regrouper les représentations des classes et des types des relations nécessaires à une représentation symbolique de la structure du corps humain. Cette représentation de la connaissance anatomique n'est pas sous une forme d'arbre, mais peut être vue de plusieurs points de vue, et avec différents niveaux de granularité. Un explorateur « Foundational Model Explorer » est disponible à cette adresse : <http://fme.biostr.washington.edu>. Une capture est présentée sur la figure 1.6. Chaque structure apparaît avec ses relations aux autres structures éventuellement à différentes granularités. Des informations sur la structure, comme sa définition, ses sous-parties, ou ses caractéristiques morphologiques sont données.

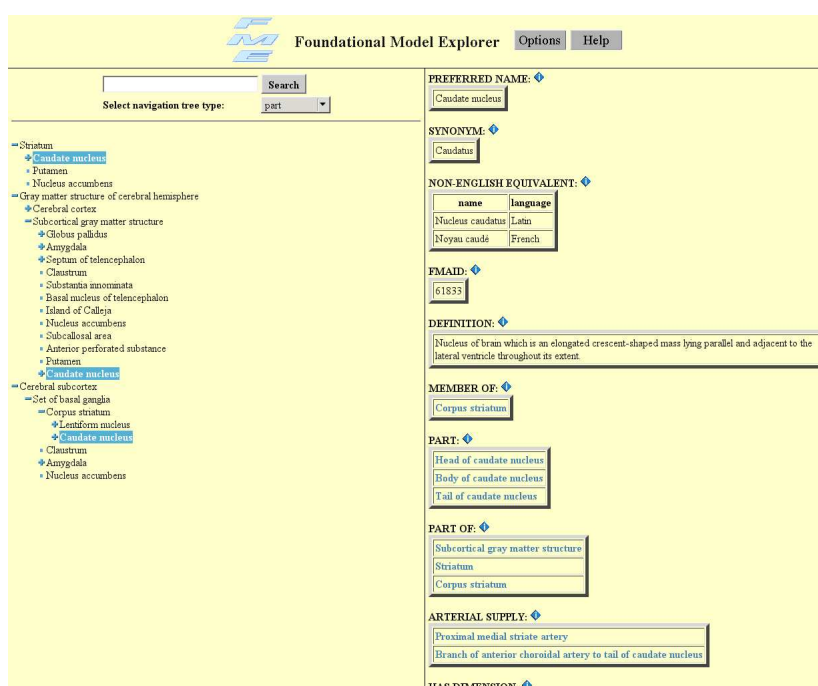


FIG. 1.6 – Une capture d'écran de la FMA disponible à l'adresse : <http://fme.biostr.washington.edu>.

Neuranat :

Neuranat (Hasboun (2005)) est un site plutôt dédié à l'enseignement de la neuroanatomie et qui ne propose pas de hiérarchie des structures. Cependant, il propose des atlas morphologiques et fonctionnels du cerveau, ainsi que des vidéos et des animations autour du sujet. Les atlas sont disponibles à cette adresse : <http://www.chups.jussieu.fr/ext/neuranat>. La figure 1.7 présente une vue de l'atlas IRM en trois dimensions.

1.2.1.1 Atlas et modèles de forme

Si les IRM cérébrales présentent une grande variabilité au niveau des caractéristiques des structures cérébrales, la structure de l'anatomie cérébrale présente une grande régularité et a permis l'élaboration d'atlas anatomiques et fonctionnels (Talairach et Tournoux (1988)).

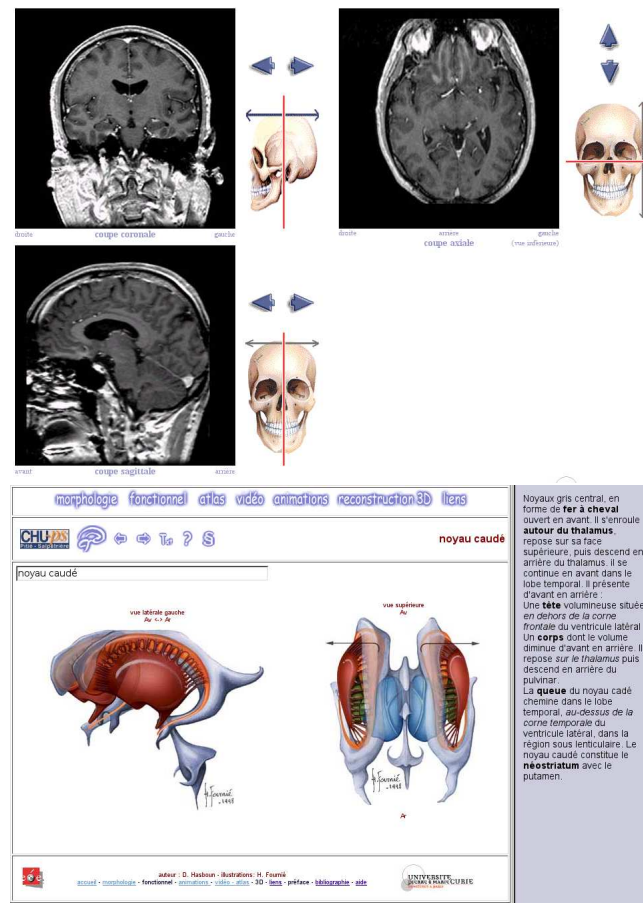


FIG. 1.7 – Deux captures d’écran du site neuranat. En haut une vue de de l’atlas 3D IRM. En bas la description d’une structure (le noyau caudé). Le site est accessible à cette adresse : <http://www.chups.jussieu.fr/ext/neuranat>.

Les atlas sont des représentations moyennes des structures anatomiques, qui peuvent être générées de différentes manières. Leur utilisation pour guider la reconnaissance des structures consiste à effectuer une mise en correspondance de l’atlas vers l’image à reconnaître. Les modèles de forme proposent un apprentissage distinct des formes de chacune des structures, afin d’être plus représentatifs de chacune.

Atlas probabilistes et atlas moyens Les premières méthodes utilisant un atlas utilisent en pratique une unique image annotée manuellement. Utiliser une unique image limite bien entendu la variabilité usuelle des structures, et empêche la représentation des particularités (par exemple, certaines circonvolutions du cortex n’apparaissent pas chez tous les sujets). Les atlas probabilistes et les atlas moyens cherchent à représenter la variabilité en fusionnant l’information provenant de différentes images annotées manuellement.

La génération d’un atlas probabiliste, tel l’atlas ICBM (Mazziotta *et al.* (1995)), consiste à générer une carte de probabilité par structure, à partir du recalage affine d’un ensemble de cas segmentés manuellement. Pour chaque carte obtenue, la probabilité reflète le nombre d’occurrences après recalage de la structure en ce point. Une image peut alors être générée représentant l’atlas. Pour l’atlas ICBM, 452 images en pondération T1 de jeunes adultes ont été utilisées. Une repré-

sensation moyenne de tous les éléments de la base permet d'accroître la représentativité de la base et de prendre en compte la variabilité normale des structures. Mais cette gestion de la variabilité s'effectue au détriment de la précision. De plus, les images moyennes pour chaque structure sont floues.

Les atlas moyens essaient de remédier à ce problème, en proposant d'effectuer un recalage de groupe (Guimond *et al.* (2000); Joshi *et al.* (2004); Bhatia *et al.* (2004); Blezek et Miller (2007)). L'objectif est d'extraire un atlas moyen d'un groupe de sujets de la base de cas annotés manuellement, composé de manière à ce qu'il minimise la déformation à réaliser pour être mis en correspondance avec tous les éléments de la base, c'est-à-dire que pour chaque élément de la base, la déformation par rapport à l'atlas moyen est minimisée.

Les atlas moyens permettent de mieux représenter la base d'apprentissage, mais il est toujours difficile, en moyennant l'information, de représenter des singularités de la base. Dans les cas sains, pour améliorer la représentativité, il est possible de ne pas extraire un unique atlas moyen, mais tout un ensemble d'atlas qui soient les plus représentatifs possibles (Blezek et Miller (2007)). Mais cette méthode est coûteuse, en particulier si le nombre d'atlas est grand.

Mise en correspondance d'atlas Dans le cas des premières méthodes utilisant comme atlas une unique image (Broit (1981); Iosifescu *et al.* (1997); Dawant *et al.* (1999b)), la mise en correspondance entre l'atlas et l'image à reconnaître peut être vue comme un problème de recalage entre deux images. Les variations n'étant pas identiques pour toutes les parties, le recalage n'est pas linéaire.

Pour les atlas probabilistes, plusieurs méthodes ont été proposées pour réaliser la mise en correspondance. Elle peut être effectuée à partir d'une classification initiale de l'image (Collins *et al.* (1999)), ou encore en utilisant une estimation du maximum a posteriori (MAP) par un algorithme de type espérance-maximisation (EM) (Pohl *et al.* (2002, 2006)). La mise en correspondance d'un atlas moyen peut être effectuée avec le même type de méthodes.

Dans les cas pathologiques, il est nécessaire d'adapter le modèle. Une première approche (Dawant *et al.* (1999a, 2002)) consiste à introduire la tumeur dans l'atlas. Cela peut être effectué en y plaçant une graine dont la radiométrie est celle de la tumeur. Dans le cas où la tumeur est refoulante, c'est-à-dire qu'elle va déplacer des structures, alors les déformations induites peuvent alors être modélisées. Une deuxième approche (Kyriacou *et al.* (1999); Mohamed *et al.* (2006); Zacharaki *et al.* (2008)) consiste à modéliser finement l'anatomie (notamment les propriétés biomécaniques de ses tissus) ainsi qu'un modèle de croissance de la tumeur, afin de proposer une modélisation des déformations subies. Dans toutes ces méthodes, l'information structurelle reste codée de manière implicite et est donc difficile à utiliser.

Modèles de formes Les modèles de forme proposent de modéliser les principaux modes de variations de chaque structure. Les premiers travaux (Cootes *et al.* (1995, 2001)) représentent les contours d'une structure par un ensemble de points. Les différents contours obtenus pour une même structure dans la base sont alignés et mis en correspondance. Il est alors possible d'effectuer une analyse en composantes principales (ACP). Les vecteurs propres obtenus représentent les différents modes de variation de la forme. Dans ces modèles, il est en général considéré que ces modes de variation suivent une loi normale multidimensionnelle (Leventon *et al.* (2000); Cremers *et al.* (2002)) et que toute forme de cette famille peut être exprimée comme une combinaison linéaire de la forme moyenne et des vecteurs propres (qui représentent l'écart-type du mode de déformation représenté par le vecteur propre). La probabilité d'une forme peut être obtenue à partir des coefficients de la combinaison linéaire. La reconnaissance peut alors être exprimée comme l'obtention

des paramètres du modèle de localisation et des coefficients associés aux composantes principales correspondant au cas à reconnaître.

La mise en correspondance de tous les contours d'une forme sur la base d'apprentissage est coûteuse. D'autres travaux ont donc substitué au contour une carte de distance signée ([Leventon et al. \(2000\)](#)). Avec ce type de représentation, le modèle de forme obtenu peut être intégré naturellement dans un modèle déformable comme une contrainte de ce modèle.

L'agencement spatial peut être pris en compte de manière non explicite en étendant l'approche précédente de manière à effectuer l'apprentissage joint de plusieurs formes ([Yang et Duncan \(2004a,b\)](#) et [Tsai et al. \(2003, 2004\)](#)). Dans ce cas, l'ACP est effectuée sur une concaténation des cartes de distance de toutes les formes prises en compte. Une formulation bayésienne est proposée dans ([Yang et Duncan \(2004b\)](#)) pour effectuer la segmentation en prenant en compte la contrainte multi-formes.

Les modèles de formes permettent d'améliorer le processus de segmentation en contraignant le résultat à correspondre à un petit nombre de formes. Cependant, s'ils peuvent prendre en compte la variabilité anatomique dans les cas sains, ces modèles peuvent difficilement être adaptés aux cas pathologiques, qui présentent une variabilité trop importante pour être correctement modélisés par ce type de modèles. De même que pour les méthodes fondées sur un atlas, l'agencement structurel reste codé de manière implicite, et reste donc difficile à utiliser.

1.2.1.2 Représentation structurelle de l'anatomie cérébrale

L'anatomie cérébrale peut être naturellement représentée de manière hiérarchique où chaque niveau correspond à une décomposition du niveau précédent. Par exemple, l'hémisphère droit contient (entre autres) le cortex droit et la matière blanche droite. Les nomenclatures présentées dans cette partie proposent en général une hiérarchie de l'information anatomique.

En particulier, l'anatomie cérébrale peut être représentée sous forme d'un graphe dont les nœuds correspondent aux structures cérébrales et dont les arcs décrivent des relations entre ces structures, à l'aide de relations spatiales.

Un premier modèle hiérarchique de l'anatomie cérébrale a été présenté par O. Colliot dans ses travaux de thèse ([Colliot \(2003\)](#)) en collaboration avec un neuroanatomiste D. Hasboun. La représentation est un hyper-graphe hiérarchique. Les relations entre deux niveaux du graphe sont des relations de composition et forment un graphe bi-partite. La structure est arborescente, le cerveau étant au premier niveau. Un hyper-graphe est utilisé afin de pouvoir représenter des relations spatiales ternaires comme la relation « Entre ».

Cette représentation a été étendue par la suite ([Hudelot et al. \(2006\)](#); [Atif et al. \(2007b\)](#)) vers une représentation appelée GRAFIP (« Graph of Representation of Anatomical and Functional data for Individual patients including Pathologies »). Cette représentation intègre, en plus des informations structurelles sur l'anatomie :

- des informations sur la composition des tissus, permettant d'en déduire des conséquences sur sa radiométrie ;
- de la connaissance fonctionnelle ;
- de la connaissance sur les pathologies issue des classifications WHO ([Smirniotopoulos \(1999\)](#)) et de l'hôpital Sainte-Anne ([Daumas-Duport \(1992\)](#)).

L'objectif est d'intégrer des connaissances issues de l'image dans une base de connaissances symboliques, l'ontologie de la FMA par exemple, ou dans un dossier patient. De plus, cette représentation permet une meilleure exploitation de la connaissance dans un processus de reconnaissance. La figure 1.10 présente le schéma des connaissances intégrées dans le GRAFIP.

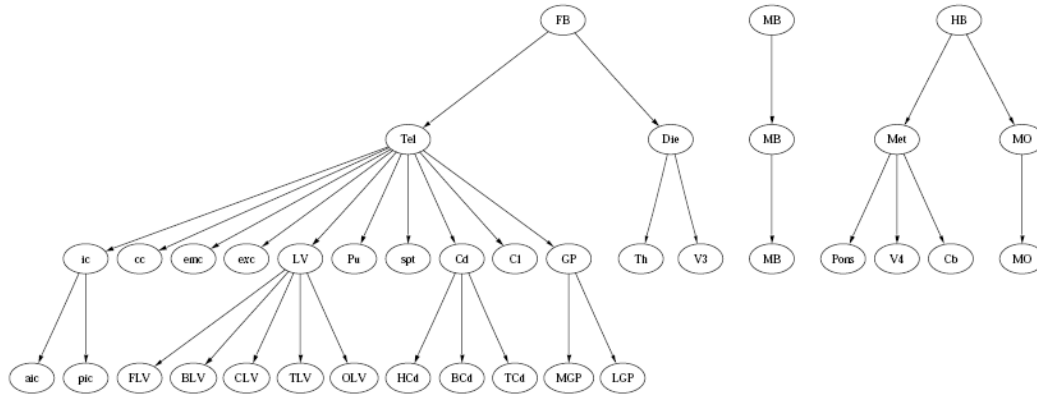


FIG. 1.8 – Les trois premiers niveaux du graphe hiérarchique proposé par Colliot (2003). Seules les relations entre niveaux sont présentées. Les structures du premier niveau correspondent au Prosencéphale (FB), Mésencéphale (MB) et au rhombencéphale (HB) [Figure extraite de Colliot (2003)].

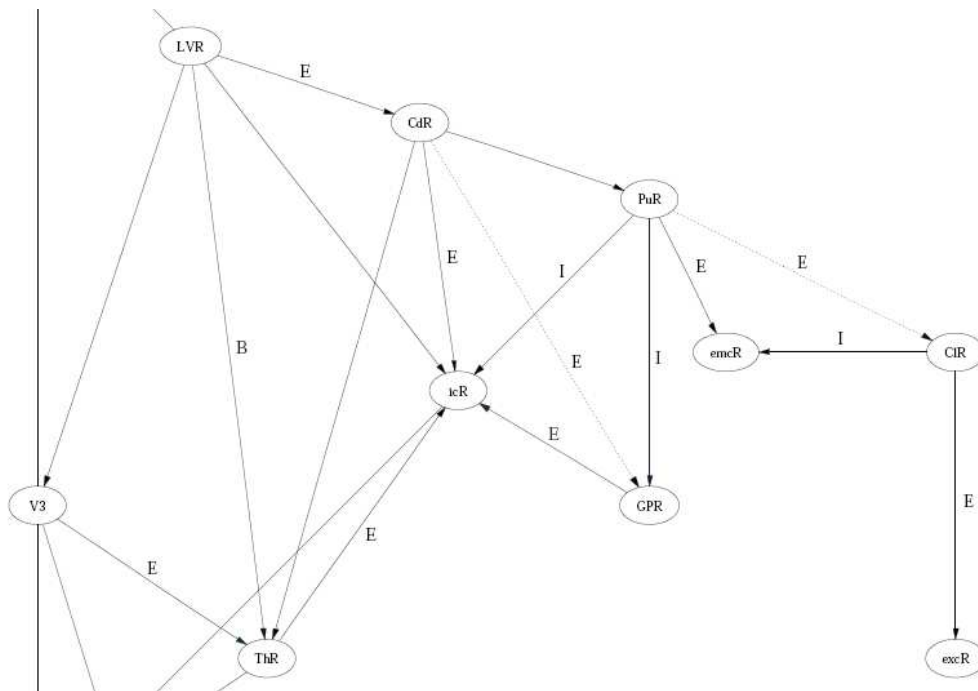


FIG. 1.9 – Un extrait du troisième niveau du graphe hiérarchique proposé par Colliot (2003). Les relations entre les structures sont les suivantes : extérieur (E), intérieur (I), haut (B), bas (B), en avant (Av), en arrière (Ar). [Figure extraite de Colliot (2003)].

Information spatiale Les hiérarchies de l'anatomie cérébrale utilisent des relations topologiques pour décrire la structure avec des relations telles que l'inclusion ou l'adjacence. Les relations entre les structures d'un niveau similaire peuvent être décrites en utilisant des relations spatiales métriques comme la direction ou l'orientation, ou encore des relations plus complexes comme la relation « entre ». L'imprécision intrinsèque des relations spatiales décrites de manière textuelle, par exemple *le noyau caudé est proche du ventricule latéral*, permet de gérer la variabilité

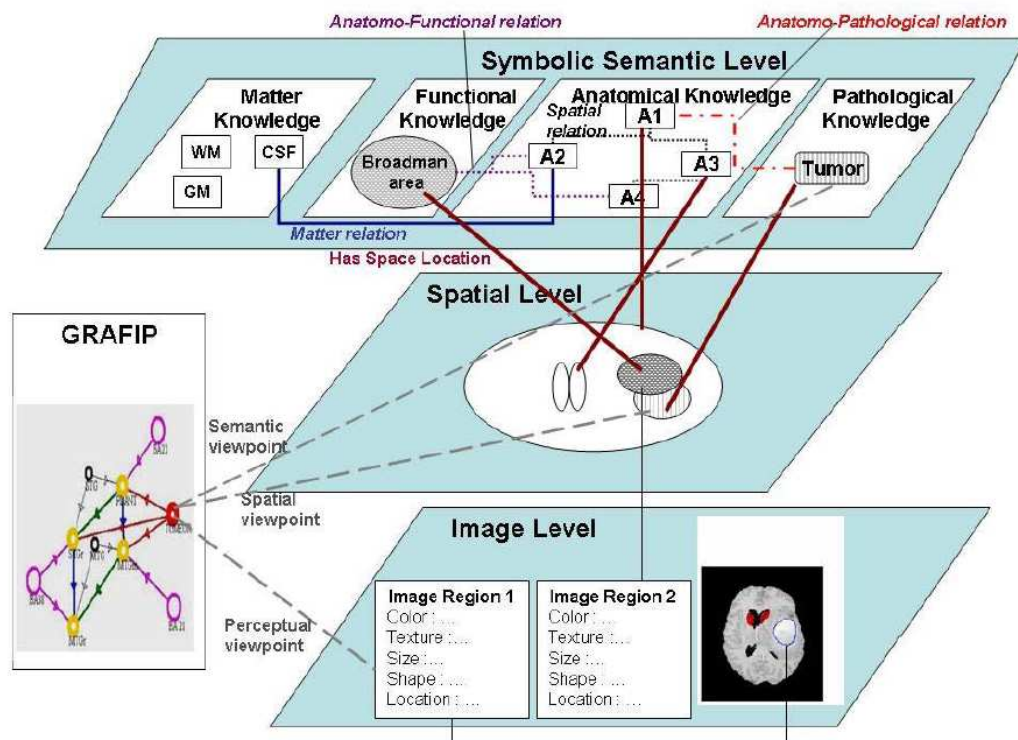


FIG. 1.10 – GRAFIP (Hudelot *et al.* (2006); Atif *et al.* (2007b)). Le modèle contient le modèle structurel de l'anatomie ainsi que des connaissances symboliques permettant d'intégrer toutes les informations dans un processus de reconnaissance [figure extraite de Atif *et al.* (2007b)].

naturelle de l'anatomie cérébrale. La figure 1.11 présente la hiérarchie proposée par Hudelot *et al.* (2008).

1.2.2 Reconnaissance avec un modèle structurel

L'agencement spatial des structures cérébrales est stable en particulier dans le cas sain. Dans le cas pathologique, l'agencement global reste relativement stable et dans ce cas, la stabilité des relations peut être estimée en fonction du type de pathologie et du type de relation (Atif *et al.* (2006a)). L'agencement spatial a donc été utilisé dans des processus de reconnaissance et de segmentation des structures sous-corticales.

Nous allons présenter trois types d'approches pour effectuer la reconnaissance des structures cérébrales en utilisant ces représentations. Le premier type de méthode correspond à une mise en correspondance d'une segmentation avec une représentation structurelle modélisée par un graphe. Le deuxième type de méthode est une approche séquentielle pour la segmentation et la reconnaissance des structures cérébrales. Nos travaux se situent dans cette deuxième catégorie. La dernière approche est une approche plus globale utilisant un réseau de contraintes dérivées du modèle structurel et qui a été proposée par Olivier Nempont dans ses travaux de thèse (Nempont (2009)).

1.2.2.1 Segmentation et reconnaissance par mise en correspondance du modèle

Le problème de reconnaissance des structures cérébrales peut être vu comme un problème de mise en correspondance entre deux graphes : la représentation structurelle modélisée par un

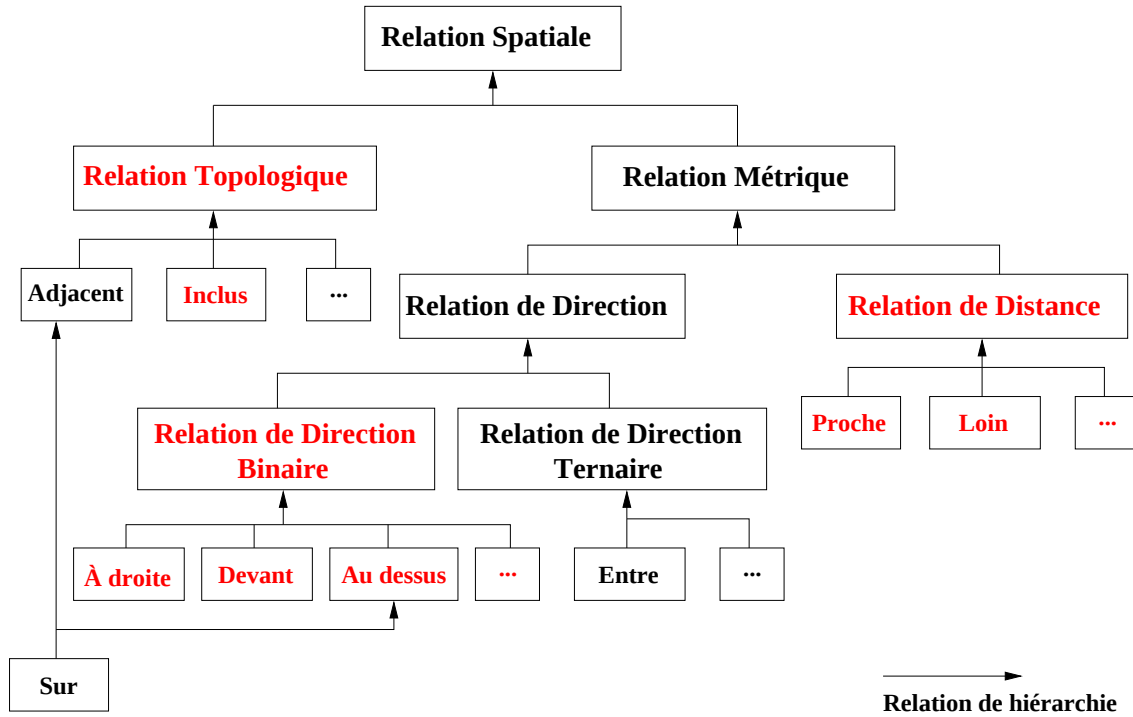


FIG. 1.11 – Une partie de la hiérarchie des relations spatiales proposée par [Hudelot et al. \(2008\)](#)[figure extraite de [Hudelot et al. \(2008\)](#)].

graphe d'une part, et une sur-segmentation de l'image à reconnaître à partir de laquelle nous pouvons extraire un graphe d'autre part. Le problème de mise en correspondance de graphes est un problème complexe qui a fait l'objet de beaucoup de travaux. On pourra se reporter à [Conte et al. \(2004\)](#); [Bunke \(2000\)](#) pour une revue de ces travaux.

Une approche par mise en correspondance de graphes a été développée par [Perchant \(2000\)](#); [Perchant et Bloch \(2002\)](#), qui proposent de trouver un morphisme flou entre un graphe modèle, créé à partir d'une image annotée manuellement, et une image à reconnaître sur-segmentée. Dans cette approche, les attributs sont représentés par des ensembles flous. La sur-segmentation comportant plus de nœuds que le graphe modèle, il s'agit d'une mise en correspondance inexacte et plusieurs nœuds de l'image sur-segmentée sont attribués à une même structure du graphe modèle. Ce problème de mise en correspondance de graphes est généralement NP-complet. Différentes approches d'optimisation, permettant de trouver une solution sous-optimale, ont été proposées par [Perchant \(2000\)](#) dont des algorithmes génétiques et une formulation bayésienne. Un algorithme d'estimation de distribution a été ensuite proposé par [Bengoetxea et al. \(2002\)](#), puis une recherche par arbre par [Cesar et al. \(2005\)](#).

Une autre approche ([Deruyver et Hodé \(1997\)](#); [Hodé et Deruyver \(2007\)](#)) utilise une sur-segmentation pour effectuer la reconnaissance des structures. Le problème est formulé comme un problème de satisfaction de contraintes à deux niveaux. Des contraintes binaires sont calculées entre les ensembles de régions regroupés dans un même nœud du modèle. D'autres contraintes sont calculées entre les régions regroupées dans un même nœud. Un algorithme de propagation adapté est proposé pour résoudre le problème bi-contraint. Une extension récente ([Deruyver et al. \(2009\)](#)) reformule le problème comme la mise en correspondance entre une image sur-segmentée et une ontologie représentée sous forme de graphe à l'aide d'un graphe conceptuel. Une extension de l'algorithme de consistance bi-contraint est proposée afin de ne plus être limité à des mises en

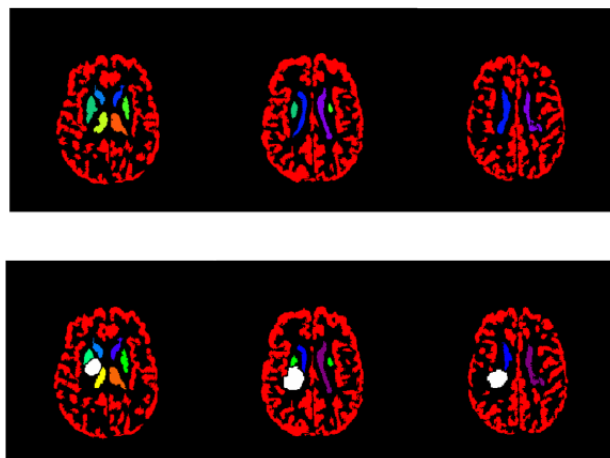


FIG. 1.12 – Résultat d’interprétation d’une image cérébrale par mise en correspondance de graphes formulée comme un problème bi-contraint proposé par [Deruyver et al. \(2009\)](#). Le résultat de la reconnaissance des noyaux gris est présenté en haut sur différentes coupes. En bas, le résultat de la reconnaissance avec l’apparition d’une tumeur qui n’était pas attendue, mais qui a été détectée par le processus et ajoutée dans le graphe. [Figure extraite de [Deruyver et al. \(2009\)](#)].

correspondance surjectives, permettant ainsi l’apparition de nouvelles structures qui n’étaient pas présentes dans le graphe conceptuel. Cette extension est particulièrement adaptée au cas de pathologies en imagerie cérébrale. La figure 1.12 présente un résultat d’interprétation d’une image cérébrale avec cette approche.

1.2.2.2 Approche itérative de la segmentation

La sur-segmentation utilisée dans certaines des approches précédentes ne garantit pas de fournir une solution initiale correcte, en particulier à cause de la radiométrie des structures cérébrales parfois difficiles à différencier de la matière qui les entoure. Les approches itératives permettent de s’affranchir de la sur-segmentation, en réalisant en même temps la segmentation et la reconnaissance des structures, et cela de manière séquentielle. Le processus débute en commençant par les structures qui sont les plus « aisées » à segmenter, comme les ventricules qui présentent un fort contraste avec les matières adjacentes, puis à chaque itération l’information spatiale et les segmentations précédentes permettent de guider le processus pour reconnaître les structures suivantes.

Une première approche a été proposée dans [Géraud et al. \(1999, 2000\)](#); [Bloch et al. \(2003\)](#) où la segmentation est effectuée dans une zone d’intérêt définie par les relations spatiales avec un processus automatique de classification des pixels de l’image, puis recalée sur un patron de la structure.

Afin de s’affranchir des patrons utilisés dans cette approche, la segmentation a été modifiée par [Colliot et al. \(2006\)](#) pour utiliser un modèle déformable utilisant les relations spatiales pour contraindre le modèle. Une extension géodésique de cette approche de la segmentation a été formulée ensuite par [Atif et al. \(2006b\)](#). Nos travaux reposent sur cette approche et utilisent cette formulation du problème de segmentation, qui est détaillée dans le chapitre 5.

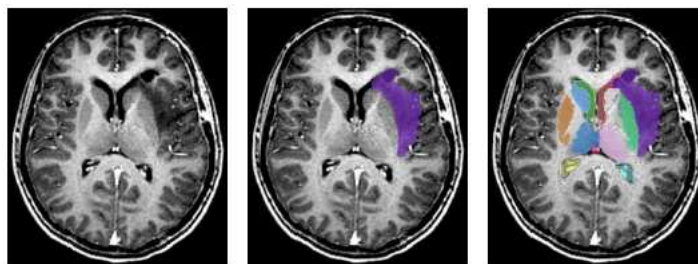


FIG. 1.13 – Résultats de la reconnaissance des structures sous-corticales par la méthode proposée par Nempont (2009) sur un cas pathologique. Le système part d'un graphe complet de contraintes qui prévoit la possibilité d'une tumeur refoulante dans l'image observée. À l'aide d'un algorithme de propagation de contraintes, les domaines de chaque structure sont réduits, et une solution peut être extraite à l'aide d'un algorithme de surface minimale. [Figure extraite de Nempont (2009)].

1.2.2.3 Approche globale par contraintes

Les travaux de thèse d'Olivier Nempont (Nempont (2009)) proposent une autre formulation du problème de la segmentation et de la reconnaissance des structures sous-corticales par un réseau de contraintes, mais sans effectuer une mise en correspondance de graphes. Ces travaux utilisent la représentation définie par (Hudelot *et al.* (2006); Atif *et al.* (2007b)). Le but recherché est d'associer à chaque structure anatomique recherchée une région de l'espace satisfaisant à l'ensemble des relations du modèle, les contraintes étant dérivées du modèle structurel.

Le problème étant trop complexe pour être directement résolu, la solution est obtenue en deux étapes. Tout d'abord, à l'aide d'un algorithme de propagation de contraintes : les bornes du domaine de chaque variable sont réduites en supprimant toutes les valeurs qui ne peuvent pas être solution du problème. Ensuite, lorsque les domaines ont été réduits grâce au réseau de contraintes, une solution approximative (au sens des contraintes) est extraite des valeurs restantes.

Les cas pathologiques ont nécessité une adaptation du processus de reconnaissance, puisque le modèle ne correspond plus à l'image. L'adaptation se limite à des pathologies peu refoulantes. Tous les modèles spécifiques (avec des pathologies) ont été pris en compte. Le processus doit alors effectuer la reconnaissance des structures comme précédemment, mais il va en plus, au cours du processus, supprimer des hypothèses sur le modèle spécifique. La reconnaissance des structures cérébrales et l'identification du modèle adapté à la pathologie sont donc effectuées de manière simultanée. Un résultat de reconnaissance avec un cas pathologique est présenté dans la figure 1.13.

1.3 Conclusion

La reconnaissance des structures cérébrales est donc une tâche complexe qui nécessite un modèle. En particulier, l'agencement spatial est une connaissance stable qui a été utilisée dans plusieurs approches. L'importance de l'information spatiale et les modèles structurels disponibles font de cette tâche un domaine d'application adéquat pour nos travaux.

Les méthodes reposant sur une mise en correspondance d'un modèle structurel avec une image représentée comme un graphe dépendent en général d'une sur-segmentation qui ne garantit pas de donner une solution initiale satisfaisante, dans le cas où une structure ne peut être différenciée de la matière qui l'entoure par exemple, comme c'est le cas du thalamus dans certaines coupes. Il est donc intéressant de se passer de cette segmentation initiale.

L'approche globale proposée par O. Nempont dans ses travaux de thèse ([Nempont \(2009\)](#)) permet de s'affranchir de cette segmentation. Cependant, le modèle utilisé possédant un grand nombre de contraintes, la complexité de la tâche est assez grande. L'utilisation d'une segmentation de structures d'une taille importante (relativement) et placée prêt du centre du cerveau comme solution initiale à cette approche peut permettre de simplifier la tâche, en diminuant beaucoup les domaines initiaux. De plus, la détection d'un cas pathologique et la segmentation au préalable de la tumeur peut également simplifier le problème.

Les approches itératives permettent également de s'affranchir de la segmentation initiale. L'approche d'O. Colliot ([Colliot \(2003\)](#)) utilise cependant une séquence de segmentation ad hoc et qui peut nécessiter une adaptation, en particulier dans les cas pathologiques. Dans tous les cas, il est intéressant de tenir compte de l'information recueillie directement dans l'image au cours du processus, pour pouvoir s'adapter au cas spécifique représenté par l'image.

Nous proposons dans ces travaux d'exploiter au mieux l'information spatiale contenue dans un modèle, tel que l'agencement spatial des structures cérébrales, pour guider la reconnaissance. L'approche que nous proposons dans le chapitre 4 nécessite un modèle structurel, ainsi qu'au moins une image annotée manuellement, afin de déterminer selon le modèle, à partir d'une structure de référence, quelle est la meilleure séquence de segmentation à effectuer pour atteindre une structure cible. Cette approche nous fournit une information supplémentaire, indépendamment de la méthode choisie pour la reconnaissance des structures par la suite.

Dans une deuxième approche, présentée dans le chapitre 5, nous proposons d'intégrer un mécanisme pré-attentionnel à un processus de segmentation séquentielle qui est vu comme une exploration progressive de l'image. L'exploration repose sur l'information spatiale, comme les approches itératives décrites précédemment, le mécanisme pré-attentionnel étant là pour guider la sélection des structures à segmenter. Cette approche peut donc être naturellement intégrée dans une approche telle que celle proposée par [Colliot \(2003\)](#).

Chapitre 2

Les mécanismes de l'attention

Dans ce chapitre, nous nous intéressons aux modélisations du système visuel humain : nous décrivons les différentes phases pendant lesquelles nous retirons des informations d'une scène, nous permettant de l'analyser, de reconnaître les objets la composent, etc. En outre, cette analyse est effectuée sans que nous ayons besoin d'y penser, de manière automatique. L'analyse d'une scène par un système dans un but d'interprétation ou de reconnaissance des objets n'est pas aussi facile.

De nombreux travaux s'inspirent du système visuel humain, que ce soit dans le but de mieux comprendre ce système, ou bien d'utiliser certains mécanismes bio-inspirés dans des tâches de reconnaissance ou d'interprétation. Nous nous intéressons également aux informations qui peuvent être extraites d'une scène, d'une image, afin de les intégrer dans un raisonnement qui utilise l'information spatiale.

Nous présentons dans ce chapitre les mécanismes attentionnels et pré-attentionnels, où et comment ils interviennent dans un système de vision, et comment nous pouvons les utiliser dans le cadre de l'application à la segmentation des structures cérébrales. L'intégration de cette information dans un système de segmentation est étudiée dans le chapitre 4.

Le chapitre commence par une présentation générale de la notion d'attention afin de comprendre le besoin d'un système pré-attentionnel, qui sera présenté dans la section 2.2. Nous présenterons alors plus en détail une implémentation du système pré-attentionnel que nous utiliserons, les cartes de saillance, dans la section 2.3 et enfin, nous détaillerons dans la section 2.4 les adaptations que nous proposons à la génération des cartes de saillance pour les calculer sur des images IRM en trois dimensions.

2.1 Qu'est ce que l'attention ?

Les mécanismes de la vision ont été étudiés depuis longtemps, à commencer par les mouvements des yeux. En effet il a été démontré que les yeux réalisent une saccade de mouvements (« overt attention » par opposition à « covert attention » où l'attention peut se déplacer sans entraîner de mouvement de l'œil). Yarbus (1967) a mis en évidence, à l'aide d'un dispositif expérimental consistant à placer un dispositif sur l'œil permettant d'en suivre les mouvements lors de l'exploration libre d'une image. Cette expérience est illustrée par la figure 2.1. Les mouvements des yeux ont également été suivis par le même dispositif, mais lors d'une exploration particulière d'une scène, où l'observateur doit remplir une tâche particulière comme compter le nombre de personnages dans une scène par exemple.

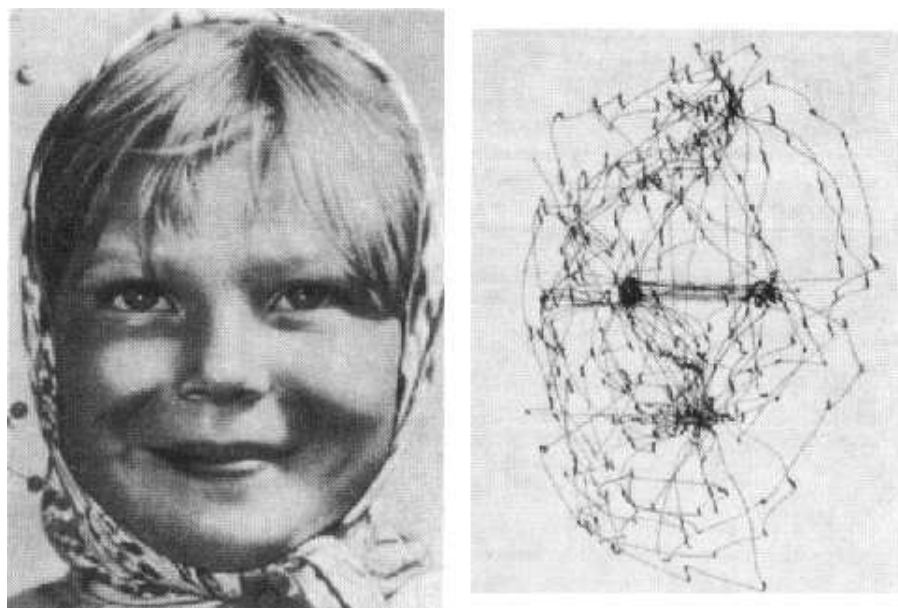


FIG. 2.1 – Mouvement des yeux lors de l'exploration d'une scène. Source : [Cotteret \(2005\)](#) d'après [Yarbus \(1967\)](#).

2.1.1 Définition de l'attention

L'attention est une notion usuelle et connue du plus grand nombre. En revanche les définitions pour la décrire restent toutefois relativement vagues et imprécises. La définition la plus classique est celle de [James \(1890\)](#) :

« Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatterbrained state which in French is called distraction, and *Zerstreuung* in German ».

[Posner \(1980\)](#) a présenté plusieurs expériences permettant de mettre en évidence l'attention visuelle et qui ont permis de définir les premières théories de l'attentionnel. Des propriétés de l'attention visuelle ont été énoncées par [Pashler \(1998\)](#) :

la sélectivité : c'est-à-dire privilégier certains stimuli au détriment d'autres ;

la limitation de capacité : c'est-à-dire comment traiter des stimuli différents en même temps ;

l'effort : une attention soutenue des mêmes stimuli visuels faisant ressentir la sensation d'un effort.

À partir de ces propriétés, nous pouvons revenir à une notion usuelle, où l'attention visuelle permet d'analyser certains stimuli de manière soutenue, ce qui peut impliquer un effort. L'analyse de certains stimuli pouvant s'effectuer en faisant totalement abstraction d'autres stimuli.

La plupart des travaux sur l'attention visuelle considèrent que l'attention porte sur une petite partie de l'image seulement : un faisceau attentionnel. Nous pouvons rapprocher cela du faisceau d'une lampe sur une scène sans lumière qui nous permettrait d'explorer la scène peu à peu, ou encore d'une lentille grossissante qui nous permettrait de ne regarder qu'une petite partie d'une image de manière détaillée. La figure 2.2 montre ce que perçoit l'œil lors d'une saccade oculaire

sur une scène donnée à cinquante centimètres de l'image. Le fait que seule une partie de l'image est observée à un moment donné implique un traitement séquentiel de la scène à analyser.

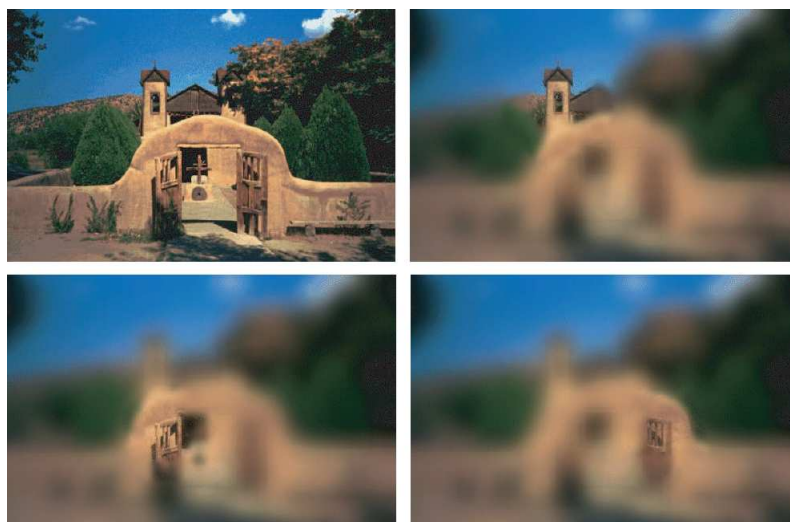


FIG. 2.2 – Ce que l’œil voit selon Machrouh. Une scène naturelle (en haut à gauche), et trois instants d’une saccade oculaire réalisée lors de la vision de cette scène. Le focus attentionnel se déplace comme un faisceau sur la scène. Source : [Cotteret \(2005\)](#) d’après [Machrouh \(2002\)](#)

2.1.2 L’unité attentionnelle

La métaphore du faisceau attentionnel nous indique que la scène est analysée de manière séquentielle, où une sélection attentionnelle est donc effectuée. La plupart des théories considèrent que cette sélection attentionnelle est spatiale. Cependant, nombre de travaux remettent en question la nature de « l’unité d’attention visuelle ».

[Posner et al. \(1980\)](#) mettent en évidence l’existence d’une sélection spatiale grâce à des signaux visuels. Ils montrent ainsi que le temps de réponse à l’apparition d’un motif particulier est d’autant plus réduit que l’on fait apparaître un signal proche de la future localisation du motif. D’autres expériences produisent des résultats similaires ([Downing et Pinker \(1985\)](#)). Mais si ces expériences montrent l’existence d’une sélection spatiale, ce n’est pas toujours le cas. Les travaux de Ulric Neisser sur le « Selective looking » ([Neisser et Becklen \(1975\)](#)) montrent que cette sélection n’est pas purement spatiale. Des résultats similaires ont été obtenus plus récemment par [Simons et Chabris \(1999\)](#). Dans cette dernière expérience, on présente aux sujets une vidéo de soixante quinze secondes dans laquelle deux équipes (en blanc ou noir) de trois personnes se passent un ballon. On demande aux participants de compter le nombre de passes. Pendant l’expérience, un événement particulier survient, une femme avec un parapluie, ou une personne déguisée en gorille passe en cinq secondes dans la scène. Le figure 2.3 montre des images de la vidéo issue de cette expérience. Quatre protocoles sont définis, avec la femme au parapluie ou le gorille comme événement, puis avec deux manières différentes de superposer l’événement avec les personnages qui se passent un ballon : dans la première les vidéos sont créées de manière indépendantes et les deux séquences sont fusionnées. Ces conditions reprennent le protocole des expériences de U. Neisser. Dans la deuxième, l’événement intervient directement au milieu des personnages. Les résultats montrent qu’une part non négligeable des observateurs, 46% toutes conditions confondues, ne détectent pas l’événement dans la scène. Cette expérience qui superpose les stimuli, tend

à montrer que si la sélection était uniquement spatiale, les événements seraient détectés.



FIG. 2.3 – Quelques images des séquences vidéos de l'expérience de [Simons et Chabris \(1999\)](#) montrant que la sélection attentionnelle n'est pas uniquement spatiale. Les observateurs des vidéos doivent compter le nombre de passes effectuées par l'une des deux équipes, blanches ou noires. On leur demande à la fin s'ils ont perçu un événement particulier pendant cette tâche. L'événement correspond soit à une femme avec un parapluie, soit à une personne déguisée en gorille. Les résultats montrent que 46% des observateurs (les 4 modalités confondues) ne parviennent pas à détecter l'événement dans la vidéo. La version en transparence vise à se rapprocher des conditions expérimentales des travaux de [Nisser et Becklen \(1975\)](#).

Le principe de superposition des stimuli permet de ne pas être dépendant de la localisation spatiale. Ce principe a également été utilisé dans les travaux de [Duncan \(1984\)](#). Ces travaux démontrent une préférence du sujet lorsqu'il doit s'intéresser à deux stimuli, si ces deux stimuli sont placés sur le même objet plutôt que dans la situation où les mêmes stimuli sont sur deux objets différents. Dans cette expérience, deux stimuli superposés sont présentés à un observateur. Le premier est une ligne qui peut être orientée de différentes manières et composée différemment (tirets, points, ...). Le deuxième est une boîte dont les dimensions varient et dont le contour est incomplet sur un côté. On demande à un sujet d'observer deux caractéristiques, soit sur le même objet, soit une caractéristique sur chacun des objets. L'expérience montre qu'il est bien plus aisé d'observer deux caractéristiques sur un même objet plutôt que sur des objets différents. La figure 2.4 présente un exemple de stimuli utilisé dans cette expérience.

On pourra se référer à [Scholl \(2001\)](#) pour avoir un panorama complet des expériences permettant la mise en évidence de l'importance de l'objet dans la sélection attentionnelle. Tous ces travaux veulent mettre en évidence le fait que l'unité attentionnelle peut être, dans certains cas, des objets discrets, et que la limite, si elle n'est pas spatiale, est plutôt le nombre d'objets qui peuvent être observés simultanément.

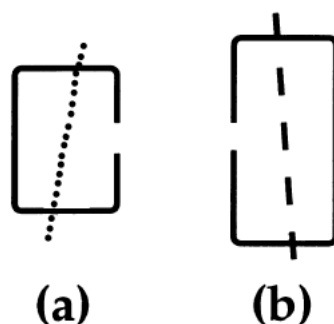


FIG. 2.4 – Les stimuli utilisés par [Duncan \(1984\)](#) pour montrer qu’il est plus aisé pour un observateur d’étudier deux caractéristiques sur un même objet plutôt qu’une caractéristique sur chacun des deux objets superposés. Chaque objet peut varier en taille ou en orientation, les motifs de la ligne également, ainsi que la position du trou dans le contour de la boîte.

2.2 Le pré-attentionnel

Les théories de l’attentionnel ont mis en évidence la sélection attentionnelle qui est effectuée sur la scène, illustrée par la métaphore du faisceau attentionnel. La sélection attentionnelle implique une exploration séquentielle d’une scène. L’étape pré-attentionnelle porte sur les mécanismes qui ont pour objectif de « guider » le faisceau attentionnel, c’est-à-dire de sélectionner dans la scène les zones qui vont être étudiées par la phase attentionnelle. Il s’agit d’une étape ascendante (« bottom-up »), c’est-à-dire guidée par les données. L’idée de mécanismes spécifiques pour guider l’attentionnel a été introduite par [Neisser \(1967\)](#). Les premiers travaux de mise en évidence expérimentale des mécanismes pré-attentionnels sont dus à Treisman pour l’identification des caractéristiques visuelles appelées « preattentive features » ([Treisman \(1985\)](#); [Treisman et Gormican \(1988\)](#); [Treisman \(1991\)](#)), ainsi que pour la gestion du pré-attentif par le système visuel ([Treisman et Gelade \(1980\)](#)).

À l’étape pré-attentionnelle, tout un ensemble de caractéristiques visuelles sont détectées de manière très rapide, sans que le nombre d’objets dans la scène influe sur le temps de recherche. La figure 2.5 illustre ce phénomène de « pop-out », où les objets qui ne diffèrent que d’une et une seule caractéristique par rapport aux autres « sautent aux yeux ». Deux exemples sont présentés. L’objet diffère dans le premier par sa couleur. Dans le deuxième, il diffère par sa forme des autres objets. Dans les expériences initiales, deux images sont présentées à l’observateur, une avec un leurre et l’autre sans. L’observateur doit indiquer s’il y a un leurre. Les temps de réponse sont alors analysés. Dans le cas présenté dans la figure 2.5.c, il y a deux objets dans la scène, des carrés rouges et des ronds bleus. Un leurre rond rouge partage donc une caractéristique avec chacun des objets. La recherche de ce type de leurre est appelée recherche conjointe, le phénomène de « pop-out » ne se produit pas et la recherche doit être effectuée de manière séquentielle en faisant appel aux mécanismes attentionnels. Dans ce cas, la recherche est plus longue et est dépendante du nombre d’objets dans la scène.

Il existe tout un ensemble de caractéristiques visuelles qui ont été identifiées dont certaines peuvent être plus difficiles à repérer que d’autres. Treisman avance l’idée que ces caractéristiques ont en commun de pouvoir être traitées en parallèle. Dans [Duncan et Humphreys \(1989a\)](#), une définition décrit ces caractéristiques comme :

« a feature or stimulus that differs from its immediate surround in some dimensions and the surround is reasonably homogeneous in those dimensions ».

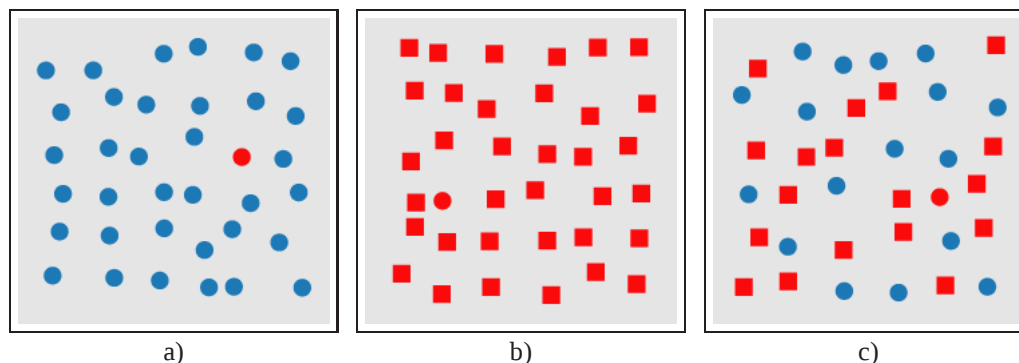


FIG. 2.5 – Illustration du phénomène de pop-out lorsqu'un leurre diffère d'une caractéristique visuelle unique des autres objets de l'image. La recherche est très rapide et n'est pas dépendante du nombre d'objets dans l'image. a) Un leurre diffère par sa couleur, b) par sa forme. c) Recherche conjointe, le leurre diffère d'un objet par sa couleur mais il est de la même couleur que l'autre objet. Inversement avec sa forme. Dans ce cas, le phénomène de pop-out n'apparaît pas, et la recherche est bien plus lente, car elle est effectuée de manière séquentielle. Source : [Healey. \(2007\)](#).

Wolfe propose une revue des caractéristiques visuelles ([Wolfe \(1998\)](#); [Wolfe et Horowitz \(2004\)](#)). La figure 2.6 présente une liste non exhaustive des caractéristiques visuelles, parmi lesquelles on trouve la couleur, la taille, la forme, l'orientation, la courbure, l'intensité lumineuse, etc. Des caractéristiques liées aux mouvements ne sont pas représentées, comme la vitesse de déplacement (si un leurre va plus vite que les objets), le sens de déplacement (s'il est différent du déplacement des autres objets). Certaines caractéristiques sont plus élémentaires, et certaines sont moins rapides à être détectées. (Voir [Wolfe \(1998\)](#) pour une discussion détaillée à propos de ces caractéristiques.)

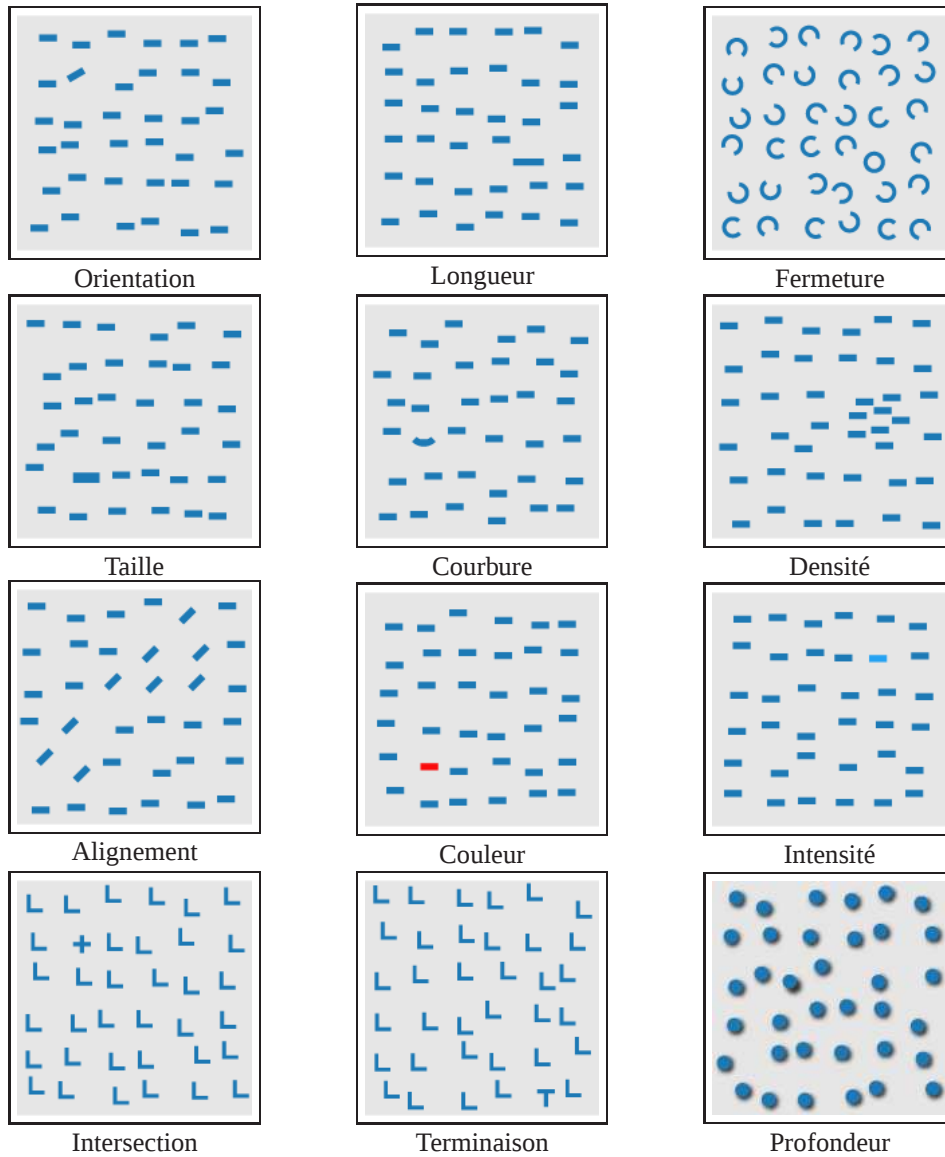


FIG. 2.6 – Illustration non exhaustive des différentes caractéristiques visuelles pré-attentionnelles (source : [Healey. \(2007\)](#)). On peut ajouter à cette liste des caractéristiques liées au mouvement : le sens de déplacement, la vitesse de déplacement et le clignotement. La différence d'intensité n'est pas visible sur une impression, mais est visible sur une version électronique de ce document. Pour la profondeur, la différence est visible sur les ombres de chaque point.

2.2.1 Les différentes théories pré-attentionnelles

Il existe plusieurs grandes théories sur la manière dont la phase pré-attentionnelle est gérée par le système visuel. Nous allons présenter les plus connues. Pour chaque théorie, l'objectif est de guider l'étape attentionnelle, en cherchant dans l'image les objets ou zones qui sont les plus saillants. [Itti \(2007\)](#) donne la définition suivante :

« The visual saliency is the distinct subjective perceptual quality which makes some items in the world stand out from their neighbors and immediately grab our attention. »

2.2.1.1 « Feature integration theory »

La « feature integration theory » a été proposée par Treisman ([Treisman et Gelade \(1980\)](#)) et a inspiré beaucoup de systèmes pré-attentionnels. Dans cette théorie, l'information provenant de chaque caractéristique visuelle est codée dans une carte dédiée, comme des oppositions de couleurs, des orientations. Chaque carte de caractéristiques est produite indépendamment et en parallèle des autres, ce qui garantit la rapidité de traitement pour une image. Si un leurre ne diffère que par une seule caractéristique, alors la carte correspondante sera activée et la recherche est aisée. Dans le cas d'une recherche en conjonction, où un leurre partage des caractéristiques avec d'autres objets, mais diffère par une autre, alors la recherche doit être effectuée en comparant des localisations sur les différentes cartes de caractéristiques. Cette tâche est effectuée en utilisant une carte de localisation (« map of locations » ou carte topologique). Cette carte est utilisée pour guider le focus attentionnel, permettant de trouver l'objectif. La figure 2.7 présente un schéma illustrant les différentes cartes de cette théorie.

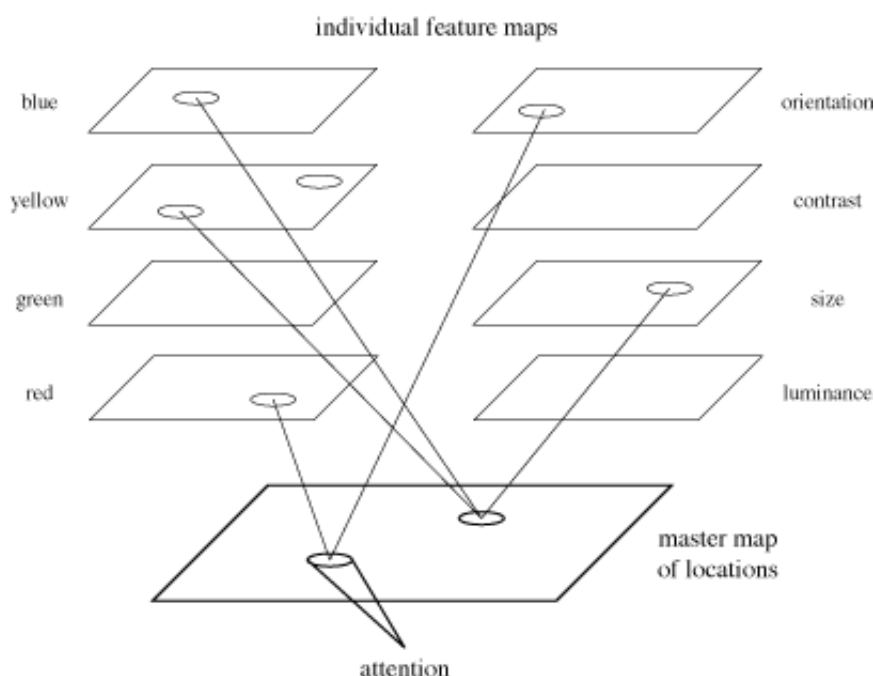


FIG. 2.7 – Schéma de principe de la « feature integration theory » ([Treisman et Gelade \(1980\)](#)). Chaque caractéristique visuelle est codée dans une carte dédiée d'une manière propre à la caractéristique. La carte topologique regroupe l'information des cartes de caractéristiques et permet de déplacer le focus d'attention pour une recherche conjointe par exemple. Source : [Healey. \(2007\)](#)

2.2.1.2 « Guided search theory »

La « guided search theory » (Wolfe *et al.* (1989); Wolfe (1994); Rodriguez-Sanchez *et al.* (2007)) est une théorie proche de la précédente mais qui prend en compte un mécanisme descendant (« top-down »). Cette approche n'est donc plus uniquement guidée par les données. Le schéma général propose une carte non plus par caractéristique visuelle, mais par type de caractéristique (comme la couleur par exemple) qui regroupera les informations de toutes les caractéristiques appartenant à cette catégorie. Toutes les cartes de caractéristiques sont regroupées dans une carte d'activation, correspondant au principe de la carte topologique de Treisman. L'intégration du processus descendant s'effectue grâce à la carte d'activation, où la saillance va être adaptée en fonction de l'objectif suivi, afin de promouvoir les caractéristiques correspondantes. Cela permet de modéliser notre habilité à rechercher de manière plus efficace des objets dont on connaît à l'avance les caractéristiques. La figure 2.8 présente le schéma général de la méthode proposée par Wolfe (1994). Du point de vue des neuro-sciences, cette habilité correspond à l'attention fondée sur les caractéristiques (« feature-based attention ») mise en évidence par plusieurs expériences (Motter (1994); Treue et Trujillo (1999); Saenz *et al.* (2002)).

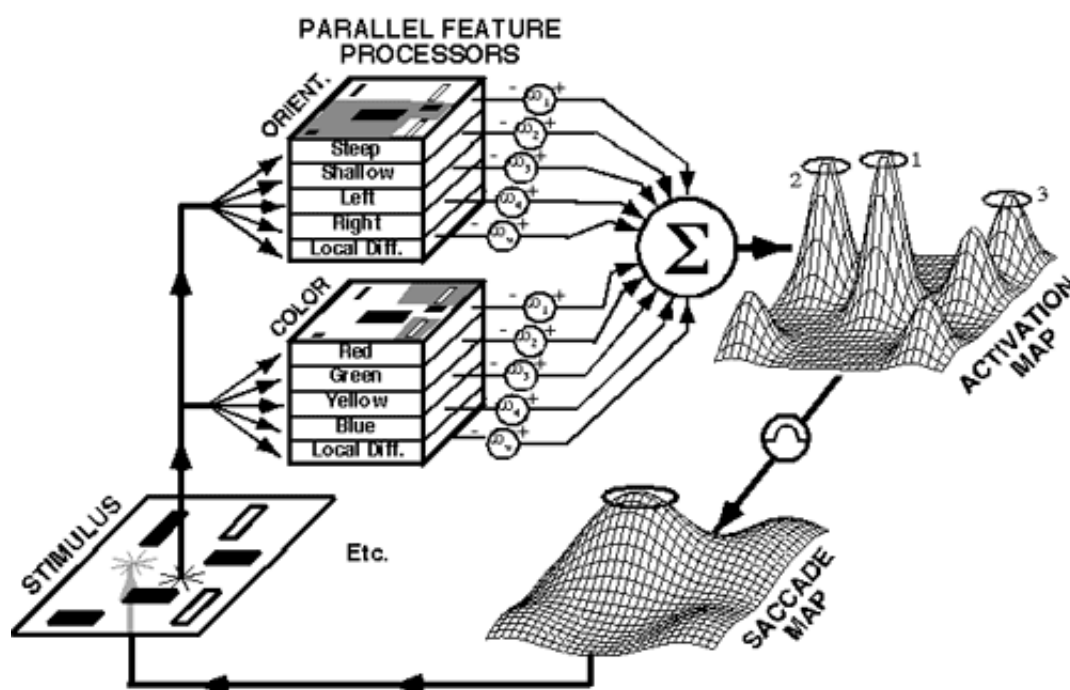


FIG. 2.8 – Schéma du principe de la « guided search theory » (Wolfe *et al.* (1989); Wolfe (1994); Rodriguez-Sanchez *et al.* (2007)). Les caractéristiques visuelles sont regroupées par catégorie et traitées en parallèle, ici la couleur et l'orientation. L'information est alors regroupée dans la carte d'activation. Le processus descendant influence sur l'activation des composantes pour modifier la recherche visuelle, afin de la faire correspondre aux caractéristiques recherchées. [Source : Itti (2007) d'après Wolfe (1994).]

2.2.1.3 « Texton theory »

La théorie des textons (Julész (1981a,b); Julész et Bergen (1983)) indique que le système visuel peut détecter de manière pré-attentionnelle des groupes de caractéristiques appelés textons,

qui sont classés en trois groupes : des formes allongées avec comme caractéristique la couleur, l'orientation ou la taille ; des terminateurs, c'est-à-dire des fins de ligne, et enfin des croisements de lignes. Comme dans la « feature integration theory », Julesz considère que la phase pré-attentionnelle est effectuée en parallèle, alors que la phase attentionnelle est séquentielle.

2.2.1.4 « Similarity theory »

La théorie des similitudes (Duncan (1989); Duncan et Humphreys (1989b); Müller *et al.* (1990)) rompt avec le schéma d'une recherche effectuée en parallèle ou de manière séquentielle. Au lieu de cela, le temps de recherche est présenté comme étant dépendant des similarités entre l'objectif et les autres objets d'une part, mais aussi de l'homogénéité des autres objets. La recherche sera d'autant plus facile que l'objet recherché est différent des autres objets. Elle sera également plus facile si tous les autres objets se ressemblent, et si les variations de l'une ou l'autre similarité ont plus ou moins d'importance en fonction du niveau de l'autre. Dans cette théorie, le champ visuel est segmenté par unités structurelles qui partagent une même caractéristique visuelle. Chaque unité structurelle peut être ensuite à nouveau subdivisée, ce qui permet d'obtenir une hiérarchie du champ visuel.

2.2.2 Conclusion sur les théories pré-attentionnelles

Toutes les théories décrites ici proposent une manière pour extraire de l'image des informations saillantes, zone ou objet de la scène. La théorie d'intégration des caractéristiques est la plus répandue et a donné naissance à de nombreuses mises en œuvre. La recherche visuelle est en quelque sorte le pendant de la première théorie pour une vision guidée par un modèle, et non plus guidée par les données uniquement. La théorie des textures découle des travaux de Julesz mais n'a pas donné lieu à d'autres développements. Quant à la théorie des similitudes, elle permet d'expliquer certains comportements mais n'a pas donné lieu à des systèmes opérationnels. Nous nous intéressons ici à une recherche d'informations guidée par les données, en vue d'explorer une scène et notamment aux cartes de saillance, que nous allons présenter plus en détail à présent.

2.3 Les cartes de saillance

Les cartes de saillance modélisent un mécanisme pré-attentionnel inspiré par le système visuel humain, mais sans toutefois chercher à être psycho-réaliste. Il correspond à un exemple de mécanisme issu de la « feature integration theory ». Les cartes de saillance ont été proposées par Itti *et al.* (1998), reprenant un modèle décrit par Koch et Ullman (1985). Ce mécanisme permet une sélection attentionnelle et spatiale qui utilise des caractéristiques facilement calculables sur tout type d'image. La figure 2.9 présente le schéma général de la méthode permettant de calculer les cartes de saillance et la figure 2.10 présente un exemple de carte de saillance calculée sur l'image « Lena ». Nous allons maintenant présenter la méthode originale. Nous présenterons les adaptations nécessaires au type d'images que nous souhaitons utiliser.

Cette approche utilise des caractéristiques visuelles courantes correspondant à des percepts neurophysiologiques : des oppositions de couleurs, différences d'intensité et d'orientation. Plus précisément, une carte de caractéristiques reflétera les oppositions d'intensité (sombre ou clair), une deuxième carte combinerà les informations issues de deux oppositions de couleurs, rouge et vert d'une part, bleu et jaune d'autre part, la couleur jaune étant obtenue à partir d'une combinaison de rouge et de vert. Une troisième carte regroupera les informations sur les orientations dans

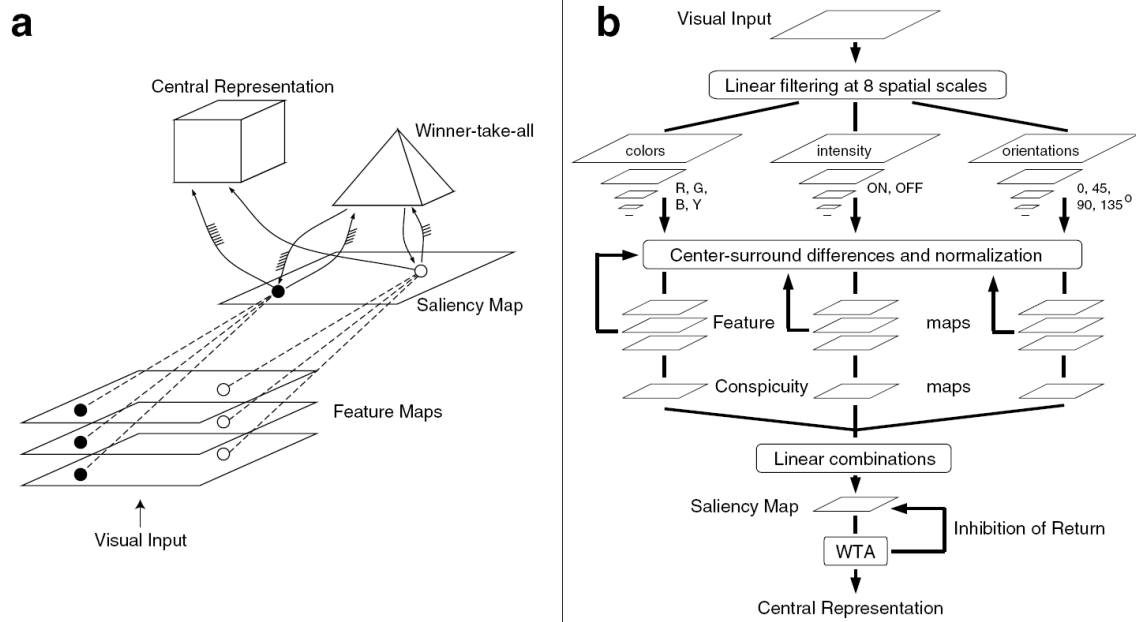


FIG. 2.9 – Schéma général pour la génération des cartes de saillance telle que décrite par [Itti et al. \(1998\)](#). Les différentes caractéristiques sont extraites et représentées sur différentes échelles dont les différences produisent des cartes de discontinuités. Les cartes sont ensuite fusionnées pour générer la carte de saillance. La zone la plus saillante est alors produite par un algorithme « Winner-take-all » et un mécanisme utilisant le phénomène d’inhibition de retour permet d’itérer sur les zones saillantes de l’image. L’inhibition de retour permet de ne pas tenir compte d’une zone saillante pendant un court moment afin de permettre l’exploration d’autres zones qui sont moins saillantes. Source : [Itti \(2005\)](#).

l’image, obtenue à partir de filtres de Gabor dans un nombre donné de directions (quatre dans la méthode originale).



FIG. 2.10 – Un exemple de carte de saillance (à droite) calculée sur l’image de gauche. Les zones sombres correspondent aux parties les moins saillantes, les zones claires aux parties de l’image les plus saillantes. Les parties claires de l’image de gauche qui sont bien contrastées avec les zones environnantes apparaissent bien saillantes dans la carte correspondante. Il faut noter que la saillance n’est pas limitée aux bords de ces zones même si ce sont les discontinuités qui sont étudiées à cause du facteur d’échelle dans la génération des cartes. À droite sur l’image, quelques structures verticales, qui sont globalement peu voyantes mais dont la géométrie attire l’œil, sont visibles sur la carte de saillance.

Pour chacune des sept sous-caractéristiques, à savoir l'intensité, deux oppositions de couleurs et quatre orientations, l'image originale est tout d'abord filtrée afin de ne conserver que l'information concernant cette caractéristique. À partir de cette image filtrée, une pyramide gaussienne est générée. La taille de toutes les images de la pyramide gaussienne est ensuite modifiée afin qu'elles possèdent toutes la même taille, qui sera la taille de la carte de saillance.

Des cartes de discontinuités sont ensuite extraites. Une carte de discontinuités représente ici les différences entre une zone et son contour immédiat, appelées différences centre-contour ($\ominus$, « center-surround difference »). En pratique, une carte de discontinuités est une différence pixel à pixel entre deux niveaux de la même pyramide, un niveau dit fin, et un niveau dit grossier. Différentes cartes sont générées avec différents niveaux fin et grossier, afin d'obtenir différents niveaux d'échelles. Les niveaux fins de la pyramide sont $c \in \{2, 3, 4\}$ et les niveaux grossiers sont $s = c + d$ avec $d \in \{3, 4\}$. Il y a donc 6 cartes de discontinuités par caractéristique. La figure 2.11 illustre la génération des cartes de discontinuités.

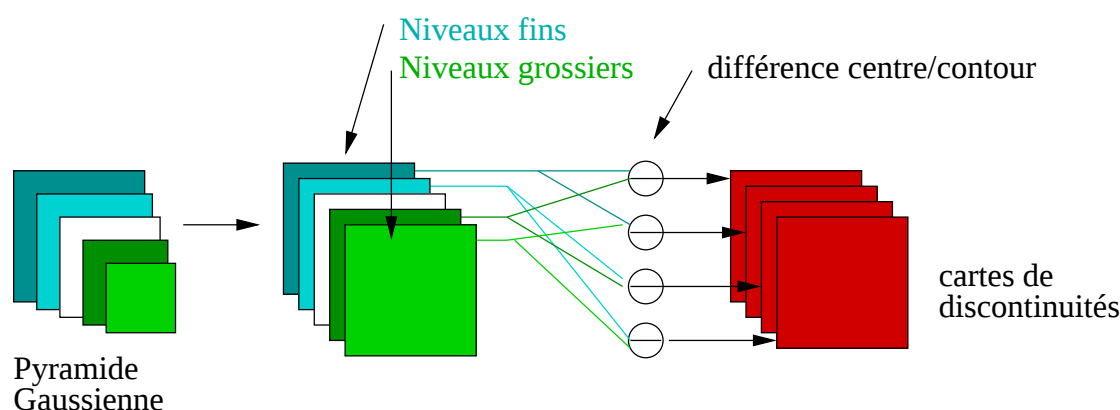


FIG. 2.11 – Traitements effectués pour chaque caractéristique. L'image originale est filtrée en fonction de la caractéristique. Une pyramide gaussienne est ensuite générée, puis remise à une même taille. Les niveaux fins comparés aux niveaux grossiers permettent de générer les cartes de discontinuités. Une fois normalisées, elles sont combinées pour former une carte unique représentant la caractéristique.

Pour chaque caractéristique, les cartes de discontinuités sont normalisées, puis fusionnées à l'aide d'un opérateur de normalisation ad-hoc, permettant de favoriser les cartes présentant des pics plus élevés que leur moyenne par rapport à une carte présentant beaucoup de pics, mais d'une hauteur similaire par exemple. Nous obtenons donc une carte pour chacune des sept caractéristiques. Les cartes correspondant à un même type de caractéristique sont fusionnées : les quatre cartes représentant les orientations sont fusionnées en une unique carte représentant toutes les orientations. Même chose pour les deux oppositions de couleurs. Il reste trois cartes représentant chaque type de caractéristique (intensité, couleur, orientation), appelées « conspicuity maps » ou carte de visibilité. Ces trois cartes sont alors fusionnées pour produire la carte de saillance.

Cette approche initiale a donné lieu à de nombreux travaux. On trouve d'autres applications dans [Itti \(2005\)](#). Dans [Walther et Koch \(2006\)](#), cette approche est utilisée pour extraire des « proto-objets ». Dans ce cas, une carte de saillance de la scène est extraite de la manière décrite précédemment. Le proto-objet extrait sera l'objet se situant à l'emplacement le plus saillant identifié par la carte de saillance. Pour extraire l'objet, il n'y a pas de segmentation, mais un proto-objet est extrait à partir des cartes qui ont été utilisées pour créer la carte de saillance. Pour cela, il est nécessaire d'identifier la caractéristique, puis la carte de discontinuité ayant le plus contribué dans

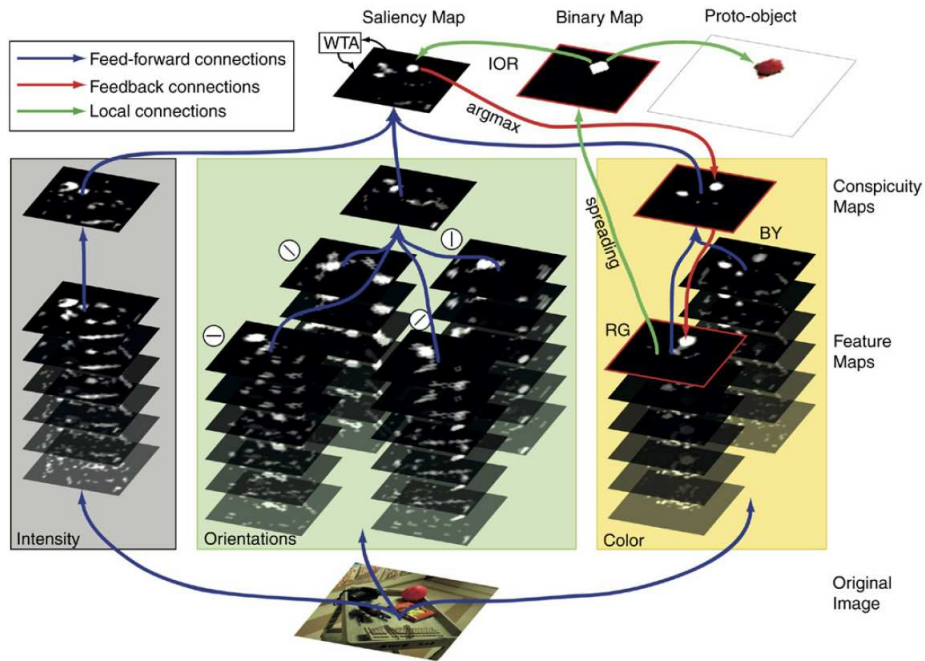


FIG. 2.12 – Méthode pour extraire des proto-objets présentée par Walther et Koch (2006) : la zone la plus saillante décrit un objet qui sera extrait par seuillage à partir de la carte de discontinuités ayant le plus contribué à sa saillance. [Source Walther et Koch (2006)]

la saillance détectée, c'est-à-dire une caractéristique et un niveau d'échelle. Une fois cette carte identifiée, nous connaissons un point du proto-objet, un seuillage de l'image est effectué, et la composante connexe qui contient le point correspond au proto-objet est extraite. La figure 2.12 illustre cette approche.

Walther et Koch (2006) présentent une méthode d'apprentissage pour adapter le processus à un problème donné. On retrouve cette notion d'apprendre les caractéristiques d'un objet pour contraindre la saillance dans Kanan et al. (2009) qui parlent de « contextual guidance » par exemple, pour rechercher de la vaisselle ou des tableaux dans une scène naturelle.

2.4 Les cartes de saillance adaptées aux images IRM

L'imagerie par résonance magnétique nucléaire (IRM) permet d'obtenir des vues en trois dimensions du corps humain d'une manière non invasive, et en particulier du cerveau humain dans notre application. Pour pouvoir calculer une carte de saillance sur ce type d'images, nous devons tenir compte des spécificités de cette modalité, et adapter le processus de génération des cartes. Nous allons à présent passer en revue les différentes étapes de la génération des cartes de saillance en précisant les principales adaptations nécessaires.

2.4.1 Pré-traitements

Nous présentons tout d'abord les traitements appliqués aux images IRM avant de commencer la génération des cartes de saillance proprement dite.

Extraction du cerveau :

Les volumes représentent le cerveau, mais également le crâne et tous les organes situés dans la tête. Étant donné que nous nous intéressons aux structures internes du cerveau humain, nous ne devons donc pas considérer toute l'image. Cela permet déjà de réduire le domaine de recherche, mais également d'éviter que les bord réguliers et bien marqués du crâne fassent apparaître de fortes valeurs de saillance tout autour du bords du cerveau. Nous allons donc utiliser un masque pour ne considérer que le cerveau, qui est segmenté à l'avance.

Résolution anisotrope :

Les images IRM ont souvent des résolutions qui sont anisotropes : la taille des voxels peut varier en fonction des directions. La génération de la pyramide dyadique implique de pouvoir redimensionner les images. Les voxels anisotropes rendent cette tâche plus compliquée. Nous utilisons donc pour le calcul des cartes de saillance des images qui ont été interpolées au préalable vers des dimensions isotropes, avec 256 voxels cubiques dans chaque direction (le choix de 256 a été guidé par les dimensions les plus fréquentes dans nos bases d'images IRM). La méthode d'interpolation utilisée est « spline resampled » proposée par [Thevenaz et al. \(2000\)](#), et qui a été adaptée pour les images IRM en trois dimensions dans le logiciel brainvisa¹.

Une fois l'image source interpolée à la taille correcte, nous pouvons passer à la génération des cartes pour chaque caractéristique. Pour cela, nous devons d'abord filtrer l'image originale en fonction de chaque caractéristique.

2.4.2 Filtrage par caractéristique

Images en trois dimensions : Les méthodes pré-attentionnelles et en particulier les méthodes bio-inspirées sont par définition en deux dimensions, ce qui est le cas de la méthode de [Itti et al. \(1998\)](#) pour les cartes de saillance, voire deux dimensions et demi pour simuler la vision stéréoscopique. Les images IRM sont des volumes en trois dimensions. Il est donc nécessaire d'adapter le processus, en particulier la notion de voisinage pour des voxels : une connexité 18 ou 26 au lieu de la connexité 4 ou 8.

Les cartes de saillance dans la méthode originale utilisent trois types de caractéristiques auxquelles le cortex humain réagit : l'intensité, les oppositions de couleurs et les orientations. Les images IRM ne possèdent qu'un seul canal (donc pas de couleurs). Les niveaux de gris qui ne représentent pas une intensité seront considérés comme tels.

Intensité : Nous considérons l'unique canal des images IRM comme une intensité. Il n'y a donc pas de filtrage nécessaire pour générer la pyramide gaussienne correspondant à cette caractéristique.

Orientations : Le calcul des cartes pour l'orientation utilise un filtre de Gabor en trois dimensions, tel que défini dans [Reed \(1997\)](#) et [Wang et Chua \(2005\)](#), de la manière suivante :

$$g(x, y, z) = \hat{g}(x, y, z) \exp(j2\pi(F \sin \theta \cos \phi x + F \sin \theta \sin \phi y + F \cos \phi z)) ,$$

avec

$$\hat{g}(x, y, z) = \frac{1}{(2\pi)^{\frac{3}{2}}\sigma^3} \exp\left(-\frac{(x^2 + y^2 + z^2)}{2\sigma^2}\right) ,$$

¹<http://brainvisa.info>

où θ et ϕ sont deux angles définissant l'orientation du filtre de Gabor, σ représente l'échelle de la fonction gaussienne et $F = \sqrt{(u_0^2 + v_0^2 + w_0^2)}$ est le paramètre correspondant à la fréquence radiale. Les valeurs de F et σ sont contrôlées par la largeur de bande B fixée à 0,55 dans nos expériences :

$$F\sigma = \frac{\alpha}{\pi} \left(\frac{2^B + 1}{2^B - 1} \right) .$$

où $\alpha = \frac{\sqrt{(2 \ln 2)}}{2}$.

En deux dimensions, les orientations choisies sont les suivantes : $\theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, 3\frac{\pi}{4}$, soit 4 orientations, les filtres étant symétriques. En trois dimensions, nous avons conservé le même ordre de grandeur entre les orientations, ce qui nous donne, exprimées en coordonnées polaires, les valeurs suivantes pour les angles θ et ϕ :

$\theta \setminus \phi$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$3\frac{\pi}{4}$	π	$5\frac{\pi}{4}$	$3\frac{\pi}{2}$	$7\frac{\pi}{4}$
0	×							
$\frac{\pi}{4}$	×	×	×	×	×	×	×	×
$\frac{\pi}{2}$	×	×	×	×				

Nous obtenons 13 orientations au total. Les filtres étant symétriques, nous avons seulement besoin d'une demi-sphère.

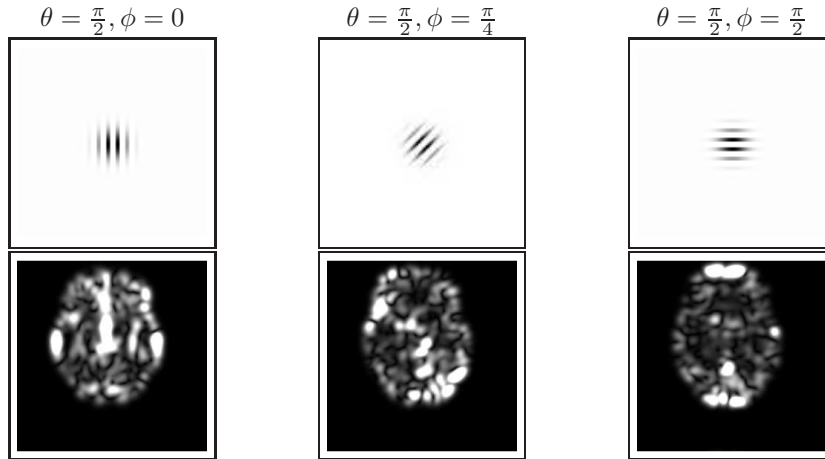


FIG. 2.13 – Filtres de Gabor. Trois exemples de filtres de Gabor avec trois orientations différentes en coordonnées polaires. Chaque colonne présente une coupe d'un filtre en trois dimensions et une coupe d'une image IRM filtrée par ce même filtre. La fréquence du filtre est de 0,2 dans cet exemple. La largeur de bande est fixée à 0,55.

2.4.3 Génération des pyramides

Contrairement à la méthode originale, nous n'utilisons pas les mêmes pyramides pour chaque caractéristique. La pyramide gaussienne de la méthode originale est utilisée pour l'intensité. Pour l'orientation, nous utiliserons une pyramide « de Gabor ».

La méthode originale utilisait des pyramides avec 8 niveaux. Dans notre cas, ce nombre de niveaux est trop élevé, considérant la taille des objets. Il nous faut donc réduire le nombre de niveaux de chaque pyramide. En considérant la taille de 256 dans chaque direction, nous limitons notre pyramide à 5 niveaux (comprenant l'image originale). Le dernier niveau a ainsi une taille de 16 dans chaque direction.

Calcul des cartes de discontinuités : les cartes de discontinuité sont générées en comparant une image de la pyramide à une échelle dite « fine » (c'est-à-dire restant proche de l'image originale), et une autre image de la même pyramide à une échelle dite « grossière ». À l'origine, la comparaison est effectuée en interpolant le niveau grossier au niveau fin, et en effectuant une soustraction point à point des deux images :

$$I(ce, co) = |I(ce) \ominus I(co)| ,$$

où ce représente le niveau fin et co le niveau grossier. La comparaison d'un pixel au niveau fin avec un pixel au niveau grossier après interpolation revient à comparer un pixel avec sa région environnante, plus ou moins grande en fonction de la différence entre les deux niveaux, d'où l'appellation de différence centre-contour.

Nous pouvons utiliser différents niveaux fins pour calculer les cartes de discontinuité. Mais l'utilisation de l'image originale (bruitée par rapport aux cartes lissées) comme un niveau fin, va représenter le bruit comme des petites discontinuités. Nous utiliserons donc comme niveaux fins les deux niveaux suivants :

$$ce \in \{1, 2\} .$$

L'intervalle permettant de calculer les niveaux fins est limité par le nombre de niveaux de la pyramide. Nous utiliserons donc les niveaux grossiers suivants :

$$co = ce + \delta, \delta \in \{1, 2\} ,$$

c'est-à-dire $1 + 1, 1 + 2, 2 + 1, 2 + 2$. Finalement, la carte de saillance résultante a une résolution de $128 \times 128 \times 128$ correspondant au deuxième niveau de la pyramide.

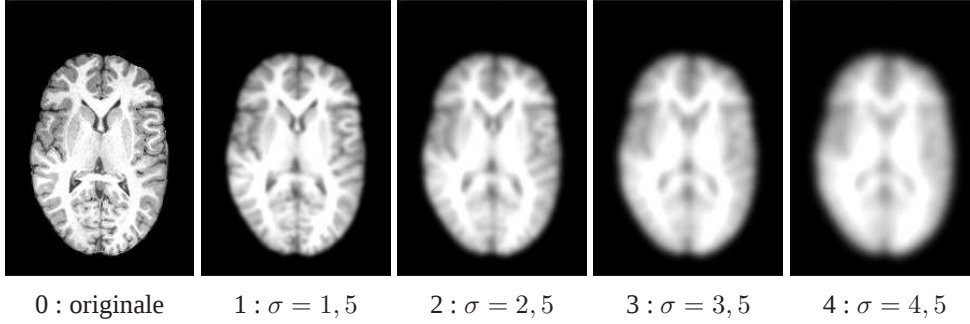
Pyramide gaussienne pour l'intensité : Nous créons pour l'intensité une pyramide en utilisant l'image originale (après interpolation). Un filtre gaussien en trois dimensions est appliqué à chaque niveau de la pyramide. Nous appliquons le filtre sans avoir redimensionné les images, mais en augmentant à chaque niveau le paramètre σ du filtre. Le niveau 0 de la pyramide est l'image originale. Pour tous les autres niveaux, le filtre utilisé a un paramètre σ correspondant au niveau de la pyramide $+ 0.5$: niveau 1, $\sigma = 1, 5$, niveau 4, $\sigma = 4.5$. La figure 2.14 présente les différents niveaux de la pyramide gaussienne, ainsi que les cartes de discontinuité dérivées.

Pyramide de Gabor pour les orientations : Pour les orientations, la méthode originale utilise également une pyramide gaussienne, comme pour les autres caractéristiques. Toutefois, les filtres de Gabor intègrent dans leur paramétrage la possibilité de faire varier le niveau d'échelle, en modifiant la fréquence du filtre par exemple. Nous pouvons donc définir une « pyramide de Gabor » où pour chaque orientation, une pyramide est générée composée d'images filtrées avec une fréquence décroissante. La fréquence initiale est de 0, 4, avec un pas de 0, 05 entre deux niveaux. La fréquence du dernier niveau est de 0, 20. Chaque image est lissée avec un filtre gaussien pour éviter le bruit. Le paramètre σ utilisé est 0, 5. La figure 2.15 montre un exemple de pyramide de Gabor obtenue pour une orientation donnée et les cartes de discontinuités dérivées.

2.4.4 Fusion des cartes de discontinuités

Pour chaque pyramide, quatre cartes de discontinuités sont générées. En considérant l'intensité et treize orientations différentes, nous avons donc 14 pyramides et 4×14 cartes de discontinuités. Toutes ces cartes doivent être fusionnées pour générer la carte de saillance. La méthode de fusion est primordiale, et en particulier l'étape de normalisation est cruciale.

Pyramide gaussienne :



Cartes de discontinuités :

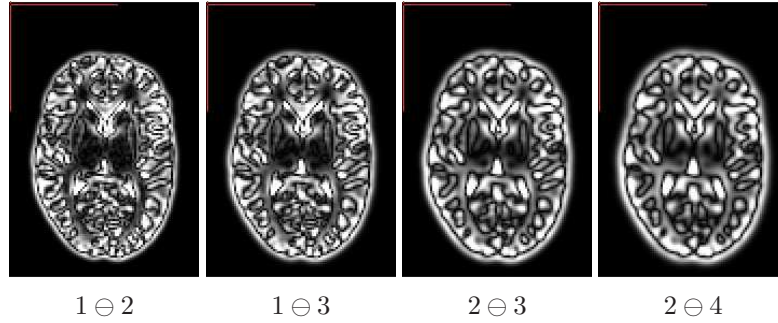


FIG. 2.14 – Les différents niveaux de la pyramide gaussienne obtenue pour l'intensité sont présentés en haut. Le niveau 0 représente l'image originale, les niveaux 1 à 4 représentent l'image originale filtrée par un filtre gaussien d'une largeur croissante de $\sigma = 1,5$ à $\sigma = 4,5$. En bas, les cartes de discontinuités obtenues en appliquant l'opérateur centre-contour $\ominus$ entre différents niveaux de la pyramide gaussienne.

Opérateur de normalisation : L'opérateur de normalisation spécifiquement défini pour les cartes de saillance est présenté notamment dans [Itti et al. \(1998\)](#). On pourra également consulter [Itti et Koch \(2001\)](#) pour une comparaison de cet opérateur avec une normalisation « naïve » ou une normalisation avec apprentissage au préalable. Que ce soit dans la méthode originale avec 42 cartes à fusionner, ou dans notre cas avec 56 cartes de discontinuité, le nombre de cartes à fusionner est suffisamment important pour qu'un pic, même important, apparaissant dans quelques cartes, soit noyé dans le bruit apparaissant dans plus de cartes.

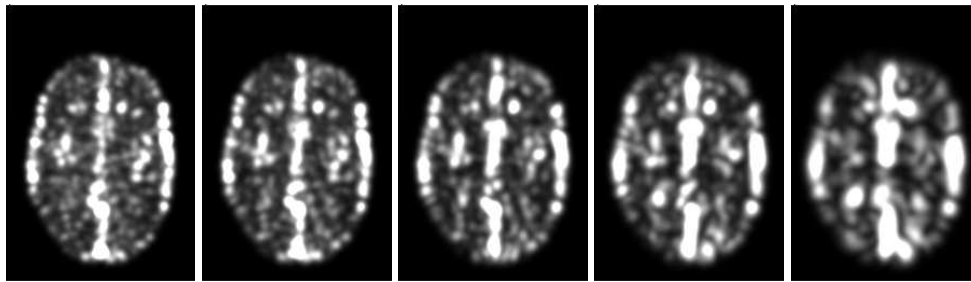
[Itti et al. \(1998\)](#) proposent donc un opérateur dénoté $\mathcal{N}$ qui permet de promouvoir les cartes dans lesquelles ne sont présents qu'un petit nombre de pics importants (zones visibles). En revanche, les cartes contenant de nombreux pics avec une même importance sont supprimées. Cet opérateur est illustré dans la figure 2.16.

La normalisation est effectuée en trois étapes :

- normalisation de la carte dans un intervalle $[0..M]$ avec un M fixe, pour supprimer les différences d'amplitude entre les différentes caractéristiques,
- Calcul de la moyenne $\hat{m}$ des maxima locaux différents de M ,
- multiplication de chaque point par $(M - \hat{m})^2$.

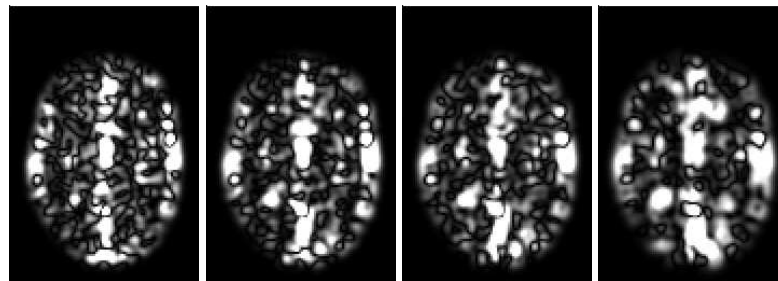
Fusion : L'objectif est de fusionner les cartes existantes pour obtenir une unique carte représentant une caractéristique. Pour chaque pyramide, cette carte est générée à partir des cartes de

Pyramide de Gabor :



0 : $freq = 0.40$ 1 : $freq = 0.35$ 2 : $freq = 0.30$ 3 : $freq = 0.25$ 4 : $freq = 0.20$

Cartes de discontinuités :



$1 \ominus 2$

$1 \ominus 3$

$2 \ominus 3$

$2 \ominus 4$

FIG. 2.15 – Les différents niveaux de la pyramide « de Gabor » obtenue pour une orientation ($\theta = \frac{\pi}{2}$ et $\phi = 0$) sont présentés en haut de la figure. On distingue nettement sur ces images le plan inter-hémisphérique. Le niveau 0 représente la fréquence la plus élevée (0, 40) et le niveau 4 la fréquence la plus faible (0, 20). Les images ont été lissées avec un filtre gaussien. Dans ce cas, le σ utilisé est de 2, 0. En-dessous, les cartes de discontinuité obtenues en appliquant l'opérateur centre-contour $\ominus$ entre différents niveaux de la pyramide « de Gabor ».

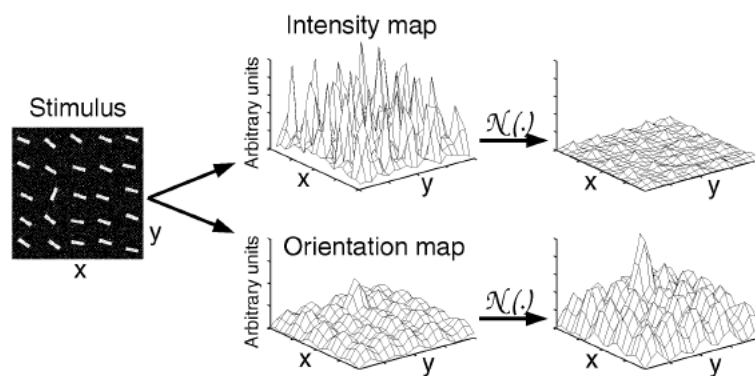


FIG. 2.16 – Opérateur de normalisation $\mathcal{N}$ [Source [Itti et al. \(1998\)](#)]

discontinuités correspondantes. Cette carte unique est une carte de visibilité (« conspicuity map »).

Pour l'intensité il n'y a qu'une unique pyramide, donc la carte de visibilité est directement obtenue après fusion des cartes de discontinuité.

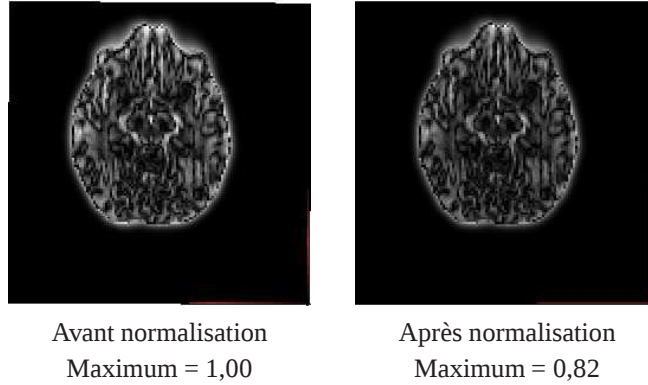


FIG. 2.17 – Effet de l’opérateur de normalisation $\mathcal{N}(\cdot)$. À gauche, avant la normalisation, à droite après. Le maximum de l’image est passé de 1,00 à 0,82 après normalisation.

$$C_{int} = \oplus \{ \mathcal{N}(I(ce, co)), ce \in \{1, 2\}, co = ce + \delta, \delta \in \{1, 2\} \} ,$$

avec $\oplus$ une addition point à point.

Pour les orientations : Pour chaque orientation, c’est-à-dire pour chaque pyramide, une carte intermédiaire est générée. Ces 13 nouvelles cartes sont ensuite normalisées avec l’opérateur $\mathcal{N}$, et fusionnées par addition point à point pour générer la carte de visibilité des orientations.

$$C_{\theta, \phi} = \oplus \{ \mathcal{N}(I_{\theta, \phi}(ce, co)), ce \in \{1, 2\}, co = ce + \delta, \delta \in \{1, 2\} \}$$

$$C_{orient} = \sum_{\theta, \phi} \mathcal{N}(C_{\theta, \phi}) .$$

2.4.5 Cartes de saillance

Dans la méthode originale, chaque caractéristique produit de manière parallèle une carte dédiée, et les trois cartes seront combinées par une moyenne pondérée. En l’absence d’une caractéristique, il manquera donc un terme à la moyenne pondérée produisant la carte de saillance, mais cette absence n’influe pas sur le calcul des autres cartes. Pour produire la carte de saillance, nous combinons donc les deux cartes. La fusion de deux cartes s’effectue au moyen d’une moyenne pondérée qui donne le même poids aux deux cartes, ne privilégiant ainsi pas une caractéristique au détriment d’une autre, comme dans l’approche originale. Elle est donnée par la formule suivante :

$$SaliencyMap = \frac{\sum_{i \in C} \mathcal{N}(C_i)}{\sum_C} .$$

où C_i représente les cartes de caractéristiques C_{int} , C_{orient} .

2.4.6 Masquage des cartes de saillance

Nous avons déjà calculé les cartes de saillance sur le cerveau uniquement, en utilisant un masque pour supprimer le crâne, entre autres. Une fois les cartes de saillance calculées, nous allons maintenant masquer les cartes de saillance, pour supprimer les zones apparaissant très saillantes aux bords du cerveau. En effet, les bords du cerveau peuvent faire apparaître des fortes saillances dues aux différences de contraste avec le fond (qui a été masqué), et aux orientations sur les bords.

Nous allons donc utiliser à nouveau le masque du cerveau, qui sera érodé, pour supprimer les fortes saillances apparaissant aux bords de l'image. La figure 2.18 illustre cette étape.

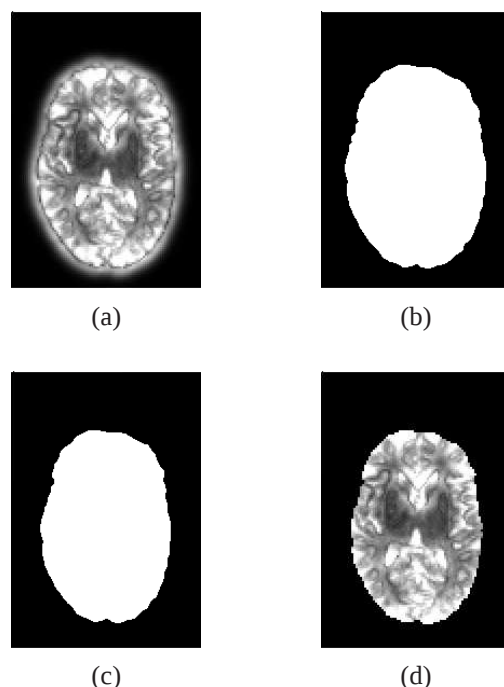


FIG. 2.18 – Utilisation d'un masque binaire du cerveau pour supprimer la forte saillance aux bords du cerveau, due aux forts contrastes des bords. (a) Une coupe non masquée d'une carte de saillance. On voit une couronne de valeurs élevées de saillance autour du cerveau. (b) Le masque correspondant à cette image du cerveau. (c) Le même masque érodé avec un élément structurant sphérique de rayon 5 pixels pour supprimer les bords du cerveau. (d) La carte de saillance masquée.

2.4.7 Résultats

La figure 2.19 présente quelques exemples de cartes de saillance calculées sur des images IRM de cerveau, ainsi que des coupes des cartes de caractéristiques. Sur une machine récente, le temps de calcul de ces cartes de saillance est de l'ordre de 15 minutes environ. Le temps de calcul est allongé par le nombre d'orientations pris en compte. Il est possible de gagner du temps en précalculant les filtres de Gabor utilisés.

Trois cartes de saillance sont présentées ; sur chacune d'entre elles on reconnaît facilement l'image originale. Les ventricules latéraux, au centre, présentent des valeurs de saillance élevées, ce qui était attendu à cause de leur différence d'intensité avec les structures avoisinantes et leur taille. Au contraire, les putamens apparaissent dans chaque image comme un trou de saillance. Sur l'image pathologique, la tumeur apparaît dans cet exemple comme très saillante.

Les cartes de caractéristiques sont assez différentes en fonction de la caractéristique concernée. La carte d'intensité présente des valeurs élevées pour le ventricule, toujours pour le contraste, ainsi que pour des régions où la frontière entre matière blanche et matière grise est très nette. Pour les orientations, les valeurs sont plutôt floues. On distingue encore des valeurs fortes sur les bords. Les ventricules présentent ici encore des valeurs élevées grâce à l'élongation de la structure.

D'autres cartes de saillance sont présentées dans l'annexe B de ce document.

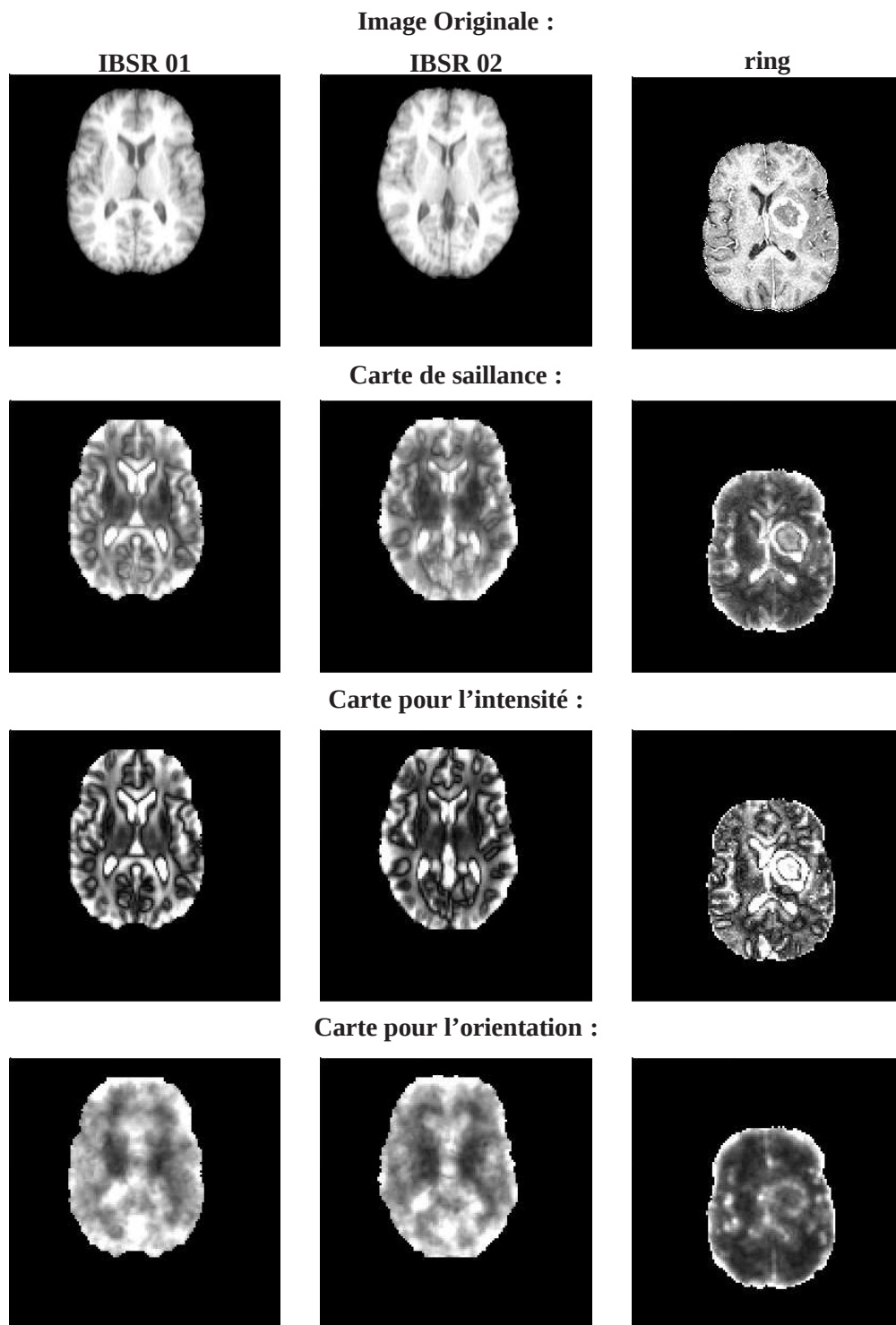


FIG. 2.19 – Quelques cartes de saillance. En haut, les images originales, la deuxième ligne présente les cartes de saillance respectives. Les deux lignes du bas présentent les « conspicuity maps », la troisième pour l'intensité et la dernière pour l'orientation.

2.5 Conclusion

Nous avons présenté dans ce chapitre la notion d'attention dans la vision, et la mise en évidence de son aspect séquentiel, guidé par une étape pré-attentionnelle détectant les parties saillantes de l'image. Nous pouvons aisément effectuer un parallèle entre un processus de traitement d'images séquentiel où la partie attentionnelle pourrait être l'étude d'une zone en particulier de l'image et la partie pré-attentionnelle la sélection d'une zone de l'image à explorer. Parmi les théories de modélisation du pré-attentionnel, nous avons plus particulièrement considéré les cartes de saillance, qui proposent de mettre en évidence les zones saillantes de l'image en utilisant des caractéristiques simples des images. Elles permettent également une analyse multi-échelles de l'image. Nous avons également présenté une série d'adaptations nécessaires permettant de calculer des cartes de saillance pour des images IRM en trois dimensions, et plus particulièrement pour les images du cerveau. Les cartes de saillance adaptées aux images IRM nous procurent une manière inédite d'obtenir de l'information dans un processus ascendant sur ce type d'image, que nous utiliserons dans le chapitre 4 pour l'exploration de ces images.

Chapitre 3

Le modèle de connaissance

Nous avons choisi dans ce travail d'exploiter la connaissance de l'imagerie cérébrale. Nous utilisons cette connaissance dans des raisonnements dont le but est la reconnaissance d'objet ou l'interprétation d'image. La reconnaissance et l'interprétation des images médicales est une tâche complexe qui nécessite l'utilisation d'une connaissance experte. En effet, les structures cérébrales sont souvent petites, leurs frontières sont souvent mal définies (comme dans le cas du thalamus), et le contraste avec la matière environnante ne permet pas toujours de les distinguer clairement. De plus, la résolution des images n'est pas très élevée. Les descriptions anatomiques usuelles telles que *neuranat*¹ ou *neuronames*² reposent principalement sur l'utilisation des relations spatiales. La figure 3.1 présente un exemple de cette connaissance. L'imprécision naturelle des relations spatiales leur permet de rester plus stables face à la variabilité inter-patients, comparé à des propriétés intrinsèques des structures anatomiques telles que leur forme ou leur taille.

De nombreux travaux ont utilisé les relations spatiales pour l'interprétation des structures cérébrales. Colliot (2003) propose d'utiliser les relations spatiales comme une force supplémentaire dans le cadre d'un algorithme de segmentation par modèles déformables. Dans Khotanlou *et al.* (2009), les relations spatiales sont utilisées dans le cadre d'une segmentation des structures cérébrales en présence de pathologies (de tumeurs cérébrales dans ce cas). Nempont (2009) propose d'utiliser les relations spatiales dans un réseau de contraintes qui, après propagation, procure les emplacements des structures. Il est alors possible de les segmenter de manière automatique. Dans cette dernière approche, il est possible de gérer les cas pathologiques en effectuant au préalable une étape de détection et de localisation de la tumeur, ce qui permet de l'inclure dans le réseau de contraintes. En revanche, si la présence de la tumeur n'a pas été détectée au préalable, le modèle ne peut s'adapter automatiquement lors de la propagation. Ici, nous nous plaçons dans le cadre d'une segmentation séquentielle, comme dans les travaux d'O. Colliot et H. Khotanlou, guidée par une représentation par graphe de la connaissance.

Dans ce chapitre, nous présentons la définition du graphe qui représente la connaissance spatiale que nous utilisons pour effectuer des raisonnements. Nous introduisons également les notations qui seront utilisées dans le reste de ce document. Nous discutons ensuite deux sources possibles de connaissances, différentes de la connaissance experte utilisée pour le raisonnement, et qui ont donné lieu à des travaux dans le cadre de la thèse. Ces travaux nous permettent de considérer des manières différentes d'obtenir un modèle et nous discuterons des conséquences pour le raisonnement spatial possible.

Dans la partie 3.1, nous décrivons quelle forme de connaissance spatiale nous utilisons et

¹<http://www.chups.jussieu.fr/ext/neuranat/>

²<http://rprcsgi.rprc.washington.edu/neuronames>

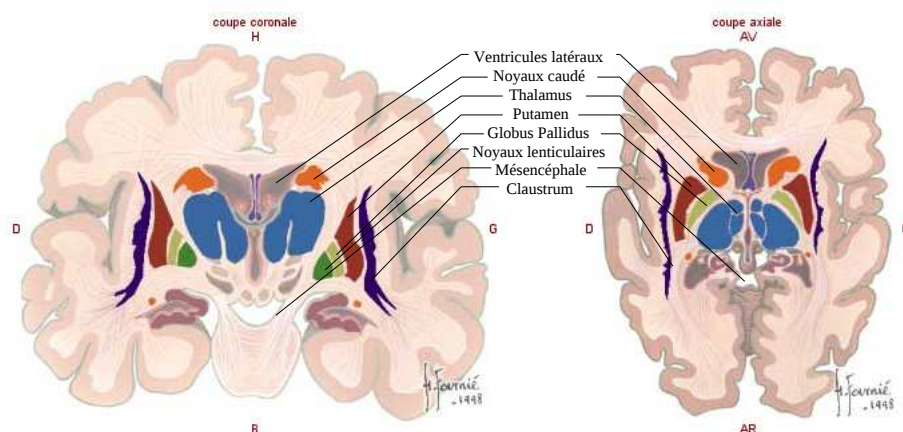


FIG. 3.1 – Exemple d’illustration provenant de l’atlas neuranat et représentant ici les structures composant les noyaux gris du cerveau.

quelle structure nous pouvons employer pour représenter la connaissance spatiale. Ensuite, nous discutons des différentes sources de connaissances possibles dans la section 3.2. Nous considérons le cas d’une connaissance experte, le cas d’une connaissance extraite automatiquement et le cas d’une connaissance extraite de manière semi-interactive. Dans chacun de ces cas, nous discutons des conséquences sur le raisonnement spatial possible avec chacune des sources. Le formalisme de représentation des relations spatiales et plus spécifiquement les relations spatiales qui seront utilisées plus tard sont présentés dans la section 3.3. Dans la partie 3.4, nous passons en revue les différentes bases de données que nous utiliserons par la suite. Enfin, dans le cadre de la connaissance experte, et plus particulièrement dans le cadre de la reconnaissance des structures cérébrales, nous verrons dans la section 3.5 comment réaliser un apprentissage des paramètres des relations spatiales.

3.1 Graphe de relations spatiales

Nous présentons dans cette partie les relations spatiales, puis le graphe qui les porte avant d’introduire les notations utilisées.

3.1.1 Les relations spatiales pour l’imagerie médicale

Une manière naturelle de décrire les relations entre les différents objets qui composent une scène est de décrire leur positions relatives, comme par exemple « *l’objet A est à droite de l’objet B* ». L’interprétation des images cérébrales appartient à un domaine où les relations spatiales sont très utilisées, comme le démontrent les livres d’anatomie tels que [Waxman \(2000\)](#). Il s’agit ici de relations spatiales textuelles.

Les relations les plus courantes sont des relations « métriques » telle que les relations directionnelles et les relations de distance, qui permettent de décrire d’une manière naturelle et imprécise les relations entre structures. Mais d’autres relations sont également bien adaptées aux structures cérébrales telle que : la symétrie, l’inclusion ou encore la relation « entre ».

Nous présentons à présent les relations utilisées ensuite dans nos applications :

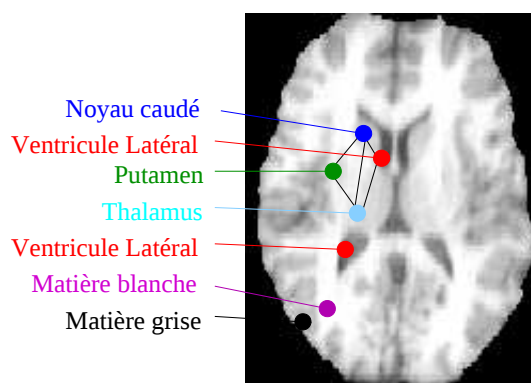


FIG. 3.2 – Une coupe d’image IRM du cerveau avec quelques structures internes étiquetées. Les structures sont présentes de manière symétrique dans les deux hémisphères. La matière blanche englobe les structures présentées. La matière grise est située plutôt sur l’extérieur du cerveau. Il faut noter que sur toutes les coupes du cerveau présentes dans ce document, l’hémisphère gauche est situé à droite de l’image. Les graphes de relations spatiales tiennent compte de cette orientation.

Orientation : Les relations d’orientation sont les relations les plus intuitives pour décrire la position relative de plusieurs structures : la structure *A* est « à droite » ou « à gauche », ou « en avant », ou bien « en arrière » de la structure *B*, ou encore, en trois dimensions, « au-dessus » ou « en-dessous ». Par exemple, dans la figure 3.2, le putamen est sur l’image à gauche du noyau caudé, lui même est à gauche du ventricule latéral. On peut noter sur cet exemple, avec cette coupe en particulier, que ces relations sont imprécises.

Distance : Différentes relations peuvent être déduites de la notion de distance, en particulier les notions imprécises telles que « loin de » ou « proche de ». Dans la figure 3.2, le noyau caudé est proche du ventricule latéral. Nous verrons dans les représentations des relations spatiales que des relations topologiques comme l’adjacence peuvent être exprimées comme une relation de distance.

La symétrie : Le cerveau possède un plan de symétrie, le plan inter-hémisphérique, et nombre de structures apparaissent de manière symétrique de chaque côté de ce plan. La symétrie peut donc être d’une grande utilité pour le raisonnement spatial dans le cerveau. Les relations directionnelles « gauche » et « droite » sont d’ailleurs souvent exprimées en fonction de ce plan de symétrie et deviennent « intérieur » (entre la structure et le plan de symétrie) ou « extérieur », ce qui permet de décrire une relation de la même manière quel que soit l’hémisphère (Colliot (2003)).

Mais bien entendu, la symétrie peut être mise à mal par la présence d’une tumeur dans un hémisphère du cerveau. L’analyse d’anomalies dans la symétrie des deux hémisphères cérébraux a d’ailleurs été utilisée comme méthode pour détecter la présence d’une pathologie (Khotanlou *et al.* (2009)).

Entre : La relation « entre » (Bloch *et al.* (2006)) est une relation ternaire permettant donc de définir l’espace se trouvant entre deux structures. L’utilisation de cette relation permet d’être plus précis que l’utilisation de deux relations de direction à partir des deux mêmes structures (ce n’est pas équivalent). La principale difficulté de cette relation est d’être ternaire, ce qui empêche sa modélisation dans le cadre d’un graphe « classique » (mais elle est possible en utilisant un hypergraphe). Néanmoins, elle peut être représentée par une relation ad-hoc utilisant deux arcs désignant

exactement la même relation, et comportant chacun les informations permettant la représentation de la relation complète.

Inclusion : La relation d'inclusion est naturelle dans une structure comme le cerveau et permet de prendre en compte le facteur d'échelle dans les structures. Toutes les structures se trouvent dans le cerveau, ce qui est implicitement pris en compte en appliquant un masque binaire de la segmentation du cerveau, sur l'image originale à segmenter. De la même manière, les deux hémisphères se trouvent dans le cerveau, etc.

3.1.2 Graphe de relations spatiales

Les graphes sont bien adaptés pour représenter une connaissance générique telle que les objets d'une scène et les relations spatiales entre ces objets. Chaque nœud du graphe représente un objet de la scène, et un arc du graphe porte la ou les relations spatiales identifiées entre deux objets de la scène. Les relations spatiales binaires sont directement intégrées dans ce modèle.

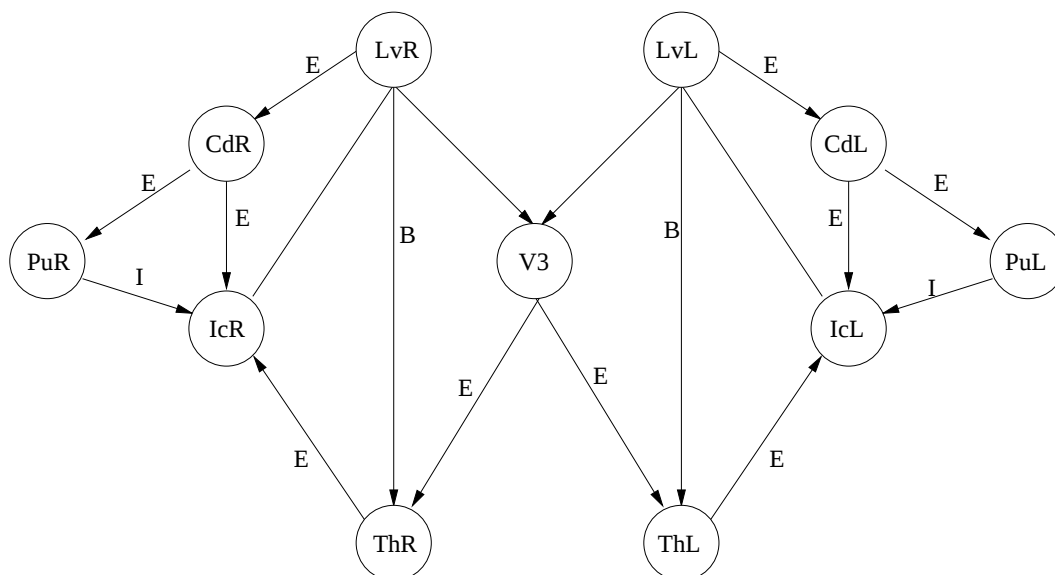


FIG. 3.3 – Extrait du troisième niveau d'un graphe des structures internes du cerveau (source Colliot (2003)). Les structures présentes sont Lv : ventricule latéral, Cd : noyau caudé, Pu : Putamen, Th : Thalamus, V3 : 3ième ventricule, Ic : capsule interne. Les relations spatiales E : extérieur, I : intérieur, B : bas.

Une modélisation par graphe hiérarchique des structures cérébrales et des relations spatiales entre ces structures a été proposée par Colliot (2003), et un extrait du troisième niveau du graphe est présenté dans la figure 3.3. La modélisation du cerveau complet est effectuée avec un graphe hiérarchique. Les différents niveaux sont reliés avec des relations d'inclusion (le cerveau est au premier niveau). Cette modélisation permet de représenter les structures de différents niveaux et les relations d'inclusion qui les relient. Nous ne nous intéressons pas à une modélisation aussi complète dans notre cas.

Les relations ternaires, comme « entre » ou les relations de symétrie qui nécessitent un lien vers le plan de symétrie, ne sont pas représentées naturellement par un graphe simple. Il est possible de les représenter avec un hyper-graphe, où les hyper-arcs notamment permettent de relier directement un nombre quelconque de structures. Néanmoins, si le pouvoir de représentation des

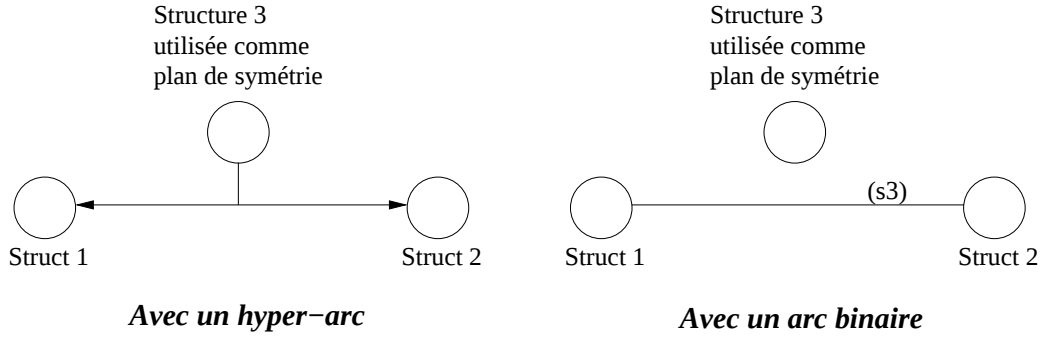


FIG. 3.4 – Exemple de modélisation d’une relation ternaire à l’aide d’un arc binaire : Nous avons une relation « La structure 1 est symétrique à la structure 2 par rapport à la structure 3 ». Avec un hyper-graphe, les trois structures peuvent être reliées à l’aide d’un hyper-arc. Avec un graphe classique, un arc simple est utilisé pour relier les deux structures symétriques, et l’axe de symétrie devient un champ de l’arc.

hyper-graphes permet de gérer les relations ternaires ou plus généralement n-aires, il est également possible de gérer les relations ternaires avec des relations binaires, et donc de conserver des graphes simples. Pour cela, il est alors nécessaire de munir un arc simple des informations manquantes. Par exemple, pour une relation de symétrie, le plan de symétrie est renseigné dans l’arc. Cet exemple est présenté dans la figure 3.4.

3.1.3 Notations

Nous introduisons ici des notations qui seront utilisées dans le reste du document, pour le graphe et les relations spatiales. D’autres notations sont présentées dans d’autres parties de ce chapitre.

Pour le graphe :

Nous utiliserons pour le graphe les notations suivantes :

V	: ensemble fini de nœuds
Σ_V	: l’ensemble des étiquettes des nœuds
L_v	: interpréteur de nœuds $L_v : V \rightarrow \Sigma_V$
E	: ensemble de couples (ordonnés) de nœuds dénomés « arcs »
Σ_E	: l’ensemble des étiquettes des arcs
L_e	: interpréteur d’arc $L_e : E \rightarrow \Sigma_E$
$G = (V, L_v, E, L_e)$	: graphe attribué avec des arcs orientés.
$\delta(v, e)$	: Pour chaque nœud $v \in V$ et chaque arc $e \in V \times V$, $\delta(v, e)$ est une fonction de transition qui retourne le nœud v' tel que $e = (v, v')$
$A(v)$	: Pour chaque nœud $v \in V$, $A(v)$ retourne l’ensemble des arcs sortants connectés à v
$p = (v_1, v_2, \dots, v_n)$	: un chemin de longueur n étiqueté $l_p = (v_1, e(v_1, v_2), v_2, \dots, v_n)$

Pour les relations spatiales :

Un arc orienté entre deux nœuds v_i et v_j comporte au moins une relation spatiale entre les deux objets représentés par les nœuds. Nous définissons une base de connaissance KB qui définit

toutes les relations spatiales existant entre les différents objets, c'est-à-dire entre les structures anatomiques dans le cas de l'interprétation des images médicales :

$$KB = \{v_i R v_j, v_i, v_j \in V, R \in \mathcal{R}\}$$

et

$$e = (v_i, v_j) \in E \iff \exists R \in \mathcal{R}, (v_i R v_j) \in KB$$

où $\mathcal{R}$ désigne l'ensemble des relations, et E l'ensemble des arcs d'un graphe.

3.2 Sources de connaissances

En fonction des domaines ou des applications, plusieurs sources de connaissances peuvent être disponibles, mais elles ne sont pas toutes équivalentes, en particulier au niveau des raisonnements possibles. Nous présentons, en plus de la connaissance experte utilisée pour l'imagerie cérébrale, deux autres sources de connaissance.

3.2.1 Connaissance experte et textuelle

Dans le cadre de l'imagerie médicale, les descriptions anatomiques fournissent souvent les relations spatiales existant entre les structures anatomiques ([Waxman \(2000\)](#)). Ces relations spatiales sont décrites d'une manière textuelle, donc sémantique, ce qui laisse la liberté du choix de formalisme de représentation et permet de conserver toute l'imprécision naturelle de ce type de connaissance. La figure 3.3 présente un extrait de graphe représentant les structures cérébrales qui illustre ce genre de connaissances.

Stabilité des relations

Cette connaissance peut être considérée comme valide et stable, tant que la variabilité des différents cas est prise en compte par la modélisation des relations spatiales utilisées. Néanmoins, les cas pathologiques sont à même d'invalider des relations, entièrement ou partiellement, voire de détruire des structures. Il est donc utile d'étudier les paramètres de ces relations pour plus de précision et pour être en mesure de détecter et de prendre en compte les cas pathologiques.

3.2.2 Connaissance extraite automatiquement

Si aucune connaissance extérieure n'est disponible, il est toujours possible d'extraire une représentation structurée d'une image. Ce type de représentation est utilisé par exemple dans des problèmes de catégorisation d'images, car les représentations structurées d'images permettent de faire apparaître les constituants de l'image et leur relations, et ainsi de les utiliser pour la catégorisation au lieu de caractéristiques globales moins pertinentes. En particulier, les relations entre les constituants peuvent être des relations spatiales ([Aldea et al. \(2007a,b\)](#)).

Dans ce type de problème, l'extraction de la sémantique de l'image est implicitement effectuée lors de l'apprentissage du modèle d'une classe à l'aide d'une base d'entraînement. C'est-à-dire que l'apprentissage des caractéristiques d'une classe d'objets structurés doit faire apparaître le motif caractéristique de la structure de la classe et ne pas tenir compte du bruit, du fond de l'image ou des variations intrinsèques. Par exemple si l'on doit apprendre automatiquement une classe de « chiens » en utilisant des images de différents types de chiens variant en couleurs, formes et tailles, et sur différents fonds possibles (forêt, champ, eau, intérieur de maison, etc.), alors le motif structurel sera par exemple la structure anatomique du chien, tête, corps, quatre pattes,

queue. Mais encore une fois, la sémantique apparaît de manière implicite, aucun des objets n'étant identifié individuellement.

Les relations spatiales apportent une connaissance stable sur une classe, et il est donc intéressant d'essayer d'extraire de manière automatique un modèle de l'agencement spatial des éléments d'une classe. Nous avons proposé en collaboration avec Emanuel Aldea (Aldea (2009)) d'extraire un tel modèle et d'effectuer un apprentissage de ce modèle pour la classification d'images. Une première approche proposée par Aldea *et al.* (2007a) est une méthode de classification d'images à partir de noyaux marginalisés pour des graphes. Dans cette approche, les images sont représentées par des graphes d'adjacence à partir d'une sur-segmentation automatique en régions. La similarité entre graphes est définie par une méthode de noyaux généralisés et permet de construire un classifieur d'images.

Nous avons étendu cette approche afin de prendre en compte non seulement des attributs intrinsèques aux régions du graphe, mais également des attributs structurels portés par les arcs du graphe. Ces attributs sont des relations spatiales représentées par des ensembles flous. Dans ce type d'approche, les graphes qui sont comparés, que ce soit pour l'apprentissage ou pour la classification, ne sont pas isomorphes et les composants du graphe ne sont pas identifiés. Il est donc nécessaire d'utiliser des relations spatiales qui permettent d'effectuer des comparaisons entre n'importe quelles composantes de l'image. Il est nécessaire dans ce cas que la comparaison de relations spatiales puisse être effectuée de manière symétrique.

Nous avons proposé d'utiliser pour cela des relations spatiales métriques, une orientation, ainsi que des relations topologiques comme une adjacence floue, et une mesure pouvant être vue comme un degré d'adjacence. Les définitions des représentations floues des relations spatiales de distance, d'orientation ou d'adjacence floue sont présentées dans la partie 3.3. La notion de degré d'adjacence est plus spécifique à ces travaux, nous allons donc la présenter brièvement.

Mesure d'adjacence fondée sur une comparaison floue

La distance et l'orientation ne sont pas toujours significatives. Par exemple, la distance ne distingue pas deux régions adjacentes par un unique pixel de deux régions imbriquées. Dans ce dernier cas, un histogramme d'angles³ n'a pas beaucoup de sens non plus. Nous proposons donc une autre caractéristique, topologique, qui estime un degré d'adjacence entre deux régions.

Nous estimons le degré d'adjacence entre deux régions en mesurant la corrélation entre la portion de l'espace « proche » de la première région dite de référence et la deuxième. Cette mesure est maximale lorsque la région de référence est imbriquée dans la région cible. Elle est nulle si les deux régions sont trop éloignées l'une de l'autre. Une valeur moyenne implique que deux régions sont adjacentes pour au moins la moitié du contour de la région de référence.

La représentation de la relation « proche de » par un ensemble flou est définie dans la partie 3.3. La figure 3.5 présente deux exemples de représentations de cette relation.

Il est nécessaire d'évaluer cette représentation avec une valeur réelle afin qu'elle puisse être utilisée dans le processus. L'évaluation est effectuée en calculant un critère de satisfaction floue (Bouchon-Meunier *et al.* (1996)) et de ressemblance floue :

$$Sat(proche(R_1), R_2) = \frac{\sum_{x \in S} \min(\mu_{proche(R_1)}(x), \mu_{R_2}(x))}{\sum_{x \in S} \mu_{proche(R_1)}(x)},$$

et

$$Res(proche(R_1), R_2) = \frac{\sum_{x \in S} \min(\mu_{proche(R_1)}(x), \mu_{R_2}(x))}{\sum_{x \in S} \max(\mu_{proche(R_1)}(x), \mu_{R_2}(x))}.$$

³Un histogramme d'angles représente, pour deux objets A et B , les angles entre le segment formé par un couple de points (a, b) , $a \in A, b \in B$ et un axe de référence

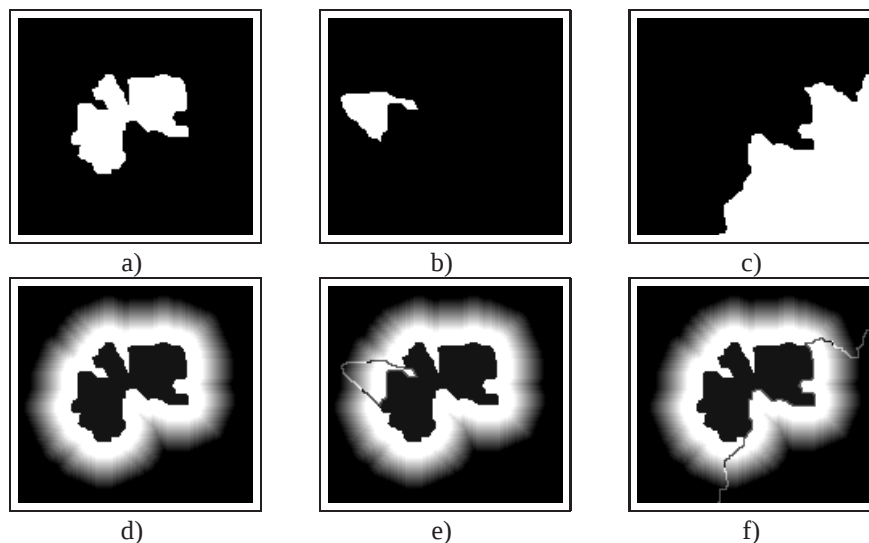


FIG. 3.5 – (a) Région 1. (b) Région 2. (c) Région 3. (d) Sous-ensemble flou correspondant à la relation « proche de la région 1 ». (e) Même chose avec la frontière de la région 2 en sur-impression. La satisfaction floue dans ce cas est de 0,06. (f) De même avec la région 3. La satisfaction floue est de 0,29.

où $\mathcal{S}$ désigne l'espace de l'image, R_1 et R_2 sont deux régions de l'image, $\mu_{proche}(R_1)$ l'ensemble flou qui représente la relation « proche de » la région R_1 , et μ_{R_2} l'ensemble flou qui représente la région R_2 . L'utilisation de ces valeurs dans le processus de classification et les résultats de classification sont présentés dans les travaux de thèse d'E. Aldea ([Aldea \(2009\)](#)).

3.2.3 Connaissance extraite de manière semi-interactive

Entre la connaissance experte et la connaissance extraite automatiquement, il existe également la possibilité d'extraire de la connaissance de manière interactive, avec un utilisateur expert ou non. Toutefois, à cause des limitations intrinsèques du mode d'interaction et des limitations dues à la pondération entre la nécessaire simplicité des interactions et à la quantité d'informations nécessaire pour obtenir un modèle générique, la connaissance extraite est forcément partielle.

Proposer à l'utilisateur de lui-même désigner les objets ou les classes de segmentation permet d'obtenir un problème de segmentation bien posé, c'est-à-dire où l'utilisateur exprime le résultat souhaité, ce qui n'est pas le cas de beaucoup de méthodes de segmentation classiques. Toutefois, il n'est pas évident de mettre en œuvre une telle interaction, à moins d'avoir déjà segmenté les objets présents sur l'image, ce qui simplifie en effet le problème. Nous allons maintenant voir comment effectuer cette interaction pour créer un modèle.

3.2.3.1 Les traces de l'utilisateur

Dans [Consularo et al. \(2007\)](#), l'utilisateur dépose des traces sur une image. Aucune contrainte n'est imposée à l'utilisateur. Chaque couleur correspond à une classe de segmentation différente, et la ou les traces correspondantes ne sont pas forcément connectées. La figure 3.6 montre une image, ainsi qu'un exemple de plusieurs ensembles de traces qui pourraient y être déposées en fonction de différents problèmes.

Il y a toutefois une contrainte implicite pour la construction du modèle et du graphe associé. Afin de pouvoir exprimer des relations spatiales, il est nécessaire d'avoir au moins deux nœuds

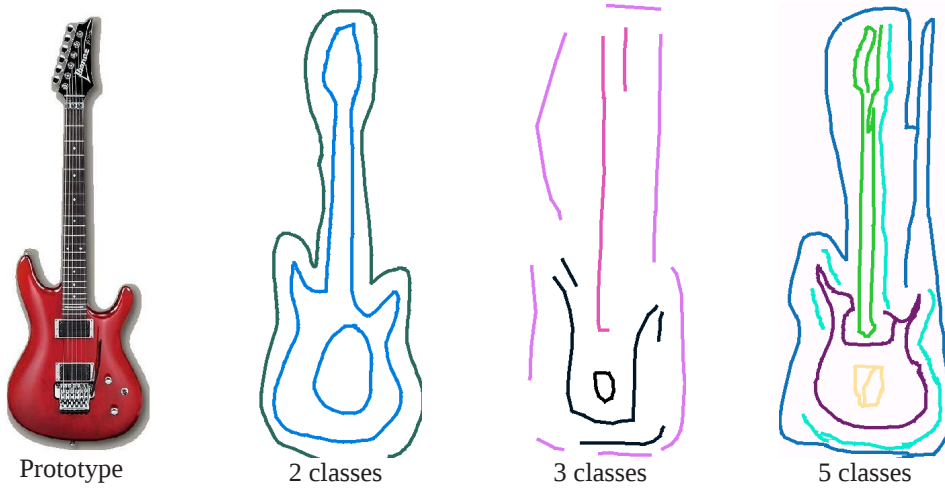


FIG. 3.6 – Exemple de traces utilisées pour construire un modèle à partir d’une image. Chaque couleur désigne une classe de segmentation différente. En fonction du problème, le nombre de classes et leur aspect changent. On peut souhaiter segmenter uniquement la guitare par rapport au fond (2 classes, image de gauche), ou segmenter la guitare en plus ou moins de constituants (3 classes au centre ou 5 classes à droite).

dans le graphe. En pratique, il sera préférable d’en avoir plus de deux, surtout si l’une des classes correspond au fond de l’image. En effet, le but est de représenter un objet sous forme d’un ensemble structuré d’objets.

3.2.3.2 Récupération des objets

L’utilisateur dessine des traces sur l’image, indiquant ainsi le nombre de classes de segmentation et les emplacements approximatifs des objets à segmenter. En revanche nous ne possédons pas de segmentation de l’image sur laquelle l’utilisateur a dessiné les traces. Nous pouvons ainsi effectuer une sur-segmentation de l’image et récupérer ainsi les régions intersectant les traces. Toutefois, les régions issues d’une sur-segmentation n’ont pas de sémantique propre. À moins d’utiliser un processus d’apprentissage tel que dans le cas d’une connaissance extraite de manière automatique (ce qui nécessiterait un gros travail de l’utilisateur), les informations recueillies ne sont pas suffisantes pour créer un modèle un tant soit peu générique d’objets structurés.

Il existe de nombreuses méthodes permettant de segmenter une image à partir de graines, comme la ligne de partage des eaux avec marqueurs ([Meyer \(2001\)](#)). Toutefois, le problème ici n’est pas d’obtenir un partitionnement de l’image, mais d’isoler les objets pointés par l’utilisateur. Le partitionnement de l’image impose que chaque partie de l’image, même ambiguë, soit attribuée à une étiquette représentant un objet. De plus, nous considérons ici un exemple pour créer un modèle, pas une image « parfaite ». Il est donc préférable de laisser les zones ambiguës attribuées à aucun objet.

Nous avons étudié une autre approche pour permettre la création du modèle. Il s’agit, avant de segmenter l’image de manière automatique, de la simplifier en effectuant une régularisation. La méthode utilisée est décrite dans [Darbon et Sigelle \(2006a,b\)](#) et permet d’optimiser de manière exacte des fonctionnelles du type :

$$F(u) = \|u - f\|_1 + \beta \int_{\Omega} |\nabla u| dx ,$$

où ∇ correspond à un gradient.⁴ Ce modèle a deux avantages : il permet de supprimer les textures, et donc cela permet de transformer une zone texturée pointée en une zone homogène et donc de la rendre facilement segmentable de manière automatique ; le deuxième avantage est de produire de larges zones homogènes de l'image en supprimant des détails, mais tout en préservant les discontinuités importantes, ce qu'un lissage classique ne permettra pas. Cette régularisation comporte cependant deux problèmes : tout d'abord il est nécessaire de fixer le paramètre β qui permet de contrôler la régularisation (plus β est élevé, plus la régularisation sera forte), et qui donc a une grande influence sur le résultat. Ce paramètre est dépendant de l'application et dans notre cas, nous avons une unique image. Il est donc très difficile d'estimer correctement ce paramètre. De plus, cet algorithme fonctionne de manière exacte (et très rapidement), mais uniquement sur des images en niveaux de gris, ce qui impose de perdre l'information de couleur. La figure 3.7 illustre l'influence du paramètre de régularisation sur le modèle résultant.

Enfin, pour récupérer les « objets », une segmentation automatique est ensuite effectuée, avec une méthode par « mean-shift » (Comaniciu et Meer (2002)) qui comporte également des paramètres comme la taille minimale des régions en sortie, mais ce paramètre en particulier n'a que peu d'influence après la régularisation.

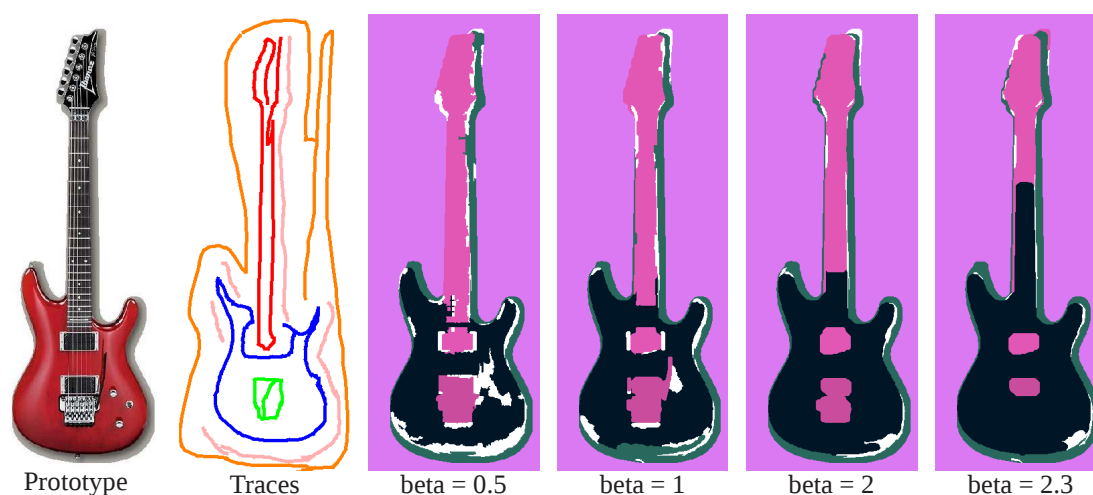


FIG. 3.7 – Influence du paramètre β de régularisation sur le modèle généré. Dans les quatre cas présentés, l'image de départ, les traces utilisées et tous les autres paramètres sont identiques. Le modèle est composé par 5 classes. Les parties apparaissant en blanc sont les zones ambiguës de l'image et n'appartiennent à aucune classe. Seul le paramètre de régularisation est modifié. Lorsque la régularisation est faible, la segmentation automatique produit beaucoup de petites régions qui sont ambiguës (principalement celles qui demeurent entre des traces différentes). Plus la régularisation est forte, et plus l'ambiguïté diminue, mais des zones comme le manche de la guitare sont fusionnées avec le corps.

3.2.3.3 Interprétation des traces

Les traces dessinées par l'utilisateur appartiennent aux objets désignés par l'utilisateur. Toutefois, ces objets ne sont pas forcément homogènes. Par exemple, l'utilisateur peut choisir en fonction du problème de désigner un personnage comme étant un objet à part entière ou désigner

⁴Les programmes correspondant sont accessibles sur cette page :
<http://jerome.berbiqui.org/total-variation-code/>

plusieurs de ses parties comme des objets. La régularisation permet de retrouver les objets pointés par l'utilisateur dans leur ensemble, qui peuvent appartenir à la même classe de segmentation.

Si l'utilisateur souhaite désigner une zone uniforme, nous effectuons les hypothèses suivantes :

- si la zone est plutôt fine, alors la trace correspondra plus ou moins au squelette morphologique de la région,
- si la zone est plutôt large, alors la ou les traces suivront les contours de la région. Dans ce cas, les régions même non homogènes comprises entre des traces d'une même classe de segmentation peuvent être considérées comme faisant partie de cette région. La figure 3.8 montre les régions regroupées de cette manière dans le modèle. La première image montre la segmentation automatique de l'image et la troisième image montre les objets initiaux, composés des régions qui intersectent les traces correspondant à cet l'objet. Sur la dernière image, de nombreuses régions notamment en bas et dans la région centrale sont regroupées, étant cernées par des régions attribuées aux même objet (ou un bord).

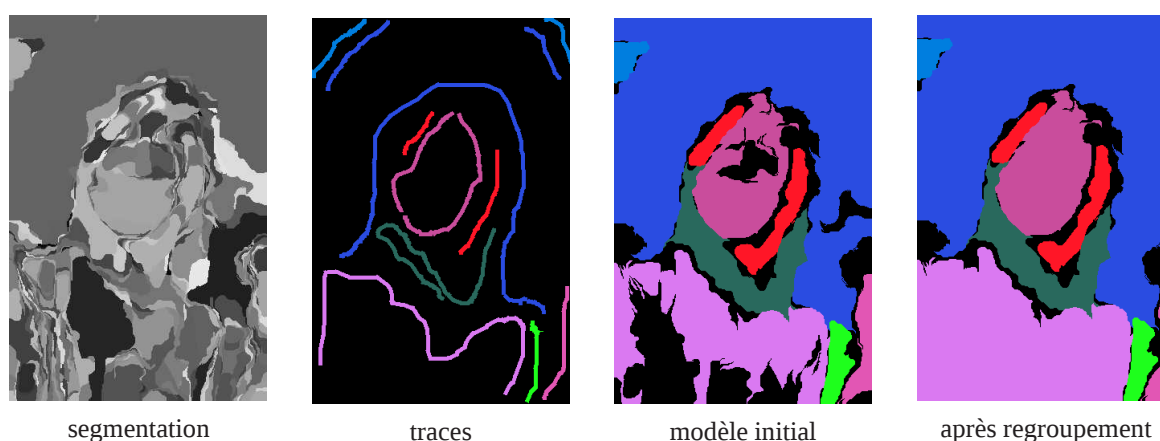


FIG. 3.8 – Regroupement de régions par déduction. L'image de gauche montre l'affectation des régions effectuée en fonction des traces uniquement. Les régions (où groupes de région) non marquées (en noir sur l'image) peuvent être affectées à une région si elles sont entourées par une unique région. Par contre, les régions se situant entre deux marques différentes sont considérées comme ambiguës et sont exclues du modèle.

Dans le cas où les traces sont utilisées pour relier plusieurs objets dans une même classe de segmentation non homogène, alors les traces ne correspondent plus aux deux cas de figure présentés. Dans ce cas, les traces vont surtout relier les différentes zones, afin de les marquer comme appartenant à une même classe sémantique.

Il pourrait être intéressant de regrouper des régions à une région adjacente en fonction de critères correspondant aux caractéristiques de la région adjacente. Par exemple si une trace intersecte une région homogène en termes de couleur, alors le critère de couleur devient plus important que d'autres pour regrouper d'autres régions avec cette première région. À l'inverse, si une région est texturée, alors le critère de couleur devient moins important.

3.2.3.4 Création du modèle

Ces hypothèses permettent de récupérer, à partir de l'image originale, une segmentation automatique et des traces de l'utilisateur, un modèle composé d'objets désignés par l'utilisateur ou des groupes d'objets regroupés dans une même classe de segmentation. En ajoutant les zones considérées comme ambiguës, le modèle est également une partition de l'image originale.

Le fait que le modèle soit composé d'objets plutôt que de régions issues d'une segmentation permet de déduire une sémantique de la structure de la scène, par exemple une relation directionnelle entre deux objets, qui serait sans signification si elle était effectuée à partir de régions d'une sur-segmentation. Une inférence vers des relations spatiales textuelles est possible, et permettrait d'ajouter de l'imprécision dans le modèle. Par exemple, à partir d'un histogramme d'angles, revenir à une direction générique (droite, gauche, haut, bas, en avant, en arrière). De même, pour une fonction de distance, connaissant les dimensions de l'image, les notions de proche ou loin peuvent en être déduites.

Les relations directionnelles ou de distance sont toujours définies quels que soient les deux objets concernés, et ces relations sont donc privilégiées. Mais il serait intéressant de déduire des relations plus spécifiques ou plus complexes, qui sont aussi éventuellement plus discriminantes ou informatives. Par exemple, si nous considérons le cas de deux régions où l'une forme un trou dans la deuxième, la distance et la position relative donnée comme une orientation seront deux relations spatiales moins pertinentes qu'une relation spécifique « entouré par ». De même dans le cas d'une région entourant une autre, ou le long d'une autre, etc.

Le graphe résultant peut être construit de plusieurs manières en fonction du nombre d'arcs souhaité. Au minimum, ce sera un graphe d'adjacence, où deux régions en contact direct ou indirect sont reliées par un arc. Une connexion indirecte serait constituée de deux régions reliées par une région classée comme ambiguë précédemment. Au maximum, le graphe peut être complet. En fonction du degré d'adjacence retenu, la première version peut omettre des liaisons importantes. La version complète peut mettre au même niveau des liaisons importantes et d'autres non significatives. Il serait nécessaire d'étudier le modèle en fonction du nombre de connexions. Ces travaux et une application sont décrits dans l'annexe C.

3.2.4 Conclusion sur les sources de connaissances

TAB. 3.1 – Les différents types de sources de connaissances, leurs formes et les conséquences sur le type des objets manipulés et le raisonnement spatial.

Type de source	Forme	Concepts manipulés	Relations spatiales
experte	textuelle : livre d'anatomie	objets identifiés	relations imposées, sémantique possible repère imposé
automatique	quelconque	régions de segmentation	pas de sémantique tout doit être relatif pas de repère
utilisateur	descripteurs visuels	classe sémantique : l'utilisateur identifie les objets	raisonnements possible sémantique possible repère imposé

Le choix d'une source de connaissances représente un compromis entre la généralité du modèle et sa précision. La connaissance experte telle qu'une relation spatiale décrite de manière textuelle va permettre par son imprécision naturelle de prendre en compte des variations naturelles. Cependant, nous avons dans ce cas un modèle qu'il est nécessaire d'instancier. De plus, il faut avoir accès à une connaissance experte, ce qui n'est pas forcément le cas dans tous les domaines d'application. Dans le cas où la connaissance est acquise de manière automatique, nous pouvons manipuler directement des régions, et donc calculer des relations spatiales de manière précise mais

elles ne sont pas identifiées. La version semi-interactive est intermédiaire, elle permet de manipuler des objets, mais qui ne sont pas identifiés. Mais dans ce cas, il faudrait que l'utilisateur soit un expert pour arriver à un modèle aussi complet que dans le premier cas. Du reste, des problèmes d'optimisation se posent pour ce genre de modèle. Dans le cadre de l'imagerie cérébrale, nous avons accès à une grande connaissance experte telle que des descriptions anatomiques, et nous allons utiliser cette connaissance par la suite.

3.3 Formalisme flou pour les relations spatiales

Les relations spatiales portées par le modèle sont représentées à l'aide d'un formalisme flou. Ce type de représentation permet de modéliser l'imprécision intrinsèque de relations telles que « proche de » ou encore « derrière », la variabilité potentielle, même si elle est plutôt réduite dans le cas d'images normales (c'est-à-dire non pathologiques dans le cas des images médicales), et la nécessaire souplesse pour effectuer un raisonnement spatial (Bloch (2005)).

Deux types de problématiques peuvent apparaître lorsque des relations spatiales sont utilisées :

- étant donné deux objets, éventuellement flous, comment déterminer le degré de satisfaction d'une relation entre ces deux objets ;
- étant donné un objet de référence, comment définir la région de l'espace dans laquelle une relation spatiale par rapport à cette référence est satisfaite à un certain degré. Nous nous intéressons ici à cette question.

Nous utilisons donc des représentations spatiales des relations, c'est-à-dire un ensemble flou dans le domaine spatial $\mathcal{S}$ définissant une région dans laquelle une relation R à un objet de référence A est satisfaite. Le degré d'appartenance de chaque point à cet ensemble flou correspond au degré de satisfaction de la relation en ce point (Bloch (2005)). La figure 3.11 illustre le type de représentation utilisé pour représenter une relation de distance.

Pour représenter les différentes relations spatiales, nous utilisons les intervalles flous, dont un exemple est présenté dans la figure 3.9. Ce sont des ensembles flous particuliers, dont chaque α -coupe (coupe de niveau) représente un intervalle classique.

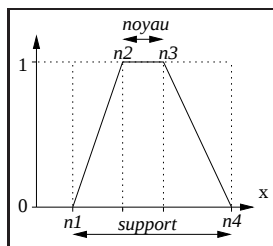


FIG. 3.9 – Un intervalle flou de forme trapézoïdale.

Nous allons à présent décrire comment calculer les représentations des principales relations spatiales utilisées dans nos expériences : une distance, une direction relative et l'adjacence (ou « très proche de »).

3.3.1 Représentation de la relation de distance

Une relation de distance peut être définie comme un intervalle flou f d'une forme trapézoïdale sur $\mathbb{R}^+$, un exemple est illustré dans la figure 3.10. Un ensemble flou μ_d de l'espace de l'image $\mathcal{S}$ peut être dérivé en combinant f et une carte de distance d_A à l'objet de référence A :

$$\forall x \in \mathcal{S}, \mu_d(x) = f(d_A(x)), \quad (3.1)$$

où

$$d_A(x) = \inf_{y \in A} d(x, y) \quad (3.2)$$

La figure 3.11 montre un exemple de représentation d'une fonction de distance utilisant cette définition.

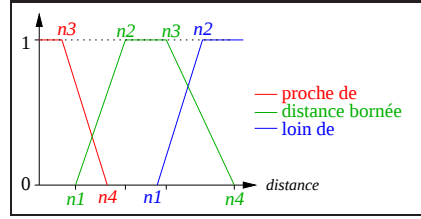


FIG. 3.10 – Intervalles flous de forme trapezoïdale illustrant trois relations spatiales de distance. Le premier (en rouge) représente une relation « proche de ». Dans ce cas, les valeurs $n1$ et $n2$ valent 0. Le deuxième nombre flou (en vert) représente une distance bornée des deux côtés. Le dernier nombre flou (en bleu) représente une relation « loin de ». Dans ces cas, les valeurs $n3$ et $n4$ sont au maximum de la distance.

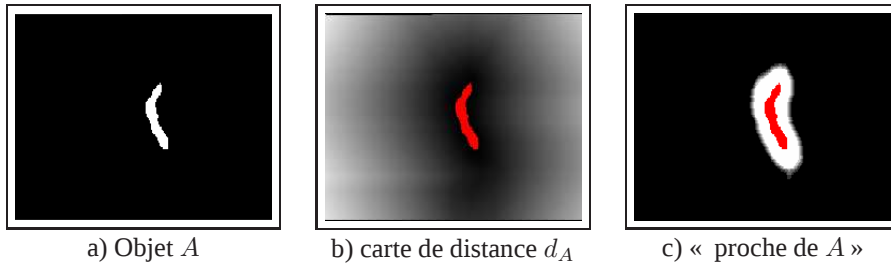


FIG. 3.11 – (a) Une coupe de la représentation binaire en 3 dimensions d'un ventricule latéral. (b) Carte de distance dérivée de A . (c) Ensemble flou correspondant à la relation « proche du ventricule latéral ».

3.3.2 Représentation de la relation d'orientation

Les relations directionnelles sont représentées en utilisant l'approche dite des « paysages flous » (Bloch (1999)). Une dilatation morphologique δ_{ν_α} par un élément structurant ν_α représentant la sémantique de la relation « dans la direction α » est appliquée à l'objet de référence A :

$$\mu_\alpha = \delta_{\nu_\alpha}(A) \quad (3.3)$$

où ν_α est défini, pour $x \in \mathcal{S}$ exprimé en coordonnées polaires (ρ, θ) , tel que :

$$\nu_\alpha(x) = g(|\theta - \alpha|) \quad (3.4)$$

où g est une fonction décroissante de $[0, \pi]$ vers $[0, 1]$ et $|\theta - \alpha|$ est défini modulo π . Cette définition est étendue en 3 dimensions en utilisant deux angles pour définir une direction. La représentation de la relation directionnelle illustrée par la figure 3.12 a été générée en utilisant cette définition.

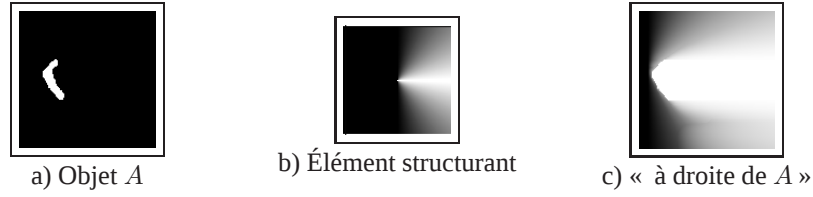


FIG. 3.12 – (a) Une coupe de la représentation binaire en 3 dimensions d'un ventricule latéral. (b) Élément structurant pour la relation « à droite ». (c) Paysage flou représentant « à droite du ventricule latéral ».

3.3.3 Représentation de l'adjacence

Une adjacence *stricte* est une relation qui est très sensible à la segmentation des objets et sa satisfaction peut dépendre d'un unique point. La figure 3.13 illustre cette sensibilité. Pour éviter une définition trop stricte (binaire), et donc n'offrant que peu de souplesse, nous avons choisi une interprétation de l'adjacence comme une relation « très proche de ». Cette relation peut donc être définie comme une fonction de la distance entre deux ensembles, donnant un degré d'adjacence plutôt qu'une valeur booléenne, en utilisant une formulation similaire au calcul de la représentation de la relation de distance décrite ci-dessus :

$$\mu_{adj}(A, B) = h_{A,B}(d(A)) \quad (3.5)$$

où h est un intervalle flou d'une forme trapézoïdale, mais dont les 3 premières valeurs sont en 0 tel que celui présenté dans la figure 3.14 et $d()$ est une carte de distance.

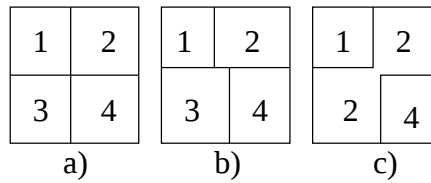


FIG. 3.13 – Illustration de la sensibilité d'une définition stricte de l'adjacence en fonction de la segmentation obtenue. (a) Quatre régions adjacentes deux à deux, une segmentation correcte donnera une image similaire faisant apparaître les régions. (b) Une segmentation possible où les régions 1 et 4 ne sont plus adjacentes à la suite du déplacement d'une frontière. (c) Dans le pire cas, il reste deux couples de régions adjacentes : (1, 2) et (2, 4) à la suite de la fusion des régions 2 et 3 de l'image originale, due à un déplacement de frontières important.

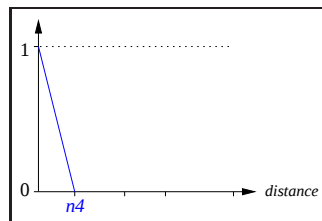


FIG. 3.14 – Nombre flou de forme trapézoïdale utilisé pour la définition de la notion d'adjacence, vue comme une relation de distance « très proche de ». Les valeurs de $n1$, $n2$ et $n3$ sont toutes égales à 0.

3.3.4 Autres relations

D'autres relations peuvent être définies d'une manière similaire (Bloch (2005)). Ces modèles sont génériques, mais pour toutes les définitions de représentations spatiales présentées ici, le degré de satisfaction d'une relation dépend d'une fonction (f , g ou h) qui est choisie comme un intervalle flou de forme trapézoïdale par simplicité. Une procédure d'apprentissage décrite dans la section 3.5 définit les paramètres de ces fonctions en fonction d'une base d'images et de la sémantique de la relation en fonction du domaine.

3.3.5 Notations des paysages flous

L'interpréteur d'arc L_e associe à chaque arc un ensemble flou μ_{Rel} , défini dans le domaine spatial $\mathcal{S}$, représentant une fusion conjonctive de toutes les représentations spatiales des relations portées par cet arc, par rapport à l'objet de référence, c'est-à-dire l'origine de l'arc. Pour qu'un arc existe entre deux nœuds, il est nécessaire qu'au moins une relation spatiale existe entre les deux objets représentés par deux nœuds, μ_{Rel} ne peut donc pas être vide. Si $\mu_{R_i}^e$, $i = 1, \dots, n_e$ sont les n_e relations portées par un arc e , alors μ_{Rel}^e s'exprime de la manière suivante :

$$\mu_{Rel}^e = \top_{i=1..n_e}(\mu_{R_i}^e) \quad (3.6)$$

avec $\top$ une t-norme (conjonction floue, voir Dubois et Prade (1980) pour une présentation des conjonctions et des disjonctions floues).

3.4 Base de données d'images cérébrales

Pour réaliser l'apprentissage décrit dans la section suivante, nous avons besoin d'une base d'apprentissage qui illustre la diversité des cas rencontrés, afin de pouvoir représenter cette diversité. Nous avons besoin :

- de cas sains qui représentent les cas « normaux »,
- de cas pathologiques, représentatifs des différentes pathologies existantes,
- que certaines structures de ces images soient segmentées afin de pouvoir calculer les ensembles flous représentant les relations spatiales et calculer leur adéquation.

Nous utilisons les structures cérébrales suivantes :

Dans les deux hémisphères :

- Ventricule latéral
- Noyau caudé
- Thalamus
- Putamen

Dans le plan inter-hémisphérique :

- Le troisième ventricule

Si l'image est pathologique :

- La tumeur

soit 9 ou 10 structures par image.

Base de cas sains :

Nous avons constitué une base de 30 images pour les cas sains. Parmi ces images, nous utili-

sons la base IBSR (Internet Brain Segmentation Repository)⁵ qui contient 18 images IRM en 3 dimensions de cerveaux humains, manuellement segmentées par des experts. Toutes les images de la base IBSR ont été recalées, ce qui diminue la variabilité de la base, mais cela n'a pas d'impact sur l'apprentissage des relations, qui est effectué de manière relative entre les structures. La figure 3.15 présente quelques coupes de volumes issus de la base IBSR.

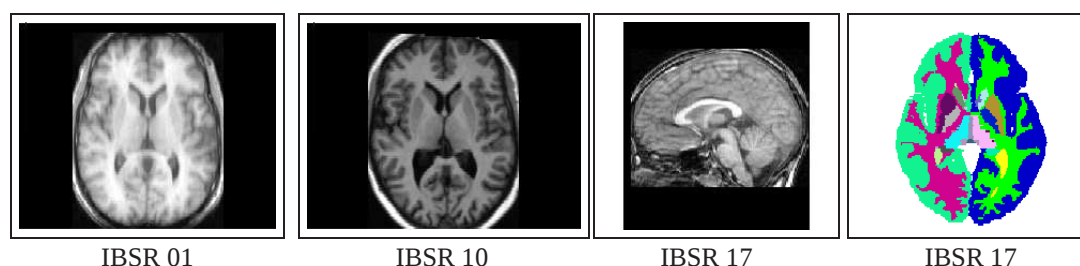


FIG. 3.15 – Trois exemples de volume de la base IBSR. Les deux premiers exemples sont des coupes axiales et la coupe présentée est la même dans les deux images (120), le troisième exemple est une coupe sagittale (coupe 127). La dernière image est une coupe d'une segmentation représentée avec une palette aléatoire.

Nous avons ajouté à ces images 11 cas provenant de la base OASIS (« Open Access Series of Imaging Studies »).⁶ Cette base contient des images de 416 sujets avec 3 ou 4 images IRM par sujet (obtenues dans une unique session). La base est présentée par [Marcus et al. \(2007\)](#). Mais cette base ne possède pas de segmentations. Nous avons donc segmenté manuellement les 11 cas que nous utilisons dans notre ensemble de cas sains.

Enfin, nous avons ajouté une dernière image, qui n'est pas accessible publiquement, et qui a été segmentée manuellement également.

Base de cas pathologiques :

Pour les cas pathologiques, nous avons constitué un ensemble de 20 images, qui ont été segmentées manuellement également et validées par des experts. Mais ces images, recueillies auprès de nos partenaires médicaux ne sont pas non plus accessibles publiquement.

La base est constituée de 16 cas, et pour deux d'entre eux, nous avons deux images à différents stades de développement de la tumeur. Pour un cas, nous avons trois images à différents moments. Nous avons donc 20 images au total, dont 16 seront utilisées pour l'apprentissage. La base contient différents types de tumeurs. La figure 3.16 présente différents exemples de pathologies issus de cette base. Les tumeurs varient :

- par leur emplacement : frontal, proche des noyaux internes, ou latérales ;
- par leur taille : plus ou moins grandes ;
- par leur type : systique, nécrotique, infiltrante, avec ou sans œdème.

Les différents types de tumeur impliquent différents comportements spatiaux, certaines vont déplacer ou déformer les structures que nous recherchons, d'autres auront peu ou pas d'impact en fonction de leur localisation. Si la tumeur produit un œdème, alors c'est l'aspect qui sera modifié plutôt que les caractéristiques morphologiques.

⁵Internet Brain Segmentation Repository. The MR brain data sets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at <http://www.cma.mgh.harvard.edu/ibsr/>

⁶<http://www.oasis-brains.org>, réalisée avec les financements suivants : Pubmed Central submission : P50 AG05681, P01 AG03991, R01 AG021910, P50 MH071616, U24 RR021382, R01 MH56584

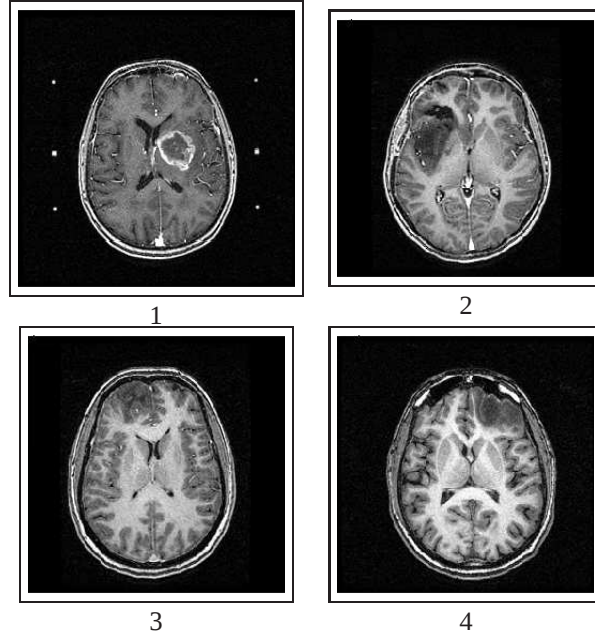


FIG. 3.16 – Quatre exemples de volumes présentant des tumeurs cérébrales. Certaines tumeurs influent directement sur les structures centrales du cerveau, comme les images du haut. D'autres tumeurs excentrées ou infiltrantes ont une influence moindre sur ces structures.

Notations :

La base d'apprentissage K , sera composée de cas sains ainsi que de cas pathologiques :

$$K = \{K^N, K^P\}$$

avec K^N l'ensemble des cas sains de la base et K^P l'ensemble des cas pathologiques de la base.

Nous pouvons dénoter par $k^i, i \in [[1, \dots, N + P]]$ un cas de la base d'apprentissage. Par simplicité, nous dénotons en pratique par $c \in K$, pour désigner un cas quelconque de la base, éventuellement en précisant un sous-ensemble K^N ou K^P . L'ensemble des objets segmentés dans c sera dénoté par O_c .

Les ensembles flous pour une relation $R \in KB$ des ensembles d'images de cas sains seront notés μ_R^N et ceux pour les images de cas pathologiques μ_R^P .

3.5 Apprentissage des paramètres des intervalles flous

Nous présentons ici de quelle manière les fonctions f , g et h (respectivement équations 4.1, 4.4 et 4.5), qui sont toutes choisies ici comme des intervalles flous de forme trapézoïdale, peuvent être apprises. L'apprentissage est nécessaire pour deux raisons :

- Permettre de prendre en compte le domaine d'application. Si certaines relations sont moins dépendantes du contexte, comme des relations d'orientation ou la relation « entre » pour laquelle c'est en partie implicite, des relations comme « proche de » ont impérativement besoin d'un apprentissage afin de pouvoir être représentées de la manière décrite dans la section 3.3.
- Si l'imprécision permet de prendre en compte la variabilité naturelle des caractéristiques telles que la forme ou la taille des structures cérébrales, les cas pathologiques apportent des variations qui peuvent être bien plus importantes. L'apprentissage va permettre de prendre

en compte les cas pathologiques, du moins ceux qui n'entraînent pas de destructions de structures. La figure 3.17 présente un exemple des effets qu'une tumeur peut avoir sur les structures cérébrales environnantes.

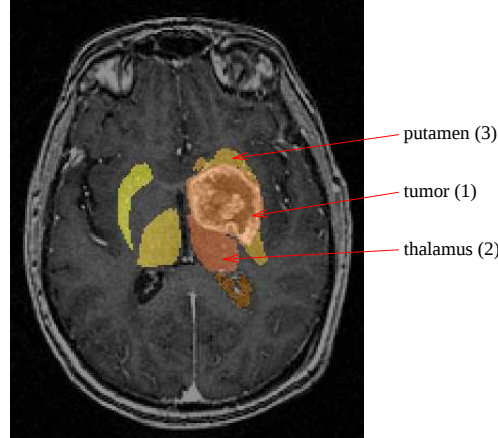


FIG. 3.17 – Exemple de l'effet d'une pathologie sur les structures cérébrales. Dans ce cas, le putamen gauche (à droite sur l'image) a été déplacé et apparaît étiré et enroulé autour de la tumeur.

3.5.1 Cadre général pour l'apprentissage des intervalles flous

Nous désirons mesurer l'adéquation d'une relation spatiale R , pour un couple de structures donné (A, B) . Dans la suite, la relation spatiale sera calculée en utilisant la structure A comme structure *de référence*, et la structure B sera la structure *cible*. L'objectif de l'apprentissage est que, pour une relation ARB donnée, l'objet B satisfasse le plus possible la relation R calculée à partir de A , il faut donc maximiser l'inclusion de l'objet B dans μ_{R_A} . Les ensemble flous μ_R sont calculés dans l'espace de l'image, ce qui nous permet de les comparer directement avec les objets de l'image.

Pour cela, nous utilisons une procédure dite « leave-one-out » qui consiste à laisser l'image qui sera utilisée pour le test, hors de la base d'apprentissage. Cette procédure impose un apprentissage différent pour chaque image, mais permet, lorsque le nombre total d'image est faible, d'avoir plus d'images pour l'apprentissage que si nous avions défini un ensemble de test séparé de l'ensemble d'apprentissage.

Pour tous les cas $c \in K$ de la base d'apprentissage, et tout couple de structures $(A_c, B_c) \in O_c$ dans l'ensemble des structures segmentées du cas concerné, nous représentons l'ensemble flou $\mu_{R_{A_c}}$ de la relation R avec A_c comme objet de référence. Dans ce cas, les paramètres des fonctions sont génériques, mais tout de même adaptés au domaine d'application.

Nous pouvons alors extraire les valeurs suivantes :

$$\min_c = \min_{x \in B_c} \mu_{R_{A_c}}(x) ,$$

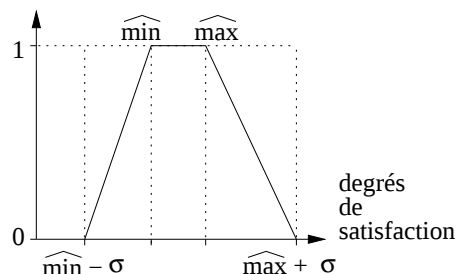
$$\max_c = \max_{x \in B_c} \mu_{R_{A_c}}(x) .$$

Ces valeurs correspondent aux degrés minimum et maximum de la relation $\mu_{R_{A_c}}$ pour tous les points de B_c

Les valeurs de satisfaction minimale $\min_c$ et maximale $\max_c$ sont calculées pour chaque instance de la base de cas et ces valeurs sont utilisées pour déterminer les paramètres des fonctions.

Nous calculons les valeurs suivantes : $\hat{min}$ est la moyenne des min_c , et σ_{min} l'écart-type de ces mêmes valeurs. Nous calculons également la moyenne $\hat{max}$ et l'écart-type σ_{max} des max_c . Les valeurs du nombre trapézoïdal apprises seront alors :

n1 :	$\hat{min} - \sigma_{min}$
n2 :	$\hat{min}$
n3 :	$\hat{max}$
n4 :	$\hat{max} + \sigma_{max}$



En fonction des relations considérées, certaines valeurs peuvent être fixées à l'avance, par exemple pour une relation « proche de » où un unique paramètre est nécessaire.

L'intervalle flou est défini d'une manière large afin de permettre de prendre en compte tous les cas de la base d'apprentissage, en particulier les cas pathologiques. Les représentations sont utilisées dans la suite de ces travaux pour estimer la localisation des objets. Il est donc nécessaire que les objets soient effectivement situés dans la représentation, au détriment de leur précision. Cependant, il est possible qu'un cas extrême ne soit pas entièrement compris dans la localisation, une valeur moyenne étant utilisée.

3.5.2 Un exemple d'apprentissage

Considérons un exemple d'apprentissage de relation d'orientation entre deux structures cérébrales : le putamen gauche est à droite du noyau caudé gauche. La structure de référence ici est le noyau caudé. Le putamen est la structure cible. La figure 3.18 présente ces deux structures. L'objectif ici est d'apprendre les paramètres de la fonction g utilisée pour représenter la relation d'orientation qui est présentée dans les équations 3.3 et 3.4

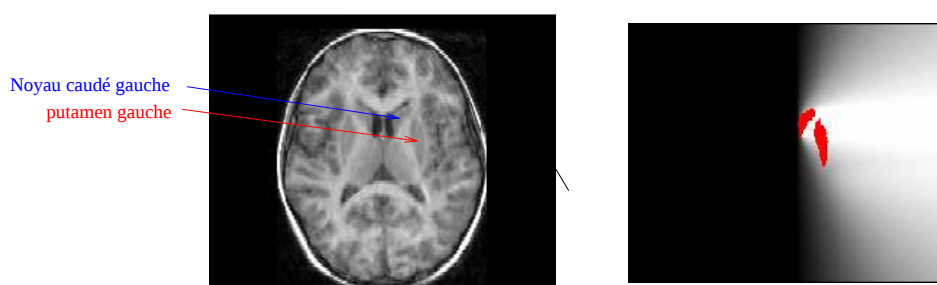


FIG. 3.18 – Apprentissage d'une relation d'orientation : les deux structures (noyau caudé et putamen) sont présentées à gauche sur une coupe de l'image ibsr 04. L'image à droite montre les deux structures en rouge en sur-impression sur une coupe de la représentation de la relation « à droite » du noyau caudé. Les valeurs de satisfaction minimale et maximale mesurées sur cette image sont respectivement de 0,37 et de 1,00.

Les valeurs minimale et maximale de satisfaction de la mesure d'inclusion pour chaque cas sont présentées dans le tableau 3.2. La moyenne des valeurs minimales est de 0,45 et l'écart type de 0,14. Les quatre valeurs du nombre trapézoïdal de la fonction g sont donc :

n1 :	$\hat{min} - \sigma_{min}$	0,31
n2 :	$\hat{min}$	0,45
n3 :	$\hat{max}$	1,00
n4 :	$\hat{max} + \sigma_{max}$	1,00

TAB. 3.2 – Valeur de satisfaction minimale et maximale obtenue pour une mesure d'inclusion I donnée entre la représentation de la relation à droite du noyau caudé gauche et le putamen gauche.

Image	Minimum ($\hat{min}$) :	Maximum ($\hat{max}$) :
cas sains		
ibsr 01	0,36	1,0
ibsr 02	0,41	1,0
ibsr 03	0,44	1,0
ibsr 04	0,37	1,0
ibsr 05	0,54	1,0
ibsr 06	0,37	1,0
ibsr 07	0,45	1,0
ibsr 08	0,42	1,0
ibsr 09	0,46	1,0
ibsr 10	0,37	1,0
ibsr 11	0,41	1,0
ibsr 12	0,38	1,0
ibsr 13	0,44	1,0
ibsr 14	0,37	1,0
ibsr 15	0,46	1,0
ibsr 16	0,40	1,0
ibsr 17	0,44	1,0
ibsr 18	0,47	1,0
cas pathologiques		
img. pat. 1	0,69	1,0
img. pat. 2	0,56	1,0
img. pat. 3	0,39	1,0
img. pat. 4	0,66	1,0
Moyenne	0,45	1,0
Écart type	0,14	0

L'intervalle flou utilisé et le résultat de l'apprentissage sont illustrés dans la figure 3.19.



FIG. 3.19 – Apprentissage d'une relation d'orientation : à gauche, nous avons le nombre trapézoïdal utilisé pour l'orientation et l'image de droite montre le résultat de l'apprentissage pour cette relation. Les valeurs sélectionnées permettent de prendre en compte l'intégralité de la structure cible dans la relation.

3.5.3 Le cas de la distance

Dans le cas des relations spatiales reposant sur une distance comme l'adjacence telle qu'elle est définie dans la partie 3.3.3, les relations « proche de », « loin de », l'apprentissage est nécessaire à deux niveaux. Il est d'abord nécessaire d'étudier dans le contexte du domaine d'application (les structures cérébrales) ce que signifie la notion de « proche de » ou de « loin de ». Une fonction f peut alors être déterminée pour chacune des relations.

Une fois ces relations connues dans le contexte du domaine d'application, nous pouvons procéder à un apprentissage tel qu'il a été décrit dans les parties précédentes, pour apprendre par exemple les paramètres de la fonction f pour une relation particulière : « le ventricule latéral est proche du noyau caudé ».

Mais en pratique, nous ne passons pas par ces deux étapes. Elles sont réalisées de manière simultanée. Il ne s'agit donc plus d'élargir la représentation floue d'une relation, mais de calculer directement les paramètres de la relation. Pour cela, un apprentissage tel que celui décrit ci-dessus est effectué. Au lieu de regarder les représentations des relations spatiales μ_R et l'inclusion avec l'objet cible de la relation, l'apprentissage est effectué sur une carte de distance calculée depuis la structure de référence. Pour une relation R et un couple de structures (A, B) , nous générons la carte de distance de la structure A :

$$d_{A_c}(x) = \inf_{y \in A} d(x, y) ,$$

Nous cherchons ensuite à extraire les valeurs suivantes de la carte de distance pour chaque cas $c \in K$:

$$\begin{aligned} dmin_c &= \min_{x \in A_c, y \in B_c} d(x, y) , \\ dmax_c &= \max_{x \in A_c, y \in B_c} d(x, y) . \end{aligned}$$

Dans ce cas, $dmin_c$ représente le minimum des distances entre un point de A et un point de B . $dmax_c$ représente le maximum de ces mêmes distances. Ces valeurs sont calculées pour l'ensemble des cas c de la base K .

Nous calculons ensuite les valeurs suivantes : $\hat{dmin}$ est la moyenne des $dmin_c$, et σ_{dmin} l'écart-type de ces mêmes valeurs. Nous calculons également la moyenne $\hat{dmax}$ et l'écart-type σ_{dmax} des $dmax_c$. Les valeurs du nombre trapézoïdal apprises seront alors :

n1 :	$\hat{dmin} - \sigma_{dmin}$
n2 :	$\hat{dmin}$
n3 :	$\hat{dmax}$
n4 :	$\hat{dmax} + \sigma_{dmax}$

Le nombre flou désigne donc un intervalle sur des distances (exprimée en mm) et non plus des degrés de satisfaction comme dans le cas général.

3.6 Conclusion

Le domaine d'application est particulier pour différentes raisons. S'il existe une variabilité inter-patients, et des modifications dues aux pathologies, nous sommes dans un cas où les objets de la scène sont connus, ainsi que leur nombre et toutes les relations qui les relient. Nous avons également la garantie que toute la scène sera visible. De ce point de vue, cette application est dans

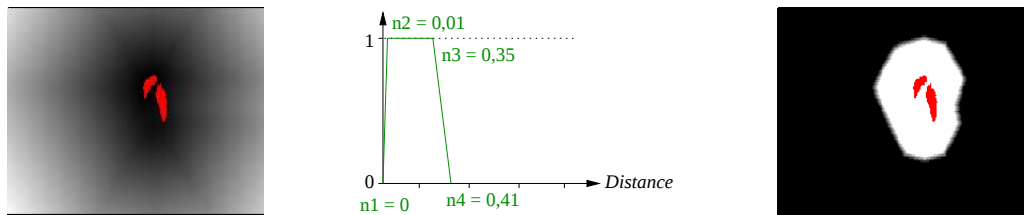


FIG. 3.20 – Exemple d'apprentissage d'une fonction de distance : « distance du noyau caudé au putamen ». Une carte de distance au noyau caudé est calculée (image de gauche). Les deux structures sont en sur-impression sur cette image. Les valeurs de l'intervalle flou g présenté au centre sont calculées sur l'ensemble des cas sains et pathologiques. Le résultat de l'apprentissage est présenté à droite, avec les deux structures en sur-impression.

un monde clos, ce qui ne serait pas le cas pour des images naturelles où la plupart des autres applications. De plus, le domaine visé est un domaine où il existe des descriptions anatomiques des relations entre les structures, à l'inverse de beaucoup de domaines où ces descriptions n'existent pas. Enfin les structures concernées ne sont pas complexes, à l'instar de l'intestin dans une étude de l'abdomen, qui nécessiterait des relations spatiales adéquates. Toutes ces caractéristiques nous permettent de faire les choix que nous avons faits : nous utilisons une connaissance experte et textuelle fournie par les références anatomiques, structurée à l'aide d'un graphe. Les relations spatiales sont représentées à l'aide d'un formalisme flou, qui permet de représenter l'imprécision du modèle. De ce point de vue, le modèle que nous utilisons ici est dédié à la reconnaissance et l'interprétation des structures cérébrales. Au contraire, les discussions sur les sources de connaissances se placent dans un cadre plus ouvert à divers types d'images traitées et de problèmes.

Chapitre 4

Optimisation avec représentation des structures

De nombreux travaux utilisent des modélisations par graphes pour guider la reconnaissance, comme par exemple dans [Colliot et al. \(2006\)](#); [Mangin et al. \(1996\)](#); [Deruyver et al. \(2009\)](#). Certains de ces travaux utilisent un processus de segmentation séquentielle. Ce type de processus permet de diviser un problème de segmentation globale en sous-problèmes, et d'ordonner la résolution de ces sous-problèmes, le principe étant de commencer par les problèmes les plus « simples », pour aller vers les plus « difficiles ». Pour la segmentation des structures cérébrales, la difficulté de segmentation d'une structure particulière est liée aux caractéristiques de la structure elle-même, comme sa forme, ou à la quantité d'information disponible sur son entourage immédiat : la structure sera plus simple à segmenter si toutes les structures avoisinantes sont déjà segmentées.

Nous avons décrit au chapitre 3 comment nous pouvons créer un modèle à partir des connaissances anatomiques sur un graphe spatial. Le processus de segmentation séquentielle nécessite d'avoir plusieurs objets à segmenter évidemment, mais également de pouvoir utiliser de l'information issue des objets segmentés pour permettre ou faciliter la segmentation des autres structures. Le modèle décrit dans le chapitre 3 est donc bien adapté à un processus de segmentation séquentielle. Le domaine d'application, la segmentation des structures cérébrales, est également bien adapté à ce type de processus car il y a de nombreuses structures à segmenter et les relations entre les structures sont décrites dans des ouvrages neuro-anatomiques comme celui de [Waxman \(2000\)](#).

Parmi les travaux cités, nous nous intéresserons particulièrement aux travaux décrits par [Colliot et al. \(2006\)](#) qui proposent d'utiliser les relations spatiales pour segmenter et reconnaître les structures cérébrales de manière progressive, en utilisant un modèle similaire au modèle décrit au chapitre 3. À chaque itération, une structure est segmentée, et sa segmentation est guidée par les relations spatiales existant entre cette structure et les structures précédemment segmentées. Toute la connaissance générique est représentée en utilisant un graphe. Une structure particulière (les ventricules latéraux) est utilisée comme structure de référence, c'est-à-dire comme point de départ de la segmentation. Enfin, une séquence ad-hoc de segmentation des structures, à partir de la structure de référence, permet de ne segmenter les structures les plus compliquées qu'une fois que suffisamment d'information est disponible. La figure 4.1 présente des résultats de segmentation obtenus grâce à ce cadre de segmentation, et avec une séquence définie de manière ad-hoc et empirique. L'objet de ce chapitre est de proposer des raisonnements permettant de remplacer la séquence de segmentation ad-hoc par une séquence de segmentation optimale par rapport à la connaissance disponible, au modèle et aux données utilisées. Une séquence de segmentation correspond ici à un chemin dans un graphe et répond donc au problème suivant : à partir d'une structure de référence,

quelles sont les segmentations successives à effectuer pour segmenter au mieux une structure objectif ?

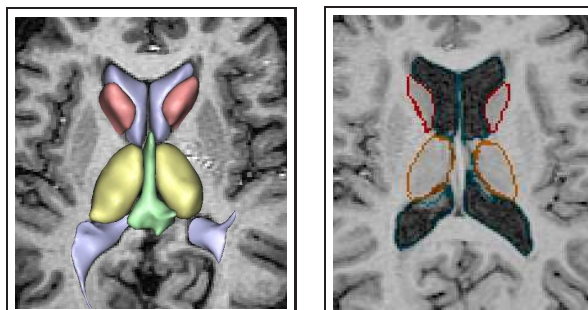


FIG. 4.1 – Les résultats de segmentation présentés par [Colliot et al. \(2006\)](#).

L'imprécision intrinsèque des relations spatiales leur confère une grande stabilité qui permet de prendre en compte la variabilité anatomique existant dans les structures cérébrales. Mais lorsqu'un cas comportant une pathologie se présente, alors la variabilité peut devenir trop importante pour pouvoir être prise en compte de cette manière. En effet, des structures peuvent être déformées, déplacées et même éventuellement disparaître. Dans ce dernier cas, la structure même de l'image est altérée. Dans ce chapitre et le suivant, nous proposons différentes méthodes permettant de déterminer une séquence de segmentation qui ne soit plus ad-hoc, mais qui soit adaptée en fonction des connaissances qui sont disponibles et des pathologies éventuelles.

Dans une première partie, nous allons présenter une approche utilisant une représentation de la forme de chaque structure, pour inférer le chemin de segmentation optimal en fonction du modèle fourni et de la connaissance a priori sur les structures utilisées. L'utilisation d'une connaissance a priori nous permet de définir un chemin complet avant de commencer à effectuer des segmentations. En revanche l'approche est dépendante de la connaissance a priori pour les formes de chaque structure, et est donc moins susceptible de s'adapter aux cas pathologiques (qui entraînent des modifications morphologiques).

Dans une deuxième partie, nous présentons donc une manière de modifier la méthode qui permet de prendre en compte les cas pathologiques. Nous considérons dans ce cas que la pathologie est connue, c'est-à-dire dans notre cas que le type de tumeur cérébrale (sa classe et ses caractéristiques) est connu.

Tous les raisonnements présentés dans les deux premières parties de ce chapitre effectuent une optimisation qui peut être qualifiée de locale, dans le sens où ils attribuent à chaque arc une mesure de manière indépendante, utilisée ensuite pour inférer le chemin. Dans une troisième partie, nous présentons une méthode permettant d'effectuer une estimation globale de la pertinence d'un chemin, c'est-à-dire permettant de faire la sélection du chemin complet à partir d'un unique critère évaluant sa pertinence.

4.1 Raisonnement avec représentation de la forme des structures

Dans cette première partie, nous proposons une méthode permettant de déterminer une séquence optimale de segmentation entre une structure de référence et une structure cible. Cette méthode utilise, pour chaque structure contenue dans le modèle, une représentation de la forme de cette structure. Nous distinguons deux cas différents. Pour commencer, le cas normal, dans le cas de la segmentation des structures cérébrales, correspond à une image sans pathologie, le cas dit

« sain ». Dans un deuxième temps, nous présentons une adaptation de cette méthode pour les cas qui présentent une pathologie.

4.1.1 Raisonnement dans le cas « sain »

Notre objectif est donc de proposer un raisonnement qui permette la sélection du « meilleur » chemin entre une structure de référence, qui sera segmentée au préalable, et une structure que nous souhaitons segmenter et reconnaître, dans une image donnée. Nous utilisons le modèle présenté dans le chapitre 3, à savoir un graphe spatial où les nœuds correspondent aux structures cérébrales et les arcs portent les relations spatiales existant entre des structures. Un chemin correspond donc à une séquence de structures à segmenter, et doit permettre de conduire à la meilleure segmentation possible de la structure objectif, en fonction de nos connaissances de la scène.

La notion de « meilleur » chemin est relative aux contraintes du processus de segmentation. Dans cette première approche, notre raisonnement doit être apte, dans le cadre du processus de segmentation séquentielle présenté par [Colliot et al. \(2006\)](#), d'être à même de remplacer le chemin de segmentation ad-hoc utilisé, le processus lui-même restant inchangé. Les contraintes du processus portent donc sur la possibilité de segmenter une structure en utilisant les relations spatiales issues des structures segmentées pour guider la segmentation des structures restantes. Notons que nous ne définissons pas, pour une structure donnée, un minimum de relations spatiales qui seraient nécessaires à sa segmentation. Cette information, plutôt empirique, pourrait être incorporée sur chaque nœud du graphe. Dans cette approche, nous ne tenons pas non plus compte de la difficulté intrinsèque de la segmentation de chacune des structures.

Dans ce cas, nous considérons des chemins simples et sans boucle et donc le nombre de chemins est borné. Ce nombre peut néanmoins être très élevé, et le problème de l'extraction d'un chemin peut rapidement devenir trop complexe pour être calculé. Mais dans notre cas, nous nous limitons à un petit nombre de nœuds, et nous évitons donc ces problèmes. Néanmoins, l'utilisation de cette méthode avec un graphe plus grand nécessiterait de résoudre cette problématique.

Nous avons donc un modèle qui porte une connaissance générique composée d'objets (des structures cérébrales) et de relations spatiales décrites de manière textuelle. Pour déterminer la meilleure séquence de segmentation, nous proposons de munir chaque relation spatiale présente dans le modèle d'une mesure de sa pertinence. Cette mesure évalue l'adéquation avec laquelle cette relation spatiale décrit l'agencement spatial entre les deux structures : la référence et la cible de la relation. Cette évaluation n'est possible qu'à partir du moment où nous avons choisi un formalisme de représentation des relations spatiales, car la comparaison va porter sur la représentation de la relation spatiale et non pas sur la relation elle-même. Nous avons décrit dans la partie 3.3 le formalisme flou utilisé pour représenter les relations spatiales. La méthode utilisée pour représenter une relation spatiale répond à la question suivante : pour une structure de référence, quels sont les lieux de l'espace où cette relation est satisfaite. Cette méthode nous permet d'obtenir une représentation dans l'espace de l'image d'une relation spatiale où chaque point de la représentation correspond à la mesure de satisfaction de la relation en ce point. Notre critère d'évaluation va donc comparer pour chaque relation un ensemble flou et une structure « cible ». Cette mesure, qui sera décrite dans la partie suivante, sera ensuite utilisée pour définir un poids sur les arcs du graphe, et également comme critère pour l'optimisation du chemin en utilisant des algorithmes dérivés de notions classiques de la théorie des graphes présentés dans la partie 4.1.3.

La structure de référence dans le cadre des images IRM du cerveau humain est le ventricule, qui présente en général un fort contraste avec les structures environnantes, et qui se situe approximativement au centre (dans une coupe axiale) du cerveau. La figure 4.1.a présente une coupe axiale d'une image IRM 3D du cerveau, et quelques structures, dont le ventricule, sont marquées.

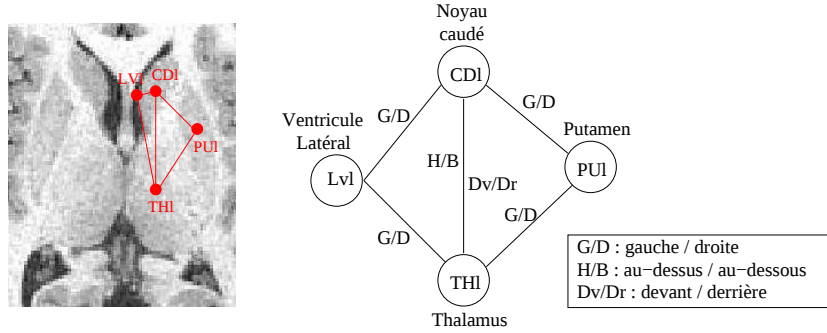


FIG. 4.2 – La connaissance et le modèle utilisés dans cette étude. L'image à gauche est une coupe d'un volume cérébral en vue axiale et montre les structures anatomiques suivantes : LVI ventricule latéral, CDI noyau caudé, THl le thalamus et PUI le putamen. À droite, les mêmes structures apparaissent dans le graphe modélisant la connaissance. Les relations spatiales entre ces structures sont portées par les arcs du graphe.

4.1.2 Évaluation de la pertinence d'un arc

Nous proposons ici deux critères permettant d'évaluer la pertinence d'un arc comme une mesure de l'adéquation entre un ensemble flou représentant la relation spatiale portée par cet arc, et de la structure cible de l'arc. Les représentations des relations spatiales sont effectuées dans l'espace de l'image.

Les notations que nous utilisons pour les graphes et pour désigner les ensembles flous ont été introduites dans le chapitre 3 dans la première partie. Nous nous contenterons ici de rappeler les notations les plus utilisées. Par la suite, nous utilisons la notation $G = (V, E)$ pour désigner un graphe relationnel attribué, avec V désignant l'ensemble des nœuds et E l'ensemble des arcs. Un interpréteur d'arc associe à chaque arc e un ensemble flou μ_{Rel} , défini dans le domaine spatial, et représentant l'ensemble (éventuellement singleton) des relations spatiales portées par cet arc, et calculé par rapport à une structure de référence qui est le nœud source de l'arc, tel que défini par Bloch (1999). De la même manière, un ensemble flou μ_{Obj} est porté par chaque nœud et correspond à la représentation de la structure cérébrale portée par le nœud.

L'ensemble flou μ_{Obj} représente la structure cible. Cette représentation peut être une simple segmentation issue de la base d'apprentissage, sous forme de carte binaire de la structure, éventuellement rendue floue à l'aide d'une dilatation floue par exemple pour accroître la prise en compte de la variabilité. Cette représentation peut également provenir d'un atlas. Il est important que les données utilisées pour calculer les représentations des relations spatiales et les représentations des structures proviennent de la même source afin que leur comparaison soit effective. Dans tous les cas, l'information provient d'une image déjà segmentée et non pas de l'image qui doit être segmentée.

Critère de pertinence

La pertinence d'une relation spatiale doit donc représenter l'adéquation entre μ_{Rel} et μ_{Obj} , c'est-à-dire le degré avec lequel les ensembles flous représentant des relations spatiales ayant une même structure pour cible donnent une localisation précise de cette structure. Si la structure cible est représentée sous forme d'un ensemble flou, alors nous avons deux ensembles flous, définis dans l'espace de l'image, et donc aisément comparables. La pertinence des relations spatiales, dans le cadre du formalisme flou de représentation choisi, nous permet de déduire deux critères de pertinence :

- la localisation de la relation,
- la précision de la relation.

Une relation spatiale fournit une indication sur la position de la structure cible par rapport à la structure de référence, la position exacte étant donnée par la connaissance a priori. Si la relation spatiale fournit une bonne localisation, alors en chaque point de la structure cible, la relation spatiale doit avoir un degré de satisfaction maximal. Plus spécifiquement, si nous comparons des ensembles flous, il est nécessaire que l'ensemble des points de la structure cible soient situés dans le noyau de la relation spatiale (c'est-à-dire un degré de satisfaction de 1).

Une bonne localisation, telle qu'elle vient d'être définie, permet de s'assurer que la relation est pleinement satisfaite à l'emplacement de l'objet. Mais ce critère n'est pas suffisant, car la taille du support de la relation spatiale n'est pas prise en compte. Le support de la relation spatiale, représentée dans l'espace de l'image, correspond à l'ensemble des points pour lesquels le degré de satisfaction de la relation n'est pas nul. Par exemple, dans un cas extrême, tous les points de l'ensemble flou peuvent satisfaire entièrement la relation. Dans ce cas, la localisation sera toujours correcte. Il est donc nécessaire de tenir compte d'un autre critère qui estime la précision de la relation. Nous la définissons comme le rapport entre la taille de l'objet et la taille du support de la relation étudiée.

Nous pouvons trouver un cadre formel approprié pour comparer des ensembles flous dans (Bouchon-Meunier *et al.* (1996)), où les auteurs proposent des mesures de comparaison ainsi qu'une classification de ces mesures. Deux mesures permettant d'estimer les critères de pertinence décrits ont été étudiées :

Mesure de satisfaction :

Le premier critère est une mesure de satisfaction (« M-measure of satisfiability (Bouchon-Meunier *et al.* (1996)) ») définie ainsi :

$$f_s(Rel, Obj) = \frac{\sum_{x \in \mathcal{S}} \min(\mu_{Rel}(x), \mu_{Obj}(x))}{\sum_{x \in \mathcal{S}} \mu_{Obj}(x)} , \quad (4.1)$$

où $\mathcal{S}$ désigne l'espace de l'image. Ce critère mesure la précision de la position de la structure dans la région où la relation est satisfaite, et sera maximale si la structure est entièrement située dans le noyau de μ_{Rel} . Cependant la taille de la région où la relation est satisfaite n'est pas prise en compte. Dans le cas extrême (mais improbable) où le support de la relation correspondrait au domaine de l'image, la relation serait alors maximale avec n'importe quel objet. Si la représentation de la structure n'est pas floue, alors cette mesure est réduite à :

$$f_{s_{crisp}}(Rel, Obj) = \frac{\sum_{x \in Obj} \mu_{Rel}(x)}{|Obj|} . \quad (4.2)$$

Mesure de ressemblance :

Le deuxième critère est une mesure de ressemblance (« M-measure of resemblance (Bouchon-Meunier *et al.* (1996)) ») définie comme :

$$f_r(Rel, Obj) = \frac{\sum_{x \in \mathcal{S}} \min(\mu_{Rel}(x), \mu_{Obj}(x))}{\sum_{x \in \mathcal{S}} \max(\mu_{Rel}(x), \mu_{Obj}(x))} . \quad (4.3)$$

Ce critère mesure l'adéquation entre la structure dans la région de l'espace où la relation est satisfaite, le maximum étant atteint si l'objet et la relation sont identiques. Cette mesure permet d'évaluer en même temps le positionnement mais aussi la précision de la relation. Il s'agit donc d'un taux de recouvrement entre μ_{Rel} et μ_{Obj} .

Comparaison des critères :

La mesure de satisfaction est maximale lorsque la localisation de la relation est correcte, mais elle ne prend pas en compte la taille du support ; la normalisation étant effectuée par la taille de l'objet cible. La mesure de ressemblance est maximale lorsque les deux ensembles flous sont identiques, permettant donc de répondre aux deux critères de localisation et de précision. Toutefois, la mesure de la précision pourrait être plus fine. En effet, les cas où l'ensemble flou correspondant à la relation spatiale est identique à l'ensemble flou représentant l'objet ($\mu_{Rel} = \mu_{Obj}$) est improbable, et pas forcément souhaitable, car une telle relation ne pourrait gérer la variabilité.

Le critère de précision pourrait être raffiné en fonction de chaque relation spatiale. Par exemple, pour une relation d'orientation, la précision concerne moins la taille du noyau que la précision des angles utilisés pour calculer la relation. D'un autre côté, le cas extrême où le support d'une relation correspond à tout l'espace de l'image ne se présente pas en pratique. Les deux critères proposés sont donc satisfaisants pour notre application. La mesure de satisfiabilité est en revanche plus simple à calculer car elle est limitée à l'ensemble flou correspondant à l'objet. La figure 4.3 illustre cela.

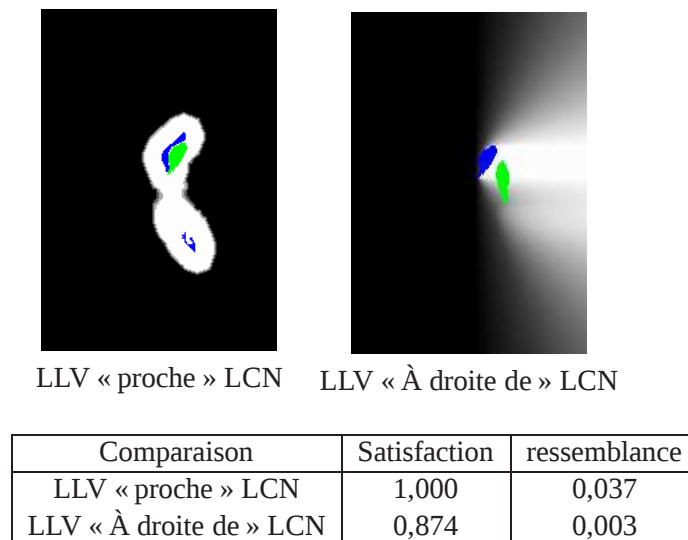


FIG. 4.3 – Comparaison des critères. La mesure de satisfaction reflète que les deux structures cibles (en vert) sont proches du noyau de la relation spatiale représentée à partir de la structure de référence (en bleu) : entièrement dans le noyau dans le premier cas (à gauche) et partiellement dans le deuxième cas. Les mesures de ressemblance ont des valeurs beaucoup plus faibles, car la normalisation est effectuée par rapport à la taille du support de la relation. Dans le premier cas, la relation « proche de » est plus précise que dans le deuxième cas où nous avons une relation d'orientation. Les valeurs reflètent principalement cette différence, plus que la position dans le noyau.

4.1.3 Sélection du « meilleur » chemin

Nous avons présenté deux mesures pouvant être utilisées pour mesurer la pertinence d'une relation (ou des relations) portée(s) par un arc du graphe. L'une ou l'autre mesure peut alors être insérée dans le graphe comme un poids. Il faut noter que chaque poids a une valeur positive et que donc il ne peut y avoir de boucle avec un poids négatif. Nous nous retrouvons donc dans une situation proche de problèmes classiques de la théorie des graphes et nous nous intéressons

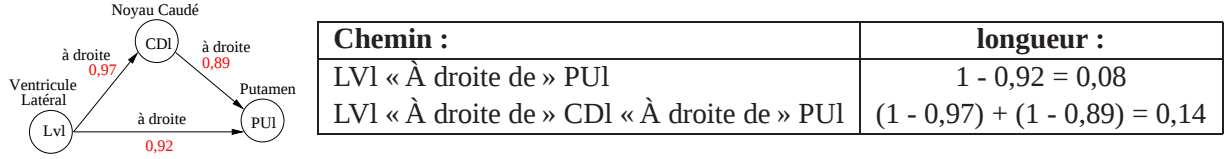


FIG. 4.4 – Un petit exemple pour illustrer le comportement du plus court chemin avec un critère de satisfaction. Nous ne considérons que des relations d’orientation dans ce cas. Nous considérons un arc qui relierait directement le ventricule au putamen, ce qui n’est pas souhaitable en pratique vu l’éloignement des structures. Le chemin direct a une longueur de 0,08 contre 0,14 pour le chemin comportant deux arcs. Pour qu’un chemin comportant deux arcs soit choisi contre un chemin comportant un unique arc même moyen, il est nécessaire que les deux arcs du chemin aient des valeurs deux fois meilleures que l’arc du chemin unique. Ce cas ne se présente pas dans nos expériences. En prenant la moyenne de satisfaction sur le chemin, le chemin plus long sera préféré.

donc aux algorithmes classiques d’optimisation pour trouver le meilleur chemin. Toutefois, nous expliquons également pourquoi les chemins qui pourraient être obtenus par ces algorithmes n’ont pas forcément les caractéristiques souhaitées pour notre problème de segmentation, nous allons donc présenter comment adapter des notions issues de ces algorithmes à notre problème.

Meilleur chemin moyen :

L’algorithme du plus court chemin effectue une optimisation globale sur le graphe. Cette optimisation peut donc accepter pour un chemin donné des valeurs assez différentes, même si le chemin est optimal. Un chemin globalement correct peut alors inclure un arc avec une faible valeur de satisfaction (ou un fort poids). De plus, cet algorithme favorise (évidemment) des chemins courts, et pas seulement des chemins avec des arcs avec un poids faible : un chemin comportant un unique arc avec un poids élevé sera préféré à un chemin comportant deux arcs avec des valeurs meilleures. Le processus de segmentation utilisant les relations spatiales pour guider la segmentation, il est important de noter que plus le chemin comporte de relations utilisables, et plus la segmentation sera encadrée. Potentiellement, un chemin plus informatif est donc plus intéressant.

Pour illustrer cela, nous pourrions envisager un chemin direct entre le ventricule et le putamen, au lieu de segmenter en premier lieu le thalamus et le noyau caudé, puis le putamen. Ce chemin serait sans doute privilégié par l’algorithme du plus court chemin, mais ne permettrait pas une meilleure segmentation à cause de l’imprécision et de l’éloignement des deux structures concernées (nous ne considérons pas une relation de distance ici, mais uniquement l’éloignement entre les deux structures qui induit une plus grande imprécision). La figure 4.4 illustre ce comportement. L’arc « direct » entre le ventricule et le putamen, grâce à sa faible longueur, est préféré à tous les autres.

Nous proposons donc une adaptation de cet algorithme en normalisant le coût de chaque chemin par sa longueur. Cette adaptation conduit à sélectionner non plus le chemin le plus court, mais plutôt le chemin qui a le plus faible poids moyen. Cette modification permet de ne pas favoriser des chemins courts par rapport aux chemins plus informatifs.

Soit $\mathcal{F}$ l’ensemble des ensembles flous sur le domaine spatial. Soit $f : \mathcal{F} \times \mathcal{F} \rightarrow \mathbb{R}$ une fonction à valeurs réelles, ici une mesure de comparaison (parmi les mesures précédemment décrites). La sélection du meilleur chemin moyen $\hat{p}$ entre deux nœuds v et v' sera le résultat de :

$$\min_{p \in \mathcal{P}} \frac{\sum_{e \in p} (1 - f(\mu_{Rel}, \mu_{Obj}))}{card(p)}, \quad (4.4)$$

où e est un arc dans le chemin p , $\mathcal{P}$ représente l'ensemble des chemins de v à v' , μ_{Obj} est l'ensemble flou représentant la structure cible de l'arc e , μ_{Rel} est l'ensemble flou représentant la relation spatiale portée par l'arc e , et $card(p)$ représente le nombre de nœuds présents dans le chemin p . Par exemple dans la figure 4.2, si v est le ventricule latéral (Lvl) et v' le putamen (PUI), un des chemins entre ces deux structures est : $Lvl - L/R - CDl - L/R - PUI$.

Plus grand flot minimal :

Le problème de l'arc « déficient », c'est-à-dire la possibilité pour un arc ne satisfaisant que peu les critères de sélection de se retrouver dans un chemin globalement bon, peut être contourné en caractérisant un chemin par son arc de flot minimal, ce qui correspond à l'arc de plus faible capacité (le poids ici) parmi les arcs du chemin. Nous proposons donc d'effectuer la sélection parmi ces arcs de plus faible capacité, en choisissant celui qui présente la plus forte valeur.

Il faut donc chercher le maximum parmi les capacités minimales de chaque chemin, et nous proposons d'optimiser le critère suivant :

$$\max_{p \in \mathcal{P}} (\min_{e \in p} (f(\mu_{Rel}, \mu_{Obj}))) \quad (4.5)$$

avec les mêmes notations que dans la méthode précédente. Nous considérons dans nos exemples des graphes avec peu de structures et de chemins possibles, nous pouvons donc effectuer l'optimisation avec une recherche exhaustive parmi tous les chemins à partir de la structure de référence vers la structure cible. Pour chaque chemin, la capacité minimale est calculée et le chemin possédant le maximum parmi ces valeurs est sélectionné.

Cette formulation permet d'éviter les chemins qui ont un arc trop faible, et donc de résoudre le problème de cet arc. De plus, n'étant pas dépendante du nombre d'arcs du chemin, elle ne favorise pas les chemins d'une longueur donnée, ce qui évite le deuxième problème souligné avec l'algorithme du plus court chemin.

4.1.4 Expériences

La figure 4.2 présente le graphe qui est utilisé dans nos expériences. Ce graphe contient 4 structures cérébrales : le ventricule latéral gauche, qui est également la structure de référence, le noyau caudé, le thalamus et enfin le putamen, qui est la structure cible dans nos expériences. Nous reprenons ici les structures et les relations spatiales utilisés par [Colliot et al. \(2006\)](#), le but étant de remplacer le chemin ad-hoc par un chemin déterminé par une optimisation du graphe. Il s'agit donc de trouver le meilleur chemin entre le ventricule et le putamen. Chacune de ces structures est présente dans les deux hémisphères du cerveau de manière symétrique par rapport au plan inter-hémisphérique. Nous ne considérons que les structures du côté gauche ici. Les expériences considèrent en outre que l'extraction de chacune des structures présente le même niveau de difficulté.

Comme source pour la représentation des structures, nous avons choisi d'utiliser un unique cas, pour lequel nous possédons une segmentation. Cependant, chacune des structures est rendue floue afin de représenter une certaine imprécision. La figure 4.5 illustre le mécanisme de « fuzzification » d'une structure anatomique. Cette « fuzzification » est effectuée en effectuant une dilatation floue de la représentation d'une structure par un élément structurant parabololoïde défini ainsi :

$$se(x, y, z) = 1 - \frac{(x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2}{\sigma^2},$$

où (x_c, y_c, z_c) représente le centre de l'élément structurant et σ est un paramètre fixé à 5 dans nos expériences.

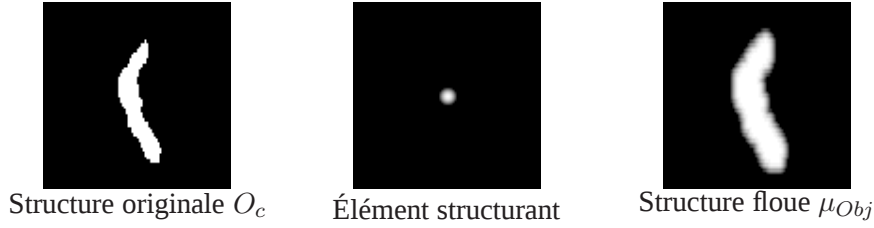


FIG. 4.5 – « Fuzzification » d’une représentation d’une structure. La structure (à gauche) provient d’une segmentation du cerveau. Une dilatation floue par un élément structurant parabolique (au centre) permet de rendre la structure floue (à droite).

Nous pouvons alors définir μ_{obj} ainsi :

$$\mu_{obj} = \delta_{se}(O_c)$$

où O_c représente une structure anatomique.

Les relations spatiales comprises dans le graphe pour nos expériences sont des relations directionnelles. Toutefois, l’extension à d’autres types de relations spatiales binaires peut être effectuée simplement. Il faut toutefois que la fusion de deux relations spatiales conserve une valeur sémantique. Les relations spatiales sont calculées à partir des représentations binaires des structures, les représentations floues sont utilisées pour calculer les critères entre la représentation de la relation et la structure cible de cette relation.

La représentation des relations spatiales est effectuée selon le formalisme flou présenté dans la partie 3.3. La figure 4.3 présente un exemple de relation spatiale d’orientation.

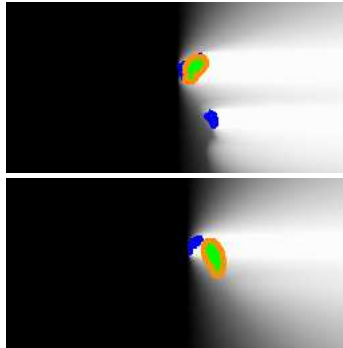
Arc			Satisf.
Ventricule latéral	« Au-dessus »	Thalamus	0,97
Ventricule latéral	« À droite de »	Noyau caudé	0,97
Noyau caudé	« En avant »	Thalamus	0,97
Thalamus	« En arrière »	Noyau caudé	0,96
Thalamus	« À droite de »	Putamen	0,92
Noyau caudé	« À droite de »	Putamen	0,89
Noyau caudé	« Au-dessus »	Thalamus	0,82
Thalamus	« Au-dessous »	Noyau caudé	0,64

TAB. 4.1 – Le classement des arcs en fonction de la mesure de satisfaction.

TAB. 4.2 – Le classement des arcs en fonction de la mesure de ressemblance.

Arc			Ressem. (x100)
Noyau caudé	« En avant »	Thalamus	0,73
Ventricule latéral	« Au-dessus »	Thalamus	0,46
Noyau caudé	« À droite de »	Putamen	0,44
Thalamus	« À droite de »	Putamen	0,42
Noyau caudé	« Au-dessus »	Thalamus	0,41
Thalamus	« En arrière »	Noyau caudé	0,40
Ventricule latéral	« À droite de »	Noyau caudé	0,32
Thalamus	« Au-dessous »	Noyau caudé	0,20

Évaluation des arcs :



Noyau caudé « À droite du » Ventricule
satisfaction : 0,97
ressemblance : 0,32

Putamen « À droite du » Noyau caudé
satisfaction : 0,89
ressemblance : 0,44

Graphe valué :

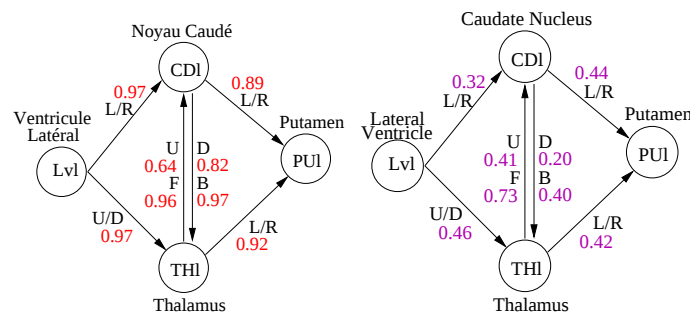


FIG. 4.6 – Les valeurs de satisfaction (en bas à gauche) et de ressemblance (en bas à droite) mesurées pour chaque arc. En haut deux exemples avec un ensemble flou μ_{rel} représentant une relation d'orientation et l'ensemble flou μ_{obj} correspondant à la cible. Les mesures de satisfaction et de ressemblance sont calculées en comparant ces ensembles flous. La structure de référence de la relation est représentée en bleu. Le support de μ_{obj} est représenté en vert orange et le noyau de μ_{obj} en vert clair.

TAB. 4.3 – Les valeurs de satisfaction moyenne et minimale pour chacun des chemins. Les valeurs en gras dans le tableau indiquent les chemins retenus.

Chemin			1 - Moyenne	Minimum
LVI « Au-dessus »	THI « En arrière »	CDI « À droite de » PUI	0,06	0,89
LVI « Au-dessus »	THI « Au-dessous »	CDI « À droite de » PUI	0,17	0,64
LVI « À droite de »	CDI « À droite de »	PUI	0,07	0,89
LVI « À droite de »	CDI « En-avant »	THI « À droite de » PUI	0,05	0,92
LVI « À droite de »	CDI « Au-dessus »	THI « À droite de » PUI	0,10	0,82
LVI « Au-dessus »	THI « À droite de »	PUI	0,06	0,92

La figure 4.6 et le tableau 4.1 présentent des mesures du critère de satisfaction obtenues pour chaque arc du graphe. Nous cherchons dans cette expérience un chemin entre le ventricule latéral et le putamen, certains arcs sont donc inutiles (les arcs qui reviennent vers le ventricule ou les arcs issus du putamen) et ne sont pas présents. Le tableau 4.3 présente les scores obtenus pour chacun des chemins selon le critère de satisfaction. Avec ce dernier, les deux méthodes proposées de sélection du chemin, le meilleur chemin en moyenne ou le chemin avec le « plus grand flot minimal », sélectionnent le même meilleur chemin qui est :

TAB. 4.4 – Les valeurs de ressemblance moyenne et minimale pour chacun des chemins. Les valeurs en gras dans le tableau indiquent les chemins retenus.

Chemin			1 - Moyenne	Minimum
LVI « Au-dessus »	THI « En arrière »	CDI « À droite de » PUI	0,57	0,40
LVI « Au-dessus »	THI « Au-dessous »	CDI « À droite de » PUI	0,63	0,20
LVI « À droite de »	CDI « À droite de »	PUI	0,62	0,32
LVI « À droite de »	CDI « En-avant »	THI « À droite de » PUI	0,51	0,32
LVI « À droite de »	CDI « Au-dessus »	THI « À droite de » PUI	0,62	0,32
LVI « Au-dessus »	THI « À droite de »	PUI	0,56	0,42

LVI « À droite de » CDI « En-avant » THI « À droite de » PUI.

Si le meilleur chemin est sélectionné en utilisant l'arc de capacité minimale, alors un autre chemin possède le même score que le premier chemin :

LVI « Au-dessus » THI « À droite de » PUI.

Ce chemin est moins intuitif que le précédent, car il implique moins de structures. En pratique, si deux chemins possèdent le même score, alors le chemin le plus long en termes de nœuds visités, qui sera donc le plus informatif, sera préféré.

La figure 4.6 (à droite) et le tableau 4.2 présentent les valeurs obtenues en utilisant le critère de ressemblance plutôt que le critère de satisfaction. Le tableau 4.4 présente les scores obtenus avec ce critère. Dans ce cas, le meilleur chemin obtenu au sens du meilleur chemin moyen est le même que le chemin obtenu avec le critère de satisfaction :

LVI « À droite de » CDI « En-avant » THI « À droite de » PUI.

En revanche, en utilisant le critère de l'arc de capacité minimale, la valeur de ressemblance portée par l'arc entre le ventricule et le noyau caudé est rédhibitoire, et le meilleur chemin est :

LVI « Au-dessus » THI « À droite de » PUI.

qui avait déjà une valeur identique en utilisant cette optimisation avec le critère de satisfaction.

Les deux critères évaluent les arcs de manières différentes et l'effet de la prise en compte de la précision par le critère de ressemblance est très visible car toutes les relations visant le noyau caudé, qui est la structure la plus petite en volume, ont les valeurs de ressemblance les plus faibles. À un arc près, les arcs sont d'ailleurs ordonnés par ordre de taille des structures cibles (hormis le ventricule qui n'est visé par aucune relation, le thalamus est le plus important, suivi du putamen et enfin du noyau caudé). La relation entre le ventricule et le noyau caudé a une valeur particulièrement faible par rapport à la robustesse de la relation, mais c'est entre ces deux structures que le rapport de taille est le plus important. Ces résultats montrent que le critère de ressemblance accorde une place trop importante à la taille des structures, et le critère de satisfiabilité sera alors préféré.

Néanmoins, les deux critères donnent des résultats relativement similaires. Ce chemin est le chemin qui avait été défini de manière empirique par Colliot *et al.* (2006), ce qui montre le potentiel de l'approche qui a permis de retrouver automatiquement un chemin qui avait été déterminé de manière ad-hoc.

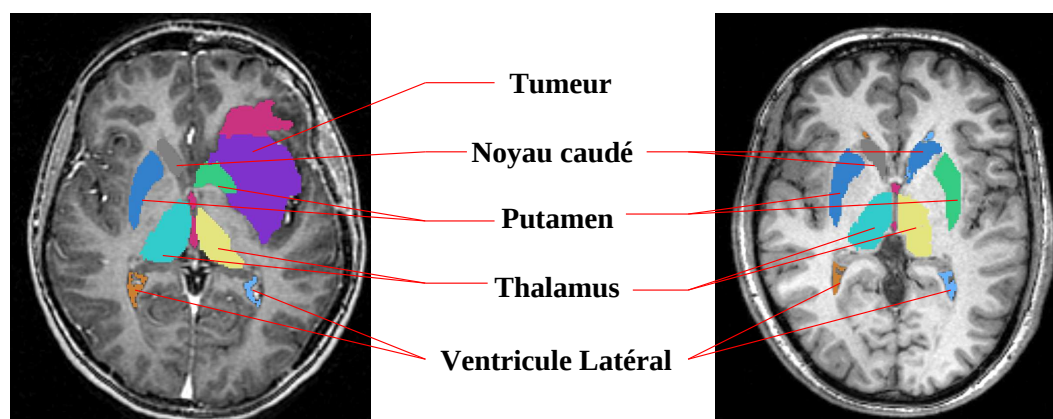


FIG. 4.7 – Deux coupes en vue axiale d’images IRM du cerveau. L’image de gauche présente une pathologie proche du ventricule latéral et des noyaux centraux. L’image de droite est un cas sain. Dans le cas pathologique, les structures ont été déplacées à cause de la tumeur. Le thalamus est écrasé, le putamen est déplacé et déformé. Le noyau caudé a également été déplacé et n’apparaît pas sur cette coupe, alors que le noyau caudé présent dans l’autre hémisphère apparaît.

4.2 Raisonnement dans le cas pathologique

L’approche que nous avons présentée dans la partie 4.1.2 n’est pas directement applicable dans le cas de la présence d’une tumeur, et nécessite une adaptation pour prendre en compte cette situation. En particulier, la présence d’une tumeur peut provoquer une importante altération de l’apparence et des caractéristiques morphométriques d’une structure. La stabilité des relations spatiales permet de prendre en compte la variabilité anatomique des structures cérébrales, mais des modifications plus profondes peuvent apparaître dans les cas pathologiques. La figure 4.7 présente un exemple de cas pathologique dans une image IRM et de l’impact de la tumeur sur les structures proches. Dans cet exemple, toutes les structures marquées ont été déplacées et éventuellement déformées comme le thalamus qui apparaît comprimé. Le putamen a également subi une grande déformation.

Nous proposons une modification de la méthode initiale qui conserve le modèle utilisé à l’identique, mais qui intègre la gestion des pathologies au niveau des poids qui sont utilisés pour optimiser le chemin de segmentation. De cette façon, il est possible de prendre en compte les modifications causées par la pathologie, mais la structure du graphe ne peut pas être modifiée, même si une structure est détruite. Il est toutefois possible avec cette nouvelle approche d’empêcher l’utilisation d’un arc si le poids prend une valeur nulle. Pour prendre en compte les effets d’un type de pathologie sur le modèle, nous exploitons une notion de degré de stabilité des relations spatiales.

4.2.1 Degré de stabilité des relations spatiales

En présence de pathologie, il a été montré par [Atif et al. \(2006a\)](#) que certaines relations spatiales sont plus stables que d’autres. Un raisonnement dépendant des pathologies a été introduit pour adapter un processus générique de raisonnement à un cas spécifique, en répondant à la question suivante : étant donné une pathologie, quelles relations spatiales demeurent stables et à quel degré ? À cette fin, un cadre pour l’apprentissage de la stabilité des relations spatiales a été mis en place par [Atif et al. \(2007a\)](#) avec une base d’images IRM composées de cas sains et de cas pathologiques manuellement segmentées représentant différentes classes de tumeurs.

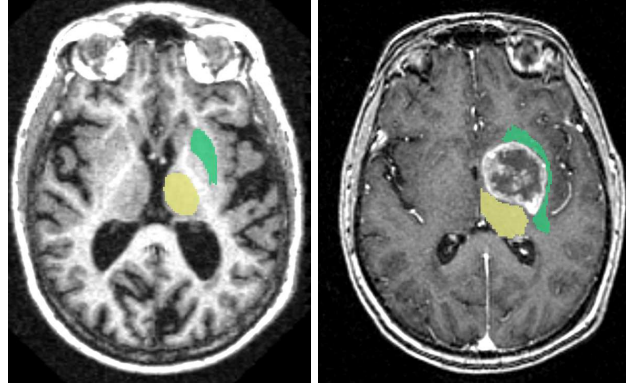


FIG. 4.8 – Étude du degré de stabilité pour la relation : thalamus « distance » putamen. L'image de gauche correspond à un cas sain, l'image de droite à un cas pathologique qui a une influence directe sur les structures, en particulier le putamen qui est déplacé et déformé.

Classification de la base d'apprentissage :

Le degré de stabilité est inféré de la comparaison (en utilisant une mesure de ressemblance) entre les relations spatiales apprises pour les cas sains et les relations apprises pour les cas pathologiques. L'apprentissage est effectué selon le protocole présenté dans la partie 3.5, mais l'apprentissage est effectué de manière distincte pour les cas sains et les cas pathologiques. Pour cela, la base de cas pathologiques est au préalable catégorisée en fonction du type de tumeur et de son impact sur les structures. La structuration de la base est effectuée à l'aide d'une classification de tumeurs cérébrales. Nous avons donc une base de cas :

$$K' = (K^N, K^{P_1}, \dots, K^{P_n}) ,$$

où K^N représente les cas sains et K^{P_i} représente les cas correspondant à une classe de pathologie P_i .

Apprentissage dans le cas sain et dans le cas pathologique :

Nous utilisons une procédure similaire à la procédure présentée dans la partie 3.5 : nous cherchons à apprendre les paramètres des fonctions f , g , et h qui modélisent respectivement les relations spatiales de distance, d'orientation ou d'adjacence respectivement pour un couple de structures (A, B) . Pour une classe de pathologie donnée, correspondant à un sous-ensemble K^{P_i} de la base d'apprentissage, nous allons effectuer pour chaque relation spatiale un apprentissage sur l'ensemble des cas sains K^N et sur le sous-ensemble K^{P_i} .

Nous pouvons considérer une relation particulière pour illustrer comment dériver le degré de stabilité d'une relation. Nous considérons ici la relation « distance » entre le thalamus et le putamen. La figure 4.8 présente ces deux structures dans un cas sain et un cas pathologique.

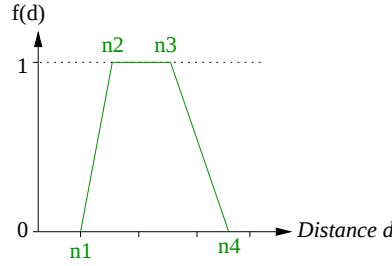
Cette relation est représentée ainsi :

$$\forall x \in \mathcal{S}, \mu_p(x) = f(d_A(x)) ,$$

où A représente le thalamus dans ce cas précis et d_A une carte de distance au thalamus définie ainsi :

$$d_A(x) = \inf_{y \in A} d(x, y) .$$

La fonction f est un intervalle flou de cette forme :



L'apprentissage consiste à apprendre les valeurs des paramètres n_1, n_2, n_3 et n_4 .

Dans le cas sain, ces paramètres sont appris sur tous les cas de l'ensemble K^N . Pour chaque cas $c^N \in K^N$, nous déterminons les valeurs suivantes :

$$\min = \min_{c^N} \min_{x \in B_{c^N}} d_{A_{c^N}}(x) ,$$

$$\max = \max_{c^N} \max_{x \in B_{c^N}} d_{A_{c^N}}(x) ,$$

où A_c représente le thalamus du cas $c \in K^N$, $d_{A_{c^N}}$ la carte de distance à A_c et B_{c^N} le putamen du cas $c \in K^N$ dans cet exemple. Ces valeurs représentent les degrés minimum et maximum de satisfaction de la relation de distance du thalamus sur tous les points du putamen.

Nous pouvons calculer la moyenne (resp. $\hat{\min}^N, \hat{\max}^N$) et l'écart-type (resp. $\sigma_{\min}^N, \sigma_{\max}^N$) des valeurs min et max calculées et ainsi dériver l'intervalle flou suivant :

- n1 : $\hat{\min}^N - \sigma_{\min}^N$
- n2 : $\hat{\min}^N$
- n3 : $\hat{\max}^N$
- n4 : $\hat{\max}^N + \sigma_{\max}^N$

Cet intervalle est défini de manière large afin de pouvoir prendre en compte tous les cas de la base d'apprentissage.

Nous considérons la classe de tumeur illustrée dans la figure 4.8, et correspondant à l'ensemble K^{P_i} . Nous effectuons une procédure d'apprentissage similaire avec cet ensemble de données. Nous obtenons ainsi les paramètres correspondant aux cas pathologiques.

Les intervalles flous de notre exemple sont illustrés par la figure 4.9. Nous avons obtenu des valeurs similaires pour le minimum de distance, et les paramètres n_1 et n_2 seront les mêmes dans le cas sain et dans le cas pathologique. En revanche, la moyenne des maximum dans le cas sain est de 31,63 et son écart-type de 4,46. Dans le cas pathologique, la moyenne est de 38.

Degré de stabilité :

La stabilité d'une relation spatiale, pour un cas de pathologie donné, est estimée en comparant les ensembles flous représentant une même relation mais avec les paramètres appris dans le cas sain d'une part, et les paramètres appris dans le cas pathologique d'autre part.

Si nous considérons dans notre exemple une image provenant de la base de cas sains $c \in K^N$, nous pouvons représenter d'une part $\mu_p^N(O_c)$ la représentation de la relation « proche de » avec les paramètres appris sur l'ensemble K^N . D'autre part nous pouvons représenter $\mu_p^{P_i}(O_c)$ la représentation de la relation « proche de » avec les paramètres appris dans sur l'ensemble de cas pathologiques K^{P_i} .

Nous pouvons à présent en déduire le degré de stabilité de la relation en comparant les deux ensembles flous ainsi obtenus. La comparaison est effectuée en utilisant une M-mesure de ressemblance (Bouchon-Meunier et al. (1996)), déjà utilisée dans l'approche initiale, qui permet de

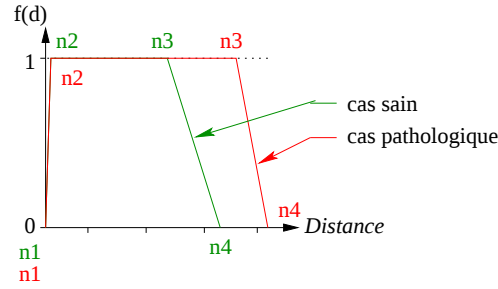


FIG. 4.9 – Intervalles flous dans le cas sain et dans le cas pathologique. Les paramètres de l'intervalle flou pour le cas sain sont les suivants : $n_1 = 0$, $n_2 = 1$, $n_3 = 32$, $n_4 = 36$. Les paramètres de l'intervalle flou dans le cas pathologique sont les suivants : $n_1 = 0$, $n_2 = 1$, $n_3 = 38$, $n_4 = 41$. Ce type de tumeur repousse le putamen, et donc la fonction est plus large.

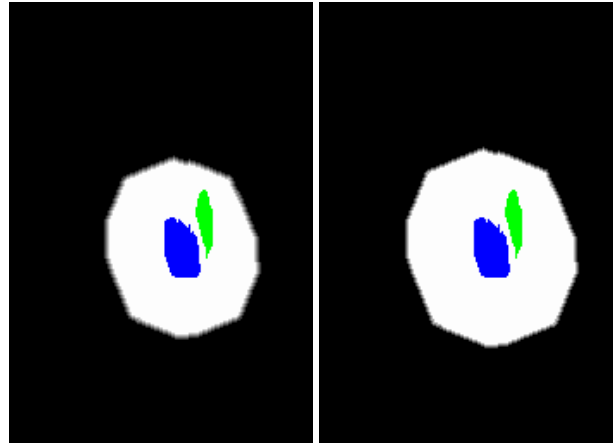


FIG. 4.10 – Les deux paysages flous représentant la relation de distance entre le thalamus (en bleu) et le putamen (en vert). À gauche, les paramètres de la fonction f ont été appris dans le cas sain. À droite, les paramètres ont été appris pour une classe de pathologie illustrée dans la figure 4.8. Le paysage flou appris dans le cas pathologique est moins précis que dans le cas sain, reflétant que le type de tumeur considéré déplace les structures considérées. La ressemblance entre ces deux paysages flous est de 0,72.

calculer la cardinalité de l'intersection de deux ensembles flous, normalisée par la cardinalité de leur réunion :

$$\mathcal{R}(\mu, \mu') = \frac{\sum_{d \in \mathcal{D}} \min(\mu(d), \mu'(d))}{\sum_{d \in \mathcal{D}} \max(\mu(d), \mu'(d))} ,$$

où $\mathcal{D}$ représente le domaine de définition des ensembles flous, par exemple l'espace des distances dans l'exemple décrit plus haut.

Nous obtenons ainsi, pour chaque relation (A, R, B) où A et B sont deux structures et R une relation spatiale, un degré de stabilité pour chaque classe de pathologie. La valeur obtenue pour notre exemple est de 0,72.

4.2.2 Adaptation de l'approche aux cas pathologiques

Nous avons estimé un degré de stabilité, qui est valable pour une classe de pathologie. Il est donc nécessaire de connaître au préalable pour quel type de pathologie l'adaptation doit être

effectuée. Pour cela, nous pouvons utiliser la classification de tumeurs cérébrales proposée par [Atif et al. \(2007a\)](#) et [Khotanlou \(2008\)](#). Une fois le type de pathologie connu et les degrés de stabilité correspondant estimés, il est nécessaire de les intégrer dans notre approche initiale. Il y a différentes manières d'intégrer cette information, nous allons en détailler deux : un élagage du graphe par le degré de stabilité, puis la pondération des poids par le degré de stabilité.

Élagage du graphe :

Le graphe original peut être filtré de telle sorte que les relations spatiales présentant un trop faible degré de stabilité soient supprimées. Ensuite, l'approche développée pour les cas sains peut être directement appliquée sur le graphe filtré. Cette approche est plutôt sévère et ne permet pas d'être souple, ce qui est primordial afin de pouvoir effectuer des raisonnements. De plus, il est nécessaire de fixer un seuil pour le filtrage du graphe. La figure 4.11 montre un exemple de filtrage de graphe, avec un seuil déterminé de manière empirique à une valeur de $T = 0,8$. Le résultat ne laisse que deux chemins possibles. Le filtrage a permis d'éliminer les arcs dont le degré de stabilité est faible, mais ne modifie pas les poids pour la suite de la méthode.

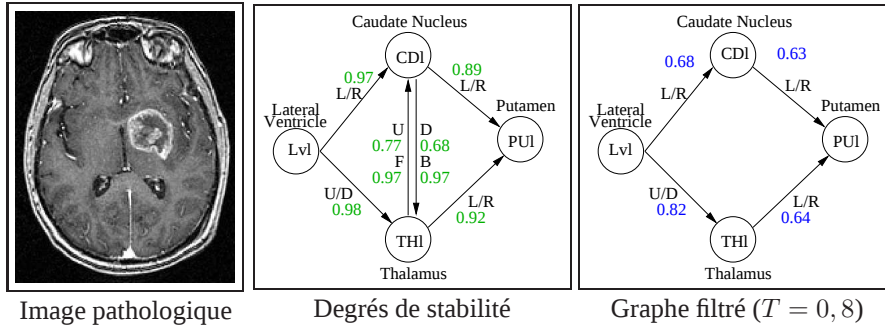


FIG. 4.11 – Élagage du graphe en fonction du degré de stabilité. Le graphe au centre montre les degrés de stabilité obtenus pour le cas de pathologie présenté sur la gauche. Le graphe de droite montre le résultat du filtrage si le seuil de stabilité est $T = 0,8$. Les valeurs de satisfiabilité des arcs restants sont notés en bleu sur ce graphe.

Pondération par le degré de stabilité :

Une autre manière de prendre en compte le degré de stabilité d'une relation spatiale, plus souple que la première approche, est de prendre en compte le degré de stabilité comme un attribut de l'arc qui porte la relation spatiale. De cette manière, le degré de stabilité peut être intégré dans le calcul du coût de l'approche initiale. De plus, cette manière d'intégrer l'information pathologique est relativement simple, si l'on considère le type de pathologie connu.

L'intégration du degré de stabilité doit être effectuée de telle manière que les chemins qui comportent des structures pathologiques ou altérées par une pathologie soient pénalisés. Pour cela, nous pouvons utiliser une t-norme (un produit par exemple) et l'intégrer directement :

$$e_w = \top(d_e, f(\mu_{Rel}, \mu_{Obj})) ,$$

où e_w représente le poids attribué à l'arc, d_e le degré de stabilité porté par cet arc et f la mesure calculée par l'approche initiale, une satisfiabilité ou une ressemblance entre la relation portée par l'arc et la structure cible. La figure 4.12.b présente les degrés de stabilité (en bleu) qui ont été appris pour chaque arc pour la classe de tumeur correspondant au cas pathologique illustré par la figure 4.12.a et les mesures de satisfaction (en rouge) qui ont été pondérées. Dans ce cas, le meilleur chemin devient le suivant :

Ventricule « Au-dessus » Thalamus « À droite de » Putamen.

La figure 4.12.c présente une segmentation du putamen qui a été effectuée en suivant cette séquence de segmentation.

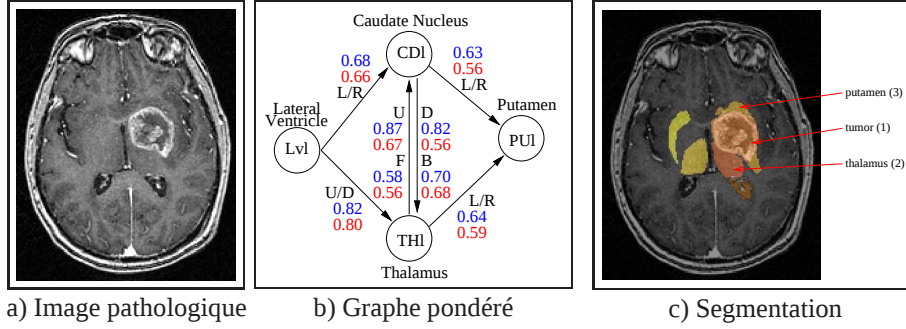


FIG. 4.12 – (a) Vue axiale d’une image IRM avec une tumeur proche du ventricule latéral et des noyaux centraux. (b) Degrés de stabilité appris avec une classe de tumeur similaire (en bleu). Les mesures de satisfaction pondérées sont en rouge. Sélection du meilleur chemin avec le critère du meilleur chemin moyen. Le meilleur chemin est dans ce cas : ventricule « Au-dessus » thalamus « À droite de » putamen. (c) Segmentation du putamen. La tumeur est extraite d’abord. Ensuite, le thalamus et finalement le putamen.

Intégration dans l’apprentissage des relations :

Nous avons vu comment réaliser l’apprentissage des relations spatiales dans le chapitre 3, dont le principe a été rappelé dans la partie précédente. Les représentations des relations spatiales sont plutôt souples dans la manière dont elles sont construites, et cette construction peut être directement adaptée en fonction du degré de stabilité. L’apprentissage proposé dans le chapitre 3 est effectué sur une base d’images saines et d’images pathologiques. Le principe est d’effectuer une extension de l’intervalle flou permettant de prendre en compte les cas pathologiques. Moins la relation sera stable et plus la relation est étendue donc floue. Si on considère un lien entre une faible ressemblance et un plus haut niveau de flou, alors la pertinence des chemins sera diminuée.

Si nous reprenons le cas de tumeur présentée dans la figure 4.12, la relation d’orientation présente entre le noyau caudé et le putamen « À droite de » peut être étendue pour gérer ce cas pathologique, en utilisant des valeurs plus larges pour les angles α_1 et α_2 utilisés pour construire la relation. Ce type d’apprentissage a été illustré dans le chapitre précédent, dans la partie 3.5.

4.3 Optimisation globale de la pertinence d’un chemin

Un chemin peut être qualifié d’une manière globale, par exemple en comptant le nombre de changements de direction, c’est-à-dire à chaque fois que deux relations d’orientation consécutives ne sont pas similaires. D’une manière plus générale, nous sommes intéressés par qualifier non seulement les différentes relations qui composent un chemin, mais aussi la fusion de toute l’information spatiale. L’approche précédente propose une optimisation locale où chaque arc est évalué de manière séparée et il n’est donc pas possible d’effectuer ce type d’évaluation. Nous proposons à présent une méthode permettant d’évaluer la pertinence d’un chemin non plus en évaluant chaque arc de manière séparée, mais en évaluant une représentation du chemin sous forme d’un unique ensemble flou. Nous allons tout d’abord voir sous quelle forme les chemins seront représentés afin de pouvoir être évalués ensuite.

4.3.1 Fusion des connaissances spatiales

Afin de pouvoir effectuer une estimation globale d'un chemin, nous allons déterminer comment représenter un chemin. Nous utilisons comme modèle un graphe dont chaque arc comporte des relations spatiales, et nous avons vu de quelle manière ces relations spatiales étaient représentées dans l'espace de l'image. Nous proposons de représenter un chemin en fusionnant l'information spatiale composant un chemin, c'est-à-dire en fusionnant les représentations des relations spatiales portées par chaque arc composant ce chemin. Chaque ensemble flou porté par un arc étant calculé dans l'espace de l'image, la fusion de ces ensembles flous est aisée et naturelle. Comme précédemment, les représentations des relations spatiales seront calculées en utilisant le formalisme flou. Nous allons donc combiner une information a priori et l'information spatiale contenue par un chemin pour représenter ce chemin.

Nous devons tout d'abord calculer les représentations des relations spatiales portées par les arcs du chemin étudié, en suivant le formalisme flou décrit dans le chapitre 3 ainsi que dans la partie précédente. Nous avons le choix entre une fusion conjonctive ou une fusion disjonctive. La figure 4.13 permet une comparaison entre une fusion conjonctive (un minimum) et une fonction disjonctive (un maximum) pour représenter le même chemin. Dans cet exemple, le chemin représenté est le suivant : LVL « à gauche » CDI « devant » THl « à gauche ». Pour chaque type de fusion, deux cas sont étudiés : dans le premier cas, uniquement des fonctions d'orientation sont utilisées et aucun apprentissage n'est effectué sur ces fonctions. Dans le deuxième cas, un apprentissage est effectué, tel qu'il est décrit dans le chapitre 3. Dans ce cas, nous pouvons ajouter sur chaque arc une relation de distance, qui sera fusionnée avec la relation d'orientation avec un minimum.

La fusion disjonctive nous permet de conserver toutes l'information apportée par les relations spatiales du chemin. Cependant, la zone représentant le chemin va vite être grande par rapport aux structures du chemin. L'information n'est donc pas pertinente. Avec un apprentissage et des relations de distance, la représentation du chemin est plus limitée, mais reste grande par rapport aux structures. Nous utilisons plutôt une mesure conjonctive, qui permet de restreindre l'information. Mais la conjonction fait disparaître beaucoup d'information. Dans le cas avec un apprentissage et des relations de distance, il reste peu d'information sur le chemin, et certains arcs n'apportent plus d'information, comme entre le ventricule et le noyau caudé dans la représentation en bas à droite de la figure. Sans apprentissage, nous avons encore une information sur l'ensemble du chemin. Nous utiliserons ce dernier cas dans nos expériences. Parmi les t-normes et t-conormes, nous avons choisi d'utiliser le minimum et le maximum dans nos expériences.

Les conjonctions de relations spatiales représentent plus la localisation de la structure visée par les arcs d'un chemin, que le chemin lui-même entre ces deux structures, ce qui explique que plus ces localisations sont précises, et moins la représentation du chemin semble correcte, car les interstices entre les structures sont moins représentés que dans le cas sans apprentissage, où les localisations sont moins précises.

La représentation du chemin est donc générée en fusionnant tous les ensembles flous obtenus en utilisant un opérateur de fusion conjonctif (une t-norme), ainsi :

$$\mu_p = \top[\mu_{Rel_i^p}, i = 1 \dots N^p] \quad (4.6)$$

où $\top$ est une t-norme et p un chemin composé de N^p relations. Dans nos expériences, nous utilisons une norme minimum. Le processus permettant la génération de la représentation d'un chemin est illustré dans la figure 4.14.

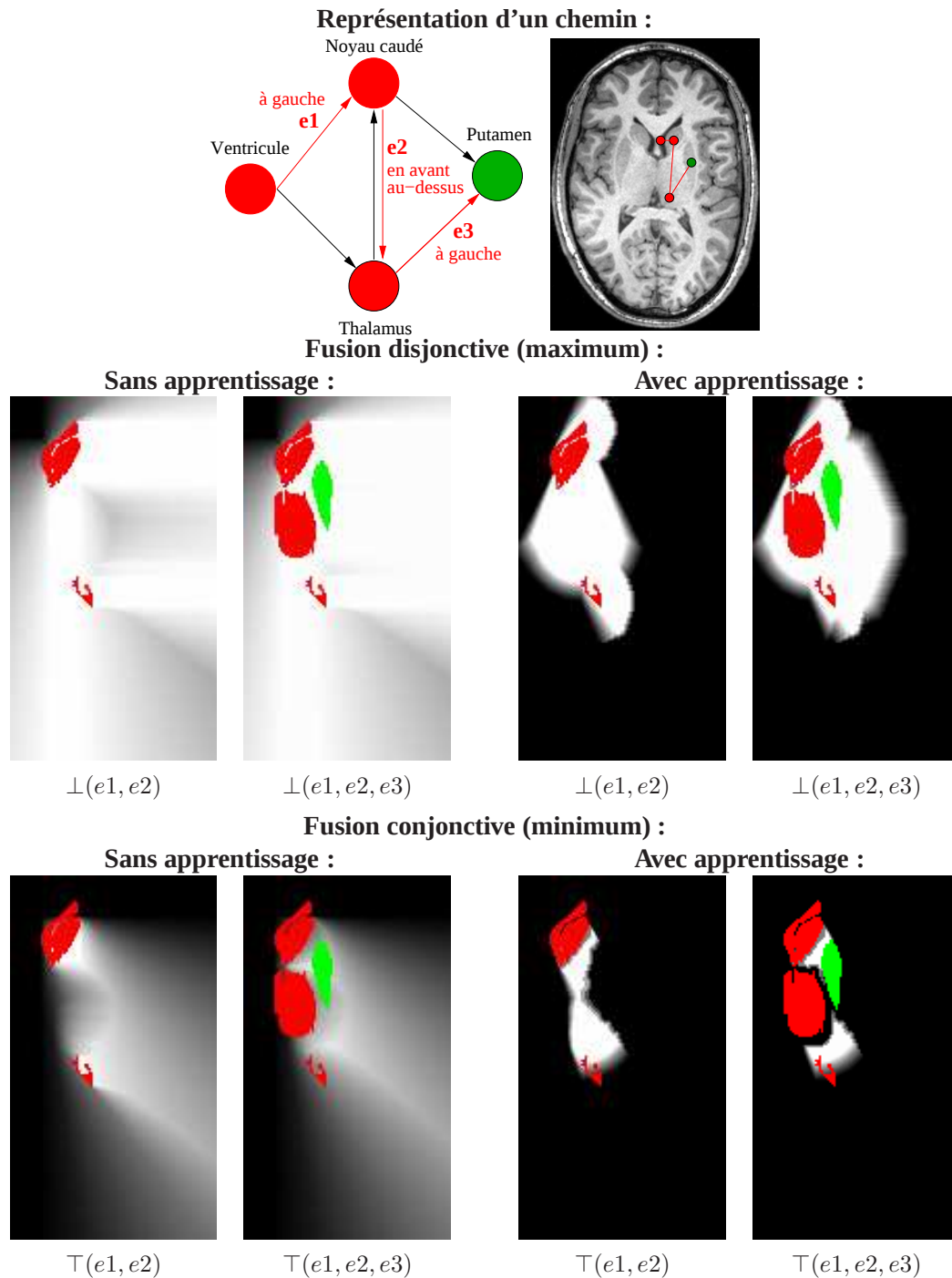


FIG. 4.13 – Comparaison entre une fusion disjonctive (au centre) ou conjonctive (en bas) pour calculer la représentation d'un chemin. La fusion disjonctive conserve toute l'information, et dans le cas de relations non bornées telles que les relations d'orientation, une large partie de l'espace est intégrée dans la représentation du chemin. Avec un apprentissage et l'ajout de relations de distances fusionnées avec les relations d'orientation, moins d'information est prise en compte, mais toujours beaucoup relativement à la taille des structures et du cerveau. En revanche, la fusion conjonctive conserve beaucoup moins d'information. La conjonction des relations spatiales visant une structure d'un chemin a plutôt représenté la localisation de cette structure, et non pas le chemin entre deux structures. Plus ces localisations sont précises, et moins la représentation d'un chemin apparaît correcte, c'est-à-dire moins les espaces entre les structures sont représentés dans la représentation du chemin. Il est important que la structure recherchée soit comprise dans la représentation, et que celle-ci ne soit pas trop étendue, ce qui est le cas de la fusion conjonctive sans apprentissage dans cet exemple.

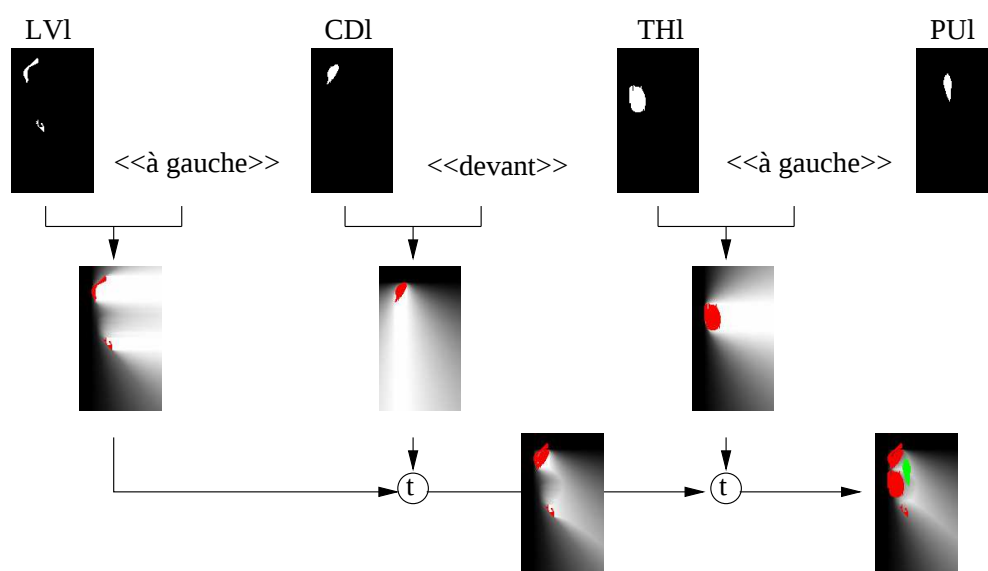


FIG. 4.14 – Génération d’une représentation pour le chemin suivant : ventricule (LVI) « À droite de » noyau caudé (CDI) « En-avant » thalamus (THI) « À droite de » putamen (PUI). Une coupe des représentations de chaque structure est présentée dans la ligne supérieure. Pour chaque relation portée par un arc d’un chemin, nous la représentons dans l’espace de l’image, en utilisant la représentation de la structure de référence de chaque relation. Ces représentations sont présentées dans la ligne du milieu, avec les structures de référence en rouge. Les représentations de chaque relation sont ensuite fusionnées en utilisant une t-norme (ici le minimum). La ligne du bas présente à gauche la fusion des deux premières relations (avec les structures de référence des deux relations en rouge), puis après fusion de la troisième relation, la représentation du chemin à droite, avec la structure cible du chemin (le putamen) en vert.

4.3.2 Évaluation du chemin par mesure de son entropie

À l'aide du processus de fusion qui vient d'être décrit, nous obtenons donc un ensemble flou décrivant chaque chemin que nous étudions. La taille du graphe étant petite, nous pouvons effectuer une exploration exhaustive des chemins du graphe. Si le graphe considéré était plus important, alors il serait nécessaire de prendre en compte la complexité de cette tâche.

Nous souhaitons sélectionner le chemin qui répond le mieux à nos attentes, c'est-à-dire qui répondra le mieux aux contraintes du processus de segmentation des structures. Pour cela, nous proposons de sélectionner le chemin qui est le plus précis possible, c'est-à-dire le chemin qui laisse le moins de doute possible sur l'emplacement de chaque structure à segmenter le long du chemin. Pour cela, nous allons donc sélectionner le chemin le « moins flou ». Comme mesure de flou, toujours dans cette optique, nous utilisons une mesure d'entropie floue [Luca et Termini \(1972\)](#) définie ainsi :

$$H(\mu_p) = -K \left(\sum_{x_i \in \mathcal{S}} \mu_p(x_i) \log \mu_p(x_i) + \sum_{x_i \in \mathcal{S}} (1 - \mu_p(x_i)) \log(1 - \mu_p(x_i)) \right), \quad (4.7)$$

où μ_p est l'ensemble flou correspondant à la fusion de toutes les relations spatiales contenue dans le chemin p et K est une constante de normalisation.

Le meilleur chemin $\hat{p}$ sera donc le chemin le « moins flou », donc avec le minimum d'entropie floue :

$$\hat{p} = \arg \min_{p \in \mathcal{P}} (H(\mu_p)). \quad (4.8)$$

Il faut remarquer que cette mesure est utilisable lorsque, comme dans notre cas, les relations sont plus floues lorsqu'elles sont moins précises. Il serait inutile de mesurer ce critère sur des régions qui ne sont pas floues, qui donneraient une valeur d'entropie nulle même si les régions sont très larges et n'apportent pas d'aide au processus de segmentation.

4.3.3 Expériences

Les mesures de l'entropie floue pour chacun des chemins sont présentées dans le tableau 4.5. Le chemin qui possède l'entropie floue la plus basse est le suivant :

LVI « down of » THl « up of » CDl « left of » PUI

Cette représentation est illustrée par la figure 4.15. Ce chemin contient plusieurs changements de direction, ce qui explique que la conjonction des représentations des relations spatiales soit très concentrée sur une petite zone uniquement, et donc présente une entropie faible. D'une manière plus générale, l'entropie floue calculée sera dépendante des changements de direction du chemin, plus que de la précision des relations elles-mêmes.

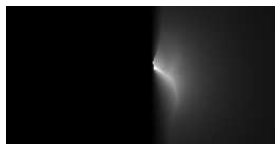


FIG. 4.15 – Une coupe d'une représentation 3D du chemin LVI « down of » THl « up of » CDl « left of » PUI après fusion des connaissances spatiales contenues dans le chemin.

TAB. 4.5 – L'entropie floue obtenue pour chacune des représentations des chemins entre le ventricule et le putamen dans notre graphe. Les structures sont les suivantes : ventricule latéral (LVI), thalamus (THI), noyau caudé (CDI), putamen (PUI).

Chemin :				Entropie floue :
LVI « Au-dessus »	THI en dessous	CDI « À droite de »	PUI	0,08
LVI « Au-dessus »	THI en dessous	CDI « À droite de »	PUI	0,17
LVI « À droite de »	CDI « À droite de »		PUI	0,26
LVI « À droite de »	CDI en avant	THI « À droite de »	PUI	0,16
LVI « À droite de »	CDI « Au-dessus »	THI « À droite de »	PUI	0,16
LVI « Au-dessus »	THI « À droite de »		PUI	0,16

4.3.4 Adaptation aux cas pathologiques

Cette approche reposant sur les représentations des relations spatiales, son adaptation aux cas pathologiques doit porter sur les paramètres de ces relations : l'influence d'une relation est diminuée en étendant sa représentation spatiale, par une dilatation floue de l'ensemble correspondant par exemple. L'extension de sa représentation spatiale correspond à rendre la relation spatiale plus permissive, et donc augmenter le degré de flou et l'entropie du chemin. Les chemins qui possèdent des structures altérées seront ainsi pénalisés.

4.3.5 Conclusion sur l'approche globale

Cette approche permet de prendre en compte un critère global sur le chemin, sa précision selon le critère d'entropie floue. Toutefois, la représentation globale n'est pas satisfaisante. Les représentations utilisant une disjonction sont trop larges et les représentations des chemins se recouvrent trop pour être discriminantes. D'un autre côté, les fusions avec un opérateur de conjonction réduisent suffisamment l'information, mais des parties du chemin ne sont pas représentées et les chemins sont dépendants des changements de direction.

4.4 Conclusion

Nous avons montré dans cette première approche que l'ordre de segmentation des structures d'un processus de segmentation séquentiel peut être déduit de manière automatique, et les résultats, limités à un petit graphe, montrent que le chemin déduit automatiquement est le même que le chemin qui avait été construit de manière ad-hoc. L'extension proposée pour les cas pathologiques nous a permis, en prenant en compte la notion de degré de stabilité d'une relation spatiale, d'adapter le processus à un type de pathologie donné pour déterminer un meilleur chemin dans ce cas.

Cette approche comporte certaines hypothèses. La pertinence des relations est estimée, mais la difficulté intrinsèque de segmentation de chaque structure n'est pas prise en compte. Les critères que nous utilisons ne permettent pas non plus de considérer la précision intrinsèque des relations. Par exemple, une relation d'adjacence sera naturellement plus précise qu'une relation d'orientation (en termes de taille de support). Même si les critères sont normalisés par la taille du support, le rapport à la taille de l'objet cible n'est pas identique.

L'approche utilisant des représentations des chemins serait prometteuse si une bonne représentation d'un chemin pouvait être déduite du chemin. En effet, cette représentation, liée par exemple à une information visuelle telle que celle étudiée dans le chapitre 2, peut permettre de détecter des

événements d'une manière plus globale qu'en raisonnant avec une relation à chaque fois. Nous verrons dans le chapitre suivant comment faire une estimation globale des chemins, mais a posteriori.

L'utilisation d'une connaissance a priori permet de réaliser une optimisation globale sur le chemin complet avant de segmenter. Cependant, l'optimisation est ici locale au sens où la pertinence de chaque arc est évaluée de manière indépendante des autres arcs. Une autre critique est que cette approche n'utilise pas d'information provenant de l'image à segmenter, mais est effectuée uniquement à partir de la connaissance a priori, à part l'adaptation aux cas pathologiques qui prend en compte le type de pathologie présente sur l'image. Une variation de cette approche serait d'effectuer l'optimisation globalement de la même manière mais de réviser le modèle à la suite de chaque segmentation de structure pour prendre en compte cette image. Nous aurions ainsi une instanciation progressive du modèle, mais dans ce cas, même si une optimisation du chemin complet est effectuée, elle est utilisée comme une manière de choisir la prochaine structure à segmenter uniquement. L'approche présentée dans la partie suivante fonctionne de cette manière.

L'objectif du prochain chapitre est de combler l'absence d'information provenant de l'image à segmenter. Nous avons présenté au chapitre 2 comment les modèles du système visuel pouvaient apporter de l'information extraite directement d'une image, via les mécanismes pré-attentionnels, et en particulier les cartes de saillance. Nous allons à présent voir comment intégrer cette information visuelle dans un processus de segmentation séquentielle.

Chapitre 5

Optimisation avec information visuelle

L'approche proposée dans la partie précédente et ses adaptations dans le cas pathologique utilisent des représentations de la forme de chacune des structures, qui proviennent d'un atlas anatomique ou simplement d'une base d'images. Cette connaissance a priori sur les objets de la scène apporte deux informations indispensables aux raisonnements proposés : la forme des objets utilisée dans le calcul des relations spatiales, et leur localisation, ou leur agencement. La prise en compte de la forme et de la taille des objets est importante dans la définition des relations spatiales, la sémantique de la relation pouvant être différente en fonction des caractéristiques morphologiques des objets. La localisation spatiale nous permet d'estimer la pertinence d'une relation spatiale par rapport à l'objet qu'elle vise.

Mais l'utilisation d'une telle connaissance générique a forcément des limites dans son exhaustivité, et plus encore en imagerie médicale, où le nombre de cas disponibles est plus limité que dans d'autres domaines. Notre base de données ne reflète pas la variabilité complète inter-patients, ni les différences qui peuvent exister avec par exemple des enfants plutôt que des adultes, ou inversement des personnes âgées, ou encore d'autres pathologies que des tumeurs cérébrales qui pourraient agir sur les structures cérébrales ou la matière. Nous avons vu que l'approche présentée ne pouvait pas prendre en compte les cas pathologiques (les tumeurs cérébrales) sans adaptation. Les adaptations proposées permettent de prendre en compte le degré de stabilité d'une relation spatiale pour un cas de pathologie donné. Mais là encore, il est difficile d'obtenir, pour chaque classe de pathologies et pour chaque étape du développement, une base d'apprentissage suffisante pour prendre en compte les différents cas possibles. Il est difficile d'être exhaustif, voire impossible et cela est encore plus vrai dans le cadre de l'imagerie cérébrale. De plus, même si nous pouvions obtenir un modèle générique de la connaissance que nous utilisons dans l'approche initiale, ce modèle ne permettrait plus forcément la reconnaissance telle qu'elle est effectuée.

Si la connaissance a priori telle que nous l'utilisons ne peut pas entièrement répondre à nos besoins, alors nous avons besoin d'aller chercher de l'information ailleurs. L'information que nous cherchons est bien entendu contenue dans l'image que nous voulons segmenter, mais inaccessible tant que le modèle n'est pas instancié pour cette image. En fait, avec un processus de segmentation séquentielle, le modèle est progressivement instancié, et nous pouvons obtenir des informations plus précises des parties déjà reconnues de l'image, de la même manière que dans les processus d'attention visuelle.

Nous proposons dans ce chapitre une méthode permettant de s'affranchir des représentations des formes des structures. Dans cette approche, nous utilisons un critère de sélection des structures qui est issu d'une information visuelle directement extraite de l'image à segmenter elle-même, permettant de prendre en compte les particularités de l'image. Nous utilisons pour cela une carte de saillance, selon un mécanisme pré-attentionnel que nous avons décrit dans le chapitre 2. Avec

cette approche, la segmentation de l'image est vue comme un processus d'exploration de l'image. Par rapport à la première méthode proposée, cette méthode ne permet pas d'évaluer un chemin complet avant segmentation. Le critère de sélection des structures permet ici de sélectionner la *prochaine* structure à segmenter uniquement. Le chemin de segmentation optimal est donc entièrement déterminé une fois toutes les segmentations effectuées.

5.1 Utilisation d'une information visuelle

Nous allons commencer dans cette partie par établir des correspondances, pour notre cas particulier, entre les mécanismes de l'attention visuelle et un processus de segmentation séquentielle tel celui que nous utilisons, et qui sera décrit dans la deuxième partie de ce chapitre. Nous allons également voir quel critère dérivé d'une information visuelle nous pouvons intégrer dans ce processus.

5.1.1 Attention visuelle et segmentation séquentielle

Nous avons décrit dans le chapitre 2 la notion d'attention visuelle. Les modèles du système visuel font en général apparaître deux étapes et deux types de mécanisme, respectivement attentionnel et pré-attentionnel. D'une manière simplifiée, l'objectif de l'étape pré-attentionnelle est de guider l'étape attentionnelle en sélectionnant les parties de l'image dites saillantes, c'est-à-dire qui « attirent l'œil ». La notion de saillance est généralement associée à la présence de discontinuités de caractéristiques de bas niveau dans l'image. La sélection qui est effectuée permet au processus attentionnel de se focaliser sur une partie restreinte de la scène. Cette partie peut être un objet ou une zone de l'image. La restriction de la phase attentionnelle à une zone réduite de l'image permet de réduire le coût de traitement de cette zone.

Nous proposons alors d'effectuer un rapprochement entre les différentes étapes du processus de segmentation séquentielle, la sélection de la séquence de segmentation et la segmentation elle-même, et les deux phases des modèles de l'attention visuelle. La phase attentionnelle où une zone restreinte de l'image est analysée avec attention correspond à la segmentation d'un objet. La sélection d'une zone à segmenter revient donc à guider l'attention visuelle, et correspond donc à l'étape pré-attentionnelle. Le tableau 5.1 présente en détail le parallèle effectué entre les deux notions.

Il existe différentes théories des mécanismes pré-attentionnels. Dans certaines approches, les interactions entre les deux étapes attentionnelles et pré-attentionnelles sont plus complexes et imbriquées. D'ailleurs, l'unité attentionnelle n'est pas toujours une région de l'image, mais parfois un objet. Le chapitre 2 présente les expériences qui ont mis en évidence une sélection autre que spatiale lorsque l'observateur a une tâche spécifique à accomplir. La tâche de l'observateur dans notre cas n'est pas comparable aux tâches de haut niveau qui peuvent être demandées à un observateur, comme de compter le nombre de personnages d'une scène, ou le nombre de passes d'un groupe de personnages jouant avec un ballon comme dans l'expérience illustrée dans la figure 2.3. Notre tâche n'est donc pas comparable, et avant de pouvoir éventuellement effectuer ce genre de tâche, il nous faut d'abord voir et reconnaître ce qu'il y a dans l'image. Nous nous intéressons donc plutôt aux processus d'exploration de l'image guidés par les données uniquement, comme la théorie très répandue de l'intégration de caractéristiques.

Dans cette théorie, le processus pré-attentionnel est un processus ascendant, c'est-à-dire uniquement guidé par les données, dont l'objectif est de sélectionner une région de l'espace qualifiée de saillante. La saillance est dérivée de caractéristiques globales de l'image. Puis l'information issue de chacune des caractéristiques est fusionnée pour donner une carte unique, représentant toutes

TAB. 5.1 – Un appariement de chaque étape d'un processus de segmentation séquentielle à une modélisation de l'attention visuelle telle que décrite par la théorie d'intégration des caractéristiques, décrite dans le chapitre 2. Dans ce cas, l'étape pré-attentionnelle est guidée par les données, ce qui est le cas des premières modélisations, mais les travaux plus récents proposent très souvent des liens descendants.

Étape :		Système visuel :	Segmentation Sé-quentielle :
Pré-attentionnelle	Objectif	Sélectionner une zone ou un objet de l'espace pour un examen attentif	Sélection de la zone de l'espace (ou de la prochaine structure) à segmenter
	Mode	Processus ascendant effectué à partir de l'image entière et où les caractéristiques sont traitées de manière parallèle	À partir de caractéristiques globales de l'image
Attentionnelle	Objectif	Examen attentif d'une petite zone de l'image	Segmentation d'une partie de l'image
	Mode	Sur une petite zone de l'image et de manière séquentielle	Dans une zone définie par des relations spatiales
Inhibition de retour	Objectif	Ne pas bloquer l'œil sur la dernière zone sélectionnée	Utiliser les objets déjà segmentés pour contraindre la recherche
	Mode	Masquage d'une zone temporairement	Masquage des zones déjà segmentées

les caractéristiques. Cette carte unique est nommée carte de saillance. Nous avons décrit dans la partie 2.3 le processus permettant d'extraire une carte de saillance d'une image, à partir des caractéristiques d'intensité, de couleur et d'orientation d'une image. Le mécanisme de création de ces cartes a été décrit par [Itti et al. \(1998\)](#) à la suite des travaux de [Koch et Ullman \(1985\)](#). Nous avons également décrit dans la partie 2.4 les adaptations nécessaires du mécanisme d'extraction des cartes de saillance aux images IRM.

Une autre étape, ou plutôt un mécanisme intégré dans l'étape pré-attentionnelle, peut trouver son équivalent dans le processus de segmentation séquentielle, il s'agit de l'inhibition de retour. Si un observateur regarde une scène fixe, alors les zones saillantes demeurent identiques au cours du temps. Mais si l'exploration de l'image est guidée par l'information de saillance, alors l'attention visuelle risque d'être bloquée sur une même zone en absence de mouvement. Il existe donc un mécanisme permettant d'inhiber pendant un bref laps de temps une zone saillante sur laquelle l'attention visuelle a été focalisée. Un problème similaire se pose si nous segmentons une zone d'une image qui se trouve à côté d'un objet qui attire le processus de segmentation (un fort contraste avec le reste de l'image par exemple). Dans [Colliot et al. \(2006\)](#), il est montré que l'utilisation des relations spatiales permet de contraindre un modèle déformable pour éviter de se retrouver sur les bords d'un objet déjà segmenté. La figure 5.1 illustre cet effet. La segmentation du noyau caudé

ne s'arrête pas sur les bords du ventricule. Nous pouvons donc, à l'aide des relations spatiales, simuler implicitement un mécanisme d'inhibition de retour.

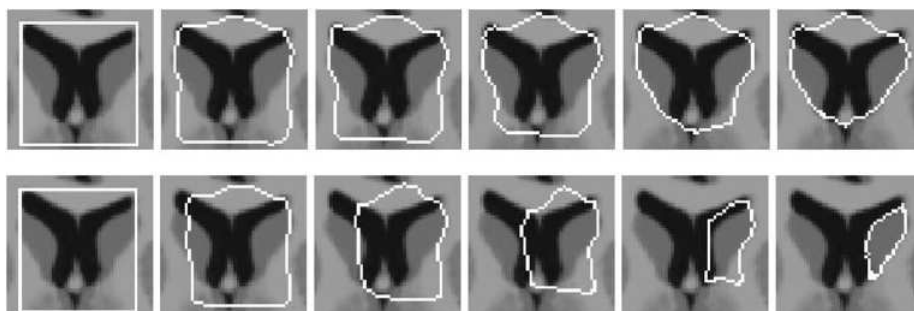


FIG. 5.1 – L'équivalent de l'inhibition de retour dans le cadre de la segmentation séquentielle. L'utilisation des relations spatiales permet d'éviter que le modèle déformable reste bloqué sur les bords du ventricule qui présentent un fort gradient d'intensité, comme c'est le cas dans la ligne supérieure de l'image. Sur la deuxième ligne, les relations spatiales entre le ventricule et le noyau caudé permettent de contraindre le modèle vers l'extérieur du ventricule et ainsi trouver le bon contour. Cet effet peut être comparé au mécanisme d'inhibition de retour qui permet à l'attention visuelle de ne pas rester bloquée sur une zone de l'image pourtant très saillante. De cette manière, toute l'image peut être explorée. [Source [Colliot et al. \(2006\)](#)].

Le cadre de segmentation séquentielle est donc vu comme un processus d'exploration et d'analyse progressive de la scène, ou de l'image. Nous proposons donc l'introduction d'un mécanisme pré-attentionnel dans le processus d'optimisation d'un chemin de segmentation pour une segmentation séquentielle d'une image. L'utilisation de cette information pour l'optimisation du chemin de segmentation doit nous permettre de nous passer de la connaissance a priori utilisée dans l'approche initiale pour optimiser le chemin de segmentation. En revanche, et contrairement à l'approche initiale, nous allons effectuer la sélection d'une zone de l'image, selon des critères reposant sur les données. Dans l'approche initiale, le chemin de segmentation était calculé à partir de l'information a priori et l'optimisation était effectuée sur l'ensemble du chemin avant segmentation.

Nous allons commencer par étudier, à l'aide d'images segmentées de la base de données, quelle est la saillance de chaque structure, c'est-à-dire étudier la saillance à l'emplacement de chacune des structures.

5.1.2 Saillance et difficulté de segmentation

Les approches précédentes ne tiennent pas compte de la difficulté intrinsèque de segmentation de chacune des structures, c'est-à-dire que la segmentation de chacune des structures est considérée avec une égale difficulté. Mais l'expérience de segmentation montre que cela n'est pas forcément vrai, et que la difficulté varie en fonction des structures et des images. Ces difficultés peuvent varier en fonction de plusieurs critères comme la forme, l'homogénéité, la texture, le contraste ou les contours d'une structure. Des règles génériques peuvent toujours être construites, par exemple : « cet objet est plus difficile à segmenter que cet autre objet » mais ce type de règle n'est pas toujours vrai, même dans un domaine d'application restreint.

Nous avons présenté dans le chapitre 2 la notion générique de saillance, et plus spécifiquement dans la partie 2.3 comment l'information de saillance est estimée par le système décrit par [Itti et al. \(1998\)](#). L'information de saillance dans ce système est dérivée de l'étude des discontinuités de certaines caractéristiques dans l'image : intensité, oppositions de couleur et orientations.

En effet, pour chaque caractéristique étudiée, les cartes de discontinuité générées reflètent la différence de niveau entre un point et son voisinage. Il s'agit donc d'une information de type gradient (ou une approximation locale du gradient de l'image filtrée pour représenter une caractéristique). Cette information est calculée selon différents niveaux d'échelles, puis fusionnée dans une carte unique. Cette carte unique représente donc les discontinuités d'une caractéristique donnée, et pour différents niveaux d'échelle. Toutes ces cartes sont ensuite fusionnées pour donner la carte de saillance.

Les algorithmes de segmentation d'image ont pour objectif de poser une frontière entre des régions d'une image, et en général cette frontière représente une discontinuité. Dans une application pour la segmentation des structures cérébrales, le problème est plutôt de savoir où placer une frontière, car les bords sont souvent flous et mal définis. Nous considérons donc que l'information de saillance est directement reliée aux difficultés de segmentation d'un objet en considérant qu'un objet avec un contour plus saillant, c'est-à-dire présentant une discontinuité plus marquée, sera plus aisé à segmenter qu'un objet comportant un contour moins saillant. Cependant, la saillance peut donner plus d'information. En effet, certaines tumeurs cérébrales par exemple sont très saillantes. Une forte saillance peut donc indiquer non seulement une zone plus aisée à segmenter, mais si nous disposons, via un apprentissage par exemple, de la distribution moyenne de saillance pour une zone, alors nous pouvons également détecter une anomalie comme une pathologie.

Nous proposons d'étudier la saillance d'une image segmentée, afin de vérifier empiriquement si le niveau de saillance d'un objet correspond à la difficulté notoire de le segmenter.

5.1.3 Apprentissage de la saillance

Nous souhaitons étudier dans cette partie les zones de la carte de saillance correspondant aux structures cérébrales comprises dans le modèle. Chaque carte de saillance est calculée sur une image complète. Mais grâce aux segmentations des images utilisées pour calculer la saillance, nous avons masqué la carte de saillance pour nous intéresser aux zones correspondant aux structures. L'objectif de cette partie est de construire un critère fondé sur la saillance qui sera utilisé dans les parties suivantes. Nous allons également effectuer un apprentissage des distributions de saillance.

En fonction de l'objectif de segmentation choisi, certaines parties d'un objet peuvent être plus intéressantes que d'autres. Si nous considérons un algorithme recherchant les contours, alors la zone la plus importante à prendre en compte est le contour de l'objet et son entourage immédiat. Mais nous allons regarder l'information de saillance, qui est calculée à différents niveaux d'échelle. L'information du contour est donc située sur une zone plus large que le contour. De plus, si nous considérons la taille des structures, petite par rapport à la taille du cerveau, il faut s'intéresser à l'information de saillance dans tout l'objet, ainsi que dans une couronne autour de l'objet, correspondant typiquement à une dilatation de l'objet par une boule unitaire en 6-connexité.

Pour chaque image $c \in K$, la carte de saillance SM_c est calculée sur l'image complète selon la méthode adaptée aux images IRM décrite dans la partie 2.4. La saillance SAL_{O_c} correspondant à un objet O_c de cette image est extraite en utilisant le masque dilaté de cet objet sur la carte de saillance :

$$\forall x \in \mathcal{S}, \quad SAL_{O_c}(x) = \min(\delta_1(O_c)(x), SM_c(x)) ,$$

où $\mathcal{S}$ représente l'espace de l'image. La saillance d'un objet est donc représentée dans cet espace.

Nous pouvons alors calculer un histogramme h de la saillance d'un objet, en calculant l'histogramme de SAL_{O_c} . Les cartes de saillance sont normalisées dans un intervalle $[0, 1]$. Le nombre N de niveaux de quantification de l'histogramme est fixé arbitrairement à 100.

$$h[i] = \sum_{x \in \mathcal{S}} \mathbb{1}_i(SAL_{O_c}(x)) ,$$

où $1(\cdot)$ est la fonction indicatrice. Cet histogramme est ensuite normalisé afin d'obtenir une fonction de densité de probabilité : pour $i = 1, \dots, N$

$$h_{O_c}[i] = \frac{h[i]}{\sum_{i=1}^N h[i]} .$$

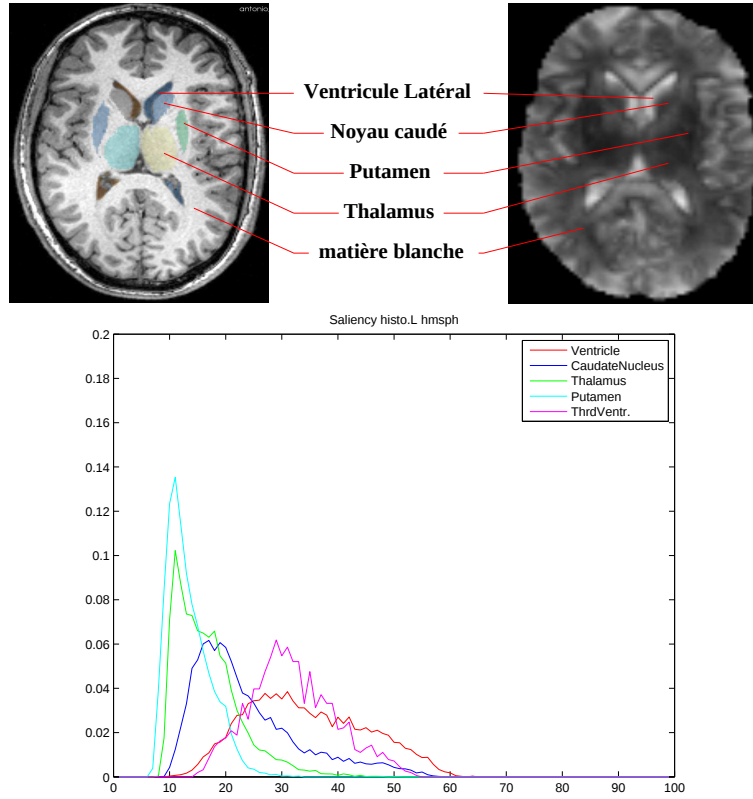


FIG. 5.2 – Les distributions de probabilité de saillance normalisées de cinq structures cérébrales de l'hémisphère gauche d'une image ne présentant pas de pathologie. En haut, l'image originale est présentée à gauche avec une segmentation en sur-impression et la carte de saillance correspondant à cette image est présentée à droite. Les structures sont pointées sur la segmentation de l'image ainsi que sur la carte de saillance. Le troisième ventricule qui apparaît sur l'histogramme n'est pas présent dans cette coupe. Les distributions du putamen et du thalamus présentent un pic pour des valeurs assez faibles de saillance. Les zones correspondantes dans la carte de saillance montrent des zones de faible saillance à l'emplacement de ces structures. La distribution correspondant au ventricule présente des valeurs plus élevées, et on distingue nettement ces plus fortes valeurs sur la carte de saillance.

La figure 5.2 présente les distributions de probabilité de saillance obtenues pour cinq structures localisées dans l'hémisphère gauche d'une image qui ne présente pas de pathologie. Nous pouvons voir, pour cette image, que les distributions du putamen et du thalamus sont des distributions mono-modales et centrées sur les valeurs faibles de saillance, et la carte de saillance présente aux emplacements de ces structures des valeurs faibles. Plus généralement, la matière blanche qui englobe ces structures présente des valeurs faibles. La distribution du ventricule est par contre plus étalée mais est centrée sur des valeurs de saillance plus importantes. Ce résultat était attendu car cette structure présente un fort contraste avec son entourage immédiat et peut être segmentée plus aisément que les autres structures présentées sur cette image.

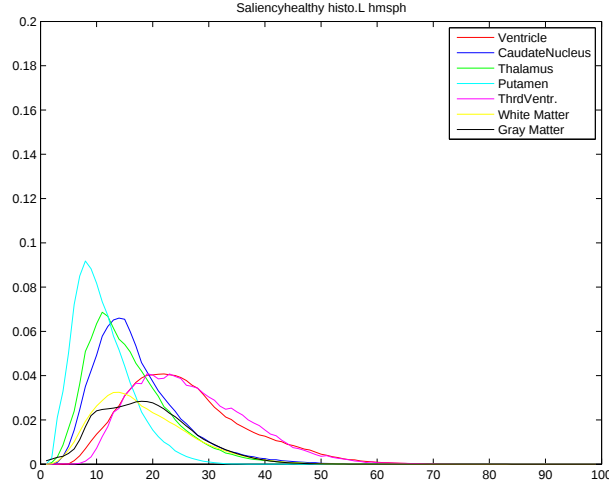


FIG. 5.3 – Distribution de saillance moyenne pour 5 structures cérébrales de l’hémisphère gauche (le troisième ventricule est en fait situé sur le plan inter-hémisphérique et n’est pas attribué à l’un ou l’autre des hémisphères, alors que les autres structures sont présentes dans les deux hémisphères). Les distributions sont calculées sur 30 cas sains (sans pathologies). Les distributions de la matière blanche (« WM ») et de la matière grise (« GM ») sont estimées sur les 18 cas de la base IBSR uniquement.

La figure 5.3 présente quelques distributions moyennes qui ont été estimées sur les 30 cas sains de la base d’apprentissage, pour des structures de l’hémisphère gauche. Pour la matière grise et la matière blanche, les distributions ont été estimées sur les 18 cas de la base IBSR uniquement, les segmentations n’étant pas disponibles pour les autres cas. L’apprentissage est effectué pour :

- les distributions de probabilité des cartes de saillance complètes ;
- chacune des structures du graphe ;
- les tumeurs dans les cas pathologiques.

La distribution de probabilité pour une structure donnée est calculée comme la moyenne des distributions de probabilité calculées pour cette structure sur chaque image :

$$\hat{h}_O[i] = \frac{\sum_{c \in K} h_{O_c}[i]}{\text{card}(K)} ,$$

où $\text{card}(K)$ représente le nombre de cas dans la base. Nous calculons ensuite la moyenne des distances EMD et la variance pour chacune des distributions, toutes les images prises en compte étant équiprobables. L’écart à la moyenne est estimé par une mesure EMD ou « Earth’s Mover Distance » (Rubner *et al.* (1998)).

La mesure EMD :

Supposons que p et q sont deux histogrammes discrets avec N niveaux de quantification, et normalisés tel que $\sum_{i=1}^N p[i] = \sum_{i=1}^N q[i] = 1$. La mesure EMD entre ces deux distributions de probabilité est définie ainsi :

$$\text{emd}(p, q) = \min_{\alpha_{i,j} \in \mathcal{M}} \sum_{i=1}^N \sum_{j=1}^N \alpha_{i,j} c(i, j) ,$$

où $\mathcal{M} = \{(\alpha_{i,j}); \alpha_{i,j} \geq 0, \sum_j \alpha_{i,j} = p[i], \sum_i \alpha_{i,j} = q[j]\}$ et où $c(.,.)$ est une distance entre les niveaux de quantification. Mais pour des histogrammes non-circulaires et en une dimension

seulement, si $c(i, j) = \frac{\|i-j\|}{N}$, alors il est établi que la mesure EMD est la différence entre les histogrammes cumulés (Villani (2003)) :

$$emd(p, q) = \frac{\sum_{i=1}^N |P[i] - Q[i]|}{N}, \quad (5.1)$$

où p et q sont deux distributions de probabilité, P et Q sont les histogrammes cumulés correspondants et N le nombre de niveaux de quantification des histogrammes. Nous utilisons cette formulation dans nos expériences. L'écart à la moyenne entre distributions de saillance est donc calculé comme la variance selon la mesure EMD :

$$V_O = \frac{\sum_{c \in K^N} emd(\hat{h}_O, h_{O_c})^2}{card(K^N)}$$

5.1.4 Un critère reposant sur la saillance

Nous souhaitons à présent utiliser la saillance dans le processus de segmentation séquentielle. Nous avons défini des histogrammes de saillance qui seront calculés dans le processus décrit dans la partie suivante. Nous avons à présent besoin d'un critère pour effectuer la sélection de la prochaine structure à segmenter en utilisant les histogrammes de saillance.

5.1.4.1 Critère simple sans apprentissage

Nous avons utilisé dans un premier temps un critère simple et qui ne nécessite pas d'apprentissage dans nos expériences. Nous avons proposé d'utiliser l'énergie des histogrammes comme critère de comparaison. L'énergie d'un histogramme H , avec N niveaux de quantification, est calculée de la manière suivante :

$$energie(H) = \sum_{n=1}^N h(n)^2,$$

où h est la fonction dénombrant le nombre d'occurrences des valeurs n dans la carte de saillance masquée. L'énergie d'un histogramme va permettre de préférer les histogrammes qui ont un support resserré, c'est-à-dire qu'un pic à base étroite mais haut sera préféré à un histogramme plus étalé. La figure 5.3 présente les distributions de probabilité pour quelques structures, ainsi que pour la matière blanche qui les englobe. Le critère sélectionné permet de préférer les histogrammes des structures par rapport à celui de la matière blanche.

Le tableau 5.2 présente les mesures de saillance pour trois structures cérébrales, le noyau caudé (« LCN ») le thalamus (« LTH ») et le putamen (« LPU »), ainsi que pour la matière blanche (« LWM ») et la matière grise (« LGM »). Ces mesures (l'énergie de l'histogramme) sont toujours plus grandes pour les trois structures anatomiques que pour les matières. Nous avons cependant laissé ce critère, qui ne permet pas de déterminer précisément l'ordre entre deux distributions de saillance. La figure 5.4 donnera un exemple de comparaison où la distribution présentant des valeurs plus hautes de saillance présente une énergie inférieure. Nous remplaçons ce critère pour un autre qui utilise une mesure fondée sur la mesure EMD définie dans la partie précédente, et que nous allons décrire à présent.

5.1.4.2 Critère utilisant une mesure EMD

Nous souhaitons comparer des régions entre elles, et pour cela nous souhaitons comparer la saillance de ces régions, afin de déterminer laquelle est la plus propice à être segmentée à un

TAB. 5.2 – Mesures de saillance (mesure de l'énergie d'un histogramme de saillance) pour trois structures anatomiques, la matière blanche (LWM) et la matière grise (LGM) pour toutes les images de la banque de données IBSR. LCN : noyau caudé, LTH : thalamus and LPU : Putamen.

LCN	LTH	LPU	LWM	LGM
0,065	0,057	0,068	0,026	0,015
0,097	0,064	0,095	0,041	0,020
0,039	0,033	0,042	0,027	0,017
0,050	0,031	0,054	0,026	0,017
0,038	0,028	0,107	0,027	0,018
0,054	0,038	0,099	0,038	0,025
0,039	0,024	0,046	0,023	0,018
0,040	0,026	0,046	0,020	0,014
0,039	0,026	0,061	0,026	0,020
0,045	0,030	0,060	0,027	0,014
0,037	0,025	0,048	0,019	0,011
0,033	0,029	0,032	0,026	0,017
0,037	0,033	0,069	0,031	0,020
0,046	0,030	0,061	0,025	0,017
0,033	0,026	0,044	0,017	0,014
0,032	0,025	0,044	0,022	0,015
0,045	0,032	0,049	0,022	0,020

instant donné du processus et selon les connaissances disponibles à cet instant. Pour chacune des structures, nous calculons un ensemble flou correspondant à sa localisation. Le processus de calcul de ces ensembles flous est détaillé dans une partie suivante.

La précision de la localisation d'une structure dépend de l'information spatiale disponible au moment où elle est représentée. Moins il y a d'information spatiale disponible, et moins la localisation est précise, et plus la localisation risque d'inclure des objets en plus de l'objet recherché. Or, nous souhaitons comparer la saillance des objets recherchés. La distribution de saillance peut donc inclure de l'information non pertinente pour juger de la saillance d'une structure. Pour cette raison, la comparaison directe de deux localisations ne permet pas de comparer la saillance des structures visées.

Notre critère sera donc fondé sur deux informations. Nous allons extraire la distribution de saillance de la localisation d'une structure, puis elle est comparée :

- à la distribution moyenne de saillance pour cette structure. Si la localisation est peu précise et que d'autres objets présentant des distributions de saillance différentes de celle de la structure visée sont inclus, alors la comparaison avec le modèle permet de pénaliser cette localisation. Elle permet donc d'estimer la précision de la localisation ;
- aux distributions de saillance des autres localisations. Le but est d'ordonner les distributions de saillance et de privilégier la distribution la plus saillante.

La comparaison entre la distribution apprise et la distribution de la localisation s'effectue avec une mesure EMD. Les valeurs sont centrées et réduites. La distance est calculée ainsi :

$$d_o(loc_o, mod_o) = \frac{EMD(loc_o, mod_o) - \hat{mod}_o}{\sigma_{mod_o}} . \quad (5.2)$$

où loc_o représente la distribution de saillance issue de la segmentation, mod_o la distribution apprise pour cette structure, $\hat{mod}_o$ la moyenne des distances EMD entre chaque cas de la base et la

distribution moyenne pour cette structure, et σ_{mod_o} l'écart-type de ces distances.

Pour les comparaisons entre les distributions des localisations, nous avons besoin d'une mesure qui permette de donner un ordre entre ces distributions, et pas uniquement la distance. Nous allons à présent définir cette mesure.

Mesure EMD signée :

La mesure EMD permet de calculer une différence entre distributions de probabilité. Dans notre cas, nous souhaitons comparer deux zones de l'image, masquées par la carte de saillance, afin de déterminer laquelle de ces zones est la plus saillante. La mesure utilisée pour l'apprentissage de saillance précédemment est une mesure non signée, c'est-à-dire qu'elle indique la différence entre deux distributions, mais ne fournit pas l'ordre. Elle ne convient donc pas à nos besoins en l'état. Nous avons présenté comment cette mesure, dans le cas de distributions de probabilité normalisées, et avec une certaine norme, pouvait être calculée en effectuant une comparaison entre histogrammes cumulés. Pour obtenir une distance qui nous fournisse l'ordre, nous proposons de conserver cette formulation et de déterminer le signe de la distance en comparant les différences entre les histogrammes cumulés sans valeur absolue. La distance EMD entre deux distributions p et q est la même formulation que précédemment :

$$emd(p, q) = \frac{\sum_{i=1}^N |P[i] - Q[i]|}{N} ,$$

nous calculons également la somme sans les valeurs absolues :

$$s(p, q) = \sum_{i=1}^N P[i] - Q[i] ,$$

et la distance signée sera déterminée ainsi :

$$emds(p, q) = \begin{cases} emd(p, q) & \text{si } s(p, q) < 0 , \\ -emd(p, q) & \text{si } s(p, q) \geq 0 . \end{cases} \quad (5.3)$$

Cette mesure nous permet de comparer plusieurs distributions, à l'aide de comparaisons deux à deux, afin de déterminer la zone la plus saillante. La figure 5.4 présente un exemple de comparaison entre la localisation d'un thalamus et la localisation d'un putamen. Sur cet exemple, nous pouvons voir sur l'histogramme de saillance de ces localisations que le pic correspondant au putamen est légèrement décalé vers des valeurs plus hautes. Cette différence se reflète sur les histogrammes cumulés. La distance EMD entre ces deux distributions est de 0,017. Si nous souhaitons estimer la saillance selon notre critère du thalamus, alors nous calculons la distance signée :

$$EMDS(th, pu) = -0,017 .$$

À l'inverse, si nous souhaitons estimer la saillance selon notre critère du putamen, alors nous calculons la distance signée ainsi :

$$EMDS(pu, th) = 0,017 .$$

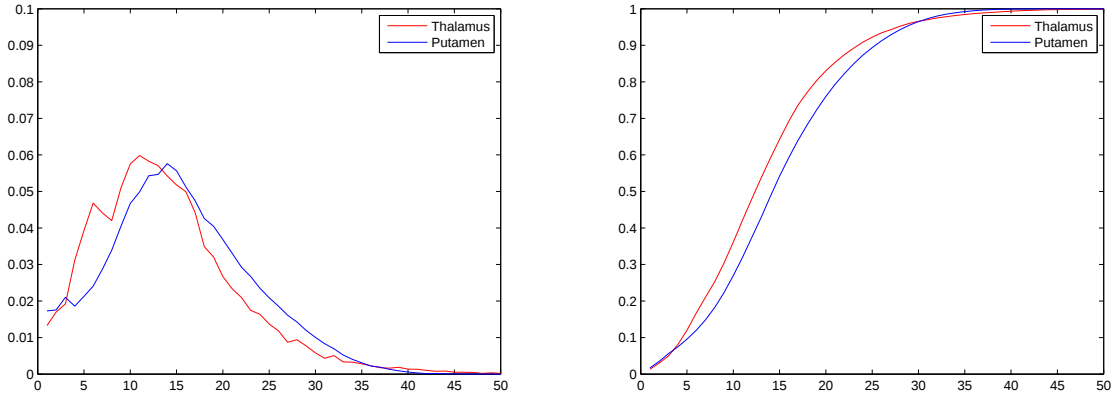


FIG. 5.4 – Comparaison entre les histogrammes de saillance des localisations de deux structures, le thalamus et le putamen au cours d’une étape. La distance EMD entre ces deux histogrammes nous donne une valeur de 0,017. La distance EMDS nous donnera une valeur de $-0,017$ pour le thalamus et une valeur de $0,017$ pour le putamen, nous permettant de déterminer la distribution « la plus saillante ». L’énergie de l’histogramme pour la distribution du thalamus est de 0,040 alors que la distribution du putamen présente une énergie de 0,037. Le critère reposant sur l’énergie donne donc dans ce cas le résultat inverse du résultat souhaité, c’est-à-dire la sélection de la distribution « la plus saillante ».

Le critère de sélection :

Cette mesure nous permet donc d’obtenir une valeur signée et de pondérer ainsi le critère de sélection c , qui est défini ainsi :

$$c_o = |d_o| - \sum_{o' \in V_c \setminus \{o\}} EMDS(loc_o, loc_{o'})$$

où V_c est l’ensemble des nœuds candidats et o désigne l’objet dont nous avons calculé la localisation. La comparaison des localisations grâce à la mesure EMDS permet de pondérer le critère par la localisation la plus saillante. Dans notre exemple, si la distance EMD d_{th} entre la distribution de saillance du thalamus avec la distribution moyenne du thalamus est la même que la distance EMD d_{pu} entre la distribution de la localisation du putamen et le modèle, alors le critère de sélection vaut pour le thalamus $c_{th} = d_{th} - 0,017$ et le critère de sélection pour le putamen vaut $c_{pu} = d_{th} + 0,017$. La mesure EMDS nous a donc permis de pondérer la comparaison avec le modèle par la localisation la plus saillante. La comparaison avec la distribution moyenne de saillance étant centrée et réduite, les valeurs additionnées ne sont pas similaires. La mesure EMDS est donc plutôt une pondération, l’évaluation de la précision de la localisation étant importante d’après nos expériences.

5.1.5 La saillance des tumeurs cérébrales

L’objectif des mécanismes pré-attentionnels en général et des cartes de saillance en particulier est de détecter dans une scène les parties qui sont saillantes, c’est-à-dire qui s’imposent à l’utilisateur lors de l’exploration libre de cette scène. Les tumeurs cérébrales n’ont pas toutes cette caractéristique. La figure 5.5 présente des images avec des pathologies qui présentent des saillances élevées d’une part, à cause notamment du contraste de la tumeur, et de la présence d’une zone nécrotique au centre, et d’autre part, des tumeurs qui présentent au contraire des valeurs de saillance

faibles, voire très faibles. L'absence de saillance de ces dernières pathologies vient de leur taille plutôt large et de leur aspect uniforme.

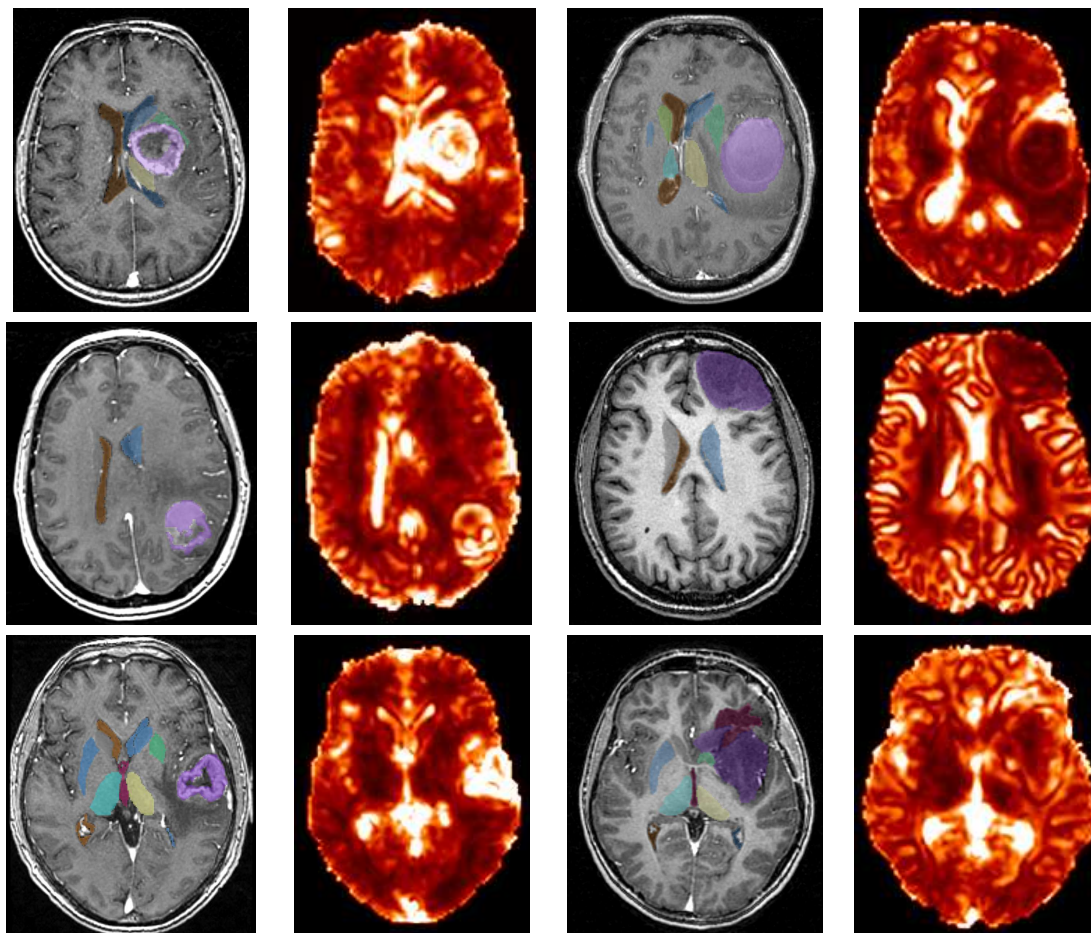


FIG. 5.5 – Deux groupes de pathologies avec des valeurs de saillance inversées. Pour chaque exemple nous présentons une coupe de l'image originale masquée par la carte d'étiquettes de la segmentation manuelle (à gauche, sur ces images, la tumeur apparaît en violet) et une coupe proche de la carte de saillance de la même image (à droite) avec une palette de couleur modifiée. Sur la colonne de gauche sont présentés trois exemples de tumeurs qui génèrent une forte saillance. Sur la colonne de droite, trois exemples de tumeurs présentant des valeurs plus faibles de saillance.

Toutes les pathologies ne partagent donc pas les mêmes caractéristiques de saillance. Parmi l'ensemble des tumeurs, certaines ont un impact immédiat sur les noyaux gris (comme les deux images présentées sur la ligne supérieure de la figure 5.5. Ces tumeurs déplacent des structures et entraînent de grandes altérations de leur morphologie. D'autres structures ont un impact beaucoup plus faible, voire nul sur les noyaux gris (il y a bien sûr un impact sur d'autres parties du cerveau). C'est le cas pour l'image à droite sur la ligne centrale de la figure 5.5. Cette pathologie est importante, mais localisée sur l'avant du cerveau. Le noyau caudé et le putamen ne sont pas vraiment affectés par cette pathologie (déformation très légère). En revanche, le ventricule de gauche est déformé, et le thalamus est écrasé.

Dans le cadre de la segmentation séquentielle, nous pouvons avoir deux objectifs par rapport à la gestion des pathologies : détecter la présence d'une pathologie d'une part, et utiliser cette connaissance pour adapter la segmentation des noyaux gris. La segmentation de la tumeur elle-

même n'est pas traitée dans nos travaux. Pour cela, nous pouvons utiliser les travaux développés par [Khotanlou \(2008\)](#).

5.2 Segmentation séquentielle avec un critère fondé sur la saillance

Si l'objectif suivi est toujours l'optimisation d'un chemin de segmentation, toutefois, l'approche proposée ici présente une optimisation a posteriori du chemin puisque les segmentations sont réalisées à chaque étape du processus. De ce point de vue, l'optimisation est effectuée localement.

Dans cette nouvelle approche, nous souhaitons garder la possibilité de mettre à jour le modèle, c'est-à-dire être capable d'ajouter ou de supprimer des structures du modèle, ce qui n'était pas possible dans la première approche. En effet, l'optimisation dans cette première approche est effectuée « off-line », sans tenir compte de l'information de l'image elle-même. De plus les représentations des relations spatiales sont toujours calculées à partir des représentations des formes d'un cas sain, même dans un cas pathologique. Les degrés de stabilité permettent une adaptation souple aux cas pathologiques. Cependant, il est nécessaire de pouvoir effectuer leur apprentissage sur une base de cas de pathologies similaires. Si nous pouvons obtenir des pathologies de même type, l'apprentissage des degrés de stabilité nécessite en outre que les localisations de la pathologie soient proches entre elles pour avoir des impacts comparables. L'apprentissage devrait donc être suffisant pour gérer tous les cas possibles.

5.2.1 Exploration progressive de l'image

Nous proposons une optimisation locale du chemin, qui tient compte de l'information disponible, à chaque étape du processus de segmentation séquentielle. Cette information provient du modèle générique et des parties de l'image déjà segmentées. Mais nous n'utilisons pas la forme des objets qui ne sont pas encore segmentés. Dans le cadre de la segmentation séquentielle, à un instant donné, nous connaissons la prochaine étape du processus. Nous voyons dans notre approche la prochaine étape comme une exploration d'une partie non connue de l'image. Seule une petite région de l'espace est analysée à un certain moment, ce qui correspond à la reconnaissance et à la segmentation d'un objet. Cette partie de l'espace est définie par les relations spatiales représentables, qui sont les relations spatiales ayant pour objet de référence un objet déjà segmenté.

Le processus est guidé en utilisant un mécanisme pré-attentionnel, ici une carte de saillance, qui indique la zone la plus saillante dans l'espace dans le domaine de recherche. Cette zone est générée en utilisant les parties déjà connues de la scène et les relations spatiales existant entre ces objets et les objets qui doivent encore être reconnus. La figure 5.6 présente le schéma général de la méthode.

Nous présentons d'abord le graphe spatial qui contient la connaissance générique et l'information de l'image extraite au cours du processus. Nous présenterons ensuite les différentes étapes de chaque étape du processus, en commençant par la manière dont le graphe est filtré pour ne conserver que l'information utile à l'étape courante. Ensuite, nous présenterons le mécanisme de sélection de la structure à segmenter, et la mise à jour du graphe après chaque segmentation.

5.2.2 Graphe spatial

Nous utilisons un graphe muni de relations spatiales tel que celui que nous avons décrit dans le chapitre 3. Les notations ont été introduites dans la partie 3.1.3. Le graphe spatial est issu de la connaissance experte et générique de la scène.

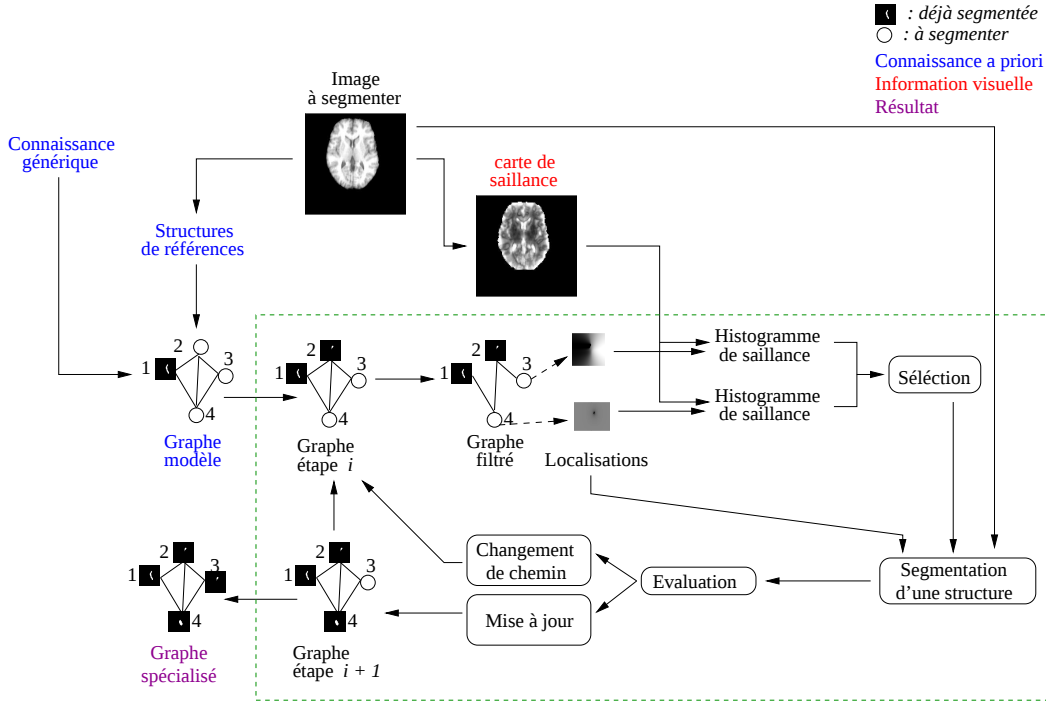


FIG. 5.6 – Schéma général de l'approche proposée qui permet l'intégration d'un mécanisme pré-attentionnel dans un processus de segmentation séquentielle. À une étape i , le graphe est filtré pour ne garder qu'un graphe bipartite entre les nœuds déjà segmentés et les nœuds à segmenter. Les relations spatiales portées par les arcs restants sont représentées dans l'espace de l'image. Elles sont fusionnées pour former le domaine de recherche et fournir la localisation de chaque structure candidate. Un critère dérivé de la saillance de ces localisations est utilisé pour sélectionner la structure à segmenter. La structure peut alors être segmentée à partir de l'information spatiale et de l'image originale. Une étape d'évaluation intervient ensuite pour détecter les éventuelles erreurs de segmentation d'une structure. Si la segmentation est suffisante, le graphe peut être mis à jour avec la segmentation de la structure. Dans le cas inverse, le graphe peut rester en l'état ou une segmentation existante peut être supprimée, avant de passer à la prochaine étape du processus.

Nous rappelons quelques notations ici. Par la suite, nous désignerons l'image originale par I . Un graphe $G = (V, E, L_e)$ est composé d'un ensemble de nœuds $v \in V$ correspondant chacun à une structure cérébrale. Il est également composé d'un ensemble d'arcs binaires $e \in E$. Chaque arc est muni d'un interpréteur permettant d'obtenir l'ensemble flou correspondant aux relations spatiales $\mu_e = L_e(e, v_1)$ portées par cet arc où v_1 est la structure de référence pour la relation. Le graphe utilisé dans nos expériences est présenté dans la figure 5.7. Il intègre 9 structures dont la plupart sont présentes de manière symétrique dans les deux hémisphères.

À l'initialisation du processus, nous avons une structure de référence. Dans le cas des structures cérébrales, le ventricule latéral peut être segmenté en utilisant une méthode de morphologie mathématique par exemple. De plus, sa position centrale et sa taille en font un bon point de référence pour les relations spatiales avec les autres structures. Nous utilisons donc les ventricules latéraux (droit et gauche) comme structures de référence, disponibles au début du processus. Le troisième ventricule est segmenté simultanément par la même procédure et est parfois connecté aux ventricules latéraux. Nous l'utiliserons comme structure de référence également.

Le choix de cette structure est cohérent par rapport à une exploration de l'image selon un

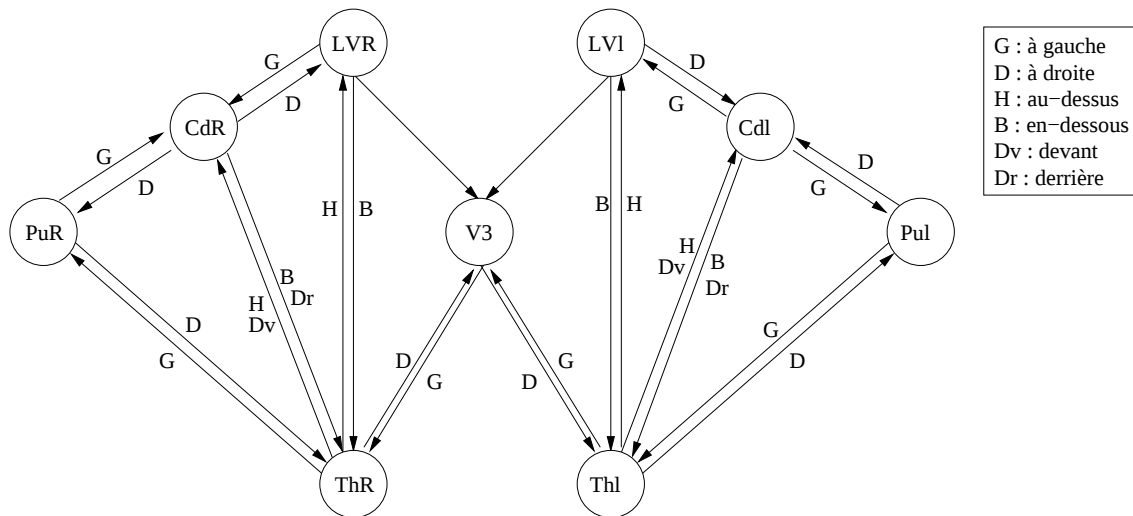


FIG. 5.7 – Le graphe utilisé dans nos expériences. Le graphe est orienté et les arcs entre deux nœuds sont doublés pour prendre en compte les différents chemins de segmentation possibles. Les relations d'orientation entre les structures sont indiquées. Nous utilisons également des relations de distance entre deux structures. Les structures présentes sont les suivantes : ventricule latéral (LV), troisième ventricule (V3), thalamus (TH), putamen (PU) et noyau caudé (CD).

critère de saillance. En effet, les ventricules sont des structures qui présentent presque toujours une forte valeur de saillance (à part dans un cas pathologique où leur grande taille diminue leur saillance). La figure 5.8 présente une image et un seuillage de la carte de saillance pour ne conserver que les plus hautes valeurs. Les ventricules apparaissent dans ces valeurs.

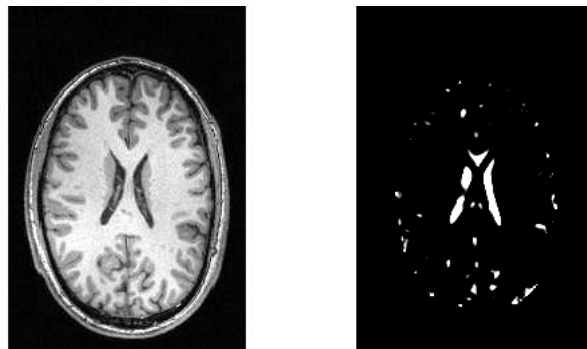


FIG. 5.8 – Une coupe d'une image (à gauche) et le seuillage de la carte de saillance de cette image. Les ventricules présentent des valeurs élevées de saillance et restent apparents même après un seuillage qui enlève la plupart de l'information de saillance. Ces structures sont donc indiquées pour débuter une exploration de l'image selon un critère de saillance.

Nous allons à présent présenter l'approche de segmentation séquentielle, en considérant une étape donnée du processus.

5.2.3 Filtrage du graphe

Nous allons à présent définir des ensembles de nœuds qui seront utilisés par la suite pour le raisonnement. Nous pouvons partitionner l'ensemble des nœuds du graphe V en deux ensembles

distincts : $V = V_{seg} \cup V_{obj}$. Nous avons d'abord l'ensemble V_{seg} des nœuds segmentés, que ce soit une structure de référence ou une structure segmentée au cours du processus. Nous avons également l'ensemble complémentaire V_{obj} des nœuds « objectifs », c'est-à-dire les nœuds qui ne sont pas encore segmentés.

À présent que ces deux ensembles sont définis, nous souhaitons maintenant exprimer la reconnaissance et la segmentation d'une structure cérébrale comme l'ensemble des opérations nécessaires pour transférer un nœud v de l'ensemble des nœuds objectifs vers l'ensemble des nœuds segmentés. À l'étape i nous avons :

$$V_{seg}^i = V_{seg}^{i-1} \cup \{\hat{v}^i\} ,$$

et

$$V_{obj}^i = V_{obj}^{i-1} \setminus \{\hat{v}^i\} ,$$

où $\hat{v}^i$ est le nœud sélectionné à l'étape i du processus.

Nous voulons préciser les liens entre le nœud sélectionné et les deux ensembles de nœuds. Puisque nous avons choisi de voir le processus de segmentation séquentielle comme une exploration progressive de l'image, l'ensemble V_{seg} correspond aux parties déjà explorées de l'image. L'ensemble V_{obj} correspond aux parties inconnues. L'exploration progressive de l'image correspond à l'extension progressive des parties connues, et de ce point de vue, à effectuer l'exploration dans une zone proche des parties déjà connues. Nous pouvons définir E_f comme l'ensemble des arcs dont la source est un nœud appartenant à V_{seg} et dont la cible est un nœud appartenant à V_{obj} :

$$E_f = \{(v_t, v_s) \mid v_t \in V_{seg}, v_s \in V_{obj}\} .$$

Les indices correspondant à l'étape ne sont pas précisés afin de simplifier l'écriture. Cet ensemble est toutefois mis à jour à chaque étape en même temps que les ensembles V_{seg}^i et V_{obj}^i .

Tous les arcs $e \in E_f$ portent des relations spatiales fournissant une information sur la zone à explorer. Si nous définissons la zone de recherche comme une fonction de l'information spatiale portée par ces arcs, alors tous les nœuds ciblés par ces arcs sont dans la zone de recherche.

Nous pouvons définir l'ensemble $V_{fo} \subseteq V_{obj}$ des nœuds ciblés par les arcs contenus dans E_f comme l'ensemble des nœuds de V_{obj} qui sont la cible d'un arc appartenant à E_f . La source de l'arc est nécessairement dans V_{seg} :

$$V_{fo} = \{v_2 \in V_{obj} \mid \exists v_1 \in V_{seg}, (v_1, v_2) \in E_f\} .$$

Nous pouvons de même définir $V_{fs} \subseteq V_{seg}$ comme étant l'ensemble des nœuds de V_{seg} qui sont extrémités d'au moins un arc dont la cible n'appartient pas à V_{seg} :

$$V_{fs} = \{v_1 \in V_{seg} \mid \exists v_2 \in V, v_2 \notin V_{seg}, (v_1, v_2) \in E\} .$$

L'exploration progressive de l'image va donc consister à rechercher le prochain nœud à segmenter parmi les nœuds de V_{fo} . La recherche utilise l'information spatiale fournie par les nœuds de V_{fs} et portée par les arcs E_f . Nous pouvons filtrer le graphe pour ne conserver que V_{fo} , V_{fs} et E_f , ce qui permet de limiter le domaine de recherche. Le sous-graphe obtenu forme un graphe bipartite. Un exemple de graphe filtré à la première étape du processus est illustré par la figure 5.9.

Une étape de la segmentation séquentielle peut donc être formulée comme fonction :

- de l'image I ;
- des segmentations précédentes V_{fs} , qui fournissent l'information spatiale permettant d'explorer l'image ;

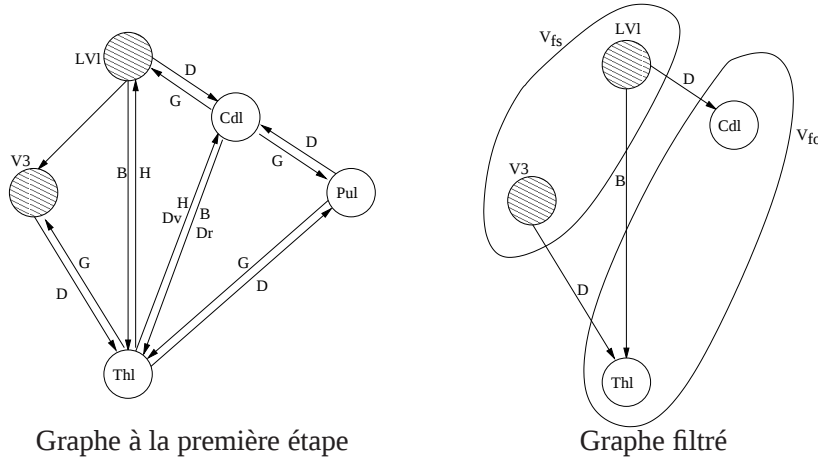


FIG. 5.9 – Le graphe filtré lors de la première étape du processus. Nous n'avons représenté ici que les structures de la partie gauche du cerveau par simplicité mais les structures de la partie droite sont présentes également à cette étape. L'ensemble V_{fs} est composé du ventricule latéral et du troisième ventricule (ces nœuds sont grisés sur la figure). L'ensemble V_{fo} est composé du noyau caudé et du thalamus. Seul les arcs issus des structures segmentées vers les structures non segmentées sont conservés et forment l'ensemble E_f .

- de l'ensemble V_{fo} des nœuds en relation avec V_{fs} ;
- des critères permettant la sélection du nœud, dérivés de la saillance, que nous allons présenter par la suite ;
- et des relations spatiales E_f existant entre les deux groupes de nœuds du graphe (segmentés et à segmenter).

Cela s'écrit :

$$V_{seg}^i = seqseg(V_{seg}^{i-1}, V_{obj}^{i-1}, sal_I, I, E_f^{i-1}) ,$$

où i indique l'étape courante.

5.2.4 Domaine de recherche

Nous avons introduit l'exploration progressive de l'image comme une extension proche des parties connues de l'image. Nous avons également spécifié que cette exploration est fonction de l'information spatiale calculée à partir des parties déjà explorées de l'image. Nous allons à présent définir le domaine de recherche, qui représente l'espace de l'image où nous cherchons la structure la plus à même d'être segmentée, selon un critère dérivé de la saillance que nous présentons dans la partie suivante.

Chaque arc de l'ensemble E_f porte des relations spatiales entre des objets déjà segmentés, et des objets à segmenter. Chaque relation spatiale est représentée dans l'espace de l'image, selon le formalisme décrit dans la partie 3.3. Pour chaque point de l'image, nous avons une valeur de satisfaction de la relation. De plus, un apprentissage des paramètres de ces relations spatiales est effectué tel qu'il a été décrit dans la partie 3.5. Cet apprentissage, effectué sur une base de cas sains et de cas pathologiques, nous permet d'effectuer l'hypothèse que la structure pointée par la relation spatiale est située dans le support de la représentation de la relation spatiale.

Chaque relation spatiale de l'ensemble E_f contribue ainsi à fournir une localisation robuste des structures recherchées. De plus, si plusieurs relations spatiales contribuent à localiser une même structure, alors la localisation peut être précisée en fusionnant ces informations. En particulier, si

nous ne disposons que d'une relation d'orientation pour une structure, alors la zone de l'image où cette relation sera satisfaite est grande par rapport à la structure. Mais si cette zone est liée à une relation de distance, alors la localisation de la structure sera beaucoup plus précise. La figure 5.10 sur la ligne du haut présente un exemple de la représentation d'une relation d'orientation (à gauche), et l'ensemble flou (à droite) correspondant à la fusion entre cette représentation et une relation de distance (au centre). La localisation de la relation est beaucoup plus précise dans ce dernier cas, même si elle reste grande par rapport à la taille de la structure.

Pour chaque arc e contenu dans E_f , un interpréteur d'arc L_e produit la représentation de chaque relation spatiale présente sur cet arc. L'interpréteur d'arc agit comme une fonction permettant d'indiquer quelles relations sont présentes sur l'arc, parmi toutes les relations spatiales possibles dans le modèle, et d'obtenir l'ensemble flou correspondant. Si la relation est présente, alors sa représentation est générée, avec des paramètres génériques pour le type de relation. L'arc contient également pour chaque relation spatiale l'intervalle flou issu de l'apprentissage pour cet arc (le couple de structures). L'intervalle flou est alors appliqué à la représentation de la relation spatiale pour générer la représentation exacte pour cette relation pour le couple de structures reliées par l'arc.

Une fois les relations spatiales portées par un arc représentées dans l'espace de l'image, nous pouvons fusionner toutes ces représentations pour obtenir un ensemble flou représentatif de l'information spatiale portée par cet arc e de manière conjonctive :

$$\mu_{Rel}^e = \top_{r \in e} \mu_e^r ,$$

où $\top$ est une t-norme (Dubois et Prade (1980)).

Pour chaque nœud candidat v , sa localisation spatiale estimée est définie par la fusion des ensembles flous représentant chaque arc ayant ce nœud pour cible. Le nœud ciblé appartient à l'ensemble V_{fo} . La localisation est calculée ainsi :

$$loc_v = \top_{e \in (A(v) \cap E_f)} (\mu_{Rel}^e) ,$$

où $\top$ est une t-norme, et $A(v)$ représente les arcs ayant le nœud v pour cible. La figure 5.10 présente le processus permettant d'obtenir la localisation de deux structures à la première étape.

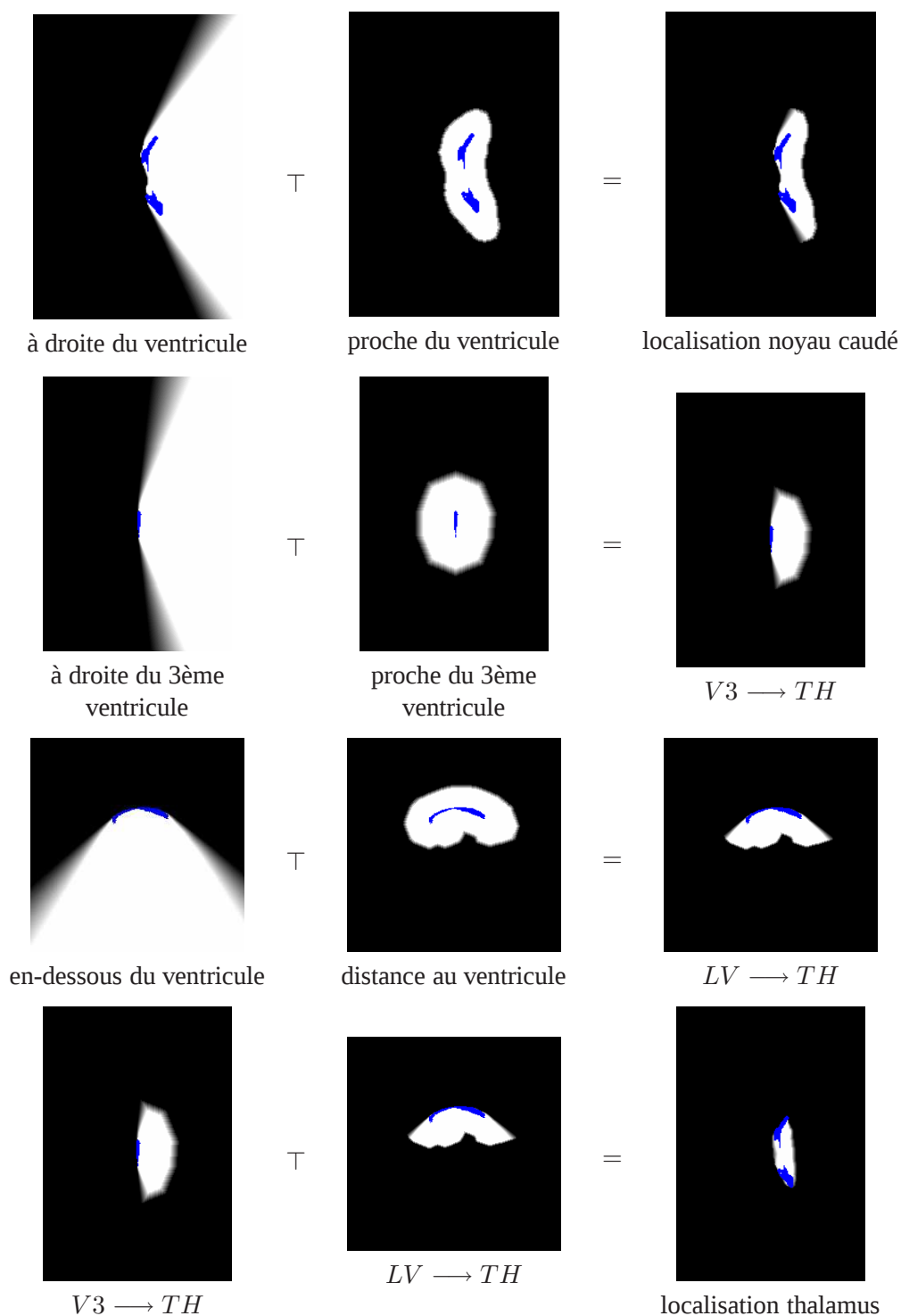


FIG. 5.10 – Représentation des relations spatiales. En haut, les relations portées par l'arc entre le ventricule (LV) et le noyau caudé (CN). La fusion (à droite) donne la localisation du noyau caudé (qui n'est pas connecté à une autre structure disponible à cette étape). Les deux lignes suivantes montrent les représentations des relations spatiales respectivement sur l'arc entre le 3ème ventricule (V3) et le thalamus (TH), et sur l'arc entre le ventricule et le thalamus (en vue sagittale). Les ensembles flous issus de ces deux arcs sont fusionnés pour donner la localisation du thalamus (ligne du bas). Le troisième ventricule est situé en-dessous du ventricule latéral.

Nous pouvons à présent définir l'ensemble flou correspondant au domaine de recherche de l'image, comme la fusion des ensembles flous portés par chaque arc composant l'ensemble E_f . Le domaine de recherche est défini ainsi :

$$\mu_{sd} = \perp_{e \in E_f} (\mu_{Rel}^e) ,$$

où $\perp$ est une t-conorme (disjonction floue) (Dubois et Prade (1980)). Le domaine de recherche indique une zone de l'espace qui inclut la localisation spatiale de tous les objets cibles (la combinaison est disjonctive). Le domaine de recherche à la première étape est illustré par la figure 5.11.

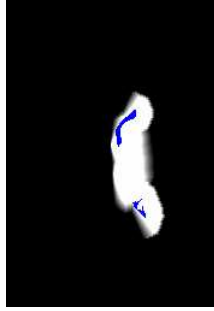


FIG. 5.11 – Domaine de recherche à la première étape du processus. Le domaine de recherche fusionne de manière disjonctive les localisations de toutes les relations spatiales de l'ensemble E_f et représente la zone de l'image où se situe l'ensemble des structures à rechercher V_{fo} . Dans cet exemple, le domaine de recherche est la fusion des relations entre le ventricule et le noyau caudé d'une part, entre le ventricule et le troisième ventricule vers le thalamus d'autre part.

5.2.5 Intégration de la saillance

La carte de saillance de l'image à explorer est calculée sur l'image complète au début du processus. La saillance est donc fixée et ne varie pas au cours des étapes du processus. Nous avons déterminé la localisation spatiale de chacune des structures de l'ensemble V_{fo} . Nous allons à présent combiner la carte de saillance avec l'espace de recherche et la localisation de chacune des structures, afin d'extraire, pour chaque structure, la saillance de la zone correspondante. Pour chaque nœud v de V_{fo} , nous déterminons :

$$saillance_v = \top(loc_v, \mu_{sd}, sal_I) ,$$

où sal_I est la carte de saillance de l'image que nous explorons. Le domaine de recherche est ici une disjonction floue des localisations, il n'est donc pas utile à cette étape. Il est possible d'appliquer une restriction sur le domaine de recherche (pour limiter de manière quantitative la zone de recherche par exemple), ce qui n'est pas fait ici. Nous calculons ensuite un histogramme de la saillance de chaque localisation, dont le nombre N de niveaux de quantification de l'histogramme est fixé arbitrairement à 100 :

$$H_v[i] = \sum_{x \in \mathcal{S}} \mathbb{1}_i(saillance_v(x)) ,$$

où $\mathbb{1}(\cdot)$ représente la fonction indicatrice. Des exemples de ces histogrammes sont présentés dans la figure 5.12.

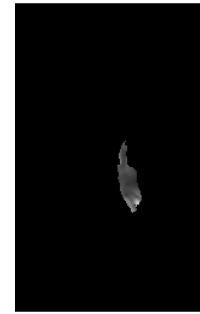
Masquage des localisations par la saillance :



Carte
de saillance

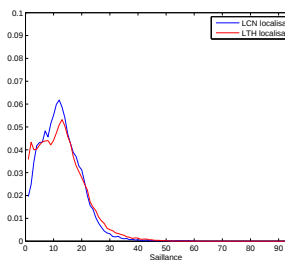


Saillance à la localisation
du noyau caudé

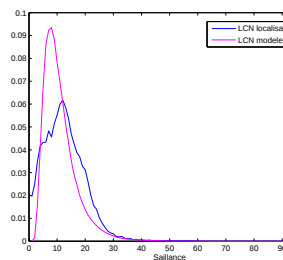


Saillance à la localisation
du thalamus

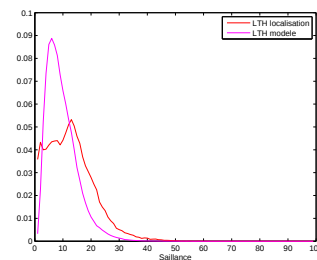
Histogrammes de saillances :



comparaison entre
localisations

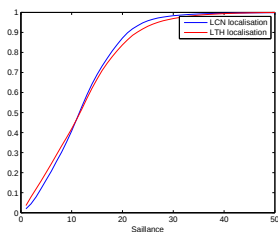


LCN : localisation et
modèle

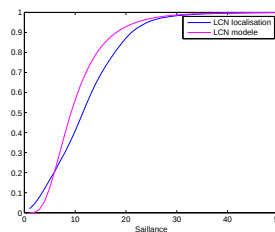


LTH : localisation et
modèle

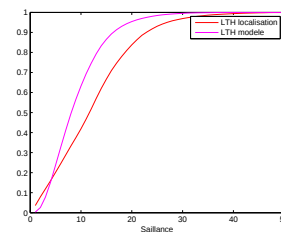
Histogrammes cumulés :



comparaison entre
localisations



LCN : localisation et
modèle



LTH : localisation et
modèle

FIG. 5.12 – Masquage des localisations par la carte de saillance (en haut à gauche) et sélection de la localisation la plus saillante selon le critère retenu. Les deux localisations (noyau caudé et thalamus) sont directement comparées (histogrammes à gauche, la localisation du noyau caudé est plus saillante selon la mesure emds avec une valeur de 0,0084). Ici, ces deux localisations sont proches et se chevauchent en partie, ce qui explique la proximité des histogrammes. Chaque histogramme est ensuite comparé au modèle appris pour cette structure (histogramme au centre pour le noyau caudé avec une valeur de $-0,089$ et histogramme de droite pour le thalamus avec une valeur de $0,791$). Dans cet exemple, le noyau caudé sera sélectionné avec une mesure de $0,076$ qui est le minimum parmi les 4 localisations concernées à cette étape du processus.

5.2.6 Sélection du prochain objet

Nous pouvons à présent effectuer le processus de sélection du prochain objet à segmenter, c'est-à-dire la sélection du nœud cible du graphe représentant un objet en particulier, que nous allons relier à une partie de l'image. La sélection est effectuée par analyse de la saillance dans le domaine de recherche. Le filtrage du graphe nous donne deux groupes de nœuds : V_{fs} et V_{fo} et la sélection est effectuée dans V_{fo} (et donc l'objet représenté du prochain nœud).

Nous avons présenté dans la partie 5.1.4 le critère que nous utilisons pour comparer les histogrammes de saillance, qui est dérivé de la mesure EMD (eq. 5.3), et qui permet de refléter non seulement la localisation la plus saillante, mais aussi la différence par rapport au modèle de saillance appris pour la structure visée. La sélection du prochain objet à segmenter est effectuée à partir de l'histogramme de saillance généré pour chaque nœud candidat en sélectionnant le nœud qui présente l'histogramme « le plus saillant » :

$$\hat{v} = \arg \min_{v \in V_{fo}} \left(|d_o| - \sum_{o' \in V_c \setminus \{o\}} EMDS(loc_o, loc_{o'}) \right) . \quad (5.4)$$

La figure 5.12 présente deux localisations dont nous avons calculé les histogrammes de saillance, représentant respectivement le noyau caudé et le thalamus gauche. Nous allons calculer le critère permettant de sélectionner une localisation parmi les deux. Dans l'exemple choisi, il y a en fait 4 structures candidates, mais nous ne présenterons que les histogrammes et les localisations de deux de ces structures, dans le même hémisphère.

Nous allons d'abord comparer chacune des distributions de saillance avec le modèle appris pour la structure concernée selon le critère présenté par l'équation 5.2 :

$$d_{mod_{lcn}} = \frac{EMD(loc_{lcn}, mod_{lcn}) - \hat{mod}_{lcn}}{\sigma_{mod_{lcn}}} = -0,089 , \quad (5.5)$$

et

$$d_{mod_{lth}} = \frac{EMD(loc_{lth}, mod_{lth}) - \hat{mod}_{lth}}{\sigma_{mod_{lth}}} = 0,791 . \quad (5.6)$$

Nous pouvons voir ici que la distribution du noyau caudé gauche est plus proche du modèle que la distribution du thalamus gauche, une fois les valeurs centrées et réduites.

Nous allons ensuite comparer les localisations entre elles selon la mesure EMDS présentée par l'équation 5.3 :

$$emds(loc_{lcn}, loc_{lth}) = 0,0084 , \quad (5.7)$$

et

$$emds(loc_{lth}, loc_{lcn}) = -0,0084 . \quad (5.8)$$

La localisation du noyau caudé gauche est donc jugée plus saillante que la localisation du thalamus selon ce critère.

Pour chacune des localisations, nous ajoutons à cette dernière valeur la comparaison aux autres structures candidates, c'est-à-dire le noyau caudé droit et le thalamus droit dans ce cas. Nous obtenons les valeurs suivantes :

$$d_{inter_{lcn}} = \sum_{v \in V_{fo} \setminus lcn} emds(loc_{lcn}, loc_v) = 0,013 , \quad (5.9)$$

et

$$d_{inter_{lth}} = \sum_{v \in V_{fo} \setminus lth} emds(loc_{lth}, loc_v) = -0,040 . \quad (5.10)$$

Nous pouvons voir ici que la localisation du noyau caudé est jugée plus saillante que la moyenne des structures candidates (valeur positive). Par contre, la localisation du thalamus est moins saillante que les autres.

La sélection s'effectue donc sur ces valeurs :

$$c_{lOC_{lcn}} = |d_{mod_{lcn}}| - d_{inter_{lcn}} = 0,089 - 0,013 = 0,076 \quad , \quad (5.11)$$

$$c_{lOC_{lth}} = |d_{mod_{lth}}| - d_{inter_{lth}} = 0,791 + 0,040 = 0,831 \quad , \quad (5.12)$$

Nous allons donc sélectionner le critère minimum, c'est-à-dire le noyau caudé dans ce cas.

La sélection du nœud permet de segmenter l'objet. La segmentation peut être exprimée comme une fonction de l'objet sélectionné selon le critère dérivé de la saillance $\hat{v}$, en fonction des relations spatiales avec les objets déjà segmentés et en relation avec le nœud à segmenter, et de l'image originale :

$$seg_{\hat{v}} = segment(\hat{v}, loc_{\hat{v}}, I) \quad .$$

5.2.7 Le processus de segmentation

Rappelons que le processus de segmentation ne fait pas l'objet de nos travaux. Les entrées sont l'image à segmenter, ainsi que l'information spatiale dont nous disposons pour cette structure. La sortie de la méthode de segmentation est une carte binaire représentant l'objet. Cependant, les résultats étant dépendants du processus de segmentation, nous allons brièvement présenter le processus de segmentation qui a été défini par [Colliot \(2003\)](#). Il est divisé en deux parties. La première partie consiste à trouver une segmentation grossière de l'objet recherché. La deuxième partie affine cette segmentation à l'aide d'un modèle déformable.

Pour obtenir une segmentation grossière de l'objet recherché, la méthode repose sur deux types d'information : la radiométrie des noyaux gris du cerveau (dont font partie les structures que nous recherchons) et la limitation de l'espace de recherche grâce à l'information spatiale.

Radiométrie des structures cérébrales :

Une analyse de l'image nous fournit les caractéristiques de la matière blanche et de la matière grise du cerveau ([Mangin et al. \(1998\)](#)), la moyenne (resp. $\hat{x}_{wm}$, $\hat{x}_{gm}$) et l'écart-type (resp. σ_{wm} , σ_{gm}). À partir de ces valeurs, il est possible d'obtenir les caractéristiques radiométriques de chaque noyau gris du cerveau. Ces travaux ont été présentés par [Poupon et al. \(2008\)](#).

Pour une structure O d'une image I donnée, nous connaissons α_o et β_o , les paramètres permettant d'obtenir les caractéristiques de la structure O à partir des caractéristiques de la matière blanche et de la matière grise du cas c . Les caractéristiques radiométriques de la structure O sont dérivées ainsi :

$$\hat{x}_o = \alpha_o \hat{x}_{wm} + (1 - \alpha_o) \hat{x}_{gm} \quad ,$$

et

$$\sigma_o = \beta_o \frac{(\sigma_{wm} + \sigma_{gm})}{2} \quad .$$

Les paramètres α et β pour chaque noyau gris sont déterminés par apprentissage sur une base de données : si nous disposons de la segmentation des structures d'une image c , alors nous pouvons calculer les paramètres α_c et β_c à partir des caractéristiques des matières de l'image, et des niveaux de gris de chacune des structures. Nous pouvons ainsi estimer les paramètres α et β pour la base complète. Dans [Poupon et al. \(2008\)](#), le paramètre α est une moyenne des α_c calculés sur chaque image. Le paramètre β est le maximum des valeurs β_c calculées pour chaque image.

Nous avons effectué une nouvelle estimation de ces paramètres avec notre base d'apprentissage. La figure 5.13 présente les valeurs obtenues pour le noyau caudé. Deux nuages de points sont affichés. Le premier correspond aux valeurs constatées avec les images originales. Le deuxième correspond aux valeurs calculées à partir d'images dont nous avons corrigé l'hétérogénéité du champ par la méthode décrite par Mangin (2000). Ces images sont utilisées dans le processus de segmentation. Les nuages de points montrent une certaine dispersion des valeurs entre 0, 1 et 0, 45 pour α , et entre 0, 45 et 1, 4 pour β .

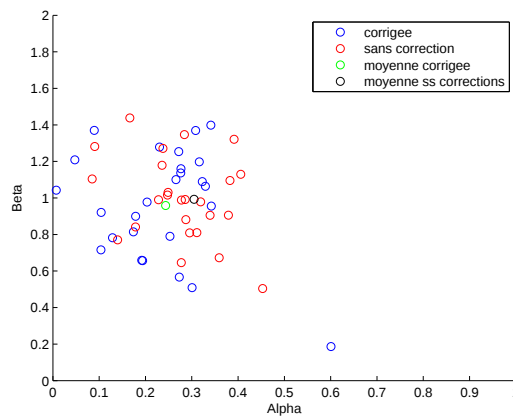


FIG. 5.13 – Les valeurs de α_c et β_c pour chaque cas sain de notre base d'apprentissage, et pour le noyau caudé. L'apprentissage a été effectué avec les images originales et avec des images dont le biais a été corrigé. Ces dernières images sont utilisées dans le processus de segmentation. Les moyennes obtenues avec notre base et avec les images originales sont proches des valeurs indiquées dans (Poupon *et al.* (2008)).

Le tableau 5.3 présente les valeurs des paramètres α et β appris sur notre base, ainsi qu'une comparaison avec les valeurs fournies par Poupon *et al.* (2008). Les valeurs apprises à partir des images originales sont assez proches de ces dernières valeurs, alors que celles calculées avec les images corrigées sont un peu plus éloignées.

TAB. 5.3 – Comparaison des valeurs α et β apprises sur notre base pour chacune des structures avec les valeurs présentées par Poupon *et al.* (2008). Les valeurs apprises avec les images originales sont assez proches des valeurs initiales. Il y a plus de différences avec les images corrigées.

	Poupon <i>et al.</i> (2008)		Img. Originales		Avec corrections	
Structure :	α	β	α	β	α	β
Noyau caudé	0,305	1,328	0,305	1,675	0,244	1,398
Thalamus	0,633	1,374	0,617	1,899	0,578	1,585
Putamen	0,508	1,072	0,539	1,210	0,527	1,023
Globus Pallidus	0,945	0,926				
Accumbens	0,265	1,016				

Nous avons également effectué un apprentissage en séparant les images de la base IBSR, les images de la base OASIS et les images des cas pathologiques. Le tableau présente les valeurs obtenues avec des images corrigées, pour les trois ensembles d'images. Nous utiliserons ces dernières valeurs dans nos expériences.

TAB. 5.4 – Comparaison des valeurs α et β apprises pour chacun des ensembles de la base (IBSR, OASIS et les cas pathologiques).

	IBSR		OASIS		Cas Pathologiques	
Structure :	α	β	α	β	α	β
Noyau caudé	0,216	1,208	0,278	1,398	0,303	1,693
Thalamus	0,557	1,586	0,606	1,152	0,592	1,483
Putamen	0,545	0,976	0,505	1,024	0,485	1,341

Une fois les caractéristiques $\hat{x}_o$ et σ_o connues, nous pouvons les utiliser pour effectuer un seuillage de l'image. Les deux seuils ont été fixés arbitrairement aux valeurs suivantes : $\hat{x}_o - \sigma_o$ et $\hat{x}_o + \sigma_o$. Nous obtenons ainsi une carte d'appartenance à une structure particulière du cerveau. La figure 5.14 présente les cartes obtenues pour les structures suivantes : noyau caudé, thalamus et putamen. Nous pouvons voir sur cette figure que la carte du thalamus en particulier ne permet pas de distinguer clairement la structure malgré la connaissance de ses caractéristiques.

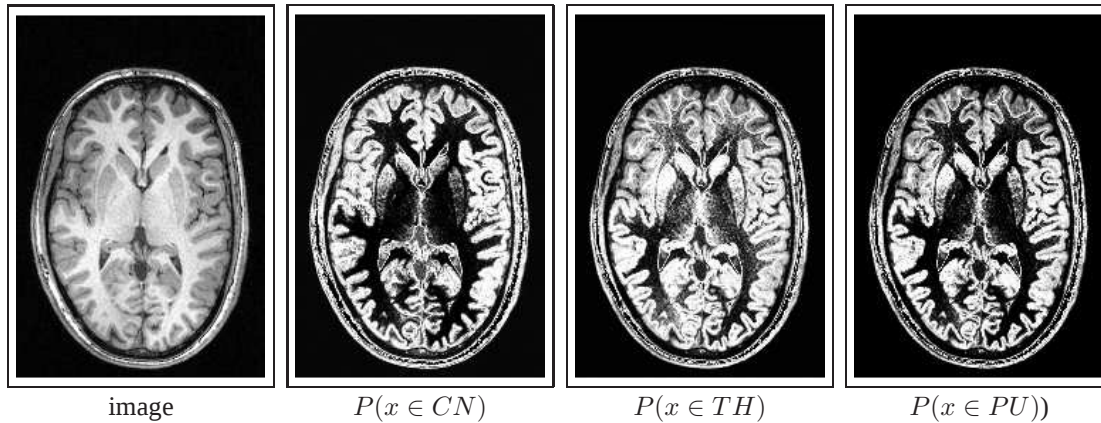


FIG. 5.14 – Carte d'appartenance à trois structures cérébrales : le noyau caudé (CN), le thalamus (TH) et le putamen (PU). Les valeurs (α, β) utilisées pour chacune des structures sont les suivantes : pour le noyau caudé (0, 305; 1, 328), pour le thalamus (0, 633; 1, 374) et pour le putamen (0, 508; 1, 072).

Identification et segmentation initiale :

Le seuillage de l'image grâce aux caractéristiques radiométriques permet d'obtenir une carte d'appartenance à la structure recherchée. Cependant, elle n'est pas suffisante pour identifier la structure. L'information spatiale permet de réduire l'espace de recherche autour de la structure. Une ouverture morphologique permet de séparer les différentes composantes restantes. La restriction de l'espace de recherche permet alors de considérer la plus grande composante restante comme étant la segmentation initiale de l'objet recherché. Cette segmentation n'a pas besoin d'être très précise, car elle sera affinée par la suite grâce au modèle déformable.

Les différentes étapes qui permettent d'obtenir la segmentation initiale sont illustrées dans la figure 5.15 en prenant l'exemple de la segmentation d'un thalamus (gauche) à partir des relations spatiales issues du ventricule latéral gauche et du troisième ventricule.

processus de segmentation du thalamus :

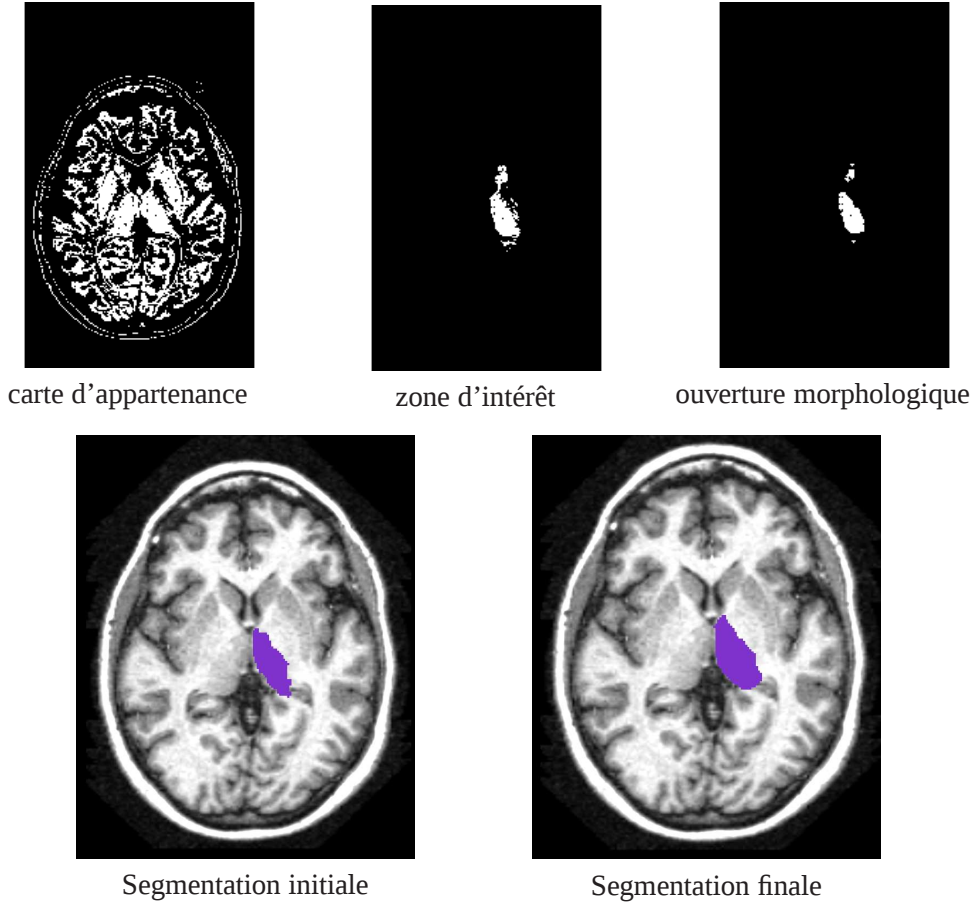


FIG. 5.15 – Les différentes étapes du processus de la segmentation du thalamus selon la méthode proposée par [Colliot \(2003\)](#). Les relations spatiales utilisées sont les relations entre le ventricule latéral, le troisième ventricule et le thalamus. La carte d'appartenance (en haut à gauche) est calculée à partir d'une analyse de l'image selon la méthode proposée par [Poupon et al. \(2008\)](#). Cette carte est masquée par la région d'intérêt, qui est une conjonction des représentations des relations spatiales ayant la structure comme cible (en haut au centre). Une ouverture morphologique permet de séparer les composantes. En particulier ici, un morceau du noyau caudé apparaissait en haut de la zone d'intérêt. La plus grande composante connexe est conservée comme segmentation initiale (après une fermeture morphologique, en bas à gauche). La segmentation initiale et les relations spatiales sont utilisées pour initialiser un modèle déformable et obtenir la segmentation finale (en bas à droite).

Modèle déformable contraint par les relations spatiales :

Le modèle déformable a été décrit en détail par [Colliot \(2003\)](#). Le modèle retenu utilise un maillage simplexe. La particularité du schéma d'évolution est d'intégrer les relations spatiales comme une force. L'évolution de la surface déformable S est décrite par l'équation suivante :

$$\gamma \frac{\partial S}{\partial t} = \mathbf{F}_{int}(S) + \mathbf{F}_{ext}(S) \quad (5.13)$$

où $\mathbf{F}_{int}$ est la force interne contrôlant la régularité de la surface, décrite ainsi :

$$\mathbf{F}_{int}(S) = \alpha \nabla^2 S - \beta \nabla^2 (\nabla^2 S) \quad (5.14)$$

et où $\mathbf{F}_{ext}$ est la force externe qui attire la surface vers les contours de l'objet recherché. La force externe est une combinaison de deux forces :

$$\mathbf{F}_{ext}(S) = \lambda \mathbf{F}_C + \nu \mathbf{F}_R \quad (5.15)$$

où $\mathbf{F}_C$ est une force d'attache aux données dérivées d'un « Gradient Vector Flow » (Xu et Prince (1998)). La force $\mathbf{F}_R$ est une force dérivée des relations spatiales utilisées pour définir la région d'intérêt. Pour chacune des structures à segmenter, nous utilisons les mêmes paramètres d'évolution pour ce modèle déformable.

Cette méthode de segmentation nécessite que la région d'intérêt, définie par les relations spatiales, soit précise. Si elle est trop restrictive, alors une partie de l'objet ne pourra pas être segmentée. D'un autre côté, si elle est trop large, alors l'identification de la composante ne peut plus s'effectuer uniquement par la taille. De tels problèmes peuvent apparaître dans notre cas lorsque nous utilisons le résultat des segmentations précédentes pour estimer la région d'intérêt et pas uniquement des structures de référence. Nous verrons dans la prochaine partie comment nous pouvons essayer de prendre en compte ce type de problème.

D'autres problèmes (décrits dans Colliot (2003)) peuvent survenir :

- lorsque la carte d'appartenance ne permet pas de faire apparaître clairement la structure, car l'ouverture morphologique risque de faire disparaître tout ou partie de la composante correspondant à l'objet recherché,
- ou au contraire, lorsque l'ouverture morphologique ne permet pas de séparer des composantes correspondant à différents objets de manière automatique (une méthode à base de ligne de partage des eaux est proposée).

Nous allons voir à présent comment nous pouvons évaluer le résultat d'une segmentation.

5.2.8 Mise à jour du graphe

Nous avons raisonné jusqu'ici à partir du modèle structurel et de l'information de saillance pour guider le choix d'une structure à segmenter. La segmentation d'une structure apporte une connaissance importante pour la suite du processus, qui permet de spécialiser progressivement le graphe vers le cas spécifique pris en compte. Cette information nous permet en outre de représenter les relations spatiales utilisées dans la suite du processus.

Il est donc important de pouvoir qualifier le résultat d'une segmentation afin de pouvoir adapter la stratégie si nécessaire. Nous présentons dans cette partie les différents cas à envisager.

Lorsqu'une structure est segmentée, le processus de segmentation utilise toute l'information spatiale disponible dans le modèle pour contraindre le processus de segmentation, en définissant une zone d'intérêt tout d'abord (la localisation de la structure), puis en contraignant le modèle déformable grâce à l'information spatiale ensuite. Cependant, lors de l'évaluation de la segmentation, nous nous limitons aux interactions entre la structure qui vient d'être segmentée et la dernière structure segmentée qui a fourni une information spatiale utilisée pour la segmentation. Cette structure est désignée comme la structure « parente » de la structure segmentée. En cas d'échec du processus de segmentation, ou si le résultat est jugé non satisfaisant selon les critères que nous présentons ensuite, les décisions concernent la structure segmentée, mais aussi la structure parente, dont la segmentation peut être supprimée.

Nous proposons principalement une stratégie en cas de problème de segmentation : nous pouvons contraindre le système à changer de chemin, c'est-à-dire à ne pas essayer de segmenter à nouveau une structure dans les conditions qui ont donné un résultat non satisfaisant. En pratique, une structure est marquée comme étant non segmentable tant qu'une structure voisine dans le graphe

n'a pas été segmentée. La segmentation d'une structure voisine permet d'utiliser l'information spatiale portée par l'arc entre les deux structures, et ainsi d'apporter de nouvelles informations. Si aucune structure ne permet d'apporter une information nouvelle, alors la structure ne peut pas être segmentée de nouveau.

À la sortie du processus de segmentation d'une structure, nous prenons en compte différents cas possibles, en fonction du résultat de la segmentation. Nous distinguons cinq cas :

- si la segmentation a échoué, c'est-à-dire que l'image produite est vide d'information, alors nous souhaitons contraindre le processus à choisir un nouveau chemin :
 1. si une structure parente existe, alors sa segmentation est supprimée et elle ne pourra être segmentée de nouveau en l'état ;
 2. sinon c'est la structure courante qui ne pourra pas être segmentée de nouveau en l'état ;
- si une segmentation a été produite, alors elle est évaluée selon un critère de cohérence spatiale et selon un critère reposant sur la saillance :
 3. soit la cohérence spatiale est trop faible, et dans ce cas la structure parente est supprimée et ne pourra être segmentée de nouveau en l'état ;
 4. soit la cohérence spatiale est suffisante mais le critère reposant sur la saillance est trop faible. Dans ce cas la segmentation est refusée et cette structure ne pourra être segmentée de nouveau en l'état ;
 5. soit la segmentation est acceptée selon les deux critères, et le graphe peut être mis à jour.

Nous allons à présent présenter ces différents cas, les critères d'évaluation, et les actions qu'ils engendrent. La figure 5.16 schématise les différents cas et les actions.

5.2.8.1 Pas de segmentation

Tout d'abord, il est possible qu'aucune segmentation ne soit possible, et cela pour deux raisons :

- soit la localisation a été mal définie, et est trop restrictive,
- soit la carte d'appartenance est insuffisante, mais ce problème est intrinsèque à la méthode de segmentation. La carte d'appartenance a été décrite dans la partie 5.2.7.

Nous n'avons pas de critère permettant de séparer ces deux cas. Une localisation est meilleure si elle est plus précise, donc plus restrictive qu'une autre. Mais si elle est trop restrictive, alors la segmentation peut échouer. Évaluer la pertinence d'une localisation nécessite un a priori sur la taille de la structure visée, ce que nous ne possédons pas. Fixer un seuil de taille peut également être hasardeux. De plus, même si la localisation a la taille de la structure, cela ne signifie pas que la structure est comprise, tout ou partie, dedans. La segmentation d'une structure est donc effectuée sans évaluation de la localisation au préalable.

Si la segmentation a échoué, alors nous pouvons uniquement émettre l'hypothèse que la segmentation d'une structure parente a donné une localisation trop restrictive. S'il y a une structure parente, alors nous supprimons sa segmentation, et nous l'empêchons d'être segmentée de nouveau en l'état.

Si la structure parente n'existe pas, alors dans ce cas, la structure a été segmentée à partir des structures de référence et le problème ne provient pas de la définition de la localisation. Dans ce cas, nous souhaitons également contraindre le processus à utiliser un autre chemin, afin de laisser au processus la possibilité de segmenter à nouveau cette structure une fois que nous aurons acquis plus d'information sur sa localisation.

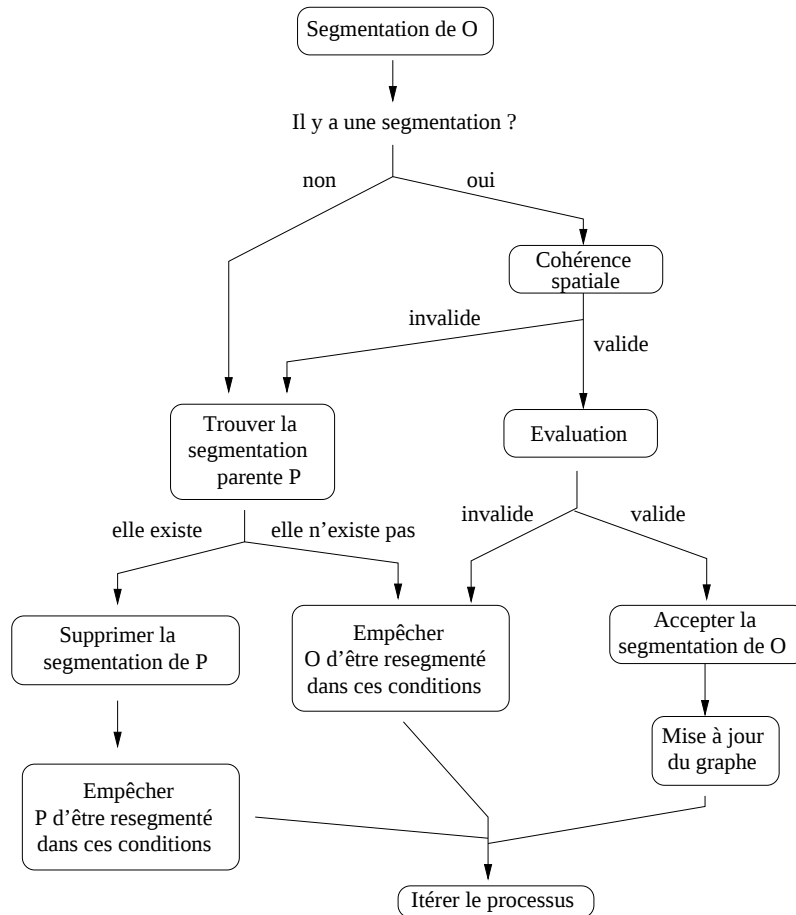


FIG. 5.16 – Procédure pour évaluer la segmentation. Si la segmentation échoue, alors le système est contraint de changer de chemin et de ne segmenter à nouveau la structure identifiée comme responsable qu’après avoir obtenu de nouvelles informations sur cette structure. Si une segmentation est produite, alors la cohérence spatiale avec la structure parente est évaluée et cette dernière peut être supprimée si elle n’est pas suffisante. Sinon, une distribution de saillance est générée et elle est comparée à la distribution apprise pour cette structure. Soit la segmentation est acceptée, soit elle ne l’est pas et dans ce dernier cas, le système est contraint à changer de chemin.

5.2.8.2 Il y a une segmentation

Si nous avons obtenu une segmentation, nous souhaitons faire une estimation de la qualité de cette segmentation. Mais, puisque nous ne souhaitons pas utiliser les représentations des structures, cette évaluation ne doit pas non plus reposer sur une comparaison avec un modèle morphologique de cette structure par exemple.

Nous proposons deux critères pour évaluer le résultat de la segmentation obtenue. Le premier critère évalue la cohérence spatiale du modèle après la segmentation. Le deuxième critère repose sur l’apprentissage de la saillance pour la structure segmentée. Ce critère est une mesure intrinsèque de la segmentation (une attache aux données).

Évaluation de la cohérence du modèle spatial :

Afin de pouvoir évaluer la cohérence du modèle, nous pouvons nous reposer sur les représen-

tations des relations spatiales qui sont déjà dans le modèle, et sur les relations que la nouvelle segmentation permet de représenter. Ces dernières visent aussi bien des structures non segmentées que des structures déjà segmentées.

La structure qui vient d'être segmentée se situe nécessairement dans la localisation définie par les relations spatiales visant cette structure. Ces relations spatiales n'apportent donc pas d'information sur la cohérence du modèle. D'un autre côté, les relations issues de la structure qui vient d'être segmentée et pointant vers des structures déjà segmentées (donc les relations inverses de celles qui ont été utilisées pour sa segmentation), peuvent nous fournir une information.

Nous proposons donc d'évaluer, à l'aide d'une mesure de satisfaction floue introduite dans le chapitre précédent, si les relations inverses sont satisfaites par la segmentation. La mesure de satisfaction f_s est définie ainsi (Bouchon-Meunier *et al.* (1996)) :

$$f_s(Rel, Obj) = \frac{\sum_{x \in S} \min(\mu_{Rel}(x), \mu_{Obj}(x))}{\sum_{x \in S} \mu_{Obj}(x)} , \quad (5.16)$$

où S désigne l'espace de l'image. Cette mesure sera maximale si la structure représentée par μ_{obj} est située dans le noyau de la relation représentée par μ_{rel} . En particulier, cette satisfaction sera très faible si une segmentation est très petite par rapport à la structure visée. Cela peut se produire lorsqu'une segmentation précédente a reconnu la mauvaise structure. Nous utilisons pour ce critère un seuil qui a été fixé expérimentalement à 0,5.

Nous illustrons cet exemple dans la figure 5.17. Dans ce cas, la segmentation du thalamus, effectuée en premier, a échoué. La carte binaire utilisée (sur l'image de gauche) inclut les autres structures qui sont de plus connectées. La segmentation sélectionnant la plus grande composante, celle correspondant au thalamus, est supprimée et le noyau caudé et le putamen sont reconnus comme étant le thalamus. La comparaison des histogrammes de saillance ne permet pas dans ce cas de détecter le problème. La segmentation du noyau caudé qui est effectuée juste après donne une segmentation quasi vide (quelques pixels). Les relations spatiales issues du noyau caudé vers le thalamus sont alors représentées, permettant de calculer la satisfaction floue entre cette représentation et la segmentation du thalamus. La satisfaction dans ce cas donne une valeur de 0, ce qui permet de détecter un problème. Dans ce cas, la segmentation du thalamus sera supprimée, ce qui permettra de segmenter le noyau caudé avant le thalamus.

Comparaison de la saillance de la segmentation :

La comparaison des histogrammes de saillance est une mesure indicative, car elle ne compare pas une caractéristique particulière de la structure. La distribution de saillance apprise pour une structure nous donne plutôt une évaluation de l'aspect visuel d'une structure au sens de la saillance. En ce sens, cela permet de gérer la variabilité naturelle des structures anatomiques. Mais cela ne garantit pas de détecter un problème de segmentation, en particulier cela ne nous donne pas une évaluation de la précision de la segmentation. Cependant, cela nous permet d'obtenir une attache aux données.

Pour cela, nous avons besoin d'un histogramme de saillance défini à partir du résultat de la segmentation. Cet histogramme est défini ainsi :

$$saillance_{seg} = \top(seg_v, sal_I) , \quad (5.17)$$

où sal_I est la carte de saillance de l'image que nous explorons et seg_v représente la carte binaire de la segmentation. L'histogramme de cette zone sera défini avec le même niveau de quantification que précédemment ($N = 100$) de cette manière :

$$H_v[i] = \sum_{x \in S} \mathbb{1}(saillance_{seg}(x)) ,$$

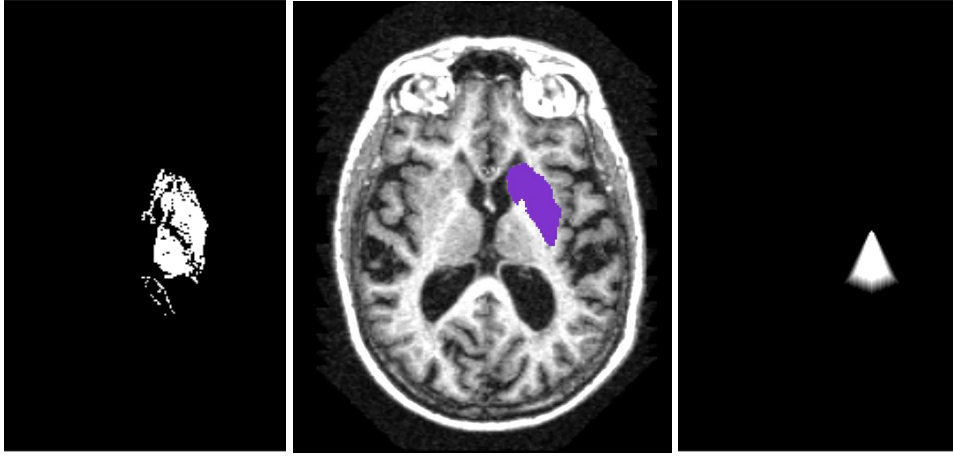


FIG. 5.17 – Illustration d’un problème de segmentation détecté grâce à la cohérence spatiale. L’estimation de la radiométrie (la carte binaire obtenue est à gauche) n’a pas permis de séparer le thalamus du noyau caudé et du putamen, et ce sont ces dernières structures qui ont été segmentées et reconnues comme étant le thalamus. La comparaison des distributions de saillance ne permet pas de détecter cette erreur. La segmentation du noyau caudé par la suite est limitée à un point. Les relations spatiales issues du noyau caudé peuvent tout de même être représentées (à droite). Le calcul de la satisfaction entre l’ensemble flou des relations issues du noyau caudé et la segmentation du thalamus est de 0 et permet donc de détecter le problème. Dans ce cas, la segmentation du thalamus doit être supprimée pour contraindre le processus à segmenter d’autres structures avant celle-ci.

où $1(\cdot)$ représente la fonction indicatrice.

Nous pouvons ensuite comparer la distribution de saillance obtenue à la distribution apprise pour le modèle à l’aide d’une mesure EMD (eq 5.1). Le modèle fournit la distribution moyenne, mais également la moyenne des distances EMD à cette distribution, ainsi que la variance de cette mesure. Nous pouvons donc centrer et réduire les valeurs. La distance entre les distributions de probabilité de la segmentation et du modèle d’une structure o est donc définie ainsi :

$$d_o(seg_o, mod_o) = \frac{EMD(seg_o, mod_o) - \hat{mod}_o}{\sigma_{mod_o}} .$$

où seg_o représente la distribution de saillance issue de la segmentation, mod_o la distribution apprise pour cette structure, $\hat{mod}_o$ la moyenne des distances EMD entre chaque cas de la base et la distribution moyenne pour cette structure, et σ_{mod_o} l’écart-type de ces distances.

Les données étant centrées et réduites, nous pouvons fixer un unique seuil pour toutes les structures $T = 2\sigma_{mod_o}$, considérant qu’un écart supérieur à deux fois l’écart type de la distribution n’est plus acceptable.

Décision :

Dans le cas où nous avons une segmentation, nous avons donc deux critères. Nous allons commencer par regarder si la cohérence spatiale est respectée. Si ce n’est pas le cas, alors nous supprimons la structure parente considérée comme responsable de l’incohérence constatée. La cohérence est mesurée uniquement sur l’arc entre la structure et sa structure parente, si elle existe. Si elle n’existe pas, alors la cohérence n’est pas prise en compte.

Nous regardons ensuite la distance entre les distributions de saillance, à l'aide du seuil que nous avons défini. Si la distance est supérieure au seuil, alors la segmentation est refusée et la structure ne pourra être segmentée de nouveau sans informations supplémentaires.

Enfin, si les distributions sont suffisamment proches, alors la segmentation est acceptée et le graphe peut être mis à jour avec cette segmentation.

5.2.8.3 Échec de la segmentation d'une structure

Si une segmentation échoue, et qu'il n'y a pas de structure parente à blâmer pour cela, alors il est possible que la structure concernée ne puisse être segmentée de nouveau qu'avec les mêmes conditions, c'est-à-dire en suivant le même chemin. Dans ce cas, la segmentation de cette structure a échoué, et il n'est plus possible de la segmenter.

5.2.8.4 Structure de contrôle

Le fait de supprimer des segmentations et de contraindre le modèle à changer de chemin peut faire boucler le processus. Par exemple, si nous segmentons des structures dans cet ordre : thalamus, noyau caudé, putamen, si la segmentation du putamen échoue, alors la segmentation du noyau caudé, sa structure parente peut être supprimée. Le processus peut alors être conduit à segmenter les structures dans cet ordre : thalamus, putamen, noyau caudé. Mais si la segmentation du noyau caudé échoue, alors la première solution sera tentée de nouveau.

Pour éviter ce problème, et pour ne pas essayer de segmenter de nouveau le même chemin, c'est-à-dire de ne pas segmenter de nouveau des structures dans les mêmes conditions, alors nous avons recours à une structure d'arbre. Une structure (anatomique) dans cet arbre aura pour « père » la structure parente que nous avons définie. À chaque nouvelle segmentation effectuée, un nouveau nœud est ajouté dans l'arbre. Si la segmentation n'est pas acceptée, l'arbre reste inchangé, mais la trace de l'échec est reportée dans le nœud correspondant. La racine de cet arbre est un nœud correspondant à toutes les structures de référence.

Outre le fait de garder une trace des segmentations déjà effectuées, nous avons introduit une procédure permettant de détecter un sous-arbre dont toutes les possibilités de chemin sont épuisées, mais qui n'a pas permis d'obtenir une segmentation acceptable de toutes les structures présentes dans ce sous-arbre. Cette procédure nous permet de considérer l'ensemble des chemins de ce sous-arbre comme défaillant et, de cette manière, de supprimer les structures segmentées de cet arbre. Nous pouvons ainsi contraindre le processus à explorer d'autres parties de l'arbre.

Grâce à cette structure, nous pouvons garantir qu'un chemin ne sera pas segmenté deux fois. De plus, nous conservons les résultats des critères d'évaluation de chaque segmentation dans la structure de l'arbre, ce qui permet une évaluation a posteriori des chemins. La figure 5.18 présente un exemple de la structure de l'arbre au cours du processus.

D'une manière plus générale, le processus d'évaluation avec les critères, notre stratégie en cas d'échec et la structure de contrôle nous permettent d'apporter une contribution à un problème inhérent aux segmentations séquentielles, qui, en utilisant l'information recueillie au cours du processus, favorisent la propagation des erreurs. Ainsi, le processus de segmentation est plus robuste aux échecs potentiels et le processus de contrôle permet de les corriger.

5.2.8.5 Mise à jour du graphe

Une fois la segmentation de l'objet validée, il faut à présent mettre à jour le graphe. Tout d'abord, il faut mettre à jour le nœud représentant la structure qui vient d'être segmentée. Le

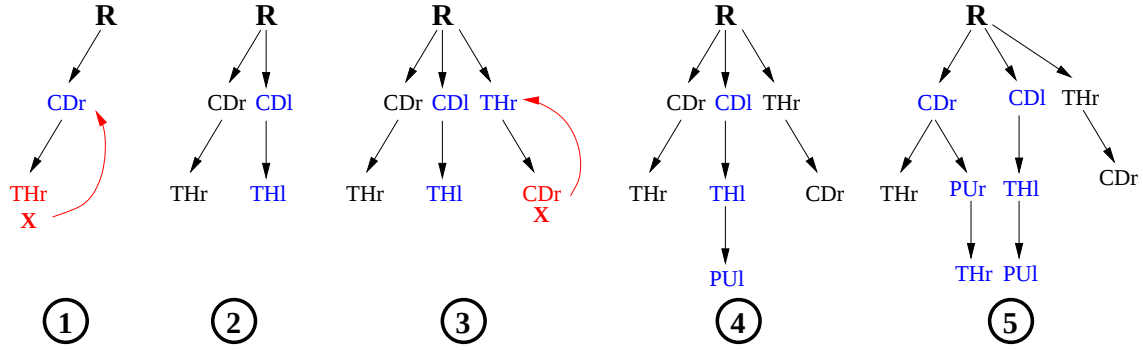


FIG. 5.18 – L’arbre de contrôle au cours du processus. Les structures notées en bleu sont les structures segmentées présentes dans l’arbre. Les structures en rouge indiquent une segmentation qui a échoué. Les structures en noir représentent les structures dont la segmentation a échoué ou dont la segmentation a été supprimée. 1) Les deux premières étapes (à gauche) : le noyau caudé droit a été segmenté, puis la segmentation du thalamus échoue à cause de la cohérence spatiale, ce qui provoque la suppression de la segmentation de sa structure parente, le noyau caudé. 2) Au cours des deux étapes suivantes, le noyau caudé gauche puis le thalamus gauche sont correctement segmentés. 3) Le thalamus droit est ensuite segmenté. Il faut noter que le noyau caudé droit n’est pas segmentable à cet instant, mais que la segmentation du thalamus va rendre possible sa segmentation. Mais la segmentation échoue et la segmentation du thalamus est supprimée. 4) Le putamen gauche est segmenté correctement. 5) La première segmentation du noyau caudé droit est rétablie, afin de permettre au processus d’explorer la branche manquante de cet arbre. La segmentation du putamen droit est effectuée correctement, puis le thalamus droit finalement.

nœud ne représente donc plus uniquement la connaissance générique mais contient également l’information de l’image.

Une fois le nœud mis à jour, nous pouvons mettre à jour les arcs issus de ce nœud. Sur chacun de ces arcs, nous générons les représentations des relations spatiales portées par cet arc. Ces représentations seront utilisées ensuite pour calculer les localisations des structures voisines (ou uniquement les préciser si elles étaient déjà connectées à un nœud précédemment segmenté. La figure 5.19 (à droite) montre comment la localisation d’une structure (le thalamus) est précisée après segmentation du noyau caudé.

Il est également nécessaire de mettre à jour les nœuds visés par ces arcs. En effet, à l’étape d’évaluation de la segmentation, nous pouvons être amenés à contraindre le modèle à ne pas segmenter de nouveau une structure tant qu’une nouvelle information (spatiale) n’est pas disponible. Si une segmentation est acceptée, alors cela constitue une information nouvelle pour les nœuds voisins. Ils peuvent donc être segmentés à nouveau, et la restriction est levée.

Enfin, il est nécessaire de mettre à jour les ensembles de nœuds utilisés au cours du processus : l’ensemble des nœuds segmentés V_{seg} reçoit le nœud $\hat{v}$ et l’ensemble des nœuds objectifs V_{obj} est privé de ce nœud. Dans la continuation de notre exemple, le graphe de la figure 5.9 mis à jour est illustré par la figure 5.19 à gauche.

Les ensembles V_{fs} et V_{fo} sont également mis à jour. D’un côté, tous les nœuds de V_{fs} qui ne sont plus connectés à au moins une structure non segmentée (dans l’ensemble V_{obj}) sont supprimés de cet ensemble. D’un autre côté, il faut ajouter dans V_{fo} tous les nœuds de V_{obj} qui n’étaient pas déjà dans cet ensemble et qui sont à présent reliés à un nœud segmenté. L’ensemble des arcs E_f est mis à jour à partir des ensembles V_{fs} et V_{fo} .

L’exploration de la scène consiste donc à sélectionner séquentiellement les emplacements pré-

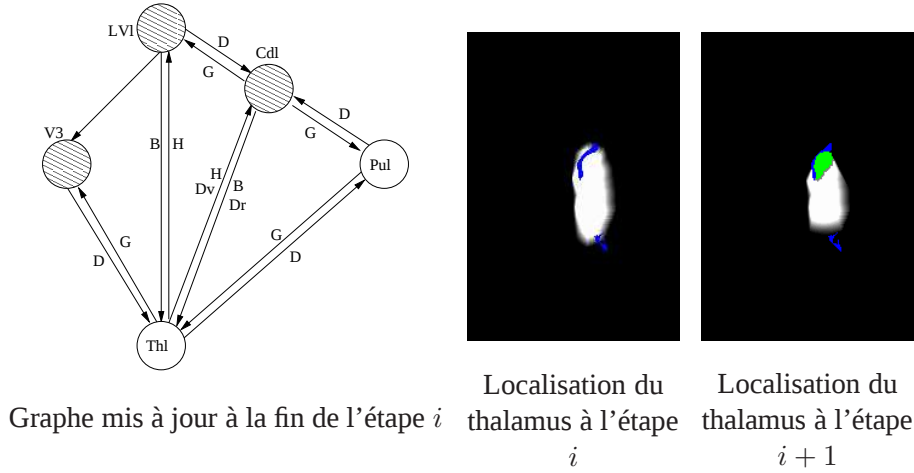


FIG. 5.19 – Mise à jour du graphe. Après segmentation du noyau caudé, il est ajouté dans les structures segmentées (les structures grisées dans le graphe à gauche). Le thalamus est toujours dans l'ensemble des nœuds V_{fo} candidats à une segmentation. Le putamen, qui n'était relié à aucune structure segmentée, est maintenant placé dans cet ensemble. La mise à jour du graphe permet de prendre en compte l'information recueillie sur l'image au cours du processus. À gauche, la localisation du thalamus à la première étape est estimée à partir des relations au ventricule et au troisième ventricule. À droite, la localisation du thalamus mise à jour après segmentation du noyau caudé (en vert) est plus précise (c'est-à-dire moins étendue) grâce à la prise en compte des relations spatiales entre le noyau caudé et le thalamus dans ce cas.

sentant les meilleures saillances au sens du critère retenu, cette sélection permet la segmentation et la reconnaissance immédiate d'un objet du modèle générique (l'objet segmenté étant identifié). Le graphe, qui ne porte au départ qu'une connaissance générique est donc progressivement spécialisé avec l'information de l'image qui est segmentée. Cette approche ne dépend pas d'une représentation des objets que nous devons reconnaître, comme c'était le cas de la première approche présentée. Enfin, cette approche nous permet de directement prendre en compte de l'information provenant de l'image à segmenter, et donc une meilleure adaptation, plutôt que de compter sur une exhaustivité du modèle.

5.3 Expériences

Nous avons effectué la segmentation des images dans le cas sain. Nous allons d'abord détailler le déroulement du processus pour un volume particulier. Nous présentons ensuite les résultats sur un ensemble d'images en nous intéressant tout d'abord aux différentes séquences de segmentation obtenues, puis nous présentons des résultats de segmentation sur la base de cas sains.

5.3.1 Déroulement du processus

Toutes les illustrations sont des coupes extraites des volumes résultats, mais tous les calculs sont effectués en trois dimensions. Les figures 5.20, 5.21 et 5.22 présentent les différentes étapes du processus. Nous présentons pour chaque image la même coupe extraite des volumes (100). La séquence de segmentation suivie est la suivante :

Noyau caudé droit
 Thalamus droit
 Putamen droit
 Thalamus gauche
 Noyau caudé gauche
 Putamen gauche

Nous pouvons remarquer que le chemin suivi dans l'hémisphère droit n'est pas le même que le chemin suivi dans l'hémisphère gauche. Cependant, comme le montrent les histogrammes de saillance présentés dans la figure 5.20, les localisations entre le noyau caudé et le thalamus produisent des distributions de saillance très proches. Le choix de l'un ou l'autre repose donc sur de petites différences. La raison principale pour laquelle les localisations entre ces deux structures sont proches est qu'elles sont en grande partie confondues car elles reposent principalement toutes les deux sur les relations issues de la même structure (le ventricule), qui est grand par rapport aux deux structures.

La figure 5.20 présente la première étape qui débouche sur la segmentation du noyau caudé droit. En haut de la figure, nous pouvons voir le graphe initial. Les nœuds candidats à la segmentation V_{fo} sont représentés en vert, et les structures de référence en bleu (V_{fs} à la première étape). Les structures de référence apparaissent également en bleu sur l'image à gauche du graphe. La connaissance spatiale utilisée à cette étape est portée par les arcs représentés en rouge.

Les localisations des quatre structures candidates sont générées. Elles sont représentées sur la deuxième ligne de la figure. Les structures de référence sont toujours représentées en bleu. La structure dont nous calculons la localisation a été ajoutée en vert sur la localisation afin de permettre une estimation de la précision de cette localisation.

Les histogrammes de saillance et les histogrammes cumulés correspondants sont ensuite calculés. La structure à segmenter est sélectionnée d'après le critère de saillance. La segmentation obtenue est présentée en rouge sur l'image en bas à gauche de la figure. Cette segmentation est correcte. Le graphe est alors mis à jour. Le noyau caudé droit est ajouté dans V_{fs} , le putamen est ajouté dans V_{fo} . L'ensemble d'arcs E_f est mis à jour en supprimant l'arc entre le ventricule et le noyau caudé à droite, et en ajoutant l'arc entre le noyau caudé et le putamen.

Les figures 5.21 et 5.22 présentent les étapes suivantes du processus. Pour chacune, nous présentons trois éléments : tout d'abord les localisations calculées ou mises à jour. Celles qui sont identiques à l'étape précédente, si aucune nouvelle information n'est intervenue, ne sont pas reportées ; ensuite, la segmentation effectuée à cette étape ; et enfin le graphe mis à jour.

Enfin, la segmentation finale obtenue est présentée en bas de la figure 5.22, et dans deux vues différentes. La segmentation des noyaux caudés est bonne. La segmentation des thalamus est presque correcte. Il manque un morceau du thalamus de gauche. La segmentation des putamens est moins correcte. Il manque dans les deux cas la queue du putamen, qui est assez fine et difficile à obtenir.

5.3.2 Les séquences de segmentation

La figure 5.23 présente de manière synthétique les différentes séquences de segmentation obtenues pour les cas sains de notre base. Nous avons séparé les structures de chaque côté de l'hémisphère pour ne laisser, de chaque côté, que quatre chemins possibles. La figure reflète le nombre d'occurrences de chaque chemin.

Un chemin en particulier apparaît dans la majorité des cas, il s'agit du chemin qui a été défini de manière ad hoc par Colliot (2003), ce qui montre la pertinence de ce choix. Dans cette

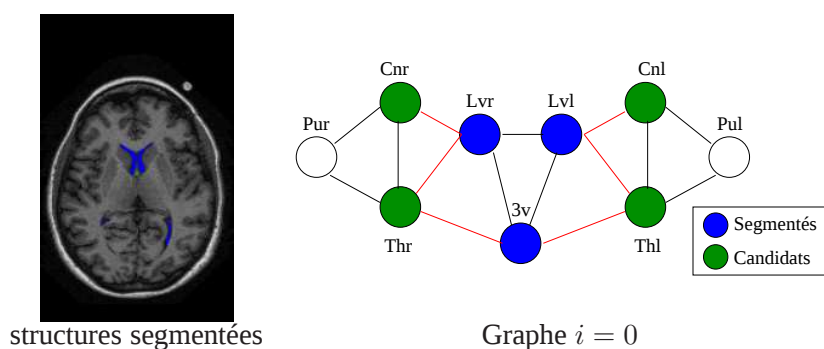
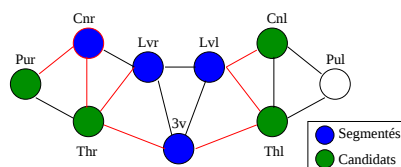
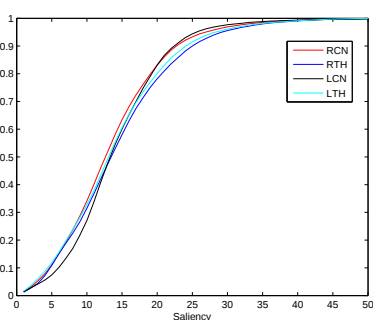
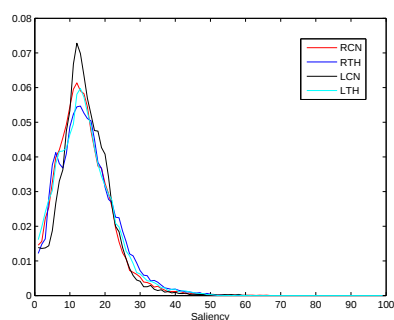
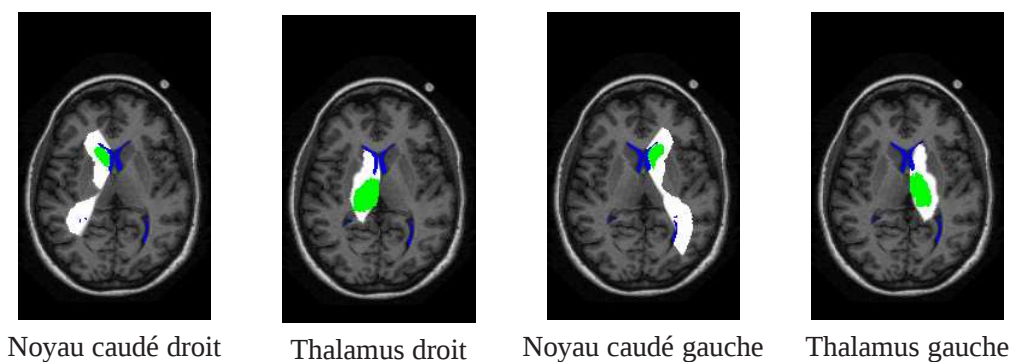
Première étape :**Localisations :**

FIG. 5.20 – Première étape du processus de segmentation séquentielle. Le graphe initial est présenté en haut. Les arcs utilisés à cette étape sont en rouge. Les localisations des structures candidates sont présentées en-dessous, l'ensemble flou correspondant apparaît en blanc, et la structure correspondante à été ajoutée (en vert). Les histogrammes de saillance sont très proches, mais les localisations se chevauchent en grande partie. La structure segmentée est le noyau caudé droit (en rouge). Le graphe mis à jour est présenté en bas. Le putamen droit est ajouté aux structures candidates.

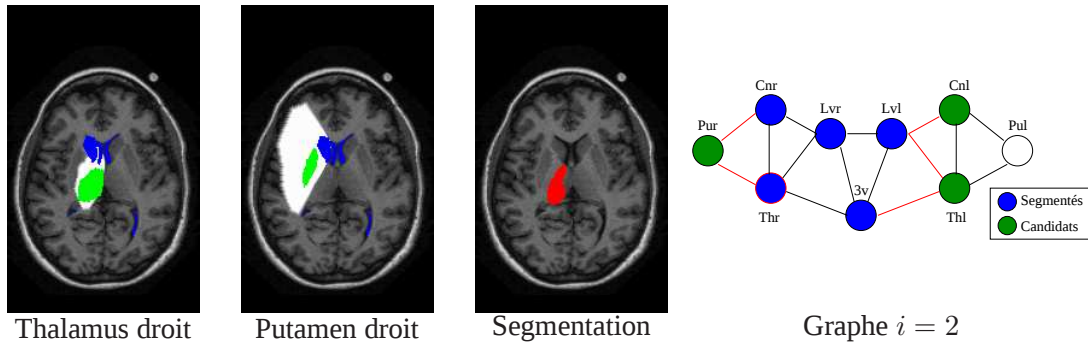
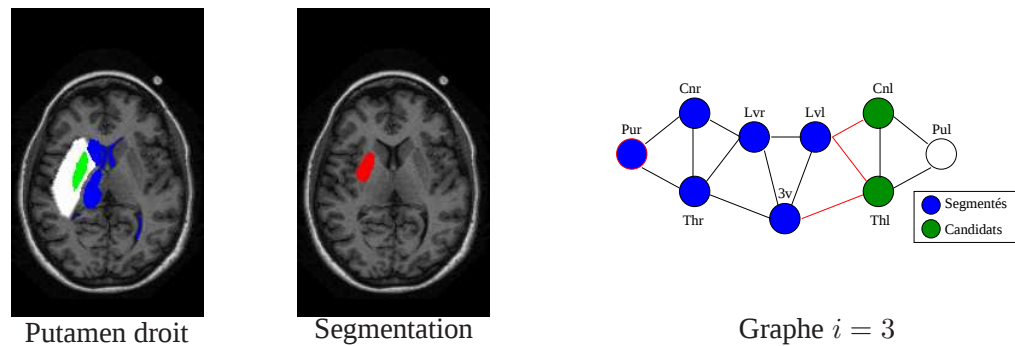
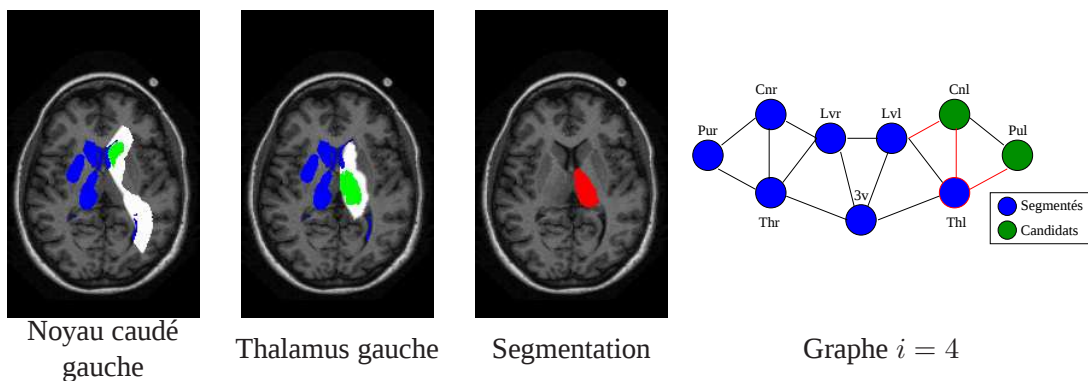
Deuxième étape :**Troisième étape :****Quatrième étape :**

FIG. 5.21 – Les étapes 2 à 4 du processus. Seuls les ensembles flous des localisations mises à jour sont présentés. Le graphe mis à jour à chaque étape est présenté sur la droite. Les structures segmentées sont le thalamus droit, puis le putamen droit et le thalamus gauche.

séquence de segmentation, la première structure reconnue est le noyau caudé (en utilisant les relations spatiales issues du ventricule latéral). La deuxième structure est le thalamus, qui profite de l'information spatiale provenant de 3 structures, le ventricule, le troisième ventricule et le noyau caudé. Enfin, le putamen, qui profite des relations issues des deux structures déjà segmentées : le noyau caudé et le thalamus. Le deuxième chemin est proche, il y a juste une inversion entre le noyau caudé et le thalamus. A chaque étape, avec ces deux chemins, l'information spatiale utilisée provient d'au moins deux structures.

Cela n'est pas le cas avec les deux autres chemins, qui sont également beaucoup moins fré-

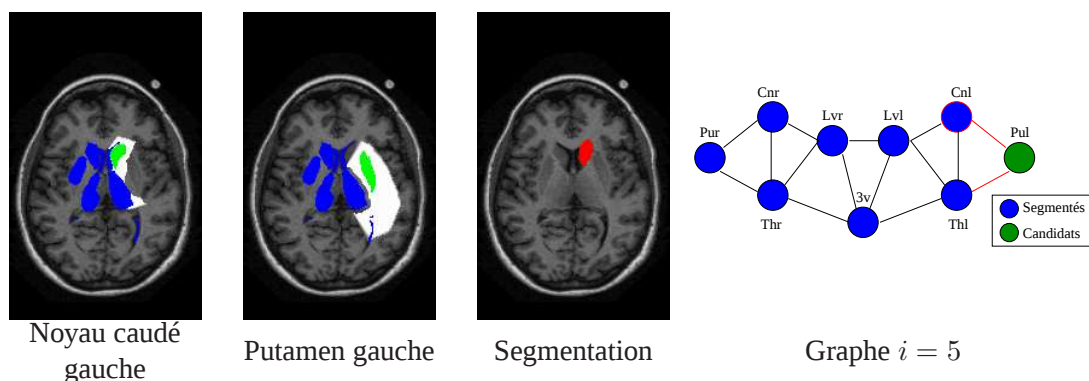
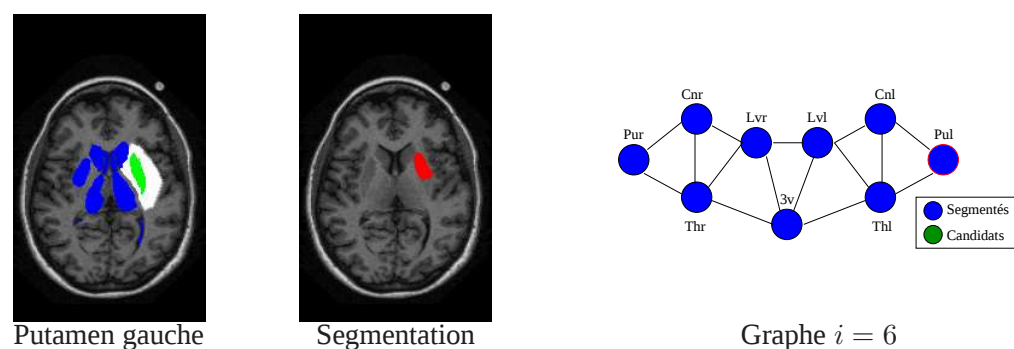
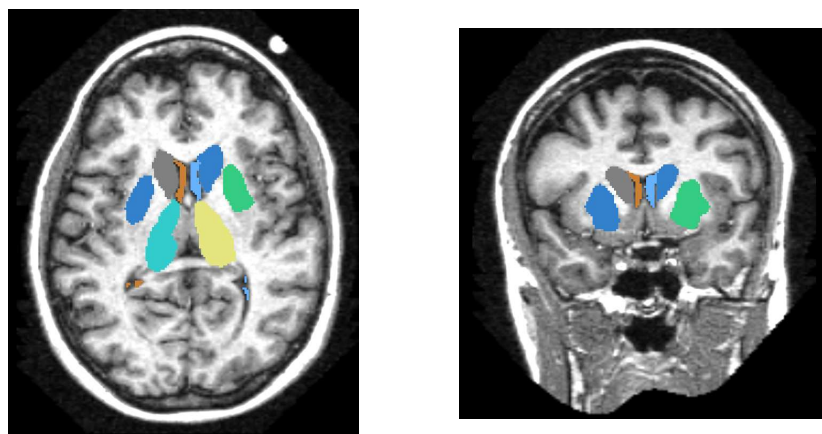
Cinquième étape :**Sixième étape :****Segmentation finale :**

FIG. 5.22 – Les deux dernières étapes (5 et 6). Les structures segmentées sont le noyau caudé gauche et le putamen gauche. La segmentation de l'image est présentée en bas dans deux vues différentes : axiale et coronale.

quents. Dans ce cas, la segmentation du putamen, en deuxième position, s'effectue avec de l'information spatiale issue d'une unique structure. Plus une localisation utilise d'information provenant de différentes structures, plus la disjonction va réduire la localisation. Dans notre cas, cela signifie que la localisation est plus précise. Les résultats montrent que les chemins plus précis sont privilégiés aux autres chemins. Cet effet n'est pas directement lié au critère de saillance. Cependant, il est préférable d'avoir des chemins plus précis, ce résultat est donc satisfaisant.

Plus généralement, nous avons une certaine variabilité dans les chemins suivis. Il y a également une variabilité entre les deux hémisphères, la première structure segmentée étant sélectionnée dans un côté ou un autre avec une fréquence similaire. Il y a principalement deux facteurs pour la sélection des structures :

- la morphologie des structures de référence, utilisées pour calculer les localisations des premières structures ;
- et la saillance de ces localisations.

Dans les deux cas, nous tenons compte des informations de l'image pour effectuer le choix, ce qui était un objectif de cette approche.

Changements de chemin au cours du processus :

Nous présentons dans les tableaux 5.6 et 5.5 les occurrences où le processus a détecté un problème au cours du processus, et a dû changer de chemin. La détection d'un problème dans l'image ne signifie pas que la segmentation finale sera erronée, mais uniquement que le chemin initial n'a pas permis d'effectuer la segmentation complète, et qu'il a été nécessaire de l'adapter au cours du processus.

Le premier tableau donne la répartition du type de problème détecté, identifié par le critère correspondant. Les chiffres proviennent de la segmentation des 30 cas sains de la base. Au cours de l'ensemble des processus de segmentation de ces images, 195 segmentations ont été initialement acceptées, alors que 38 segmentations ne l'ont pas été. Parmi les segmentations acceptées, certaines seront supprimées a posteriori si elles sont désignées responsable de l'échec d'une segmentation ultérieure sur laquelle elles ont une influence. Dans la grande majorité des cas, c'est le critère de cohérence spatiale qui a été utilisé pour rejeter une segmentation. Cependant ce critère est le premier critère testé et s'il n'est pas satisfait, alors le critère de saillance n'est pas testé. Ce résultat montre tout de même la pertinence de ce critère. Le critère sur les distributions de saillance est ensuite peu utilisé.

TAB. 5.5 – La répartition des problèmes détectés au cours du processus et qui ont mené à un changement de chemin. Dans la grande majorité des cas, c'est le critère de cohérence spatiale qui a détecté le problème. Le critère de saillance n'est presque jamais utilisé. Il y a peu de cas où aucune segmentation n'est produite.

segmentation initialement acceptée		195
segmentation refusée	critère de saillance	2
	pas de segmentation	5
	cohérence spatiale	31

Le deuxième tableau indique une répartition des images en fonction du nombre de changements de chemin effectués, et cela en différenciant les images de la base IBSR et celles de la base OASIS. Pour la plupart des images, il n'y a pas ou peu (1) de changements de chemin nécessaires. Pour certaines images, le chemin nécessite plus d'adaptations. Le nombre de changements effectués permet ainsi de mesurer la difficulté de segmentation d'une image en particulier, sans en donner les raisons de manière explicite. Les résultats confirment la difficulté de segmentation de la base IBSR par rapport à la base OASIS.

5.3.3 Les résultats de segmentation

Nous présentons dans les figures 5.24 et 5.25 les résultats de la segmentation sur les images de cas sains de notre base. Comme précédemment, les expériences sont réalisées sur les volumes

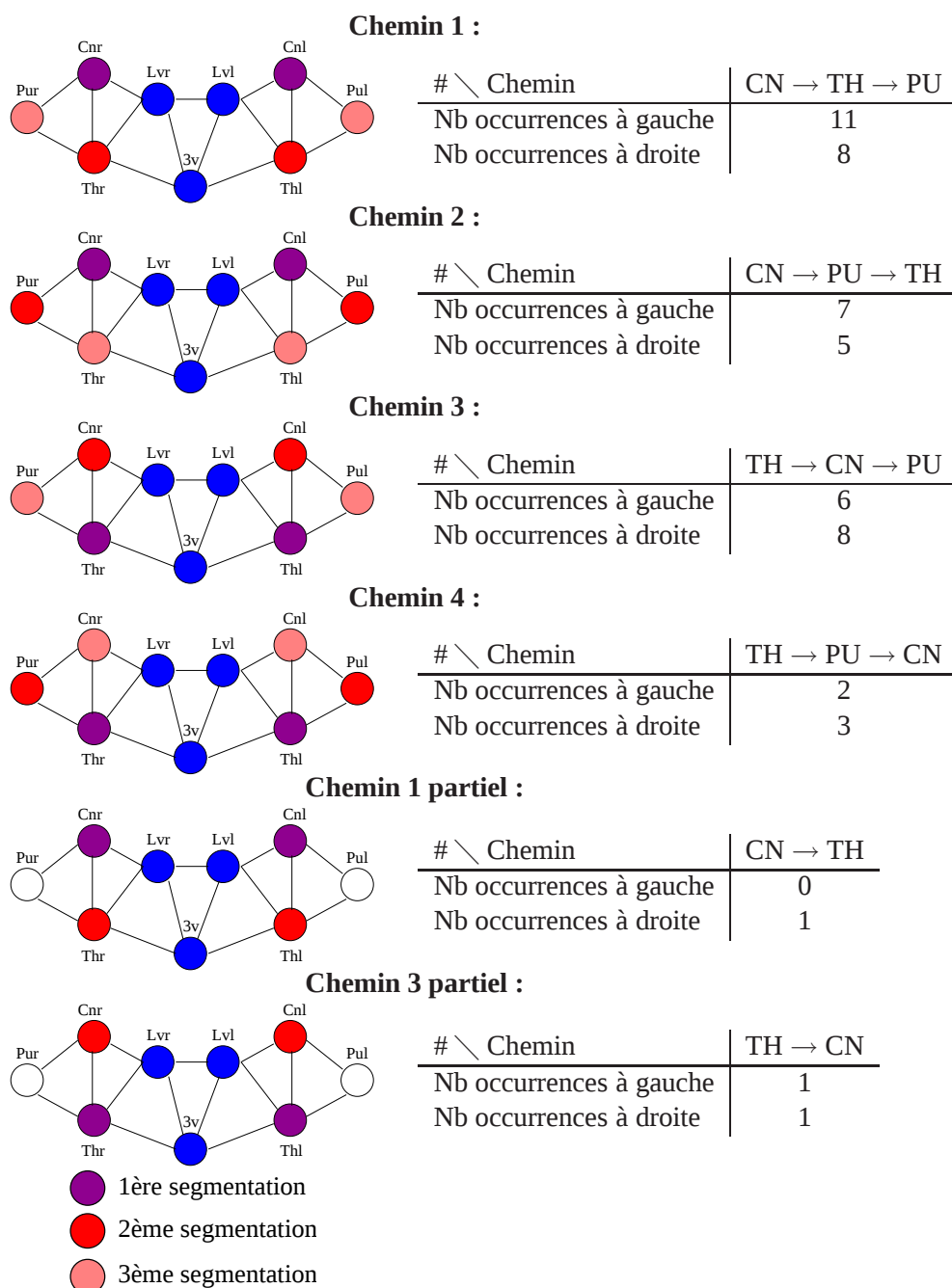


FIG. 5.23 – Les chemins de segmentation présentés de manière synthétique. Ces schémas ne reflètent que les occurrences de chaque chemin, dans chaque hémisphère. Mais le chemin suivi dans l'hémisphère droit et celui suivi dans l'hémisphère gauche peuvent être différents. Les chemins les plus fréquents sont les chemins privilégiant les structures proches des structures de référence, qui permettent d'utiliser au mieux l'information spatiale. Lorsque le putamen est segmenté en deuxième position, les relations spatiales qui permettent sa localisation ne sont issues que d'une seule structure. Le chemin le plus utilisé est le chemin ad-hoc qui était utilisé précédemment par Colliot (2003).

TAB. 5.6 – Répartition des images en fonction du nombre de changements de chemin effectués par le processus au cours de leur segmentation. Sur la plupart des images, il y a peu de changements (0 ou 1). D'autres images nécessitent plus d'adaptations au cours du processus.

	IBSR	OASIS
Aucun changement	6 (35%)	7 (64%)
1 changement	2	2
2 changements	5	2
3 changements et plus	4	0
Total	17	11

en trois dimensions, mais seule une coupe est présentée ici. Le temps de calcul pour le processus complet, sans changement de chemin, est de l'ordre de 75 minutes sur une machine récente, la calcul de la carte de saillance étant effectué à part. La grande majorité de ce temps est pris par le processus de segmentation d'une structure. Le calcul des paysages flous est également coûteux, mais l'utilisation d'une approximation de ces paysages permet de réduire le temps de calcul à environ 30 secondes par paysage.

Sur ces images, les structures de référence ont été indiquées en bleu clair, les noyaux caudés sont en jaune, les thalamus en magenta et les putamens en bleu foncé. Pour simplifier, nous avons attribué les mêmes couleurs aux structures des deux côtés de l'hémisphère (la même couleur pour les deux noyaux caudés par exemple). Sur la plupart des images, les structures ont été correctement reconnues, même si la segmentation est parfois imprécise. D'autres images présentent des structures manquantes, ou des structures qui n'ont pas été reconnues correctement, c'est-à-dire qu'elles ont été segmentées mais que leur étiquette n'est pas correcte. C'est le cas par exemple pour l'image en haut au centre de la figure 5.25 où le thalamus droit (à gauche sur l'image) a une étiquette correspondant au noyau caudé droit. Ce cas de figure se présente sur plusieurs images. La prise en compte de la cohérence spatiale a néanmoins permis de diminuer ce type d'erreur. Le seuil de la cohérence spatiale a été fixé relativement bas (0, 5), ce qui explique que ce type d'erreur puisse se produire encore.

Les segmentations sur la base IBSR sont moins correctes en général. Les images de cette base ont été recalées et les images sont souvent floues. Lorsque les frontières des structures sont fines, comme c'est souvent le cas pour les structures sous-corticales, ce flou rend le problème de la segmentation plus difficile.

Mauvaise reconnaissance :

Une mauvaise reconnaissance est une conséquence d'une localisation imprécise de la structure (si le thalamus est inclus dans la localisation du noyau caudé dans notre exemple), conjointement avec une mauvaise estimation de la radiométrie des différentes structures qui ne permet pas de les différencier (du noyau caudé dans ce cas), c'est-à-dire que les valeurs μ_s et σ_s qui estiment la radiométrie de la structure s ne sont pas adéquates. Nous avons montré dans une partie précédente que les paramètres α et β utilisés pour estimer les valeurs μ_s et σ_s étaient une moyenne pour α et un maximum pour β de valeurs relativement dispersées. Ce genre d'imprécision n'est donc pas imprévu.

Imprécision des segmentations :

Nous retrouvons les problèmes de segmentation du putamen déjà évoqués dans la présentation du déroulement complet du processus. Le putamen est une structure qui s'étire et dont la pointe est

difficile à récupérer lors de la segmentation. En particulier lors du seuillage de la région d'intérêt, l'effet de volume partiel sépare le corps du putamen de la queue, ce qui empêche une bonne segmentation. La forme de la pointe est ensuite difficile à récupérer avec un modèle déformable. Un autre problème avec le putamen se pose dans les coupes basses, où il se confond avec la matière grise environnante.

Il y a d'autres imprécisions pour la segmentation du thalamus, dont les contours ne sont presque pas visibles dans certaines coupes, sa radiométrie se confondant avec celle de la matière blanche. Dans les deux cas, avec le putamen ou avec le thalamus, le problème se pose au niveau de la segmentation initiale. Le modèle déformable permet de récupérer une meilleure segmentation, si la solution initiale est suffisante.

5.3.4 Résultats dans les cas pathologiques

Les expériences réalisées avec des images présentant des pathologies ont été effectuées dans des conditions similaires à celles des expériences précédentes. En particulier, l'apprentissage des relations spatiales est effectué sur la même base d'apprentissage, contenant des images saines ou pathologiques. Les seuils utilisés dans nos expériences sont également les mêmes. La connaissance de la pathologie n'est donc pas utilisée ici. Nous avons toutefois effectué un apprentissage particulier des informations a priori radiométriques pour la base de cas pathologiques. Cependant, si les bases IBSR et OASIS sont relativement homogènes, la base de cas pathologiques l'est moins, et la moyenne des valeurs est donc moins pertinente.

La figure 5.26 présente quelques résultats de segmentation dans des cas pathologiques. Si les putamens sont souvent manquants, les noyaux caudés et les thalamus sont par contre reconnus correctement dans la plupart de ces cas. Le putamen, par sa position et sa forme allongée, est une structure plus sensible aux déformations que d'autres structures.

Les trois cas sur la première ligne de la figure 5.26 présentent de fortes déformations des structures. Dans ces cas, les structures qui sont moins touchées ont pu être reconnues, alors que les structures les plus déformées ne le sont pas. Ces résultats nous donnent une piste afin de détecter la présence d'une pathologie dans le modèle. Si une pathologie est détectée, alors nous pouvons la segmenter de manière indépendante grâce aux travaux de thèse de H. Khotanlou ([Khotanlou \(2008\)](#)) et ajouter un nœud correspondant dans le modèle, relié aux autres structures. Toutefois, une estimation de la déformation et de l'impact sur les relations spatiales environnantes serait nécessaire pour adapter le modèle.

5.4 Conclusion

Nous avons présenté une approche qui intègre dans un processus de segmentation séquentielle un critère fondé sur la saillance de l'image que nous souhaitons segmenter et reconnaître. Cette approche n'a plus besoin des représentations des objets avant qu'ils ne soient segmentés pour procéder à l'optimisation du chemin, permettant une plus grande adaptation à l'image à segmenter. La variabilité dans les chemins de segmentation obtenus montre que nous tenons compte de la variabilité des images dans le processus d'optimisation.

L'approche itérative présente un avantage certain en permettant d'effectuer conjointement la segmentation et la reconnaissance des structures. Cela permet d'exploiter au mieux l'information spatiale du modèle au cours du processus. Nous avons introduit un processus de contrôle de la segmentation séquentielle utilisant l'information spatiale et l'information visuelle qui permet de rendre le processus de segmentation plus robuste aux échecs éventuels et de les corriger.

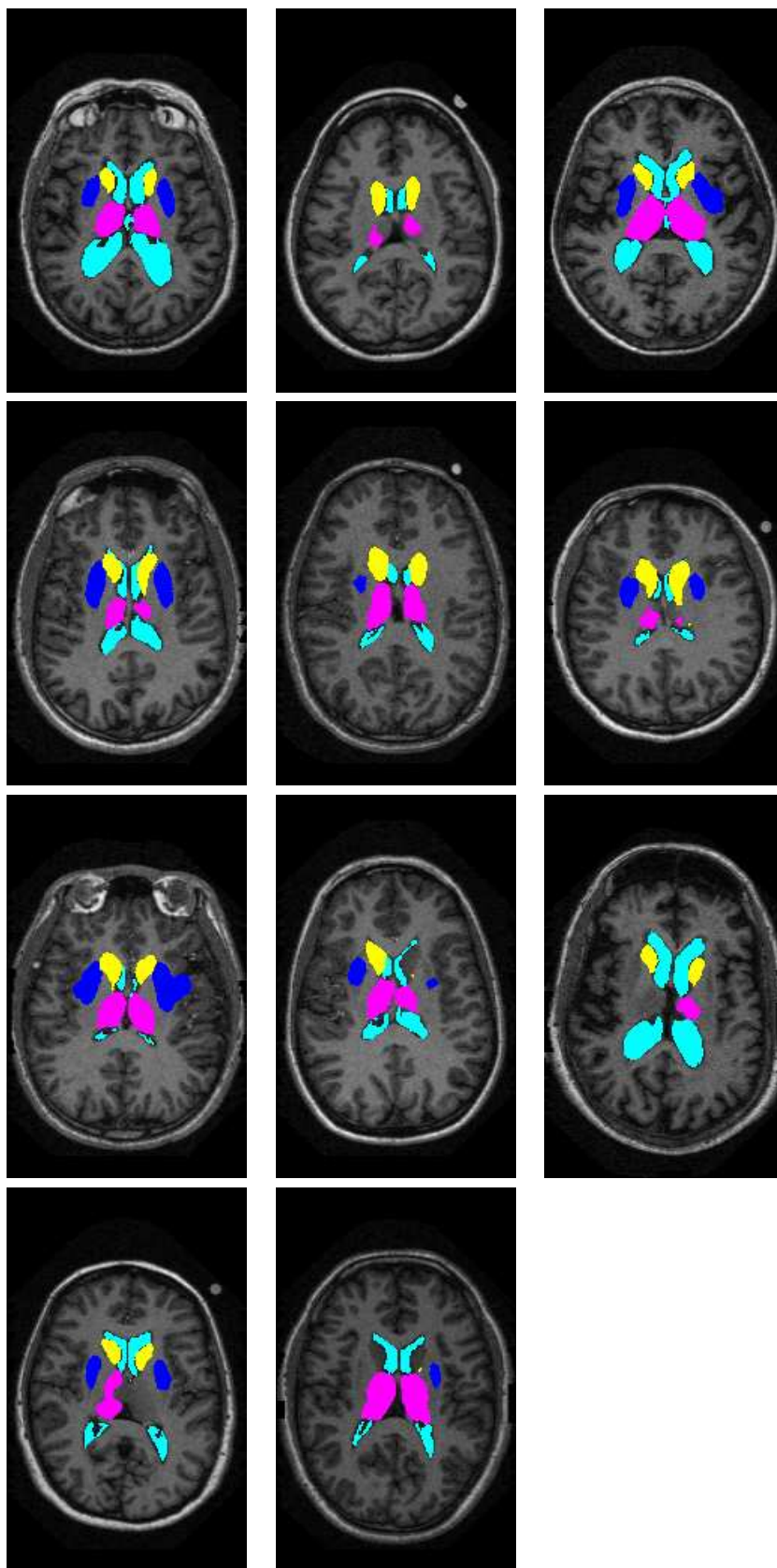


FIG. 5.24 – Résultats de segmentation dans le cas sain sur les images de la base OASIS présentes dans notre base.

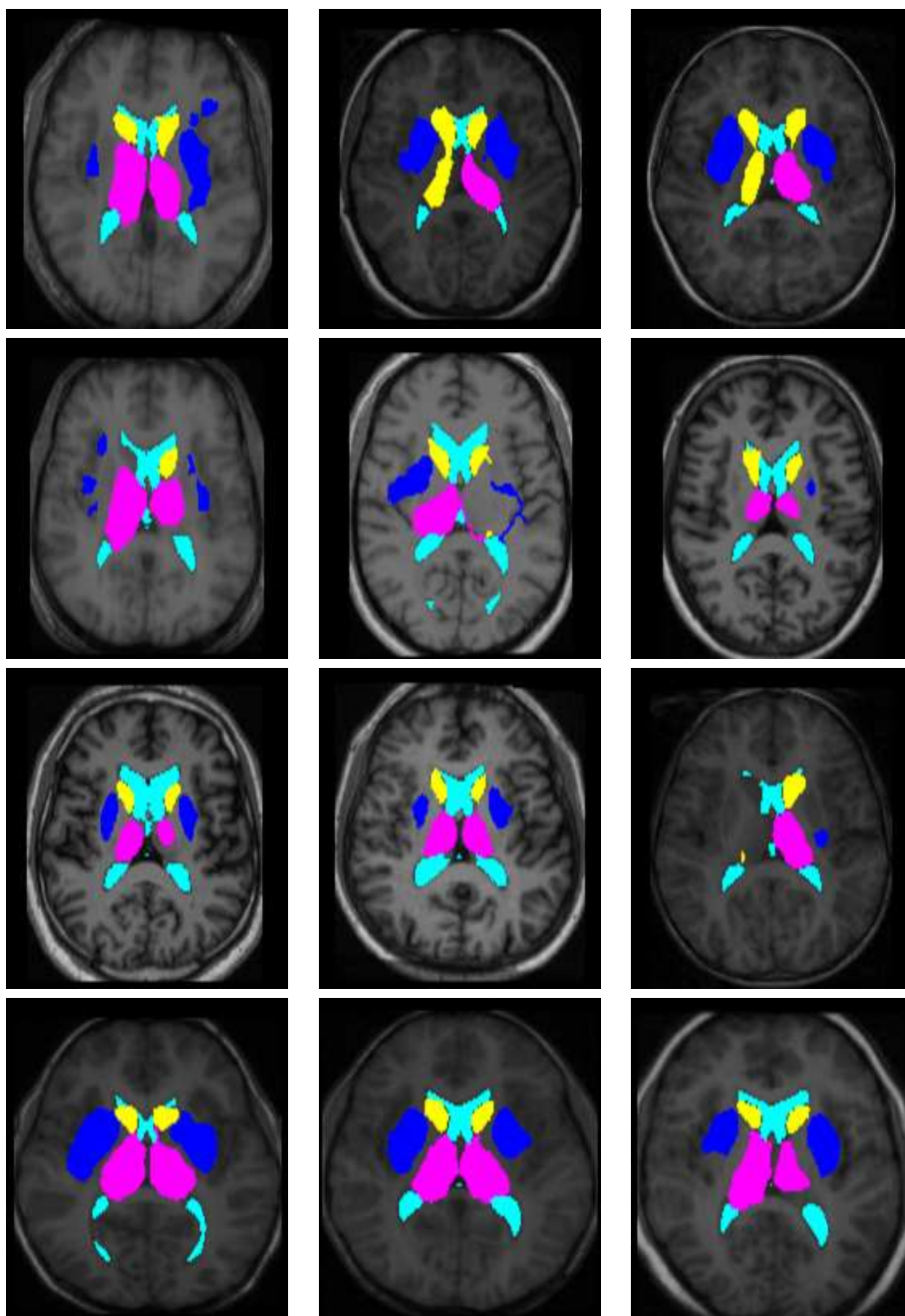


FIG. 5.25 – Résultats de segmentation dans le cas sain sur les images de la base IBSR présentes dans notre base.

Nous avons effectué la segmentation et la reconnaissance des images de notre base. Si la segmentation est parfois imprécise, la reconnaissance des diverses structures est, le plus souvent, correctement effectuée, en particulier grâce au critère de cohérence de l'information spatiale du modèle. Ces résultats montrent l'intérêt d'utiliser l'information spatiale pour segmenter ce type de structure. Les segmentations obtenues ne sont toutefois pas toujours correctes, en particulier nous avons soulevé deux problèmes : la mauvaise reconnaissance d'une structure (identifiée comme une autre structure) et l'imprécision de la segmentation.

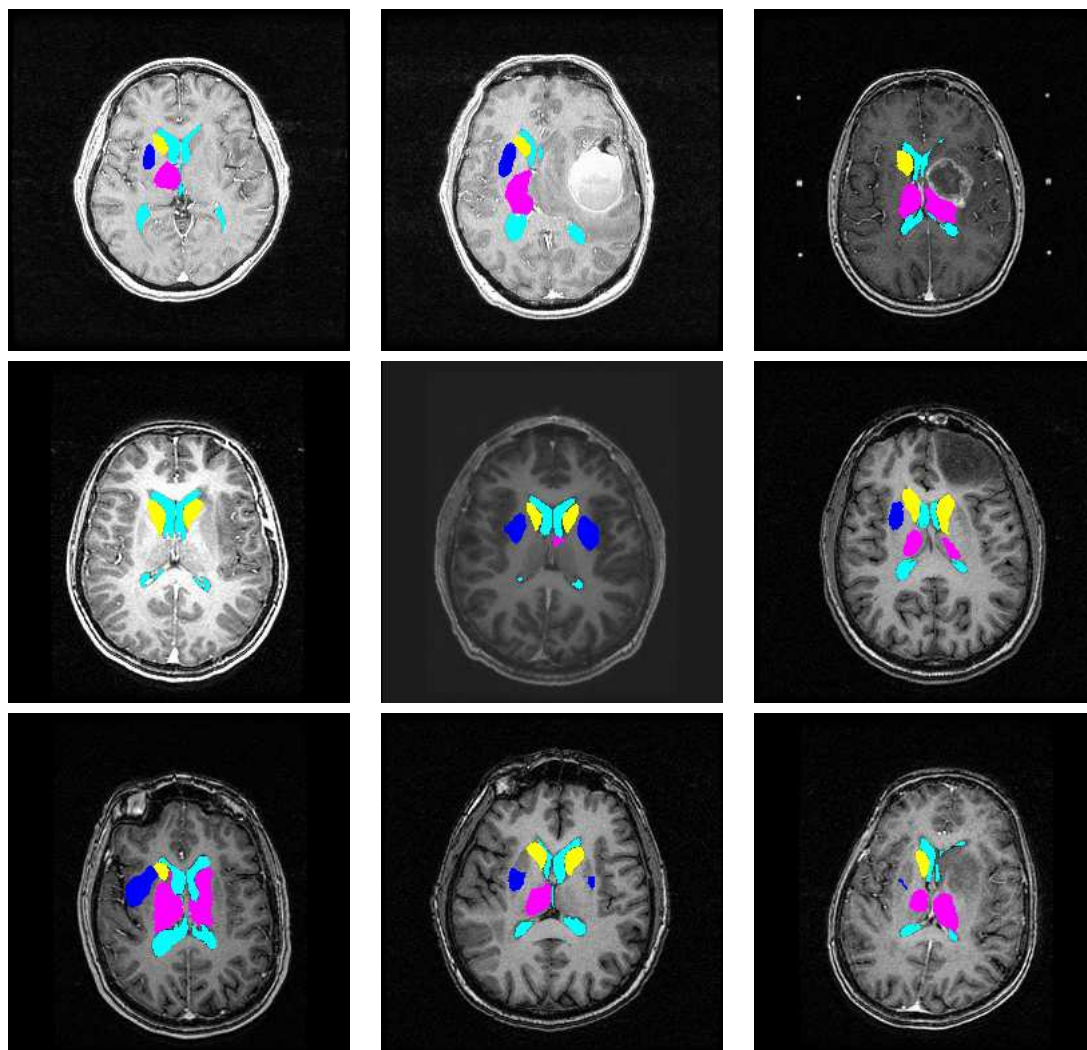


FIG. 5.26 – Résultats de segmentation dans les cas pathologiques. Le contraste des images a été augmenté pour une meilleure visibilité.

Nous avons également effectué la segmentation et la reconnaissance des cas pathologiques. L'apprentissage effectué sur la base prend en compte les cas pathologiques, en particulier dans l'apprentissage des relations spatiales, mais le processus est le même dans les cas normaux et les cas pathologiques. Dans ces cas, il manque des structures, surtout lorsque l'image subit une grande déformation, mais cela nous fournit une piste pour détecter ces cas.

La saillance est issue des travaux sur les mécanismes pré-attentionnels bio-inspirés. L'apport d'un critère fondé sur la saillance est de chercher à détecter ce qui est saillant dans une image, c'est-à-dire ce qui « accroche l'œil » à l'étape pré-attentionnelle. L'apprentissage de la saillance confirme certaines intuitions sur les structures : la visibilité du ventricule, la difficulté de voir des structures comme le thalamus, dont les valeurs sont proches des valeurs de la matière environnante.

L'approche précédente permettait de déterminer le chemin complet avant de commencer les segmentations. Cette approche effectue une optimisation locale uniquement, au sens où uniquement la prochaine structure à reconnaître est choisie à chaque étape. D'un autre côté, la première approche ne permettait pas de prendre en compte l'information issue de l'image, alors que cette approche permet d'intégrer naturellement l'information recueillie au cours du processus.

Chapitre 6

Conclusion et perspectives

6.1 Synthèse des contributions

Nous avons présenté deux types d'approches permettant d'optimiser des chemins de segmentation à partir d'un modèle structurel d'une scène. La première approche utilise l'information spatiale contenue dans le modèle ainsi que des représentations des structures issues d'une base d'apprentissage pour effectuer l'optimisation. Cette approche permet d'effectuer une optimisation complète d'un chemin avant segmentation. La deuxième approche intègre un critère reposant sur la notion de saillance dans un processus de segmentation séquentielle pour optimiser le chemin, permettant de prendre en compte l'information provenant de l'image à segmenter dans le processus.

Nous allons à présent détailler les contributions et discuter chacune de ces deux approches.

6.1.1 Optimisation de chemin avec représentation des structures

Dans cette première partie, nous avons proposé plusieurs approches utilisant l'information spatiale et une représentation des structures pour optimiser une séquence de segmentation. Il s'agit d'une contribution directe de nos travaux et, à notre connaissance, originale.

L'utilisation d'une représentation floue des structures permet de gérer la variabilité normale des structures mais se heurte aux mêmes problèmes de représentativité que les méthodes reposant sur un atlas. L'optimisation proposée consiste à estimer la pertinence des représentations floues des relations spatiales par rapport aux structures qu'elles visent. L'apprentissage des relations spatiales est effectué de manière à gérer la variabilité normale des structures cérébrales. Cela signifie que plus une structure peut varier dans une base, plus la relation spatiale doit être définie d'une manière large. Nous considérons donc que l'ensemble flou fusionnant les relations spatiales portées par un arc est d'autant plus pertinent que sa représentation est proche de la représentation de la structure visée par ces relations spatiales.

La pertinence d'une relation est estimée avec deux types d'approches. La première approche est locale et consiste à évaluer chaque arc de manière séparée, puis à optimiser le chemin sur le graphe obtenu à l'aide de méthodes reposant sur des approches classiques de la théorie des graphes, modifiées pour mieux correspondre à notre problématique. Cette approche est limitée par l'évaluation séparée de chaque arc. L'information spatiale prise en compte pour la segmentation d'une structure provient de toutes les structures du chemin déjà visitées, et pas uniquement de l'arc précédent du chemin. Par exemple, si nous considérons trois structures A , B et C et que toutes ces structures sont reliées entre elles, alors l'évaluation du chemin A, B, C dans notre cadre ne tient

pas compte de l'arc (A, C) , même si l'information qu'il porte est utilisée dans la segmentation de C .

La seconde approche est globale et consiste à représenter un chemin sous la forme d'un unique ensemble flou, permettant d'effectuer l'optimisation en calculant une valeur représentant le chemin. Comme nous l'avons souligné dans le chapitre 4, cette approche pose le problème de la bonne représentation d'un chemin. En particulier dans le cadre des structures sous-corticales considérées, où ces structures sont proches les unes des autres et où des parties de ces structures sont souvent adjacentes (les boîtes englobantes des structures ne sont pas du tout séparées), la représentation d'un chemin est un problème difficile. Il est nécessaire d'avoir suffisamment d'information pour représenter toutes les parties d'un chemin. Il est également nécessaire que la représentation d'un chemin ne couvre pas un espace trop important. Les représentations proposées permettent que la structure cible du chemin soit comprise dans la représentation et de ne pas couvrir trop d'espace, ce qui sont, de notre point de vue, des caractéristiques importantes. Nous avons néanmoins proposé une manière originale de représenter un chemin et d'effectuer son évaluation.

Malgré ces limitations, les approches proposées permettent néanmoins d'optimiser un chemin et de proposer un chemin intuitif. L'exemple proposé dans nos expériences permet ainsi de retrouver le chemin défini de manière ad hoc, ce qui est un bon résultat pour cette approche.

Nous avons proposé une extension de ces approches aux cas présentant des pathologies, en prenant en compte les degrés de stabilité des relations spatiales. L'utilisation de ces degrés permet à notre approche de rester générique en reportant sur la définition des degrés de stabilité la gestion des différents modèles de pathologies, la variabilité des pathologies étant trop importante pour être gérée dans notre modèle. Avec cette extension, nous avons présenté un exemple où nous obtenons un chemin de segmentation adéquat vis-à-vis de la pathologie prise en compte.

6.1.2 Optimisation de chemin avec saillance

Dans une deuxième partie, nous proposons d'optimiser un chemin de segmentation à partir de l'information spatiale contenue dans le modèle et en utilisant une information visuelle, une carte de saillance, pour optimiser le choix des structures à segmenter. L'utilisation des cartes de saillance conjointement avec un modèle de l'agencement spatial d'une scène est une contribution nouvelle de nos travaux. L'optimisation est en réalité effectuée uniquement sur la prochaine structure à segmenter et non pas sur le chemin, ni sur le reste du chemin à parcourir pour atteindre une structure particulière. L'objectif est donc un peu modifié. Il ne s'agit plus ici de déterminer la meilleure séquence de segmentation entre une structure de référence et une structure cible, mais plutôt d'effectuer la segmentation et la reconnaissance de toutes les structures du graphe en suivant une séquence de segmentation optimale. La séquence optimale est déterminée a posteriori à la fin du processus.

Nous avons formulé le problème de la segmentation séquentielle comme une exploration d'une image guidée par l'information structurelle. Cette exploration s'effectue à chaque itération dans un domaine de recherche constitué de zones proches des zones déjà explorées. Cette proximité spatiale est nécessaire afin de pouvoir profiter de l'information spatiale du modèle, les représentations des relations nécessitant une structure de référence pour être générées.

Nous avons également dressé un parallèle entre l'exploration de l'image et l'exploration d'une scène selon un modèle bio-inspiré : un mécanisme pré-attentionnel nous indique la zone à explorer, et le mécanisme attentionnel effectue la reconnaissance de cette zone. Ce dernier est remplacé dans notre système par la segmentation et la reconnaissance d'une structure. Il faut noter que dans ce cas, l'unité attentionnelle s'apparente à un objet plutôt qu'à une zone de l'espace. La comparaison entre l'exploration d'une image par l'œil humain et notre processus de segmentation itérative met

toutefois en évidence une différence fondamentale : nous ne recherchons pas les zones les plus saillantes sur l'image complète, mais uniquement sur une restriction de l'espace au domaine de recherche. Cela signifie que si la zone la plus saillante d'une image n'est jamais incluse dans le domaine de recherche, alors cette zone n'est jamais visitée. Cependant, nous avons montré que nos structures de référence sont parmi les zones les plus saillantes de l'image, et qu'elles constituent un point de départ cohérent pour le processus.

Les cartes de saillance sont un processus bio-inspiré qui, même si il n'est pas psycho-réaliste, cherche à modéliser les mécanismes de l'attention visuelle. Nous avons adapté le processus de génération des cartes de saillance aux images IRM que nous utilisons pour la reconnaissance des structures sous-corticales. Ce faisant, nous avons adopté un autre point de vue en considérant les cartes de saillance comme une manière d'agrégier les indices visuels d'une scène quelconque, et pas forcément selon un observateur de la scène.

L'utilisation de l'information visuelle dans le processus de segmentation séquentielle nous permet d'intégrer de l'information provenant de l'image à reconnaître à différents niveaux. Les segmentations effectuées au cours du processus, qui sont utilisées pour représenter les relations spatiales utilisées dans les itérations suivantes, représentent une information très localisée de l'image. L'information visuelle est, elle, calculée sur des caractéristiques plus globales. L'utilisation d'une information de saillance telle que nous l'avons définie a certaines limites : le processus de génération des cartes de saillance permet d'obtenir une information représentant différentes échelles grâce aux pyramides utilisées. Cependant, dans notre approche, les structures sont petites par rapport à la taille des images. Seules des petites échelles nous apportent donc une information directement reliée à ces structures. L'information obtenue à des échelles plus grossières, où elle est lissée, apporte une information plus générale sur la scène. Une autre limitation est que la comparaison de l'information est effectuée au niveau de la localisation d'une structure. Cette localisation peut être définie par une grande région par rapport à la structure et inclure des structures autres que la structure visée. Si la saillance des structures incluses est différente de la saillance de la structure recherchée, alors l'estimation effectuée sur la localisation peut donner des résultats contre-intuitifs. Par exemple, la localisation du putamen (structure peu saillante) peut inclure des sillons du cerveau (beaucoup plus saillants).

Nous avons construit les bases d'un système d'interprétation d'images, capable de faire des choix grâce au critère dérivé de l'information visuelle, d'effectuer la segmentation d'un objet de la scène ainsi que de sa reconnaissance, et enfin d'être critique vis-à-vis de l'information recueillie et permettant de changer la stratégie si nécessaire. Le système a en outre été rendu plus robuste aux échecs potentiels. Cela est possible grâce à l'évaluation des segmentations. Elle est effectuée par une structure de données permettant d'effectuer un contrôle du processus de segmentation en utilisant l'information spatiale et l'information visuelle. Ce faisant, nous avons permis d'automatiser une procédure qui était ad hoc.

Les résultats présentés montrent une bonne reconnaissance des structures, avec des segmentations qui sont souvent imprécises, en particulier pour le putamen dont la forme est moins propice à une segmentation par un modèle déformable. Le modèle intègre peu de structures, car il est nécessaire de pouvoir étudier la saillance de ces structures. Or si une structure est trop petite, elle apporte peu d'information à la carte de saillance, en particulier à cause des différents niveaux d'échelle. Nous avons donc choisi de nous limiter à des structures qui présentent une taille suffisante. Il est de plus nécessaire qu'elles soient reliées dans le modèle par des relations spatiales. Les résultats dans les cas pathologiques montrent que le modèle peut s'adapter aux déformations qui ne sont pas trop importantes. Les grandes déformations empêchent cependant la reconnaissance de certaines structures. Cependant, l'échec de la segmentation dans ces cas, détecté par le processus, peut nous fournir un moyen de détecter la présence d'une pathologie, ce qui n'a pas été investigué pour le

moment.

6.2 Perspectives

6.2.1 Optimisation avec représentation des structures

Nous avons souligné dans les conclusions la limitation principale de l'approche permettant d'optimiser un chemin évaluant la connaissance spatiale de chaque arc de manière séparée. La prise en compte de toute l'information spatiale utilisée dans une séquence de segmentation peut-être effectuée en fusionnant, au niveau de chaque nœud du graphe, toute l'information spatiale utilisable pour ce nœud, c'est-à-dire l'ensemble des relations spatiales visant ce nœud et utilisant comme structure de référence une structure du chemin déjà visitée. Mais dans ce cas, il n'est plus possible d'effectuer une optimisation globale dans le graphe telle que nous la proposons, car l'évaluation d'un arc (où d'un nœud en fonction de l'emplacement où nous choisissons de disposer l'information) n'est plus indépendante, mais dépend à présent du chemin suivi pour arriver jusqu'à cet arc (ou le nœud). Il serait donc nécessaire d'évaluer chaque chemin de manière séparée.

Nous avons considéré dans nos travaux des graphes se composant de peu de nœuds et donc de chemins. Les optimisations peuvent être effectuées de manière exhaustive dans ce cas, la liste des chemins étant réduite. Dans nos expériences, nous calculons l'évaluation de chaque chemin. Mais il n'est pas nécessaire de connaître l'évaluation de tous les chemins si notre objectif est d'obtenir le meilleur chemin uniquement. Dans le cas où l'évaluation de chaque arc est effectuée de manière indépendante, des algorithmes classiques de la théorie des graphes peuvent nous permettre de limiter le coût de l'optimisation. Dans le cas où chaque chemin doit être évalué de manière séparée, il serait nécessaire d'utiliser la programmation dynamique pour réduire la complexité. L'utilisation d'une structure de référence unique est primordiale dans ce cas.

Dans l'approche globale, où les chemins sont représentés sous la forme d'un ensemble flou unique, nous avons souligné le problème de la bonne représentation d'un chemin. Nous avons présenté des représentations utilisant des fusions conjonctives ou disjonctives. Cependant, nous avons toujours utilisé le minimum et le maximum, qui sont respectivement la plus optimiste des t-normes et la plus pessimiste des t-conormes. Or, il existe de nombreux opérateurs, comme la norme de Lukasiewicz par exemple. Il serait intéressant de déterminer quelles sont les propriétés souhaitées pour notre fusion d'informations et quels opérateurs permettent d'y répondre au mieux.

6.2.2 Optimisation avec information visuelle

Sur la reconnaissance des structures sous-corticales, nous utilisons des structures de référence qui sont segmentées au préalable dans notre approche. Ces structures ne sont pas difficiles à segmenter en soi, mais l'automatisation du processus pose problème tout de même. En particulier, il est parfois difficile de séparer les deux ventricules. Le troisième ventricule peut également être segmenté simultanément et être connecté aux ventricules latéraux. Il serait intéressant de segmenter ces structures de manière automatique, en définissant un cadre pour veiller à leur bonne reconnaissance respective.

Notre approche peut éventuellement être utilisée comme une initialisation pour une méthode telle que celle proposée par O. Nempont dans ses travaux de thèse ([Nempont \(2009\)](#)). Notre approche permettrait de réduire la complexité en fournissant des structures déjà segmentées, permettant de réduire beaucoup les domaines dès le début du processus, et en supprimant les opérations liées aux structures segmentées.

Des modifications sont nécessaires afin de mieux prendre en compte les cas pathologiques. Les expériences réalisées ne tiennent pas compte de la connaissance de la pathologie. Deux voies sont possibles, la première consiste à essayer de déterminer si un cas présente une pathologie à l'aide de ce processus de segmentation. L'objectif serait ici d'inclure la pathologie dans le modèle structurel. Cela implique que la pathologie a un impact sur le modèle structurel de l'image pour pouvoir être détectée. La deuxième voie serait d'adapter le processus de segmentation en sachant qu'une image est pathologique et éventuellement en utilisant une segmentation préalable de la tumeur. L'objectif serait donc d'adapter la connaissance spatiale (ou sa représentation) à la pathologie.

La notion de saillance est une notion bio-inspirée que nous adaptons à nos besoins dans ces travaux. Les caractéristiques d'un système pré-attentionnel consistent à calculer des caractéristiques globales de l'image qui « sautent aux yeux » (intensité, couleur, orientation) et de manière parallèle. Les caractéristiques de l'image ont été choisies pour la réaction qu'elles produisent sur le cortex visuel. Dans nos travaux, la tâche du cortex est remplacée par notre méthode de segmentation et de reconnaissance. Les caractéristiques des images pourraient donc être adaptées par rapport à la méthode de segmentation. Par exemple, la radiométrie des structures sous-corticales étant située entre la radiométrie de la matière blanche et de la matière grise du cerveau, la carte reflétant les intensités peut être adaptée pour réagir aux discontinuités dans cet intervalle. Dans cet exemple précis, la radiométrie des matières est fournie par une analyse de l'image, mais également par une connaissance a priori sur l'image. L'utilisation de cette connaissance a priori pour calculer la carte de saillance fait que le processus n'est plus strictement guidé par les données dans ce cas.

Nous avons introduit l'utilisation d'une carte de saillance conjointement avec un modèle structurel et des représentations floues de relations spatiales, et nous avons discuté dans cette conclusion de la problématique que cette approche a ouverte, c'est-à-dire la difficulté de comparer la saillance sur des régions et non pas d'une manière globale, par rapport à un modèle de la saillance attendu pour une structure. Dans ces régions, la saillance de la structure recherchée est mêlée à la saillance de structures environnantes, entre autres. La recherche est donc dépendante de la précision de la localisation. Des travaux, introduits au chapitre 3, proposent de modifier la saillance pour chercher un type d'objets spécifiques. Mais dans notre cas, les structures ont des caractéristiques assez proches. Il faudrait donc plutôt étudier l'influence de la taille de la région. Une autre piste est de modéliser non seulement la saillance de la structure, mais également de son environnement.

Le processus de contrôle que nous avons introduit se contente, pour des raisons de complexité, de regarder les interactions entre la structure segmentée et sa structure parente. Cependant, les mesures d'évaluation sont présentes dans le graphe, et la cohérence spatiale est mise à jour à chaque itération sur chaque arc du graphe. Il serait donc possible d'effectuer une optimisation globale a posteriori de la qualité de la segmentation. Nous pouvons par exemple optimiser à l'aide d'une coupure un graphe où les arcs portent leur évaluation de la cohérence spatiale et où l'attache aux données est estimée par le critère de saillance. L'optimisation consiste ici à déterminer quels nœuds sont considérés comme valides, et lesquels sont considérés comme invalides (et itérer le processus dans ce cas). Avec un tel processus, nous pouvons prendre en compte la cohérence spatiale du modèle complet, et pas uniquement entre deux structures. Cela pourrait en outre permettre de supprimer les seuils utilisés sur les critères.

L'utilisation conjointe de la notion de saillance et des relations spatiales peut être appliquée dans un autre cadre que l'imagerie médicale. Le modèle que nous utilisons ne décrit pas toute la scène et il peut correspondre à un motif particulier dans une scène. Par exemple dans le cadre de l'imagerie satellitaire, la description d'une structure complexe telle qu'un aéroport peut être effectuée par un modèle structurel. Si nous connaissons une structure de référence appartenant au motif décrivant l'aéroport, alors nous pouvons utiliser notre système pour segmenter et reconnaître les autres parties du modèle. Dans le cas de l'imagerie satellitaire, il serait bien sûr nécessaire de

définir une méthode de segmentation adéquate, avec les informations a priori de radiométrie nécessaires. Il faut noter que notre approche, en délimitant une zone d'intérêt, permet d'utiliser des informations a priori radiométriques qui ne sont pas nécessairement suffisantes pour une segmentation globale.

Annexe A

Liste des publications

Publications liées aux travaux de thèse :

- *Geoffroy Fouquier*, Jamal Atif and Isabelle Bloch
Sequential spatial reasoning in images based on pre-attention mechanisms and fuzzy attribute graphs.
In the proceedings of the 18th European Conference on Artificial Intelligence (**ECAI'2008**).
 - *Geoffroy Fouquier*, Jamal Atif and Isabelle Bloch
Incorporating a pre-attention mechanism in fuzzy attribute graphs for sequential image segmentation.
In the proceedings of the 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (**IPMU'2008**).
 - Emanuel Aldea, *Geoffroy Fouquier*, Jamal Atif and Isabelle Bloch
Kernel Fusion for Image Classification Using Fuzzy Structural Information
In the proceedings of the 3rd International Symposium on Visual Computing (**ISVC'2007**).
 - Emanuel Aldea, *Geoffroy Fouquier*, Jamal Atif and Isabelle Bloch
Classification d'images par fusion d'attributs flous de graphes, relations spatiales et noyaux marginalisés.
Dans les actes des rencontres Francophones sur la Logique Floue et ses Applications (**LFA'2007**).
 - *Geoffroy Fouquier*, Jamal Atif and Isabelle Bloch
Local reasoning in fuzzy attribute graphs for optimizing sequential segmentation
In the proceedings of the 6th IAPR TC-15 Workshop on Graph-based Representations in Pattern Recognition (**GBR'2007**).
 - Jamal Atif, Céline Hudelot, *Geoffroy Fouquier*, Isabelle Bloch and Elsa Angelini
From Generic Knowledge to Specific Reasoning for Medical Image Interpretation using Graph based Representations.
In the proceedings of the Twentieth International Joint Conference on Artificial Intelligence (**IJCAI'2007**).
-

Autres publications :

- Helin Dutağacı, Geoffroy Fouquier, Erdem Yörük, Bülent Sankur, Laurence Likforman and Jérôme Darbon
Hand Recognition **Book chapter**
dans "Guide to Biometric Reference Systems and Performance Evaluation". Springer-Verlag, 2009. Éditeurs : D. Petrovska-Delacrétaz, G. Chollet, B. Dorizzi et A.K. Jain
 - *Geoffroy Fouquier*, Laurence Likforman, Jérôme Darbon and Bulent Sankur
The Biosecure Geometry-based System for Hand Modality
In the proceedings of the 32nd IEEE International Conference on Acoustics, Speech, and Signal Processing (**ICASSP'2007**).
 - Thierry Géraud, *Geoffroy Fouquier*, Quoc Peyrot, Nicolas Lucas and Franck Signorile
Document Type Recognition Using Evidence Theory.
In the proceedings of the Fifth IAPR International Workshop on Graphics Recognition (**GREC'2003**).
 - Alexis Angelidis and *Geoffroy Fouquier*
Visualization Issues in Virtual Environments : From Computer Graphics Techniques to Intentional Visualization.
In the proceedings of the 9th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (**WSCG'2001**).
-

Annexe B

Cartes de saillance

Nous présentons dans cette annexe des résultats de génération de cartes de saillance selon la méthode présentée dans le chapitre 3.

Dans une première partie, nous présenterons les cas sains de notre base de données et dans une deuxième partie, les cas pathologiques. Pour tous les volumes en trois dimensions, nous illustrons les résultats sur trois coupes extraites de manière automatique. Le choix des coupes est effectué à partir du masque du cerveau de chaque image, en ajoutant un nombre arbitraire de coupes dans une direction, à partir de la première coupe non vide dans une vue donnée. Les coupes ne sont donc pas comparables entre les différentes images.

Pour chaque ensemble (de cas sains et de cas pathologiques), nous présentons des cas avec plus de détails. Pour les autres cas, nous présentons uniquement les coupes de l'image originale et les cartes de saillance correspondantes. Pour les cas détaillés, nous présentons les figures suivantes :

- l'image originale ;
- la carte de saillance. La génération de la carte de saillance est détaillée dans la partie 2.4 ;
- les histogrammes de saillance calculés sur la segmentation manuelle de cette image, et qui sont utilisés pour l'apprentissage des distributions de saillance. Ces histogrammes sont définis dans la partie 5.1.3 ;
- la carte de visibilité correspondant à l'intensité ;
- la carte de visibilité correspondant à l'orientation. Les cartes de visibilité sont définies dans la partie 2.4.

Bases de données

Notre base de données est décrite dans la partie 3.4. Cette base est constituée des ensembles suivants :

- Les 18 cas de la base IBSR (« Internet Brain Segmentation Repository »)¹
- 11 cas provenant de la base OASIS (« Open Access Series of Imaging Studies »).²
- Des cas pathologiques, fournis par des hôpitaux partenaires. Certaines images ont été recueillies lors d'un projet financé par l'INCA (PL005-2005). Les hôpitaux partenaires sont les suivants :
 - L'hôpital Sainte-Anne ;
 - L'hôpital du Val-de-Grâce ;

¹Internet Brain Segmentation Repository. The MR brain data sets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at <http://www.cma.mgh.harvard.edu/ibsr/>

²<http://www.oasis-brains.org>, réalisée avec les financements suivants : Pubmed Central submission : P50 AG05681, P01 AG03991, R01 AG021910, P50 MH071616, U24 RR021382, R01 MH56584

– L'hôpital de la Pitié-Salpêtrière.

B.1 Les cas sains

Nous avons trente cas sains dans notre base. Nous allons présenter en détail le premier cas de la base IBSR, ainsi que le deuxième cas de la base OASIS.

B.1.1 IBSR 01

Image originale :

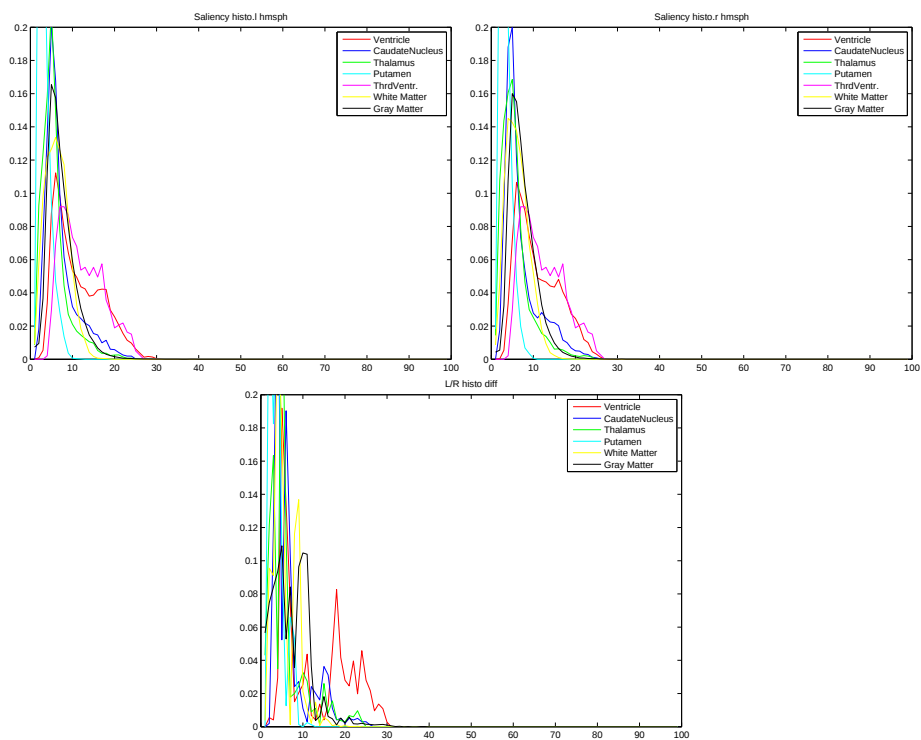


Carte de saillance :

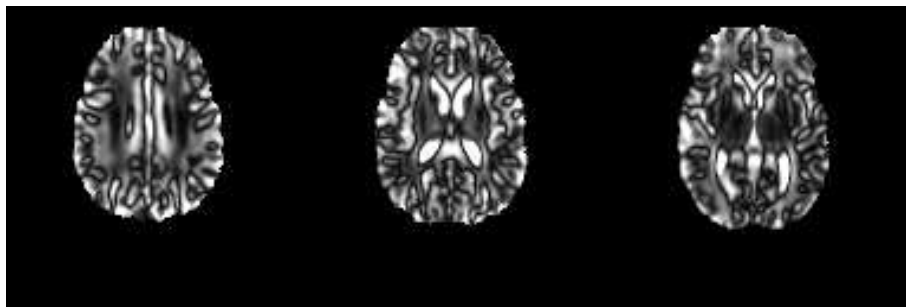


Énergies des histogrammes de saillance			
Structure :	Énergie :	Sal. min :	Sal. max :
CDl	0,112	0,02	0,25
CDr	0,113	0,02	0,26
GMI	0,106	0,01	0,39
GMr	0,107	0,01	0,35
LVl	0,061	0,02	0,3
LVr	0,063	0,02	0,27
PUI	0,236	0,01	0,12
PUr	0,233	0,01	0,1
THl	0,126	0,01	0,24
THr	0,115	0,01	0,24
V3	0,061	0,04	0,26
WMI	0,103	0,01	0,19
WMr	0,108	0,01	0,23

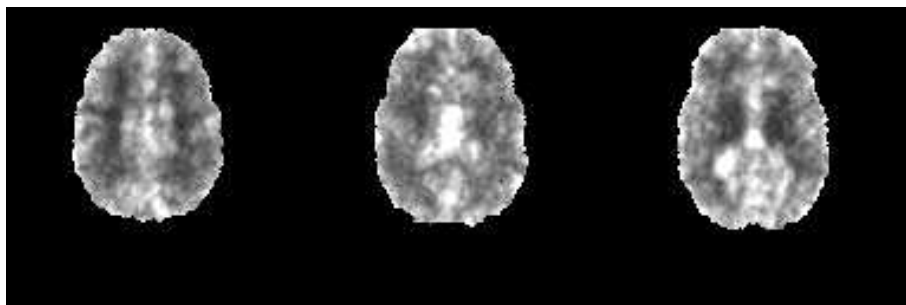
Histogrammes de saillance :



Carte de visibilité pour l'intensité :

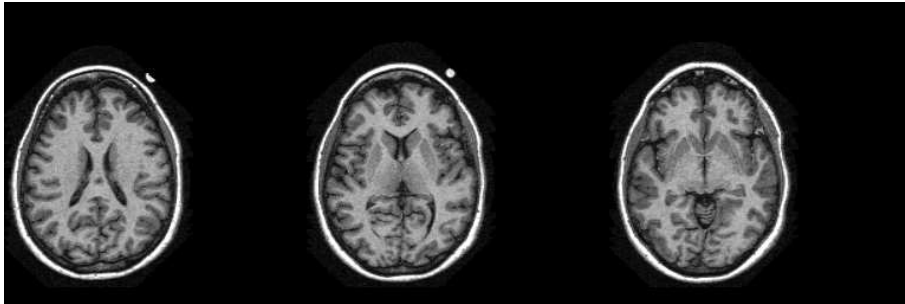


Carte de visibilité pour l'orientation :



B.1.2 Oasis 02

Image originale :

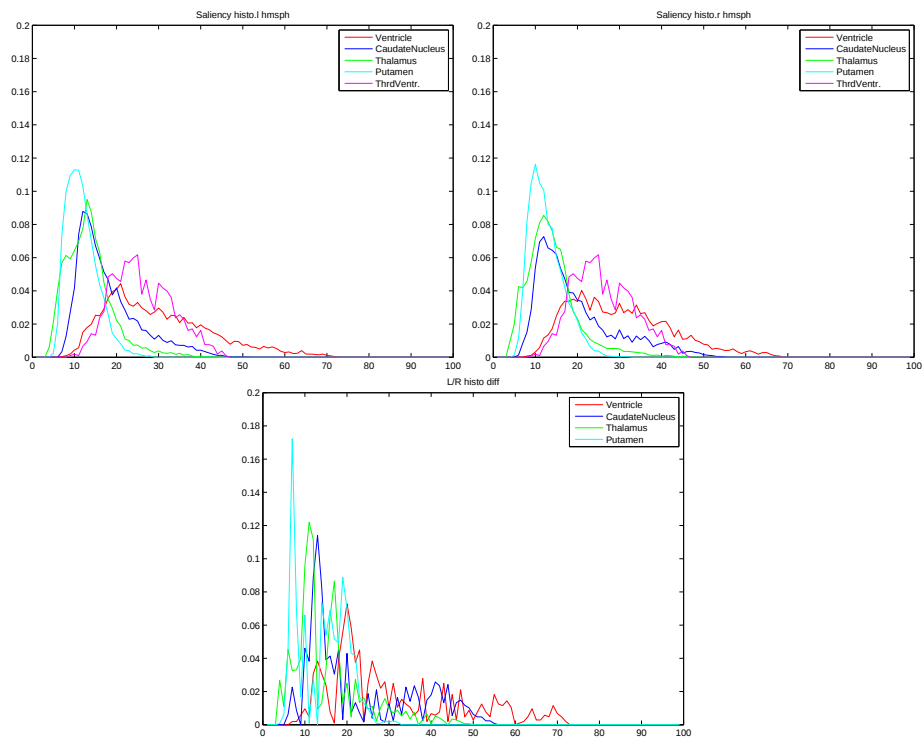


Carte de saillance :

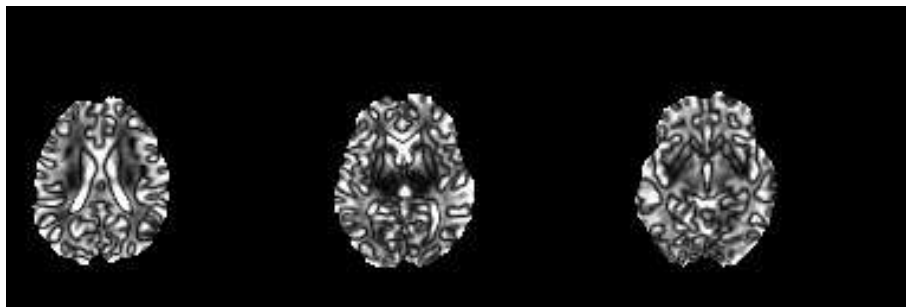


Énergies des histogrammes de saillance			
Structure :	Énergie :	Sal. min :	Sal. max :
CDl	0.051	0.07	0.51
CDr	0.041	0.06	0.55
LVl	0.025	0.07	0.72
LVr	0.025	0.08	0.68
PUl	0.084	0.04	0.28
PUr	0.077	0.05	0.32
THl	0.059	0.04	0.48
THr	0.056	0.04	0.49
V3	0.040	0.09	0.46

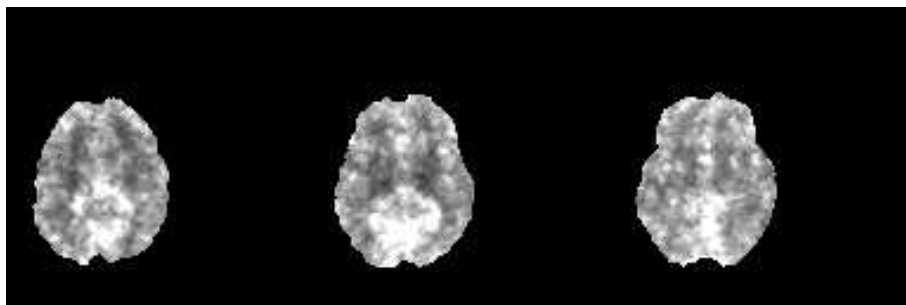
Histogrammes de saillance :



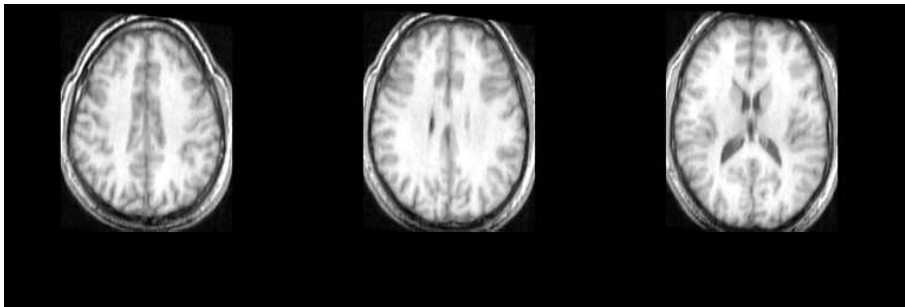
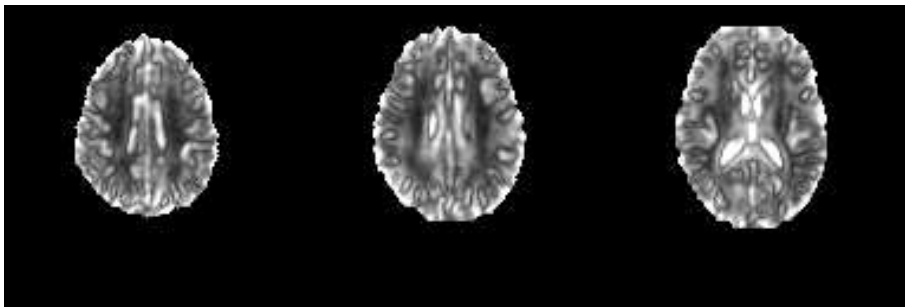
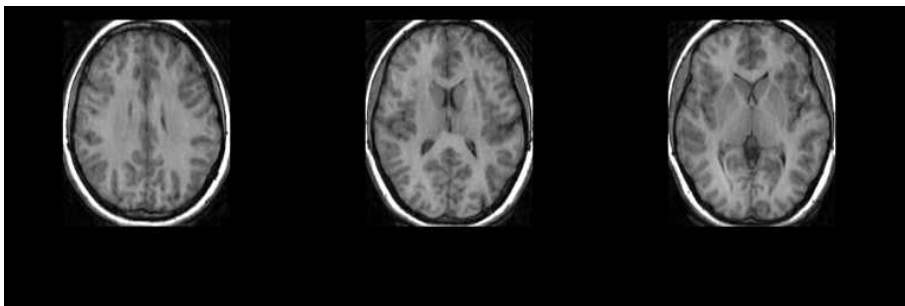
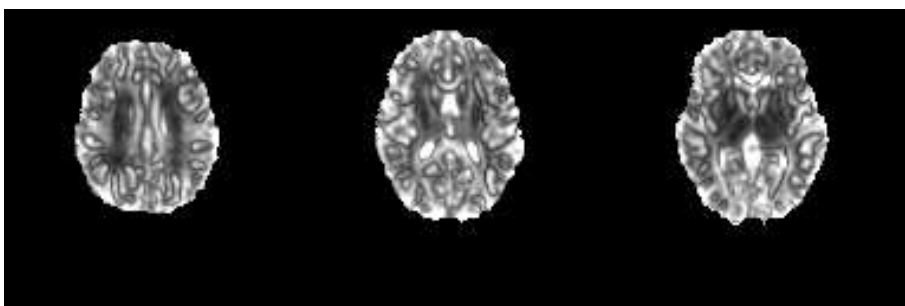
Carte de visibilité pour l'intensité :



Carte de visibilité pour l'orientation :



B.1.3 Les autres cas sains

IBSR 02**Image originale :****Carte de saillance :****IBSR 03****Image originale :****Carte de saillance :**

IBSR 04

Image originale :



Carte de saillance :

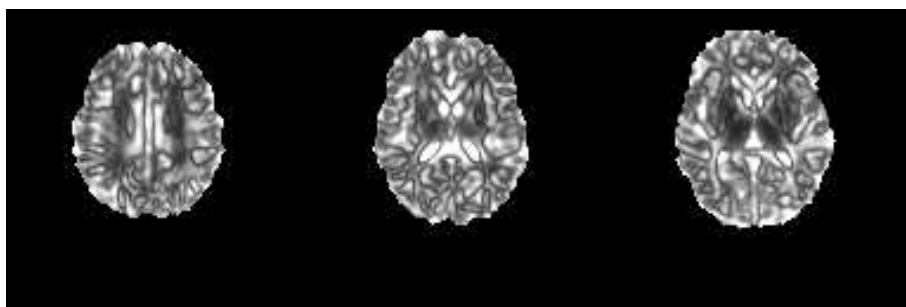
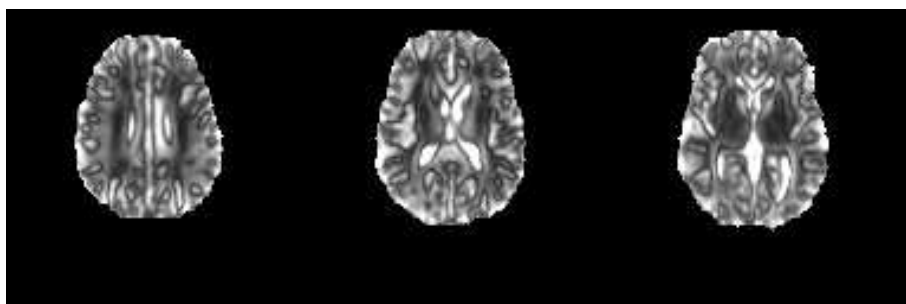
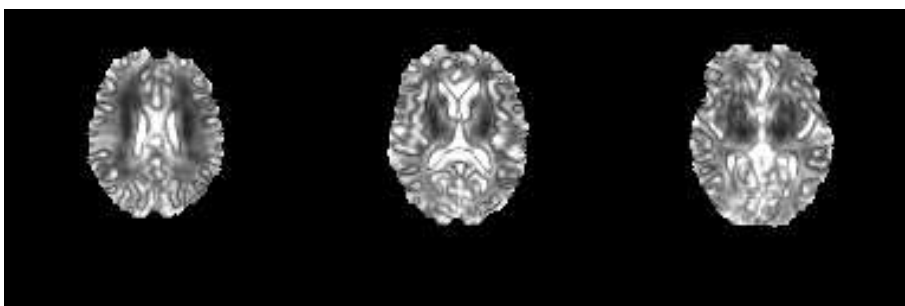
**IBSR 05**

Image originale :



Carte de saillance :



IBSR 06**Image originale :****Carte de saillance :****IBSR 07****Image originale :****Carte de saillance :**

IBSR 08

Image originale :



Carte de saillance :

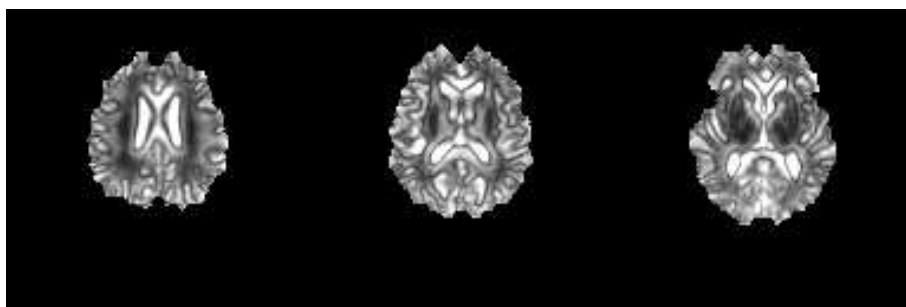
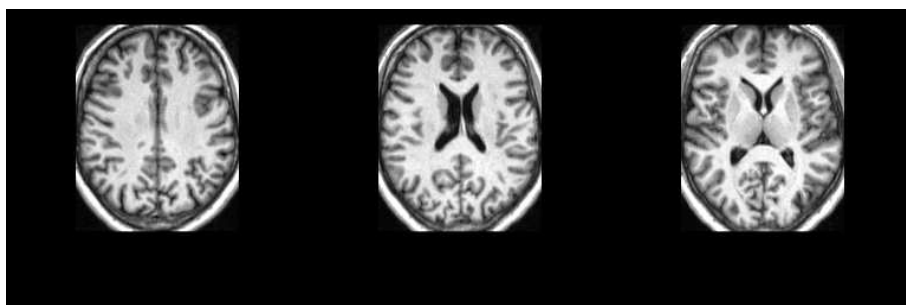
**IBSR 09**

Image originale :



Carte de saillance :



IBSR 10

Image originale :



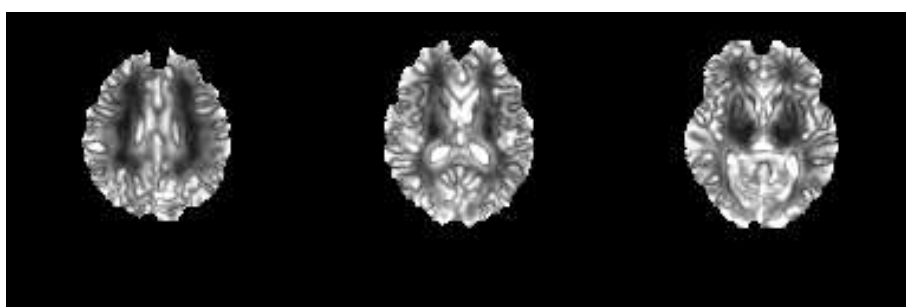
Carte de saillance :

**IBSR 11**

Image originale :



Carte de saillance :



IBSR 12

Image originale :



Carte de saillance :

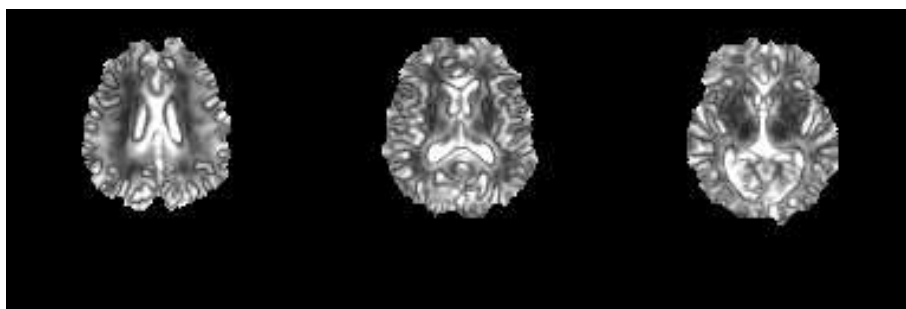
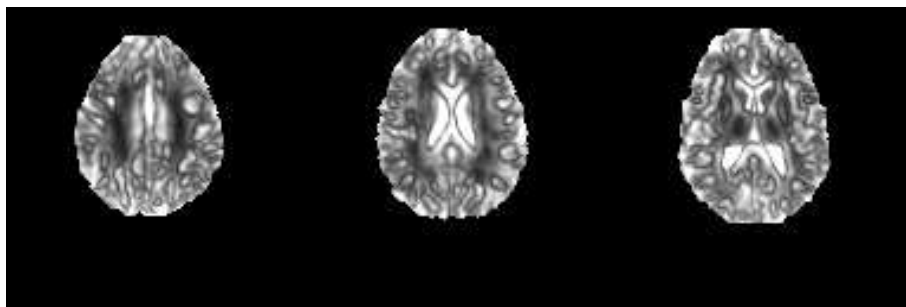
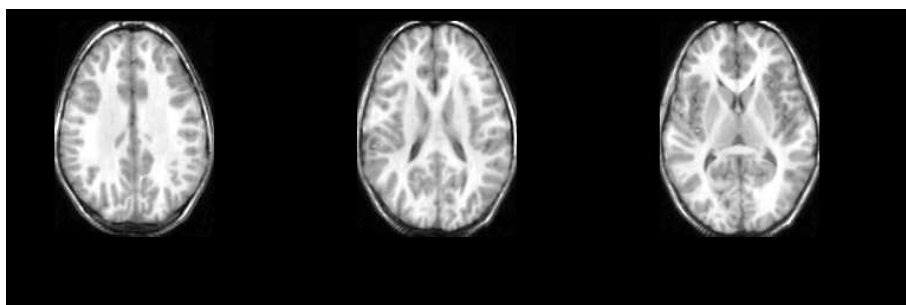
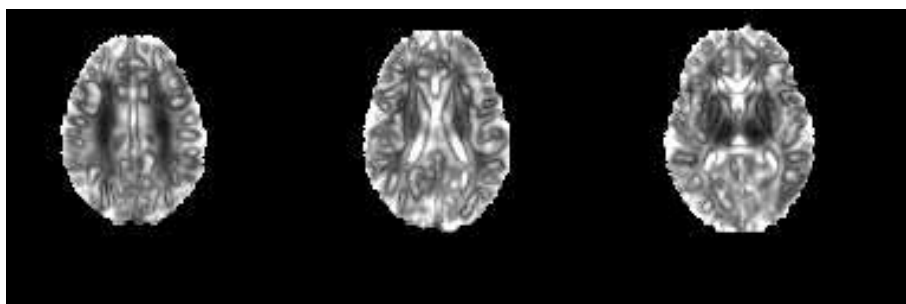
**IBSR 13**

Image originale :



Carte de saillance :



IBSR 14**Image originale :****Carte de saillance :****IBSR 15****Image originale :****Carte de saillance :**

IBSR 16

Image originale :



Carte de saillance :

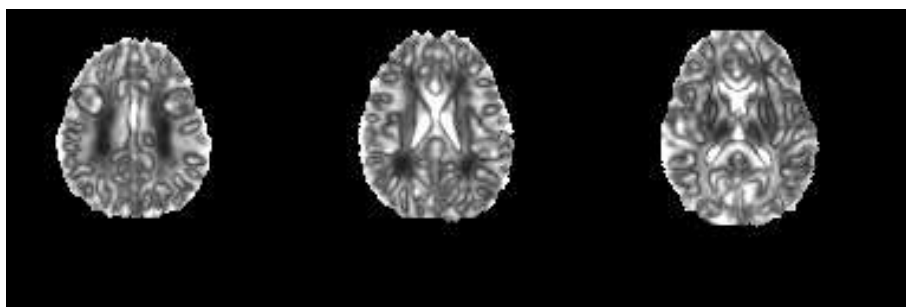
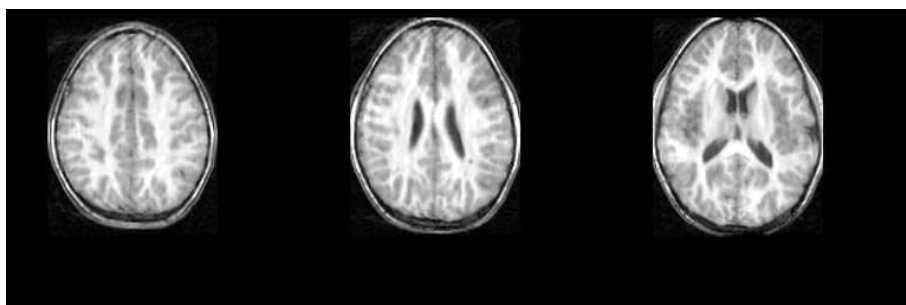
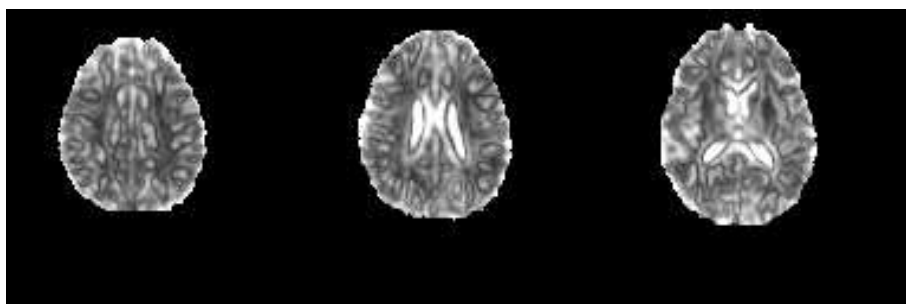
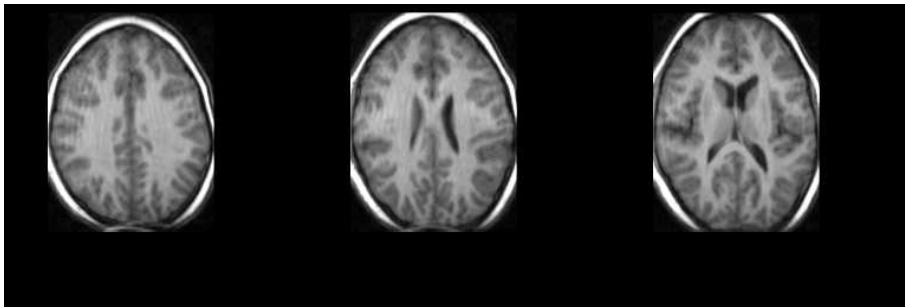
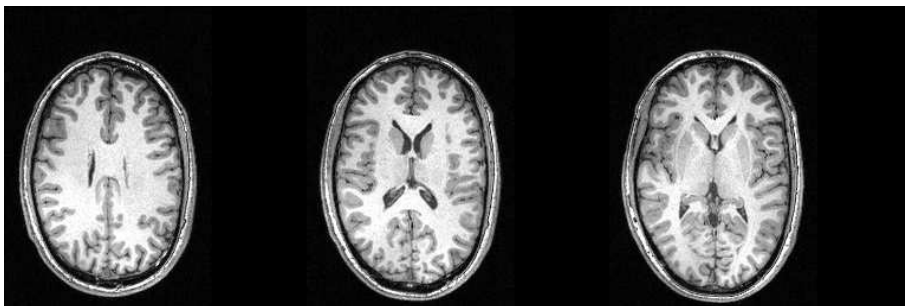
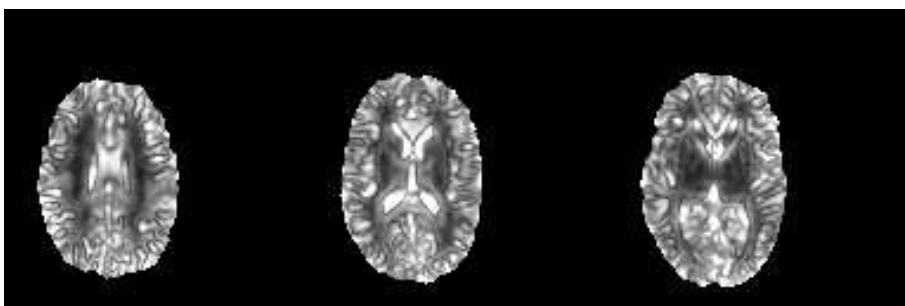
**IBSR 17**

Image originale :



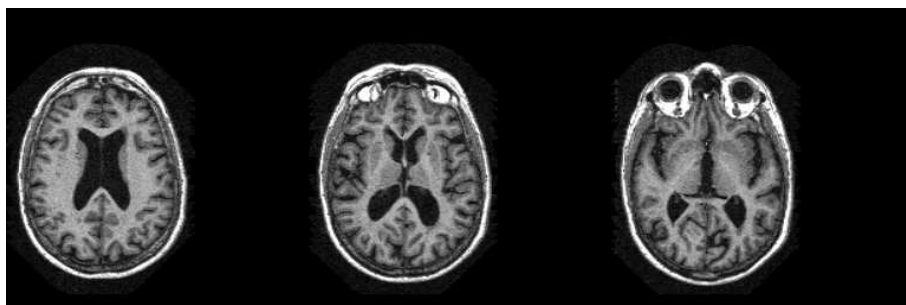
Carte de saillance :



IBSR 18**Image originale :****Carte de saillance :****cas sain****Image originale :****Carte de saillance :**

oasis 01

Image originale :



Carte de saillance :

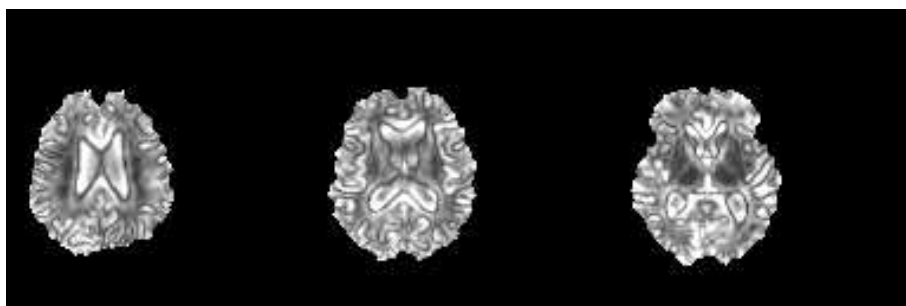


oasis 03

Image originale :

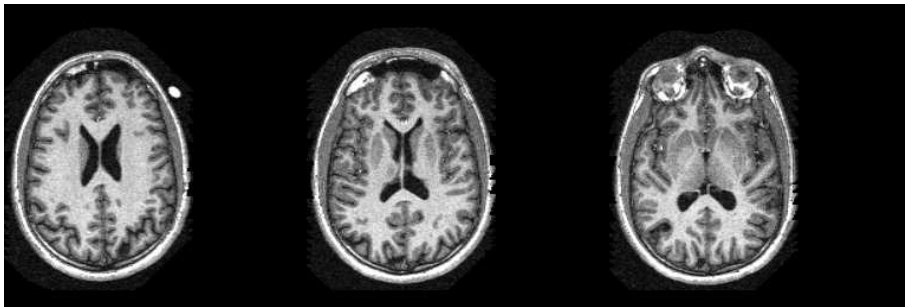


Carte de saillance :



oasis 04

Image originale :

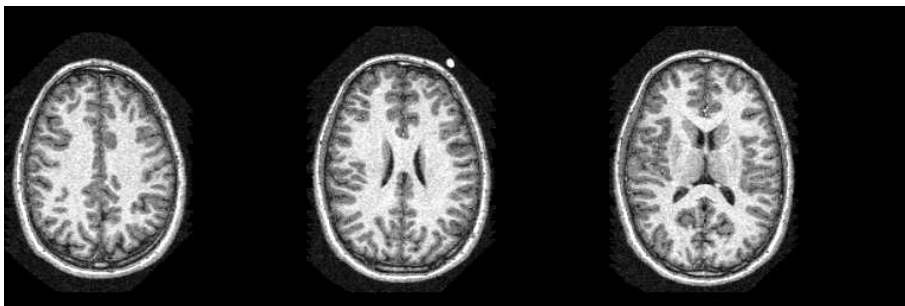


Carte de saillance :



oasis 05

Image originale :

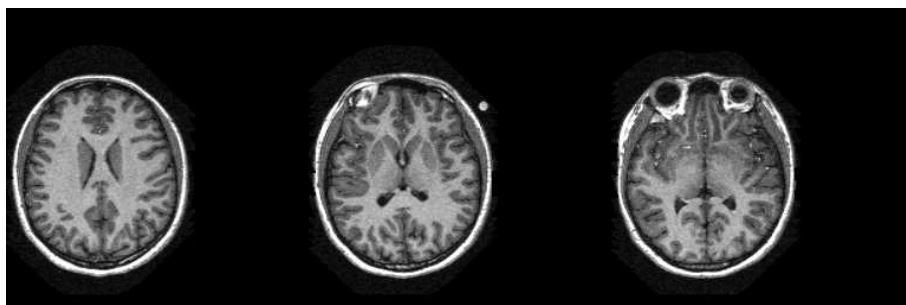


Carte de saillance :



oasis 06

Image originale :

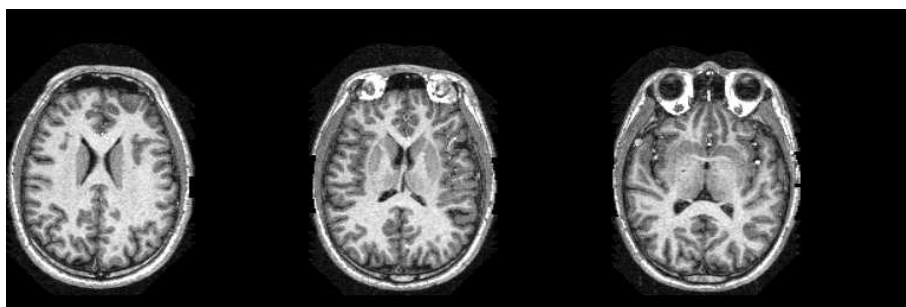


Carte de saillance :



oasis 07

Image originale :

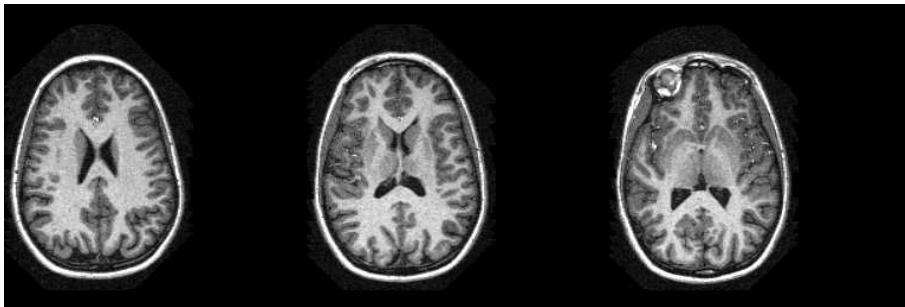


Carte de saillance :

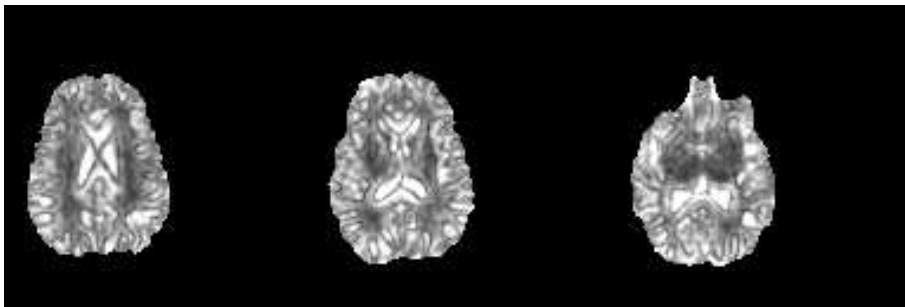


oasis 09

Image originale :

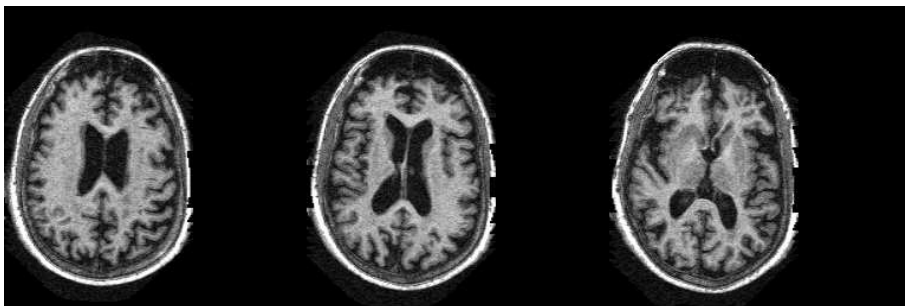


Carte de saillance :



oasis 10

Image originale :

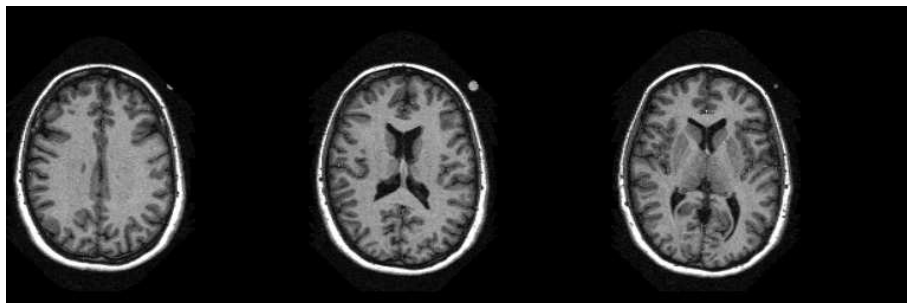


Carte de saillance :



oasis 11

Image originale :



Carte de saillance :



oasis 12

Image originale :



Carte de saillance :

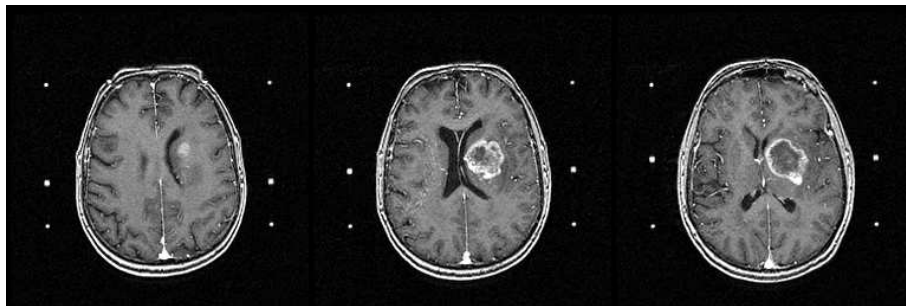


B.2 Les cas pathologiques

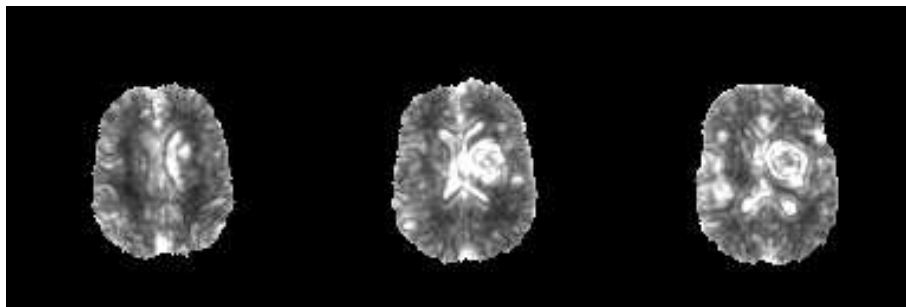
Nous présentons dans cette partie les cartes de saillance pour les vingt cas pathologiques de notre base.

B.2.1 Cas 1

Image originale :

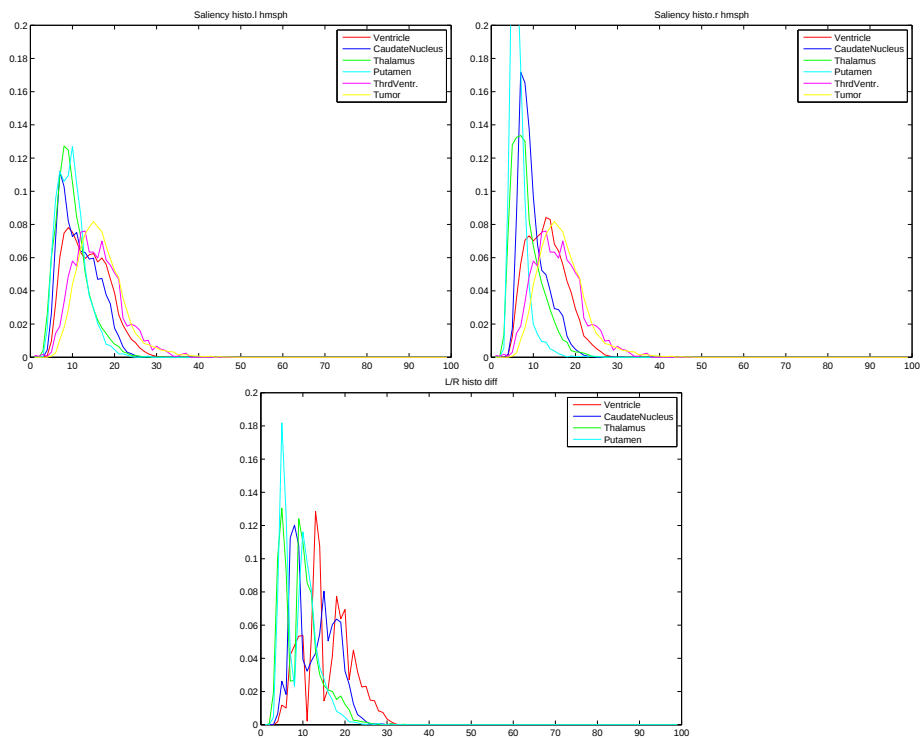


Carte de saillance :

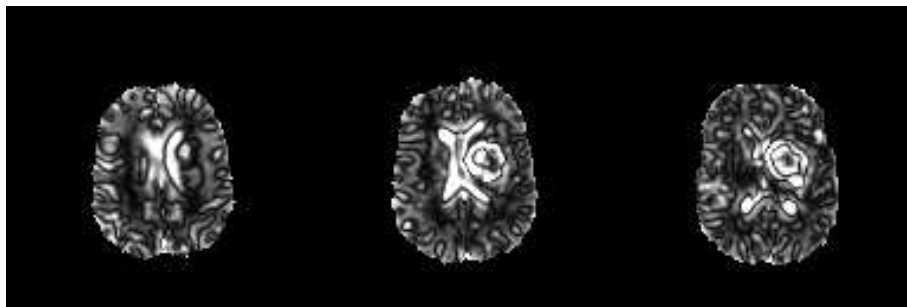


Énergies des histogrammes de saillance			
Structure :	Énergie :	Sal. min :	Sal. max :
CDl	0,069	0,04	0,27
CDr	0,107	0,04	0,31
LVl	0,057	0,04	0,32
LVr	0,062	0,04	0,28
PUl	0,089	0,03	0,28
PUr	0,183	0,03	0,26
THl	0,084	0,03	0,34
THr	0,094	0,02	0,27
V3	0,052	0,01	0,37
tumor	0,057	0,05	0,48

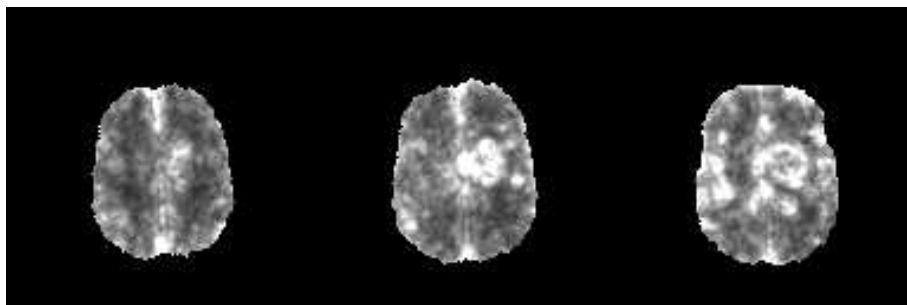
Histogrammes de saillance :



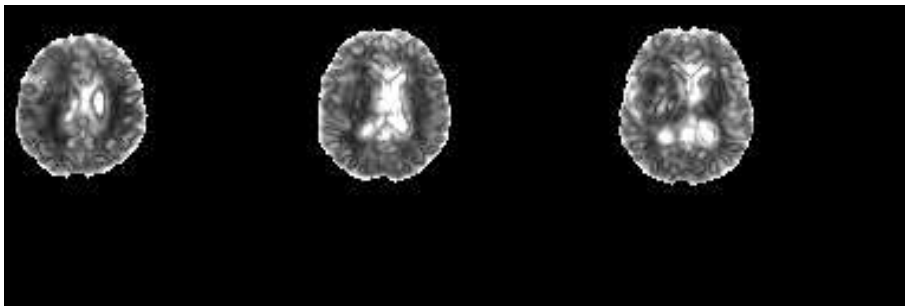
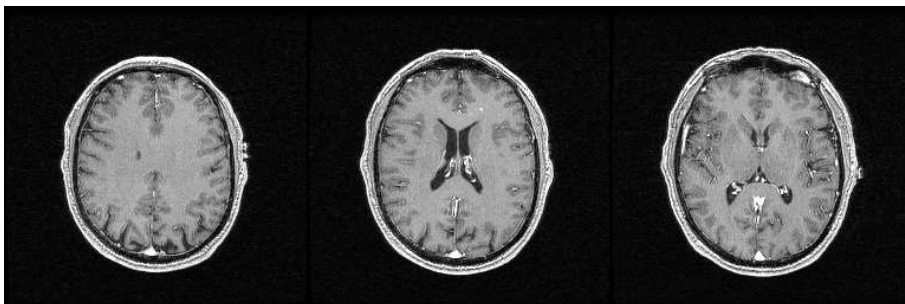
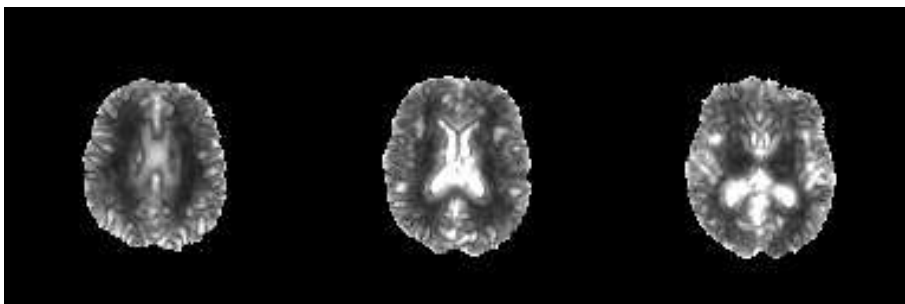
Carte de visibilité pour l'intensité :

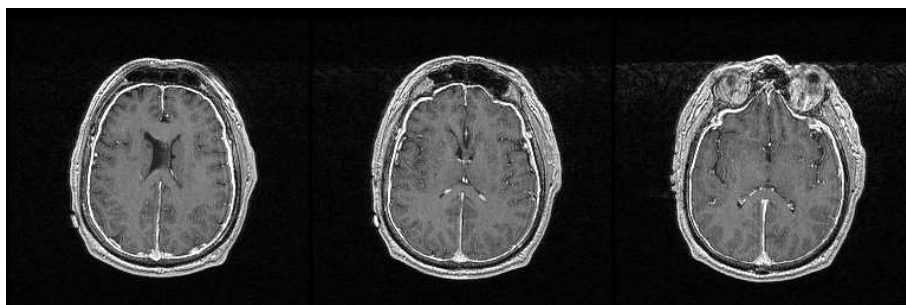
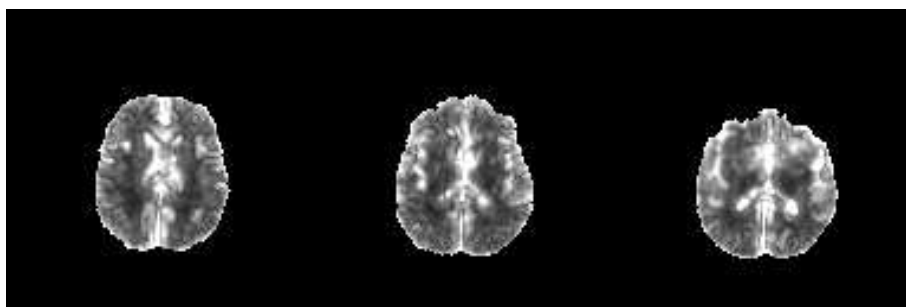
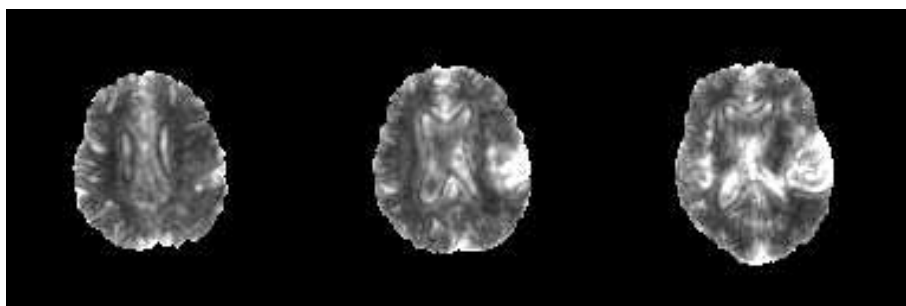


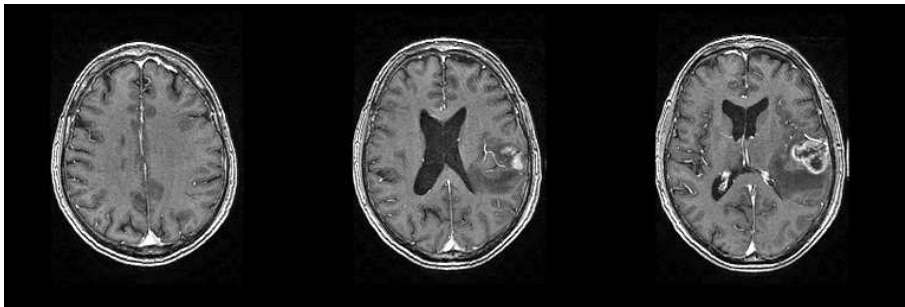
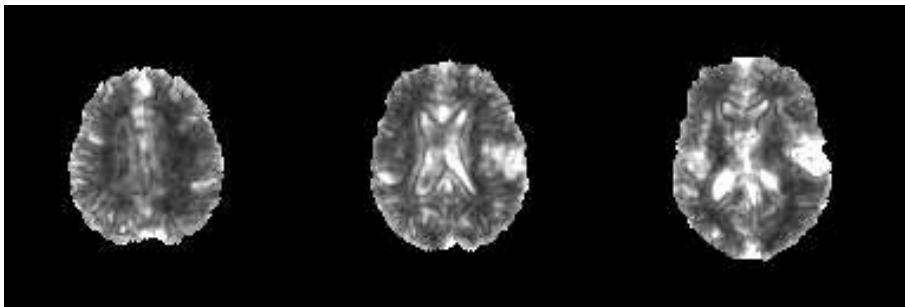
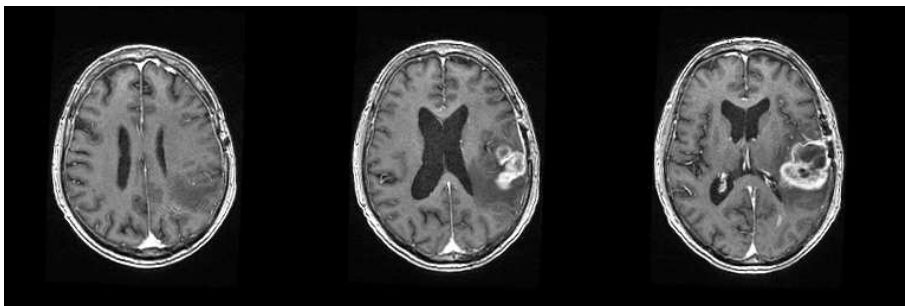
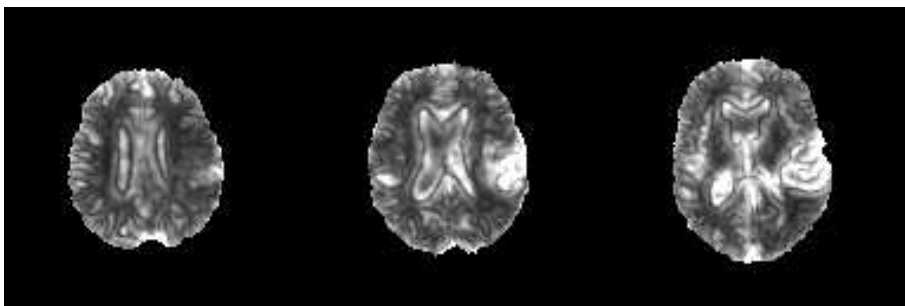
Carte de visibilité pour l'orientation :

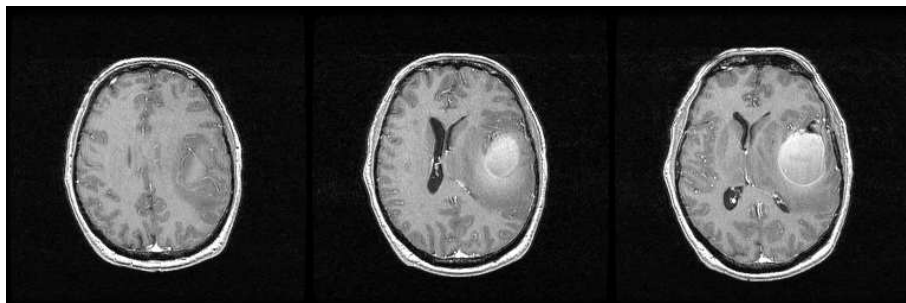
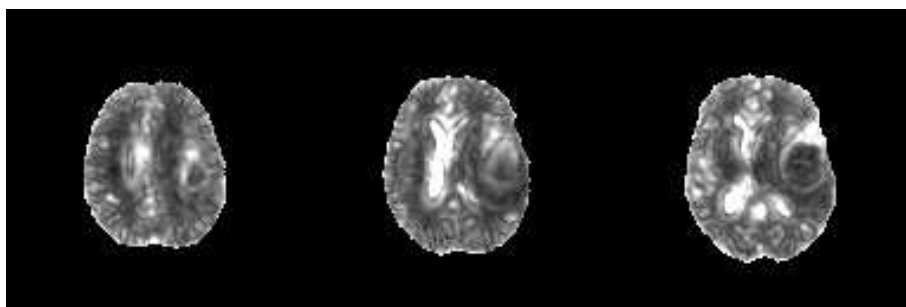


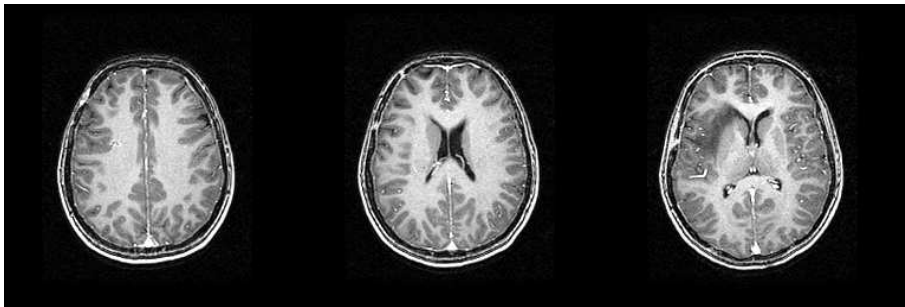
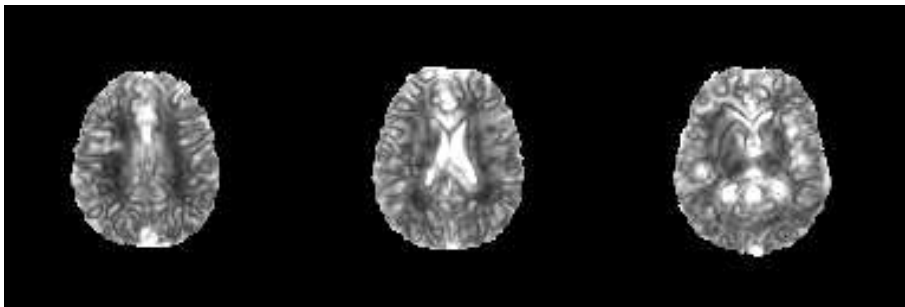
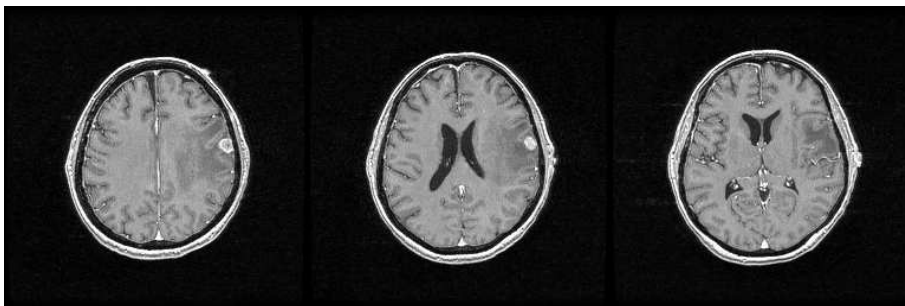
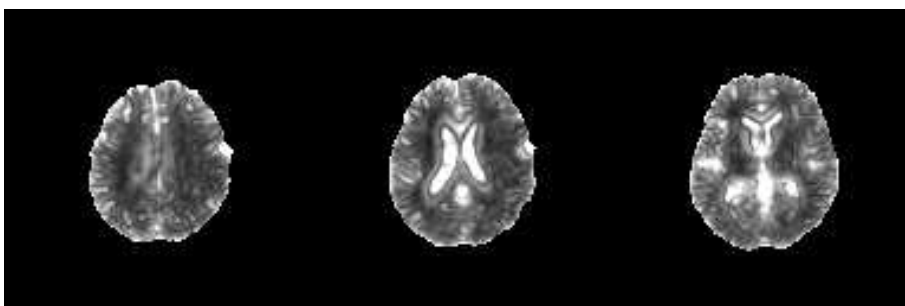
B.2.2 Les autres cas pathologiques

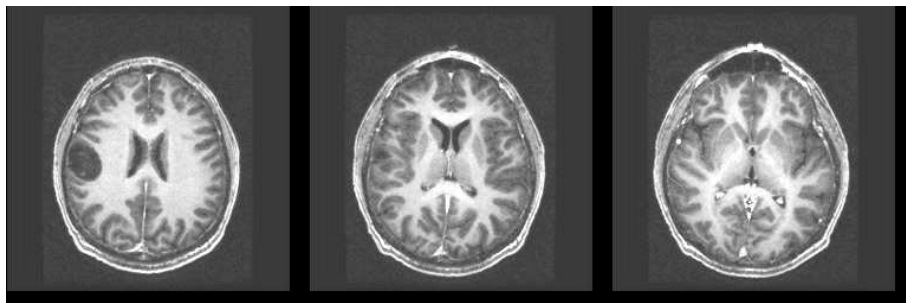
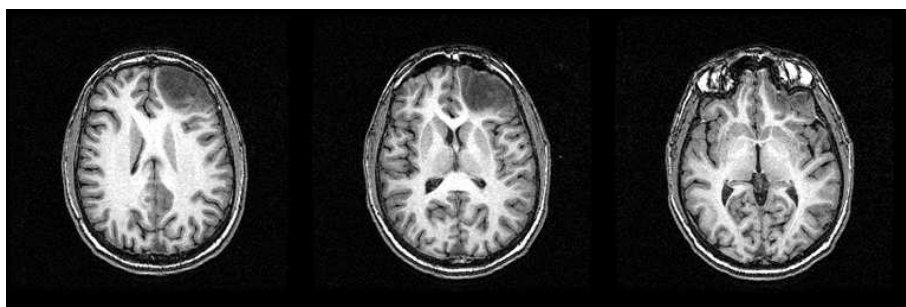
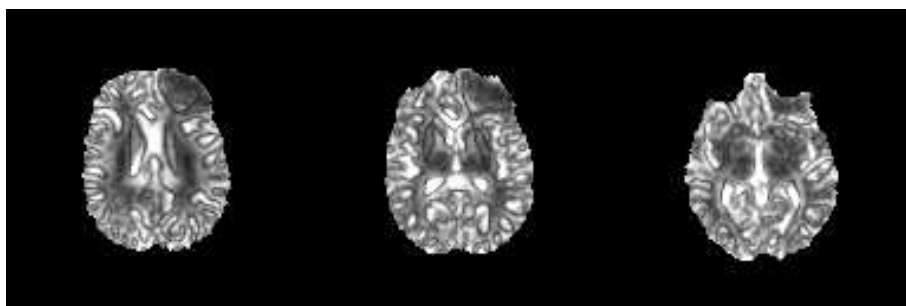
Cas 2**Image originale :****Saliency map :****Cas 3****Image originale :****Carte de saillance :**

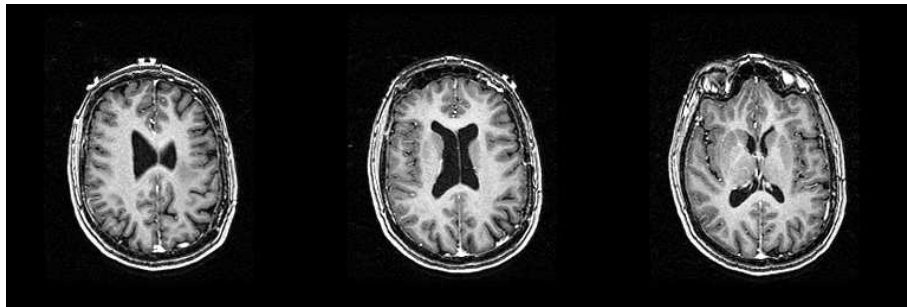
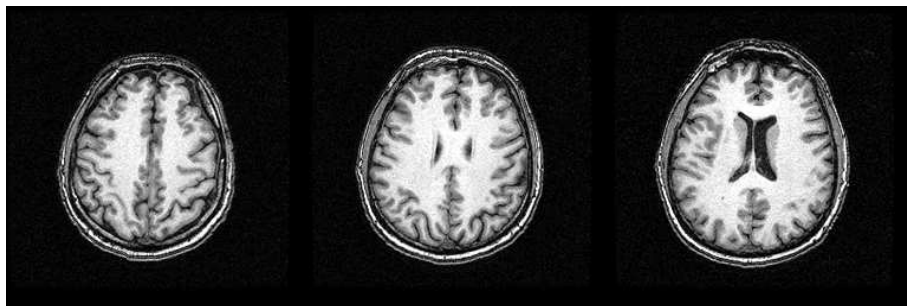
Cas 4**Image originale :****Carte de saillance :****Cas 5 / 1****Image originale :****Carte de saillance :**

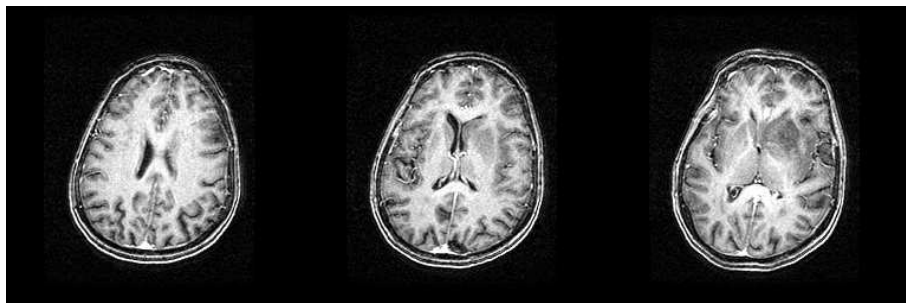
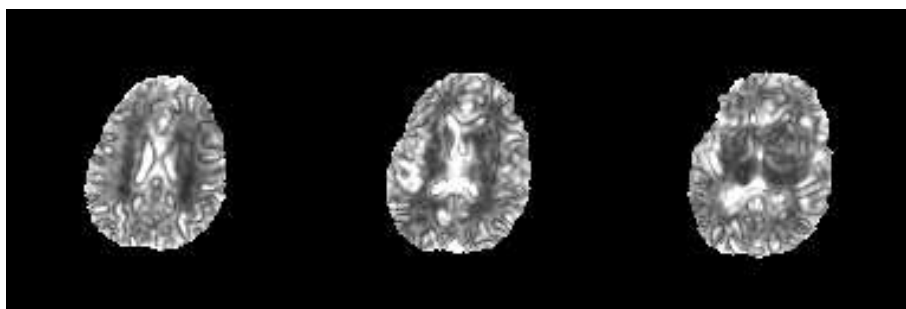
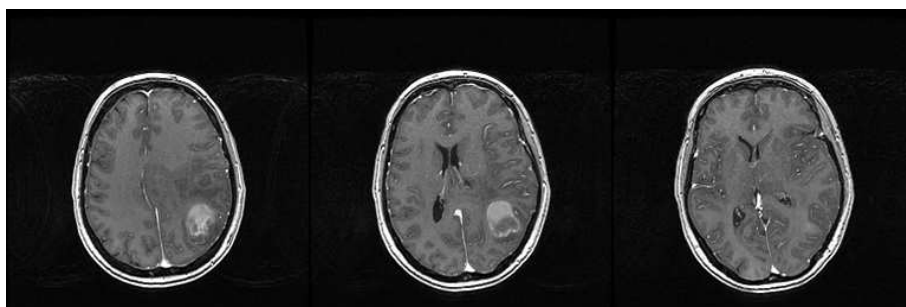
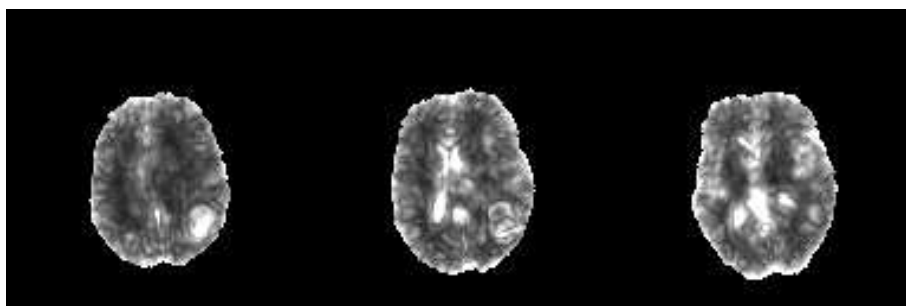
Cas 5 / 2**Image originale :****Carte de saillance :****Cas 5 / 3****Image originale :****Carte de saillance :**

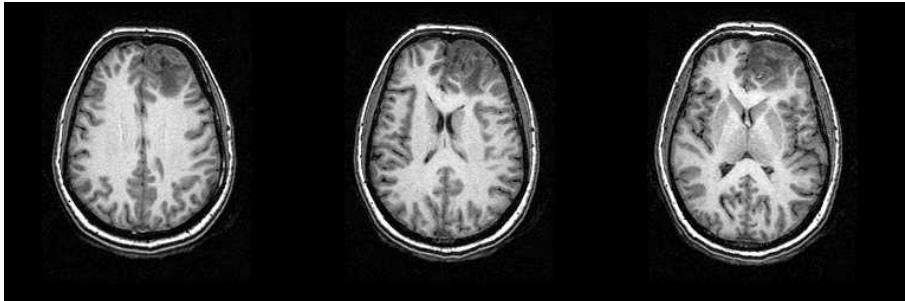
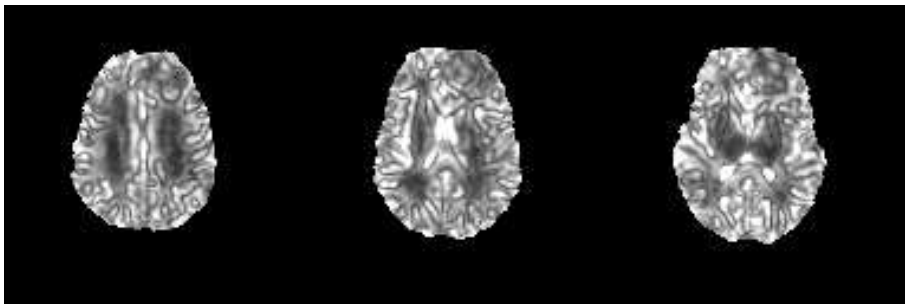
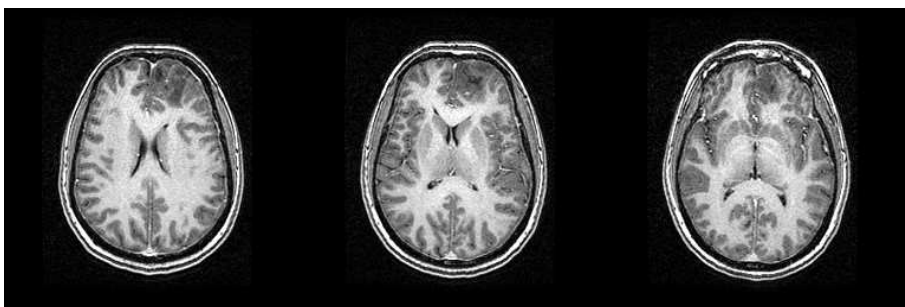
Cas 6**Image originale :****Carte de saillance :****Cas 7 /1****Image originale :****Carte de saillance :**

Cas 7 / 2**Image originale :****Carte de saillance :****Cas 8****Image originale :****Carte de saillance :**

Cas 9**Image originale :****Carte de saillance :****Cas 10****Image originale :****Carte de saillance :**

Cas 11**Image originale :****Carte de saillance :****Cas 12****Image originale :****Carte de saillance :**

Cas 13**Image originale :****Carte de saillance :****Cas 14****Image originale :****Carte de saillance :**

Cas 15 / 1**Image originale :****Carte de saillance :****Cas 15 / 2****Image originale :****Carte de saillance :**

Annexe C

Image segmentation as inexact graph matching using high-level attributes

Liste des auteurs :

Geoffroy Fouquier : Télécom ParisTech Dpt TSI, CNRS UMR 5141, Paris, France.

Roberto M. Cesar-Jr : Institute of Mathematics and Statistics USP, São-Paulo, Brazil.

Isabelle Bloch : Télécom ParisTech Dpt TSI, CNRS UMR 5141, Paris, France.

C.1 Abstract

This paper proposes a new method for model-based segmentation using a graph matching approach. The model is based both on a prototype image and on user's input, which allows deriving a segmentation where no homogeneity criterion is explicitly defined, and which is driven by the user's intention. As another contribution, an intermediate graph structure is involved in order to solve the difficult problem where no isomorphism can be expected between the model graph and the graph extracted from an over-segmentation of the image to be processed. Geometrical, topological and structural information is incorporated in a cost function, which is optimized to lead to the final result.

keywords : graph matching ; image segmentation ; spatial relations

C.2 Introduction

As shown in numerous works, structural information contained in images is an important feature for guiding different tasks such as segmentation, recognition, higher level interpretation and spatial reasoning [Bloch \(2005\)](#); [Miyajima et Ralescu \(1994\)](#). Graph representations are well adapted to encode this structural information, along with lower level information. Typically, vertices may represent regions or objects, with attributes extracted from the image data, while edges may represent relations between them (e.g. comparison of region attributes and spatial relations.)

A lot of work has been dedicated to graph matching, where two graphs to be matched are either built from two images, or from a model (or several models) and an image [Felzenszwalb et Huttenlocher \(2005\)](#); [Conte et al. \(2004\)](#); [Bunke \(2000\)](#); [Cross et Hancock \(1998\)](#). Here we consider the latter case and propose a new approach for segmenting an image based on a model built from both a

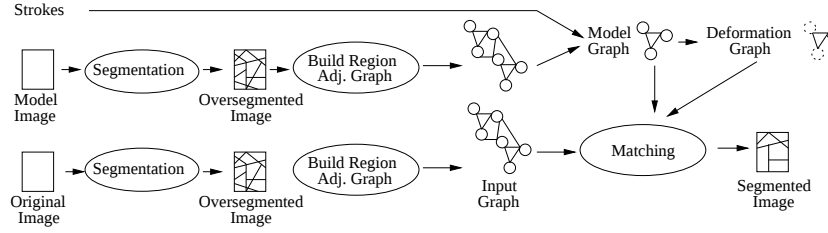


FIG. C.1 – General scheme for model-based image segmentation.

prototype image and from user's input. Our approach differs from usual segmentation tasks where the segmentation criterion is expressed as the homogeneity of some features computed in each segmented object or region, since it allows grouping into a single object regions that may be heterogeneous but that best correspond to the user's input. This input implicitly defines the segmentation criterion, which makes a major difference with respect to methods that rely on explicit criteria.

The proposed method, detailed in the next sections and illustrated in Figure C.1, proceeds as follows. The user is asked to draw strokes on a prototype image that is used to create a model. These strokes provide information on the objects the user is interested in, such as the number of classes or objects and approximate shape and colors of classes. For instance, the user may indicate a person as one class (hence the corresponding stroke will overlap regions with different local properties), or distinguish different classes like face, body, hair. This approach provides, with reduced user interaction, very strong information that alleviates the ill-posed nature of most segmentation problems. Here the problem becomes well-posed and allows segmenting what the user wants to get. The model graph is built according to this information. The segmentation of an input image is obtained by matching the model graph and a region adjacency graph (RAG) built from an input image. Usually the image graph contains many more regions than the model graph, which calls for inexact graph matching methods. Here we address this issue by using an additional graph, called deformation graph (Figure C.2), and introduced in Noma *et al.* (2009), which has the same topology as the model graph, and where each vertex corresponds to a union of regions of the input image graph. This structure provides a direct isomorphism with the model graph. Thus the segmentation is now achieved by finding the matching between the input image graph and the deformation graph which minimizes a cost function computed between the model graph and the deformation graph. This function includes comparison of vertices attributes and comparison of edges attributes.

In our previous work, some steps of this method were already described Consularo *et al.* (2007). Here the main contributions with respect to this earlier work include (i) the idea of segmenting objects based on an implicit criterion instead of an explicit one relying on region homogeneity, which allows segmenting potentially very heterogeneous areas as one object ; (ii) building a model graph which includes most of the prototype image information (not only the local information provided by the strokes) ; (iii) proposing new cost functions according to these new features of the method, taking also shape information into account, and adapting the graph matching algorithm accordingly.

This paper is organized as follows. In Section C.3 we discuss the graph constructions, while the attributes and cost functions for the optimization procedure are detailed in Section C.4. Section C.5 presents the matching algorithm and experimental results are described in Section C.6.

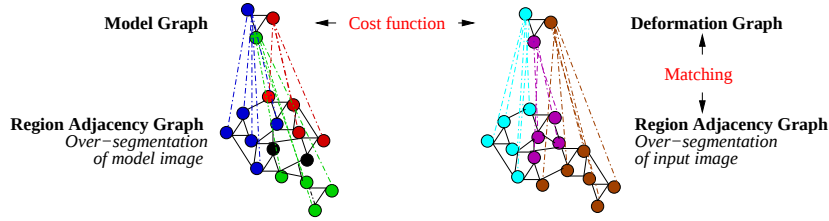


FIG. C.2 – General scheme for graph matching. The cost function is computed between the model graph and the deformation graph, which share the same structure. The matching between the input graph and the deformation graph is then derived.

C.3 Model, image and deformation graphs

In this section we describe the construction of the graphs involved in the proposed method.

Model graph G_m The model graph is built from a model (or prototype) image and from the user's input. It should represent the segmentation classes (or objects), according to the user, and the structural information (relations between classes). From our experiments with several users, it appears that usually the user draws strokes either on the border of a (generally large) region, or in the middle of the region (similar to a skeleton of the region). From these strokes, regions from an over-segmentation are grouped together to provide large and robust regions, according to the strokes. An example is illustrated in Figure C.3. The main steps of the proposed procedure are as follows :

- The user draws strokes on this image, the labeling is encoded using colors (1 color per class), with potentially several strokes for one class or one object ;
- The model image is segmented using any over-segmentation method : in our experiments we used a mean-shift approach, applied to the grey levels (or the intensity channel in case of color images) after a regularization step using a minimal total variation criterion with a L1 norm [Darbon et Sigelle \(2006a\)](#) which allows removing texture. This leads to large homogeneous regions and provides a less over-segmented image than the one used in our previous work based on watersheds ;
- A RAG is built from the segmentation ;
- A model graph, with one vertex per class is built and populated with all marked regions from the RAG (i.e. intersected by a stroke). This is the most original step of this procedure. All unmarked regions surrounded by a unique class are added to this class : this corresponds to the idea that these regions cannot represent a different class since the user did not draw a stroke over them. Unmarked regions surrounded by different classes are not included in the model ;
- Finally, an edge is created for each pair of vertices, and edge and vertex descriptors are computed in the model graph.

Input graph G_i For any image to be processed, an image (or input) graph is created. The RAG resulting from the over-segmentation is directly used. A set of features is computed on all regions. However, parameters for regularization and segmentation are less restrictive than for the model, in order to obtain smaller regions.

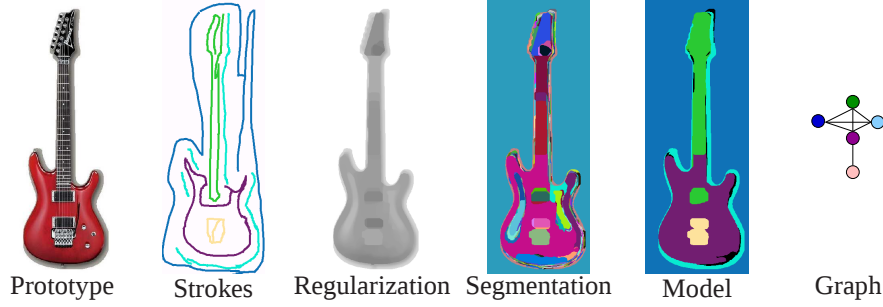


FIG. C.3 – Model generation. The minimal region size in the segmentation is 20. The resulting model is displayed with random colors. Black regions are excluded from the model.

Deformation graph G_d Our approach involves an intermediate structure, called the deformation graph, which has the same topology as the model graph but is populated from the image graph : several vertices can be merged into one in this deformation graph. This is another original feature of the proposed approach. The matching process then aims at finding the best merging of regions such that the deformation graph optimizes a cost function. This also provides the best matching between the model and the image, but without having to handle directly the problem of matching graphs with different topologies. Vertices and edges carry the same attributes as for the model graphs.

C.4 Attributes and cost function

Let Σ_V, Σ_E be the sets of vertex labels and edge labels, respectively. Let V be a finite non-empty set of vertices, L_v be a vertex interpreter $L_v : V \rightarrow \Sigma_V$, E be a set of ordered pairs of vertices called edges, and L_e be an edge interpreter $L_e : E \rightarrow \Sigma_E$. Then $G = (V, L_v, E, L_e)$ is a labeled graph with directed edges. For $v \in V$ and $e \in V \times V$, $\delta(v, e)$ is a transition function that returns the vertex v' such that $e = (v, v')$. For $v \in V$, $A(v)$ returns the set of edges adjacent to v .

C.4.1 Vertex cost : intrinsic features for each class of the model

The cost function associated with each vertex includes intensity, shape and surface information.

Intensity As mentioned above, a class may be composed of regions with non-homogeneous intensity, depending on what the user considers as being one object. Each vertex (representing a class) in both model and deformation graphs may then be composed of a set of smaller regions resulting from the initial over-segmentation, each of them being characterized by its average intensity. Note that regularization of images allows removing texture and therefore average intensity becomes more relevant. In order to take into account the potential intensity inhomogeneity inside a class of the model, the intensity cost is computed between both sets of regions R_d and R_m composing the compared vertex $V_m \in G_m$ and $V_d \in G_d$. The cost for a region from V_d is defined as the minimal grey level difference among all regions of V_m . Then, the intensity cost is defined as the average of these minimal distances for all regions composing V_d :

$$Cv_{intensity}(V_m, V_d) = \frac{\sum_{r_d \in R_d} \min_{r_m \in R_m} d(r_d, r_m)}{|R_d|}$$

where $d(r_d, r_m)$ represents the absolute difference between average grey levels of both regions, and $|R_d|$ is the number of regions composing V_d .

Shape Shape information is not meaningful for regions of an over-segmentation. However the process used for building the model makes shape information relevant for the model vertices and thus for the deformation graph, which is another advantage of using this intermediate structure. It follows that, as opposed to methods relying on a matching between the model and the input, shape information can really be involved in the proposed approach. This is a new contribution with respect to previous works.

Among the numerous existing shape descriptors, we prefer those that can be easily updated when the regions associated to each vertex of the deformation graph change. This is motivated by the number of computations of the attributes involved in our iterative matching scheme. Therefore we have chosen affine invariant moments proposed by Flusser et al in [Flusser et al \(1993\)](#). The invariant moments are a combination of central moments (shift invariant) defined as : $\mu_{pq} = \sum_x \sum_y (x - \hat{x})^p (y - \hat{y})^q$

Here is the definition of the two first invariant moments :

$$I_1 = \frac{\mu_{20}\mu_{02} - \mu_{11}^2}{\mu_{00}^4}$$

$$I_2 = \frac{\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{21}^3\mu_{03} - 3\mu_{21}^2\mu_{12}^2}{\mu_{00}^{10}}$$

Shape descriptors are computed on each vertex after applying a morphological closing in order to smooth noisy boundaries due to the segmentation process or to occlusion. We keep all normalized moments for $p, q \in [0..3]$ with $p + q > 1$ thus 13 moments.

The cost function for shape information is defined as the absolute difference between the vectors of central moments of vertices in G_d and G_m :

$$Cv_{shape}(V_m, V_d) = \frac{\sum_{i=1}^M |m_{d_i} - m_{m_i}|}{M}$$

where M is the number of moments.

Area Since area is an important feature, not taken into account with normalized moments, it is additionally included in the following cost function :

$$Cv_{area}(V_m, V_d) = \|a_m - a_d\|.$$

where a_m (a_d) is the area of V_m (V_d), normalized with respect to the model (resp. input) image size.

C.4.2 Edge cost : reflecting the structure

Spatial relations provide an important information carried by the edges, to compare the structures of the graphs. Again the use of the deformation graph makes this information relevant.

In both distance and orientation cost functions, we compare an edge from the model graph $E_m \in G_m$ with an edge of the deformation graph $E_d \in G_d$. Since model and deformation graphs have the same structure, both edges connect the same vertices and thus represent the same spatial relation.

Distance Let us consider an edge between two objects A and B . We denote the corresponding vertices in the model graph by A_m and B_m and the ones in the deformation graph by A_d and B_d . The edges between these vertices are denoted as E_m and E_d , respectively. In order to compare the relative distances, carried by these edges, we proceed as follows :

- we first compute the distances $d(x, B_d)$ for all points x of the contour of A_d and the cumulative histogram¹ of the obtained values.
- we compute the distances $d(x, B_m)$ for all points x of the contour of A_m and the cumulative histogram.
- the distance $dh_{A \rightarrow B}$ is evaluated as the distance between these cumulative histograms.

Finally, a symmetric distance is defined as :

$$C_{e_{dist}}(E_m, E_d) = \frac{\|dh_{A_d \rightarrow B_d} - dh_{A_m \rightarrow B_m}\|}{2} + \frac{\|dh_{B_d \rightarrow A_d} - dh_{B_m \rightarrow A_m}\|}{2}$$

Orientation Several methods have been proposed to define the directional relative position between two objects, which is an intrinsically vague notion. Particularly, fuzzy methods are appropriate, and here we choose to represent this information using histograms of angles [Miyajima et Ralescu \(1994\)](#). This allows representing all possible directional relations between two regions. If R_1 and R_2 are two sets of points $R_1 = p_1, \dots, p_n$ and $R_2 = q_1, \dots, q_n$, the relative position between regions R_1 and R_2 is estimated from the relative position of each point q_j of R_2 with respect to each point p_i of R_1 . The histogram of angles $H_{R_1 R_2}$ is defined as : $H_{R_1 R_2}(\theta) = \left| \{(p_i, q_j) \in R_1 \times R_2 / \angle(\vec{i}, \vec{p_i q_j}) = \theta\} \right|$ where $\angle(\vec{i}, \vec{p_i q_j})$ denotes the angle between a reference vector $\vec{i}$ and $\vec{p_i q_j}$. In order to speed up this computation, we compute histograms on the boundary of the objects. The histogram is normalized such that $\sum_{\theta} h[\theta] = 1$ in order to use the circular earth mover's distance (CEMD) defined in [Rabin et al. \(2008\)](#), i.e. the distance between normalized cumulative histograms derived from the angle histograms with a parameter μ to cope with periodicity. The CEMD is defined as : $cemd(f, g) = \|F - G - \mu\|_1$ where f and g are two histograms and F and G are cumulative histograms derived respectively from f and g . As shown in [Rabin et al. \(2008\)](#), μ is chosen as the median of the values $F(i) - G(i)$. The orientation cost is then defined as the absolute differences of CEMD :

$$C_{e_{orient.}}(E_m, E_d) = \|cemd(E_d) - cemd(E_m)\|$$

C.4.3 Connectivity

The previous features are more meaningful when a vertex in G_d represents regions forming a unique connected component. Therefore the edges between input graph vertices composing a vertex in the deformation graph should be taken into account too. In order to favor compact regions and to reduce the number of connected components, we derive a criterion based on the distance between all connected components present in a vertex of G_d :

$$C_{connectivity}(V_d) = \frac{\sum_{c_i \in V_d} (\sum_{c_j \in V_d, i \neq j} d(c_i, c_j))}{N_{cc}}$$

where c_i and c_j represent connected components in V_d , $d(c_i, c_j)$ is the maximal distance between c_i and c_j (symmetric), and N_{cc} is the number of connected components in V_d .

¹A cumulative histogram is computed as : $hc[i] = \sum_{j=1}^i h[j]$.

C.4.4 Cost function

Edge and vertex cost functions average all criteria as follows :

$$CV(V_m, V_d) = (Cv_{intensity} + Cv_{shape} + Cv_{area}) / 3$$

$$CE(E_m, E_d) = (Ce_{dist} + Ce_{orientation}) / 2$$

Finally, the cost function is a weighted mean between the vertex cost function, the edge cost function, and a connectivity cost defined as :

$$C = \alpha \sum_{V_d \in G_d} CV(V_m, V_d) + \beta \sum_{V_d \in G_d} C_{connectivity}(V_d) + \gamma \sum_{E_d \in E_d} CE(E_m, E_d) \quad (C.1)$$

where V_m is the vertex in G_m related to G_d , and $\alpha + \beta + \gamma = 1$.

C.5 Matching algorithm and optimization

As mentioned earlier, image segmentation is achieved by matching the input image (to be segmented) and the model. The input graph G_i is mapped onto the deformation graph G_d and the cost function for a given mapping is evaluated between G_d and G_m .

An initial mapping is mandatory to compute attributes carried by vertices and edges of G_d . This initial matching may be a random matching, but in order to reduce the computation time, initialization of G_d may also be carried out by applying a modified version of the segmentation method described in [Noma et al. \(2009\)](#). This initialization is achieved by matching each vertex of G_i to a vertex of G_m . The cost function evaluates the deformation between a vertex of G_m and the same vertex deformed by the candidate vertex of G_i . But since a region produced by the over-segmentation is directly compared with the model, this matching process only uses a simple image-based criterion based on a distance between grey levels (as in Section C.4.1) and a structural cost taking into account the centroids of the compared regions

The subsequent iterations minimize the cost function between G_d and G_m based on the high-level criteria explained in Section C.4. The search for better solutions is carried out by re-assigning each G_i vertex to different vertices of the deformation graph G_d in an attempt to reach lower cost values. For each re-assignment, the corresponding attributes in G_d (i.e. those associated to vertices and edges involved in the re-assignment) are recalculated, as well as the cost function. In order to speed up the computation, connected components may be re-assigned as a whole instead of a single region. When considering to move a region of G_i , the current matching of the region is a vertex of G_d . If this vertex has more than one connected component, then the whole connected component is changed. In both cases, all descriptors of the modified vertices are recalculated, as well as all edges connected to an updated vertex.

Two different optimization schemes may be used according to the initialization. With a random initialization, optimization is achieved by a simulated annealing algorithm. A vertex from G_i is selected randomly as well as the new matching which is accepted if the global cost decreases, or accepted with a probability depending on the temperature parameter otherwise. This parameter is decreased after N vertex selections, where $N = |V_i|$ is the number of vertices of the input graph G_i . In the case of a non-random initialization, optimization is achieved by an ICM scheme, i.e. a vertex from G_i is still selected randomly, but all possible matchings in G_d are computed and the

best matching is then kept. The process finishes if after N vertex selections, the energy remains the same.

Figure C.5 presents a summary of the implemented matching algorithm, where map represents the mapping between G_i and G_d (initially, to G_m , in order to initialize G_d). It is worth noting that this mapping actually represents the sought solution, i.e. each possible mapping defines a possible labelling of G_i (hence, a possible segmentation of the input image). The **while** loop implements the simulated annealing search

```

MATCHINGALGORITHM( $G_i, G_m$ )
1   $t \leftarrow \text{INITTEMPERATURE}()$ 
2   $map \leftarrow \text{INITIALMAP}(G_i, G_m)$ 
3   $G_d \leftarrow \text{INITDEFORMATION}(G_i, G_m, map)$ 
4   $c \leftarrow \text{COST}(G_d, G_m)$ 
5   $stopFlag \leftarrow \text{CONVERGENCETEST}()$ 
6  while ( $stopFlag \neq \text{TRUE}$ )
7      do
8          for ( $i \leftarrow 1$  to  $N$ )
9              do
10                  $map1 \leftarrow \text{CHANGESOLUTION}(map)$ 
11                  $G_d \leftarrow \text{UPDATEDEFORMATION}(G_i, G_d, map1)$ 
12                  $c1 \leftarrow \text{COST}(G_d, G_m)$ 
13                 if ( $\text{ACCEPTSOLUTION}(C, C_1, t)$ )
14                     then
15                          $map \leftarrow map1$ 
16                          $c \leftarrow c1$ 
17                     else
18                          $G_d \leftarrow \text{UPDATEDEFORMATION}(G_i, G_d, map)$ 
19   $t \leftarrow \text{UPDATETEMPERATURE}(t)$ 
20   $stopFlag \leftarrow \text{CONVERGENCETEST}()$ 

```

FIG. C.4 – The matching algorithm.

C.6 Experiments

Figure C.5 (first line) presents results between two close images of guitars. The model is composed by 5 classes. Two experiments have been performed, one with an input image where the guitar is approximatively at the same location as in the model, and another one where it is shifted to the right. The initialization already gives good results (but less in the shifted case). The first result corresponds to the input image, the second one to its shifted version. The method does not take into account centroids nor any absolute position attributes, being thus translation independent. However, the area is computed relatively to the image size, thus in this case, the values are different from the ones in the model. The second line of Figure C.5 presents another result between two images of motorcycles. The model is composed by 4 classes.

In both cases the initialization allows using an ICM optimization scheme. Results are not exactly the ideal segmentations, but all regions of the model are correctly found and the results

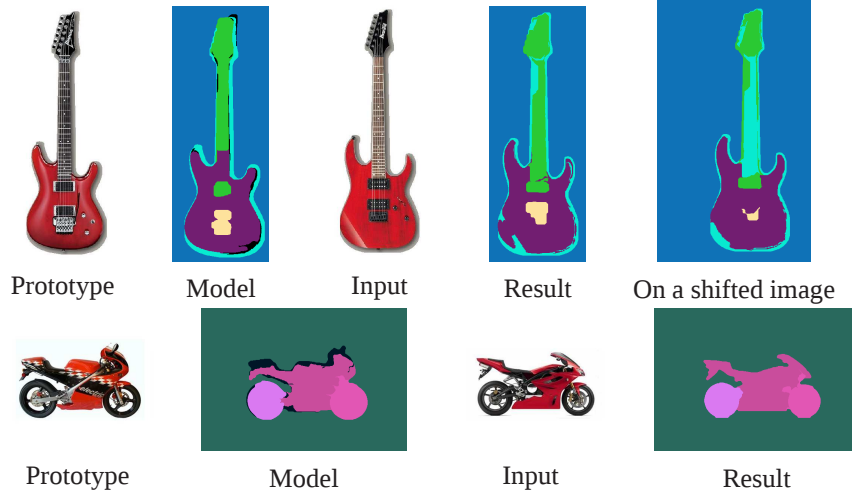


FIG. C.5 – Illustration of the method on two examples.

capture the user's intention. An ideal segmentation would give a lower cost than these results, meaning that the optimization process reaches a local minimum, which is however close to the desired one.



FIG. C.6 – Several experiments with model reuse. A guitar model is generated for the first guitar image and then used in the segmentation of other guitars of various shapes and colors.

Figure C.6 presents more results with a simpler model than the one present in Figure C.5 (there is no class for the shadow). In this experiment, the following parameters for the cost function are : $\alpha = 0.45$, $\beta = 0.35$, $\gamma = 0.20$. Results present the same difficulties with the guitar neck, which is not well defined by the model. When colors and shape differ from the prototype (the second line), the results are worst which illustrate the limitation of building a model based on a single image. In these case, the optimisation can only rely on the structural cost.

The cost function is designed by aggregating costs between many attributes and with different dynamics. There are many ways to combine them which deserve further analysis. A hint is to combine attributes like distance and orientation before cost computation using fuzzy representations of spatial relations [Bloch \(2005\)](#). The choice of the attributes carried by the vertices depends of the application. In our case, intensity and shape cost allows recognizing objects with similar features. The weight may be changed according to the image to segment. The process used for building the model produces large regions thanks to regularization, which are well suited for our purpose, but it is also driven by parameters which need to be set accordingly. A weaker regularization gives smaller regions and does not guarantee to approximate the corresponding objects. However, in all our experiments, the same values of these parameters give good results.

C.7 Conclusions

We proposed a method to segment an image using a model built according to the user's intention and able to merge inhomogeneous regions into a single segmentation class. The proposed model takes into account the structure of the prototype marked by the user. The introduction of the deformation graph allows separating the two problems of the inexact graph matching and of finding the best isomorphism with the model graph. It also allows computing high-level attributes like spatial relations and shape features.

C.8 Acknowledgment

This work has been partially funded by CAPES, COFECUB (546/07), CNPq, FAPESP and FINEP grants

Bibliographie

- A. ALBOODY, F. SEDES et J. INGLADA : Post-classification and spatial reasoning : new approach to change detection for updating gis database. *In 3rd International Conference on Information and Communication Technologies : From Theory to Applications (ICTTA)*, p. 1–7, April 2008.
- E. ALDEA : *Apprentissage de données structurées pour l'interprétation d'images*. Thèse de doctorat, Télécom ParisTech, Décembre 2009.
- E. ALDEA, J. ATIF et I. BLOCH : Image Classification using Marginalized Kernels for Graphs. *In 6th IAPR-TC15 Workshop on Graph-based Representations in Pattern Recognition, GbR'07*, vol. LNCS 4538, p. 103–113, Alicante, Spain, jun 2007a.
- E. ALDEA, G. FOUQUIER, J. ATIF et I. BLOCH : Kernel Fusion for Image Classification Using Fuzzy Structural Information. *In 3rd International Symposium on Visual Computing ISVC07*, vol. LNCS 4842, p. 307–317, Lake Tahoe, USA, nov 2007b.
- J. ALOIMONOS : Purposive and qualitative active vision. *In in the proceedings of the 10th International Conference on Pattern Recognition*, vol. 1, p. 346–360, Jun 1990.
- J. ALOIMONOS, I. WEISS et A. BANDYOPADHYAY : Active vision. *International Journal of Computer Vision*, 1(4):333–356, Jan 1988.
- J. ATIF, C. HUDELLOT, G. FOUQUIER, I. BLOCH et E. ANGELINI : From Generic Knowledge to Specific Reasoning for Medical Image Interpretation using Graph-based Representations. *In International Joint Conference on Artificial Intelligence IJCAI'07*, p. 224–229, Hyderabad, India, jan 2007a.
- J. ATIF, C. HUDELLOT, O. NEMPONT, N. RICHARD, B. BATRANCOURT, E. ANGELINI et I. BLOCH : GRAFIP : A Framework for the Representation of Healthy and Pathological Cerebral Information. *In IEEE International Symposium on Biomedical Imaging (ISBI)*, p. 205–208, Washington DC, USA, apr 2007b.
- J. ATIF, H. KHOTANLOU, E. ANGELINI, H. DUFFAU et I. BLOCH : Segmentation of Internal Brain Structures in the Presence of a Tumor. *In MICCAI Workshop on Clinical Oncology*, p. 61–68, Copenhagen, oct 2006a.
- J. ATIF, O. NEMPONT, O. COLLIOT, E. ANGELINI et I. BLOCH : Level Set Deformable Models Constrained by Fuzzy Spatial Relations. *In Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU*, p. 1534–1541, Paris, France, 2006b.
- R. BAJCSY : Active perception. *Proceedings of IEEE*, 76(8):996–1005, 1988.
- D. BALLARD et C. BROWN : Principles of animate vision. *CVGIP : Image Understanding*, 56(1):3–21, 1992. ISSN 1049-9660.
-

-
- E. BENGOTXEA, P. LARRANAGA, I. BLOCH, A. PERCHANT et C. BOERES : Inexact Graph Matching by Means of Estimation of Distribution Algorithms. *Pattern Recognition*, 35:2867–2880, 2002.
- K. BHATIA, J. HAJNAL, B. PURI, A. EDWARDS et D. RUECKERT : Consistent groupwise nonrigid registration for atlas construction. In *Biomedical Imaging : Macro to Nano*, p. 908–911, 2004.
- D. BLEZEK et J. MILLER : Atlas stratification. *Medical Image Analysis*, 11(5):443–457, 2007.
- I. BLOCH : Fuzzy Relative Position between Objects in Image Processing : a Morphological Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(7):657–664, 1999.
- I. BLOCH : Fuzzy Spatial Relationships for Image Processing and Interpretation : A Review. *Image and Vision Computing*, 23(2):89–110, 2005.
- I. BLOCH, O. COLLIOT et R. CESAR : On the Ternary Spatial Relation Between. *IEEE Transactions on Systems, Man, and Cybernetics SMC-B*, 36(2):312–327, apr 2006.
- I. BLOCH, T. GÉRAUD et H. MAÎTRE : Representation and fusion of heterogeneous fuzzy information in the 3d space for model-based structural recognition - application to 3d brain imaging. *Artificial Intelligence*, 148:141–175, 2003.
- B. BOUCHON-MEUNIER, M. RIFQI et S. BOTHOREL : Towards general measures of comparison of objects. *Fuzzy sets and Systems*, 84(2):143–153, 1996.
- D. BOWDEN et M. DUBACH : Neuronames 2002. *Neuroinformatics*, 1(1):43–59, 2003.
- D. BOWDEN et M. DUBACH : *Neuroanatomical Nomenclature and Ontology*, chap. Databasing the Brain. John Wiley and Sons, Inc., 2005.
- C. BROIT : *Optimal registration of deformed images*. Thèse de doctorat, University of Pennsylvania, Philadelphia, 1981.
- H. BUNKE : Recent developments in graph matching. In *Int. Conf. Pattern Recognition*, p. 2117–2124, 2000.
- R. CESAR, E. BENGOTXEA, I. BLOCH et P. LARRANAGA : Inexact Graph Matching for Model-Based Recognition : Evaluation and Comparison of Optimization Algorithms. *Pattern Recognition*, 38:2099–2113, 2005.
- D. COLLINS, A. ZIJDENBOS, W. BAARE et A. EVANS : Animal+ insect : Improved cortical structure segmentation. In *Information Processing in Medical Imaging*, vol. 1613, p. 210–223, Visegrád, Hungary, 1999. Springer.
- O. COLLIOT : *Representation, évaluation et utilisation de relations spatiales pour l'interprétation d'images. Applications à la reconnaissance de structures anatomiques en imagerie médicale*. Thèse de doctorat, ENST, 2003.
- O. COLLIOT, O. CAMARA et I. BLOCH : Integration of Fuzzy Spatial Relations in Deformable Models - Application to Brain MRI Segmentation. *Pattern Recognition*, 39:1401–1414, 2006.
- D. COMANICIU et P. MEER : Mean shift : A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002. ISSN 0162-8828.
-

-
- L. A. CONSULARO, R. M. CESAR et I. BLOCH : Structural Image Segmentation with Interactive Model Generation. *In IEEE International Conference on Image Processing (ICIP 2007)*, vol. 6, p. 45–48, San Antonio, Texas, USA, sep 2007.
- D. CONTE, P. FOGGIA, C. SANSONE et M. VENTO : Thirty years of graph matching in pattern recognition. *Int. J. Pattern Rec. and Art. Intell.*, 18(3):265–298, 2004.
- T. COOTES, G. EDWARDS et C. TAYLOR : Active appearance models. *Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.
- T. COOTES, C. TAYLOR, D. COOPER et J. GRAHAM : Active shape models-their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.
- G. COTTERET : *Extraction d'éléments curvilignes guidée par des mécanismes attentionnels pour des images de télédétection : approche par fusion de données*. Thèse de doctorat, University Paris XI, Orsay, France, 2005.
- D. CREMERS, F. TISCHHÄUSER, J. WEICKERT et C. SCHNÖRR : Diffusion snakes : Introducing statistical shape knowledge into the mumford-shah functional. *International Journal of Computer Vision*, 50(3):295–313, 2002.
- D. CREVIER et R. LEPAGE : Knowledge-based image understanding systems : A survey. *Computer Vision and Image Understanding*, 67(2):161 – 185, 1997. ISSN 1077-3142. URL <http://www.sciencedirect.com/science/article/B6WCX-45M8S4T-5/2/37d8bb1c2654ec6>
- A. CROSS et E. HANCOCK : Graph matching with a dual step EM algorithm. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(11):1236–1253, 1998.
- J. DARBON et M. SIGELLE : Image restoration with discrete constrained total variation. part i : Fast and exact optimization. *Journal of Mathematical Imaging and Vision*, 23(3):261–276, 2006a.
- J. DARBON et M. SIGELLE : Image restoration with discrete constrained total variation part ii : Levelable functions, convex and non-convex cases. *Journal of Mathematical Imaging and Vision*, 23(3):277–291, 2006b.
- C. DAUMAS-DUPORT : Histological grading of gliomas. *Current Opinion in Neurology and neurosurgery*, 5:924–931, 1992.
- B. DAWANT, S. HARTMANN et S. GADAMSETTY : Brain atlas deformation in the presence of large space-occupying tumors. *In Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, p. 589–596, Cambridge, UK, sep 1999a. Springer-Verlag London, UK.
- B. DAWANT, S. HARTMANN, S. PAN et S. GADAMSETTY : Brain atlas deformation in the presence of small and large space-occupying tumors. *Comput. Aided Surg.*, 7(1):1–10, 2002.
- B. DAWANT, S. HARTMANN, J. THIRION, F. MAES, D. VANDERMEULEN et P. DEMAEREL : Automatic 3-d segmentation of internal structures of the head in mr images using a combination of similarity and free-form transformations. i. methodology and validation on normal subjects. *IEEE Transactions on Medical Imaging*, 18(10):909–916, 1999b.
- A. DERUYVER, Y. HODÉ et L. BRUN : Image interpretation with a conceptual graph : Labeling over-segmented images and detection of unexpected objects. *Artif. Intell.*, 173(14):1245–1265, 2009. ISSN 0004-3702.
-

-
- A. DERUYVER et Y. HODÉ : Constraint satisfaction problem with bilevel constraint : application to interpretation of over-segmented images. *Artificial Intelligence*, 93(1-2):321–335, 1997.
- A. DESOLNEUX, L. MOISAN et J.-M. MOREL : *Gestalt Theory and Image Analysis : A Probabilistic Approach*. Springer-Verlag New York Inc., 2008.
- C. DOWNING et S. PINKER : The spatial structure of visual attention. *Attention and performance*, 1985.
- D. DUBOIS et H. PRADÉ : *Fuzzy Sets and Systems : Theory and Applications*. Academic Press, New-York, 1980.
- J. DUNCAN : Selective attention and the organization of visual information. *Journal of Experimental Psychology : General*, 113:501–517, 1984.
- J. DUNCAN : Boundary conditions on parallel search in human vision. *Perception*, 18:457–469, 1989.
- J. DUNCAN et G. HUMPHREYS : Visual search and stimulus similarity. *Psychological Review*, 96:433–458, 1989a.
- J. DUNCAN et G. HUMPHREYS : Visual search and stimulus similarity. *Psychological Review*, 3(96):433–458, 1989b.
- P. F. FELZENSZWALB et D. P. HUTTENLOCHER : Pictorial structures for object recognition. *Int. J. Comput. Vision*, 61(1):55–79, 2005.
- J. FLUSSER et T. S. SUK : Pattern recognition by affine moment invariants. *Pattern Recognition*, 26(2):167–174, January 1993.
- T. GÉRAUD, I. BLOCH et H. MAÎTRE : Atlas-guided Recognition of Cerebral Structures in MRI using Fusion of Fuzzy Structural Information. In *CIMAF'99 Symposium on Artificial Intelligence*, p. 99–106, La Havana, Cuba, 1999.
- T. GÉRAUD, I. BLOCH et H. MAÎTRE : Reconnaissance de structures cérébrales à l'aide d'un atlas à par fusion d'informations structurelles floues. In *RFIA 2000*, vol. I, p. 287–295, Paris, France, 2000.
- G. GRANLUND : Does vision inevitably have to be active ? In *Proceedings of the SCIA99, Scandinavian Conference on Image Analysis*, 1999.
- A. GUIMOND, J. MEUNIER et J. THIRION : Average brain models : A convergence study. *Computer Vision and Image Understanding*, 77(2):192–210, 2000.
- D. HASBOUN : Neuranat. <http://www.chups.jussieu.fr/ext/neuranat/index.html>, 2005.
- C. HEALEY. : Perception in visualization. Disponible en ligne : <http://www.csc.ncsu.edu/faculty/healey/PP/index.html>, 2007.
- Y. HODÉ et A. DERUYVER : Qualitative spatial relationships for image interpretation by using semantic graph. In *Graph-Based Representations in Pattern Recognition, GbRPR*, p. 240–250, Alicante, Spain, 2007.
-

-
- C. HUDELLOT, J. ATIF et I. BLOCH : Fuzzy Spatial Relation Ontology for Image Interpretation. *Fuzzy Sets and Systems*, 159:1929–1951, 2008.
- C. HUDELLOT, J. ATIF, O. NEMPONT, B. BATRANCOURT, E. ANGELINI et I. BLOCH : GRAFIP : a Framework for the Representation of Healthy and Pathological Anatomical and Functional Cerebral Information. In *Human Brain Mapping*, Florence, Italy, jun 2006.
- D. IOSIFESCU, M. SHENTON, S. WARELD, R. KIKINIS, J. DENGLER, F. JOLESZ et R. MCCARLEY : An automated registration algorithm for measuring mri subcortical brain structures. *Neuroimage*, 6(1):13–25, 1997.
- L. ITTI : Models of bottom-up attention and saliency. *Neurobiology of Attention*, 2005.
- L. ITTI : Visual salience. *Scholarpedia*, 2(9):3327, 2007.
- L. ITTI et C. KOCH : Feature combinaison strategies for saliency-based visual attention systems. *Journal of Electronic Imaging*, 10(1):161–169, 01 2001.
- L. ITTI, C. KOCH et E. NIEBUR : A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, Nov. 1998.
- W. JAMES : *The Principles of Psychology*, vol. 1. Dover Publications, 1890.
- S. JOSHI, B. DAVIS, M. JOMIER et G. GERIG : Unbiased diffeomorphic atlas construction for computational anatomy. *Neuroimage*, 23:151–160, 2004.
- B. JULÉSZ : Textons, the elements of texture perception, and their interactions. *Nature*, 290:91–97, 1981a.
- B. JULÉSZ : A theory of preattentive texture discrimination based on first-order statistics of textons. *Biological Cybernetics*, 41:131–138, 1981b.
- B. JULÉSZ et J. BERGEN : Textons, the fundamental elements in preattentive vision and the perception of textures. *Bell System Technical Journal*, 62(6):1619–1645, 1983.
- C. KANAN, M. TONG, L. ZHANG et G. COTTRELL : Sun : Top-down saliency using natural statistics. *Visual Cognition*, 17(6), 979-1003 2009.
- H. KHOTANLOU : *Segmentation 3D de tumeurs et de structures internes du cerveau en IRM*. Thèse de doctorat, ENST, 2008.
- H. KHOTANLOU, J. ATIF, E. ANGELINI, H. DUFFAU et I. BLOCH : Adaptive Segmentation of Internal Brain Structures in Pathological MR Images Depending on Tumor Types. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, p. 588–591, Washington DC, USA, apr 2007.
- H. KHOTANLOU, O. COLLIOT, J. ATIF et I. BLOCH : 3D Brain Tumor Segmentation in MRI Using Fuzzy Classification, Symmetry Analysis and Spatially Constrained Deformable Models. *Fuzzy Sets and Systems*, 160:1457–1473, 2009.
- C. KOCH et S. ULLMAN : Shifts in selective visual attention : towards the underlying neural circuitry. *Human Neurobiology*, 4(4):219–227, 1985.
-

- S. KYRIACOU, C. DAVATZIKOS, S. ZINREICH et R. BRYAN : Nonlinear elastic registration of brain images with tumor pathology using a biomechanical model [mri]. *Medical Imaging*, 18 (7):580–592, Jul 1999.
- F. Le BER, J. LIEBER et A. NAPOLI : Les systèmes à base de connaissances. In J. AKOKA et I. Comyn WATTIAU, édés : *Encyclopédie de l'informatique et des systèmes d'information*, p. 1197–1208. Vuibert, 2006. ISBN 978-2-7117-4846-4. URL <http://hal.inria.fr/inria-00201566/en/>.
- F. LE-BER et A. NAPOLI : The design of an object-based system for representing and classifying spatial structures and relations. *Journal of Universal Computer Science*, 8(8):751–773, 2002.
- M. LEVENTON, W. GRIMSON et O. FAUGERAS : Statistical shape influence in geodesic active contours. In *Computer Vision and Pattern Recognition*, vol. 1, p. 316–323, 2000.
- C. LIPSCOMB : Medical subject headings (mesh). *Bull Med Libr Assoc*, 88(3):265–266, Jul 2002.
- A. D. LUCA et S. TERMINI : A definition of non-probabilistic entropy in the setting of fuzzy set theory. *Information and Control*, 20:301–312, 1972.
- J. MACHROUH : *Perception attentive et vision en intelligence artificielle*. Thèse de doctorat, University Paris XI, Orsay, France, 2002.
- J.-F. MANGIN, O. COULON et V. FROUIN : Robust brain segmentation using histogram scale-space analysis and mathematical morphology. In *Medical Image Computing and Computer-Assisted Intervention*, p. 1230, 1998.
- J.-F. MANGIN, V. FROUIN, J. RÉGIS, I. BLOCH, P. BELIN et Y. SAMSON : Towards better management of cortical anatomy in multi-modal multi-individual brain studies. *Physica Medica*, 12 (Supplement 1):103–107, June 1996.
- J. MANGIN : Entropy minimization for automatic correction of intensity nonuniformity. In *Mathematical Methods in Biomedical Image Analysis*, p. 162–169, Hilton Head Island, South Carolina, USA, 2000.
- D. MARCUS, T. WANG, J. PARKER, J. CSERNANSKY, J. MORRIS et R. BUCKNER : Open access series of imaging studies (oasis) : Cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *Journal of Cognitive Neuroscience*, 19:1498–1507, 2007.
- D. MARR : Early processing of visual information. *Philosophical Transactions of the Royal Society of London*, 275:483–524, 1976.
- D. MARR : *Vision*. W. H. Freeman and Company, New York, 1982.
- T. MATSUYAMA et V. S.-S. HWANG : *SIGMA : a knowledge-based aerial image understanding system*. Plenum press, 1990.
- J. MAZZIOTTA, A. TOGA, A. EVANS, P. FOX et J. LANCASTER : A probabilistic atlas of the human brain : Theory and rationale for its development the international consortium for brain mapping (icbm). *Neuroimage*, 2(2PA):89–101, 1995.
- F. MEYER : An overview of morphological segmentation. *International journal of pattern recognition and artificial intelligence*, 15(7):1089–1118, 2001.
-

-
- K. MIYAJIMA et A. RALESCU : Spatial organization in 2d segmented images : representation and recognition of primitive spatial relations. *Fuzzy Sets and Systems*, 65(2-3):225–236, 1994. ISSN 0165-0114.
- A. MOHAMED, E. ZACHARAKI, D. SHEN et C. DAVATZIKOS : Deformable registration of brain tumor images via a statistical model of tumor-induced deformation. *Medical Image Analysis*, 10(5):752–763, 2006.
- B. MOTTER : Neural correlates of attentive selection for color or luminance in extrastriate area v4. *The Journal of Neuroscience*, 14(4):2178–2189, Apr 1994.
- H. MÜLLER, G. HUMPHREYS, P. QUINLAN et M. RIDDOCH : Combined-feature coding in the form domain. *Visual Search*, p. 47–55, 1990.
- U. NEISSER : *Cognitive psychology*. Appleton-Century-Crofts, 1967.
- U. NEISSER et R. BECKLEN : Selective looking : attending to visually specified events. *Cognitive Psychology*, 7:480–494, 1975.
- O. NEMPONT : *Modèles structurels flous et propagation de contraintes pour la segmentation et la reconnaissance d’objets dans les images. Application aux structures normales et pathologiques du cerveau en IRM*. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications, Mars 2009.
- A. NOMA, A. B. V. GRACIANO, R. M. CESAR-JR, L. A. CONSULARO et I. BLOCH : Inexact graph matching for segmentation and recognition of object parts. Rap. tech., São Paulo : MAC-IME-USP, 2009.
- e. a. P. AUER : A research roadmap of cognitive vision. ECVision : European Network for Research in Cognitive Vision Systems, 2005.
- H. PASHLER : *The psychology of attention*. MIT Press, 1998.
- A. PERCHANT : *Morphisme de graphes d’attributs flous pour la reconnaissance structurelle de scènes*. Thèse de doctorat, Ecole nationale supérieure des télécommunications, Paris, France, 2000.
- A. PERCHANT et I. BLOCH : Fuzzy Morphisms between Graphs. *Fuzzy Sets and Systems*, 128(2):149–168, 2002.
- K. POHL, J. FISHER, W. GRIMSON, R. KIKINIS et W. WELLS : A bayesian model for joint segmentation and registration. *Neuroimage*, 31(1):228–239, 2006.
- K. POHL, W. WELLS, A. GUIMOND, K. KASAI, M. SHENTON, R. KIKINIS, W. GRIMSON et S. WARELD : Incorporating non-rigid registration into expectation maximization algorithm to segment mr images. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, p. 564–572, Tokyo, Japan, 2002. Springer.
- M. POSNER : Orienting of attention. *The Quarterly Journal of Experimental Psychology*, 32(1):3–25, 1980.
- M. POSNER, C. SNYDER et B. DAVIDSON : Attention and the detection of signals. *Journal of Experimental Psychology : General*, 1980.
-

- F. POUPON, J.-F. MANGIN, D. HASBOUN, C. POUPON, I. MAGNIN et V. FROUIN : Multi-object deformable templates dedicated to the segmentation of brain deep structures. *In Medical Image Computing and Computer-Assisted Intervention*, vol. 1496, p. 1134, 2008.
- J. RABIN, J. DELON et Y. GOUSSEAU : Circular earth mover's distance for the comparison of local features. *In 19th Int. Conf. on Pattern Recognition*, Tampa, FL, USA, dec 2008.
- T. R. REED : Motion analysis using the 3-d gabor transform. *IEEE*, p. 506–509, 1997.
- A. RODRIGUEZ-SANCHEZ, E. SIMINE et J. TSOTSOS : Attention and visual search. *Int. J. Neural Systems*, 17(4):275–88, Aug 2007.
- C. ROSSE et J. L. MEJINO : *Anatomy Ontologies for Bioinformatics : Principles and Practice*, chap. The Foundational Model of Anatomy Ontology, p. 59–117. Springer, 2007.
- Y. RUBNER, C. TOMASI et L. GUIBAS : A metric for distributions with applications to image databases. *In Sixth International Conference on Computer Vision*, p. 59–66, Bombay, India, 1998.
- M. SAENZ1, G. BURACAS1 et G. BOYNTON1 : Global effects of feature-based attention in human visual cortex. *Nature Neuroscience*, 2002.
- B. SCHOLL : Objects and attention : the state of the art. *Cognition*, 80:1–46, 2001.
- D. SIMONS et C. CHABRIS : Gorillas in our midst : sustained inattention blindness for dynamic events. *Perception*, p. 1059–1074, 1999.
- J. SMIRNIOTOPOULOS : The new who classification of brain tumors. *Neuroimaging Clin N Am*, 9(4):595–613, 1999.
- J. TALAIRACH et P. TOURNOUX : *Co-Planar Stereotaxic Atlas of the Human Brain 3-Dimensional Proportional System : An Approach to Cerebral Imaging*. Thieme, 1988.
- P. THEVENAZ, T. BLU et M. UNSER : Interpolation revisited. *IEEE Transactions on Medical Imaging*, 19(7):739–758, 07 2000.
- A. TREISMAN : Preattentive processing in vision. *Comput. Vision Graph. Image Process.*, 31(2):156–177, 1985. ISSN 0734-189X.
- A. TREISMAN : Search, similarity, and integration of features between and within dimensions. *Journal of Experimental Psychology : Human Perception and Performance*, 17(3):652–676, 1991.
- A. TREISMAN et G. GELADE : A feature-integration theory of attention. *Cognitive Psychology*, 12:97–136, 1980.
- A. TREISMAN et S. GORMICAN : Feature analysis in early vision : Evidence from search asymmetries. *Psychological Review*, 95(1):15–48, 1988.
- S. TREUE et J. M. TRUJILLO : Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399(575-579), June 1999.
- M. TRIVEDI et A. ROSENFELD : On making computers see. *SMC*, 19(6):1333–1335, 1989.
-

-
- A. TSAI, W. WELLS, C. TEMPANY, E. GRIMSON et A. WILLSKY : Coupled multi-shape model and mutual information for medical image segmentation. *In Information Processing in Medical Imaging*, p. 185–197, Ambleside, UK, jul 2003. Springer.
- A. TSAI, W. WELLS, C. TEMPANY, E. GRIMSON et A. WILLSKY : Mutual information in coupled multi-shape model for medical image segmentation. *Medical Image Analysis*, 8(4):429–445, 2004.
- J. TSOTSOS : There is no one way to look at vision. *CVGIP : Image Understanding*, 60(1):95–97, 1994.
- D. VERNON : *Cognitive Vision Systems : Sampling the Spectrum of Approaches*, chap. The space of cognitive vision, p. 7–26. Springer, Heidelberg, 2006.
- D. VERNON : Cognitive vision : The case for embodied perception. *Image and Vision Computing*, 26(1):127 – 140, 2008. Cognitive Vision-Special Issue.
- C. VILLANI : *Topics in optimal transportation*. American Math. Soc., 2003.
- D. WALTHER et C. KOCH : Modeling attention to salient proto-objects. *Neural Networks*, 19 (9):1395–1407, Nov. 2006.
- Y. WANG et C. CHUA : Face recognition from 2d and 3d images using 3d gabor filters. *Image and Vision Computing*, 23:1018–1028, 2005.
- S. WAXMAN : *Correlative neuroanatomy*. McGraw-Hill, New York, 2000.
- J. WOLFE : Guided search 2.0 : A revised model of visual search. *Psychonomic Bulletin and Review*, 1(2):202–238, 1994.
- J. WOLFE : Visual search. *Attention*, p. 13–73, 1998.
- J. WOLFE, K. CAVE et S. FRANZEL : Guided search : An alternative to the feature integration model for visual search. *Journal of Experimental Psychology : Human Perception and Performance*, 15(3):419–433, 1989.
- J. WOLFE et T. HOROWITZ : What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, p. 495–501, June 2004.
- C. XU et J. PRINCE : Snakes, shapes, and gradient vector flow. *Image Processing, IEEE Transactions on*, 7(3):359–369, Mar 1998. ISSN 1057-7149.
- J. YANG et J. DUNCAN : 3d image segmentation of deformable objects with joint shape intensity prior models using level sets. *Medical Image Analysis*, 8(3):285–294, 2004a.
- J. YANG et J. DUNCAN : Joint prior models of neighboring objects for 3d image segmentation. *In Computer Vision and Pattern Recognition*, vol. 1, p. 314–319, Washington, DC, USA, Jul 2004b.
- A. YARBUS : *Eye movements and vision*. Plenum, New York, 1967.
- E. I. ZACHARAKI, D. SHEN, S.-K. LEE et C. DAVATZIKOS : Orbit : A multiresolution framework for deformable registration of brain tumor images. *Medical Imaging*, p. 1003–1017, Aug 2008.
-