



HAL
open science

Utilisation d'informations géométriques pour l'analyse statistique des données d'IRM fonctionnelle

Guillaume Flandin

► **To cite this version:**

Guillaume Flandin. Utilisation d'informations géométriques pour l'analyse statistique des données d'IRM fonctionnelle. Interface homme-machine [cs.HC]. Université Nice Sophia Antipolis, 2004. Français. NNT : 2004NICE4019 . tel-00633520

HAL Id: tel-00633520

<https://theses.hal.science/tel-00633520>

Submitted on 18 Oct 2011

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.

THÈSE

pour obtenir le titre de

Docteur en Sciences
de l'Université de Nice-Sophia Antipolis
Spécialité : Informatique

présentée et soutenue par

Guillaume FLANDIN

Utilisation d'informations géométriques pour l'analyse statistique des données d'IRM fonctionnelle

Thèse dirigée par : Nicholas AYACHE

Soutenue publiquement le 1^{er} Avril 2004

Composition du jury

<i>Président :</i>	Michel COSNARD	<i>Professeur</i>
<i>Rapporteurs :</i>	Christian BARILLOT Line GARNERO	<i>Directeur de Recherche</i> <i>Directrice de Recherche</i>
<i>Examineurs :</i>	Xavier PENNEC Jean-Baptiste POLINE	<i>Chargé de Recherche</i> <i>Ingénieur de Recherche</i>

Institut National de Recherche en Informatique et Automatique



Mis en page avec la classe thloria.

Remerciements

Une thèse est un travail de recherche qui se nourrit d'interactions, de collaborations et de rencontres avec une foultitude de personnes et c'est un grand plaisir pour moi de pouvoir exprimer ici ma gratitude à tous ceux qui ont fait que le mémoire que vous tenez entre les mains (ou lisez sur votre bel écran) ait pu voir le jour.

Je prie par avance ceux que je vais oublier de bien vouloir m'en excuser !

Tout d'abord je tiens à remercier chaleureusement Michel Cosnard pour m'avoir fait l'honneur de présider mon jury de thèse ainsi que Christian Barillot et Line Garnero pour avoir accepté d'en être les rapporteurs et pour le temps qu'ils ont consacré à la lecture et l'évaluation de ce manuscrit malgré les courts délais que je leur avais imposés.

J'adresse mes sincères remerciements à Nicholas Ayache, qui a dirigé ces travaux et m'a permis, en m'accueillant au sein du projet Epidaure à l'INRIA Sophia Antipolis, de faire partie d'une équipe de recherche de haut niveau, que ce soit d'un point de vue scientifique, matériel ou humain.

Je tiens à vivement remercier Xavier pour son encadrement. Ses capacités à formaliser proprement mes idées embrouillées m'ont été fort utiles et il a toujours répondu présent lorsque j'avais besoin de ses conseils éclairés. J'adresse également mes plus vifs remerciements à Grégoire – un vrai Tonton Flingueur ! Ses conseils sont toujours d'une rare pertinence... Un merci aussi tout particulier à Isabelle, pour sa sympathie et son soutien logistique sans faille. Enfin, merci à tous les membres de l'équipe Epidaure, anciens ou présents, pour le plaisir que j'ai eu à travailler avec eux à Sophia Antipolis: Eric, Jean-Didier, Jonathan, Olivier C & C, Maxime, Marc, Céline, Clément, Valérie, David, Sébastien G & O, Hervé, Alexis, Sylvain, Radu, Miguel, Guillaume DP & P, Alain, Vincent, Pascal, Mauricio, Rasmus, Françoise, Stéphane, Pierre-Yves, Oliver, Walid, Gérald, Cédric, Thibaud. Je n'oublie pas Janet et Rupert pour leur délicieux accent. Et merci Eve pour ta joie de vivre !

J'en arrive au Service Hospitalier Frédéric Joliot... Vaste sujet... J'y ai passé ma première année de thèse et j'en garde un souvenir tout simplement inoubliable et merveilleux. Quel regroupement de talents et de personnalités si attachantes ! Merci tout d'abord, bien sûr, à Jean-Baptiste qui m'a co-encadré – nombre des idées présentées dans ce mémoire sont le fruit de ses réflexions. Son soutien infaillible tant au niveau scientifique que personnel, ses qualités humaines, son amitié m'ont touché et me touchent encore énormément. Merci JB, tu devrais être le premier cité sur cette page ! Merci tout particulier à mon comparse parisien Ferath ; on a refait le monde (et SPM) plusieurs fois... Et merci du fond du cœur à toute la *dream team* du SHFJ, soit dans le désordre le plus absolu : Jeff, Philippe, Denis R & LB, Dimitri (*mais y en a marre!*), Yann (*BrainVisa, c'est au-dessus de tout ça!*), Alexandre, Isabelle, Gwénaelle, Pierre-Jean, Stanislas, Olivier, Arnaud, Jessica, Steren, Carlos, Vincent F & L, Anne-Lise, Julien, Renaud, Franck, Thierry, Paola, Giorgia, Régine, Christophe, Nicolas.

Merci à Matthew Brett (malgré les "*petites cellules qui bougent*") pour sa sympathie, ses talents comiques et pour l'hébergement lors de mes escapades londoniennes. Merci aussi à Jean-Luc Anton pour mes séjours à Marseille.

Je remercie également Will Penny, John Ashburner et Karl Friston du Functional Imaging Laboratory de Londres pour leur accueil et leur bienveillance. Attention, j'arrive....

Merci à Christian Raffaele et Thierry Viéville pour leur aide très appréciée au moment de ma soutenance.

Je dédie cette thèse à mes parents, mes grands-parents et ma sœur, que je remercie énormément. Se savoir soutenu et aimé est un bien précieux.

Enfin, merci Emma pour les tous bons moments partagés et pour ton soutien dans les moments difficiles.

*À mes parents,
À mes grands-parents,
À ma sœur*



La Joie de Vivre ou Antipolis, 1946 © Succession Picasso.

Table des matières

Chapitre 1 Introduction

Chapitre 2 L'imagerie cérébrale fonctionnelle

2.1	La cartographie fonctionnelle	12
2.1.1	Introduction	12
2.1.2	Bref historique	13
2.2	L'anatomie cérébrale	16
2.2.1	Architecture macroscopique	16
2.2.2	Architecture microscopique	18
2.3	Les techniques d'imagerie fonctionnelle	20
2.3.1	Les techniques électromagnétiques	21
2.3.2	Les techniques métaboliques et hémodynamiques	21
2.4	L'imagerie par résonance magnétique fonctionnelle (IRMf)	22
2.4.1	Principe physique de l'IRM	22
2.4.2	Bases physiologiques du signal en IRMf	26
2.4.3	Caractéristiques et artefacts des images IRMf-BOLD	28
2.4.4	Protocole expérimental	31

Chapitre 3 Méthodes d'analyse des images fonctionnelles

3.0	Notations	34
3.1	Prétraitements	34
3.1.1	Correction des distorsions géométriques	35
3.1.2	Correction du délai temporel intercoupe (<i>slice timing</i>)	36
3.1.3	Correction du mouvement	37

3.1.4	Mise en correspondance des images anatomiques et fonctionnelles	40
3.1.5	Lissage spatial des données	40
3.2	Normalisation spatiale des images	41
3.2.1	Présentation du problème	41
3.2.2	Corrélation anatomo-fonctionnelle	42
3.2.3	Algorithmes de normalisation spatiale	45
3.2.4	Labellisation des activations fonctionnelles	53
3.3	Analyse statistique des données fonctionnelles	56
3.3.1	Problématique	56
3.3.2	Le modèle linéaire général	58
3.3.3	Inférence statistique	66
3.3.4	Méthodes exploratoires	67
3.3.5	Les analyses de groupe	69
3.4	Prise en compte d'une information spatiale	70
3.4.1	Analyse par région d'intérêt	70
3.4.2	Modèles régularisés spatialement	72
3.4.3	Analyse sur la surface corticale	72
3.5	Conclusion	76

Chapitre 4

Classification de données spatiales

4.1	Cluster Analysis	78
4.1.1	Unsupervised learning	78
4.1.2	Parametric Models	79
4.1.3	Self-Organizing Maps	80
4.1.4	Hierarchical Clustering	80
4.2	Mixture models	81
4.2.1	Maximum Likelihood and EM algorithm	81
4.2.2	Gaussian Mixture Models	84
4.2.3	Model Order Selection	96
4.2.4	Classification	96
4.2.5	Incorporating Spatial Priors	98
4.3	Clustering methods in fMRI data analysis	100
4.3.1	Definition of the feature space	100
4.3.2	Introduction of spatial information	101

4.4	Conclusion	101
-----	----------------------	-----

Chapitre 5

Parcellisation anatomique du cortex

5.1	Motivations	104
5.2	Segmentation des structures cérébrales	105
5.2.1	Segmentation du néocortex à partir d'IRM T_1	106
5.2.2	Segmentation du cerveau à partir d'IRM T_2^*	107
5.2.3	Obtention du domaine à parcelliser	111
5.3	Parcellisation d'un domaine convexe	112
5.3.1	Parcellisation et critère à optimiser	112
5.3.2	L'algorithme K-Means	114
5.3.3	Transformations de distance	116
5.3.4	Mise à jour des graines	117
5.3.5	Résultats obtenus	118
5.4	Prise en compte de distances géodesiques	121
5.4.1	Diagramme de Voronoï géodésique	122
5.4.2	Calcul du centre de masse géodésique	125
5.4.3	Exemples de parcellisations obtenues	127
5.5	Introduction de contraintes spatiales	130
5.5.1	Introduction de l'information <i>sillons</i>	130
5.5.2	Imposition de la position de certaines graines	133
5.6	Conclusion	135

Chapitre 6

Applications des parcellisations anatomiques

6.1	Introduction	138
6.2	Détection d'activations sur une parcellisation	138
6.2.1	Préambule à propos de la résolution spatiale des images fonctionnelles	138
6.2.2	Jeux de données, paradigme et modèle d'analyse	139
6.2.3	Méthode de parcellisation, affectation du signal fonctionnel et détection	140
6.2.4	Comparaison entre la détection à la résolution voxel et celle à la résolution parcelle	143
6.2.5	Discussion	147

6.3	Sélection de modèle par parcelle	149
6.3.1	Motivations	149
6.3.2	Une modélisation régionale des données	150
6.3.3	Application sur un jeu de données	151
6.4	Mesures robustes de la connectivité fonctionnelle	152
6.4.1	La connectivité fonctionnelle	153
6.4.2	Problème de dimensionnalité	153
6.4.3	Etude de la nature des interactions entre parcelles	154
6.4.4	Une nouvelle mesure de connectivité	155
6.4.5	Discussion	155
6.5	Conclusion	156

Chapitre 7

Modélisation anatomo-fonctionnelle des données IRMf

7.1	Motivations	160
7.2	Un modèle de mélange spatio-temporel	160
7.2.1	Le modèle mathématique général	161
7.2.2	Adaptations pratiques	162
7.3	Exemples d'applications	164
7.3.1	Classification anatomique pure	164
7.3.2	Classification fonctionnelle pure	164
7.3.3	Classification spatio-temporelle	167
7.4	Extension multi-sujets	172
7.4.1	Motivations et modèle proposé	172
7.4.2	Classification fonctionnelle multi-sujets	173
7.4.3	Modèle anatomo-fonctionnel multi-sujets	174
7.5	Discussion	175

Chapitre 8

Conclusions et Perspectives

Annexes

Annexe A

L'inégalité de Jensen

Annexe B

L'algorithme EM vu comme une minimisation alternée

Annexe C

Dérivation matricielle

Annexe D

Développements logiciel sous MATLAB

Bibliographie

193

Chapitre 1

Introduction

Organisation du mémoire

L'imagerie par résonance magnétique fonctionnelle (IRMf) est une modalité récente permettant de mesurer *in vivo* l'activité neuronale de sujets sains ou de patients et de localiser les différentes aires mises en œuvre lors de tâches cognitives. Dans ce document, nous nous intéressons à l'utilisation de l'information anatomique cérébrale pour analyser les données fonctionnelles. Après avoir présenté dans le **chapitre 2** une synthèse des buts et des modalités d'acquisition de l'imagerie fonctionnelle, nous développons dans le **chapitre 3** les méthodes permettant de les analyser. Nous détaillons en particulier l'ensemble des étapes de la chaîne de traitement qui mettent en évidence les activations liées au paradigme expérimental. Il apparaît que l'analyse univariée classique ne prend pas en compte la corrélation spatiale et que la correction potentielle effectuée a posteriori réduit nettement la sensibilité des tests statistiques. De plus, dans le cas d'analyse multi-sujets – nécessaire pour obtenir des résultats significatifs au niveau d'une population – la normalisation spatiale forcément imparfaite conduit à une perte de sensibilité qui n'autorise que la détection des activations les plus importantes avec une localisation restreinte. L'alternative actuelle consiste en une analyse par régions d'intérêt. Celle-ci ne pose pas ces problèmes de sensibilité mais les résultats sont extrêmement grossiers car limités de par la définition même des régions considérées. Il y a donc nécessité d'une approche intermédiaire et qui puisse inclure de plus une information anatomique.

Le **chapitre 4** introduit les techniques de classification de données que nous utiliserons par la suite, en particulier les algorithmes K-Means et EM. Nous présentons ensuite un état de l'art de leur utilisation en analyse exploratoire multivariée pour l'IRMf où l'on s'aperçoit que l'aspect temporel est traité de manière tout à fait privilégié par rapport à l'aspect spatial.

Nous nous attachons donc dans le **chapitre 5** à réaliser une parcellisation anatomique permettant d'obtenir une résolution intermédiaire ajustable entre le voxel et la région d'intérêt. Nous avons développé pour cela une méthode de parcellisation reposant sur des diagrammes de Voronoï centroïdaux obtenus grâce à un algorithme de type K-Means. Une de nos contributions principales est de modifier cet algorithme pour prendre en compte une distance géodésique modélisant l'information anatomique : les cellules sont contraintes à

rester dans l'épaisseur du cortex et à ne pas traverser les sillons. Une seconde contribution concerne l'implémentation, qui s'avère particulièrement efficace par rapport à l'état de l'art grâce à l'usage de la géométrie discrète. La portée de ces contributions est très générale puisque ce type de parcellisation est utilisé dans de nombreux autres domaines, comme par exemple en géométrie algorithmique.

Après avoir encodé l'information anatomique grâce à cette parcellisation, nous affectons un signal issu de l'IRM fonctionnelle à chaque parcelle, dans le **chapitre 6**. Ceci apporte une réduction importante de la quantité de données et réalise un lissage spatial intrinsèque dont le support est anatomiquement valide, contrairement au filtrage Gaussien isotrope. Cet effet est illustré dans la première application, avec une détection classique des activations, mais effectuée sur les parcelles. Même en utilisant une correction très conservative pour les comparaisons multiples, on observe une augmentation de la sensibilité, ce qui met en évidence la compacité et la concision de notre représentation. Dans une seconde application, nous mettons en place une modélisation des données qui varie spatialement. Le meilleur modèle local est sélectionné lors d'une première session (phase d'estimation), ce qui permet d'augmenter considérablement la sensibilité et la précision de la détection lors d'une seconde session. L'utilisation d'une parcellisation permet ici une implémentation particulièrement simple et efficace de la notion de localité. C'est d'ailleurs ce qui nous a permis de réaliser cette expérience démontrant l'inadéquation d'un modèle global unique. Avec la dernière application, nous abordons les questions de connectivité fonctionnelle. L'introduction des parcelles permet ici de travailler sur l'ensemble des données tout en conservant une notion de localité spatiale, ce qui n'était pas le cas auparavant avec une réduction de type ACP ou une sélection manuelle des zones d'intérêt. Au travers de toutes ces applications, on peut voir que la parcellisation spatiale est un outils puissant pour découvrir et développer de nouvelles méthodologies d'analyse des séquences IRMf.

Au **chapitre 7**, nous revenons sur la classification des signaux IRMf avec des modèles de mélanges, résolus par un algorithme de type EM. Ceci nous permet ainsi d'explorer l'espace des réponses à plusieurs tâches en évitant l'explosion combinatoire générée par l'ensemble des contrastes de tâche possibles. Nous augmentons ensuite ce modèle de mixture avec les coordonnées spatiales pour obtenir une technique qui généralise à la fois l'approche purement signal et la parcellisation anatomique du chapitre précédent. La forme des parcelles s'adapte ainsi non seulement à l'anatomie, mais aussi au signal fonctionnel, afin de rendre compte de l'ensemble des données observées de manière la plus homogène et parcimonieuse possible. Ce cadre s'adapte naturellement aux analyses multi-sujets en fournissant des parcelles homologues non seulement spatialement mais aussi fonctionnellement. Les premiers résultats illustrent qu'une même "aire d'activation" n'est pas exactement située au même endroit chez tous les sujets dans un repère standard, prouvant ainsi que notre méthode est potentiellement capable de contourner les hypothèses de corrélation anatomo-fonctionnelle et d'homologie purement anatomiques. Ceci ouvre de nombreuses perspectives, que nous évoquons dans le **chapitre 8**, et qu'il s'agit maintenant d'explorer.

Contributions

Nous proposons dans cette thèse une approche novatrice par parcellisation pour l'analyse des données IRMf qui s'est résolument placées à la frontière de deux domaines : les méthodes géométriques, plutôt développées en imagerie médicale, et les méthodes statistiques, particulièrement évoluées en neuro-imagerie et neurosciences. C'est donc naturellement que notre travail s'est déroulé au sein de deux équipes dont ce sont les axes de recherche privilégiés : l'équipe Epidaure à l'INRIA Sophia Antipolis, et l'unité de neuro-anatomie fonctionnelle du CEA-SHFJ à Orsay. Il en est résulté une méthodologie simple et efficace qui comble un vide du côté de l'utilisateur en lui permettant d'adapter la résolution de l'analyse au niveau auquel il évalue et interprète les résultats (l'amas ou *cluster* d'activation). Cette approche s'avère également être un méta-outil puissant pour découvrir et développer de nouvelles méthodologies d'analyse des séquences IRMf. De manière plus précise, on pourra trouver au fil du manuscrit les contributions suivantes.

- Dans le chapitre 5, l'idée même d'utiliser une parcellisation anatomique est une première contribution. Une seconde contribution importante repose dans la mise en œuvre pratique, en particulier avec l'inclusion des distances géodésiques permettant d'encoder l'information anatomique. Ces travaux ont été publiés dans [Flandin *et al.*, 2002c; 2002b].
- Avec la première application du chapitre 6, la détection d'activations, nous démontrons une augmentation de la sensibilité qui met en évidence la pertinence de notre représentation [Flandin *et al.*, 2002a]. Du point de vue de l'utilisateur, nous pensons que c'est cette contribution qui peut avoir le plus d'impact, en permettant des analyses à une résolution intermédiaire entre voxel et région d'intérêt.
- Avec la seconde application du chapitre 6, nous utilisons la parcellisation comme un méta-outil pour démontrer qu'un modèle spatialement adaptatif décrit bien plus précisément la variabilité du signal BOLD au sein du cerveau que le modèle global classique [Roche *et al.*, 2003]. Cette expérience repose également sur des travaux antérieurs concernant la sélection de modèles en analyse multivariée [Kherif *et al.*, 2002c] qui ne sont pas décrits dans ce manuscrit.
- La dernière application du chapitre 6 utilise encore la parcellisation anatomique comme un méta-outil, mais cette fois-ci pour l'étude de la connectivité fonctionnelle : l'analyse de la nature des interactions temporelles entre les signaux spatialement distants dans le cerveau nous a conduit à proposer une nouvelle mesure de similarité fonctionnelle [Lahaye *et al.*, 2003b; 2003a]. La contribution majeure de notre outil est ici de permettre de travailler sur l'ensemble du cerveau, ce qui n'était pas possible avant.
- Enfin, le chapitre 7 généralise la parcellisation dans un cadre spatio-temporel, réalisant ainsi pleinement la fusion géométrie/statistiques [Flandin *et al.*, 2003]. La contribution sera ici majeure pour l'utilisateur puisqu'en neurosciences, les questions se posent au niveau des populations. A l'heure actuelle, les analyses multi-sujets retournent souvent des résultats statistiquement peu significatifs, vraisemblablement à cause du problème de la normalisation spatiale. Notre méthodologie permettant

de contourner ces hypothèses d'homologie entre voxels, nous nous attendons à des développements importants et fructueux dans les prochaines années.

- Pour finir, nous voudrions souligner quelques contributions logicielles, récapitulées très rapidement dans l'annexe D. Il s'agit essentiellement de **Parcel It Up!**, une bibliothèque de fonctions MATLAB implémentant une grande partie de nos algorithmes de parcellisation, actuellement utilisée au CEA-SHFJ et destinée à être intégrée dans BrainVisa. Citons également **XMLTree** et **M2HTML**, deux boîtes à outils MATLAB en libre accès sur le site de MathWorks et téléchargées plus de 4000 fois au total en février 2004.

Certaines de nos contributions n'ont pas trouvé place dans ce manuscrit. Il s'agit essentiellement d'une collaboration avec F. Kherif sur la définition de mesure de similarité spatio-temporelles sur les IRM fonctionnelles, reposant sur des méthodes d'analyse multivariées, et dans le but de sélectionner les sujets dont les réponses sont comparables lors d'analyses de groupe [Kherif *et al.*, 2002a; 2002b; 2003]. On peut citer également notre participation à une expérience de rétinotopie, à la fois comme sujet, ce qui nous a permis de mieux appréhender le fonctionnement pratique d'un protocole d'imagerie mentale [Klein *et al.*, 2003], et comme analyste, ce qui a amené à comprendre les mécanismes d'analyse tout au long de la chaîne de traitement des données.

Chapitre 2

L'imagerie cérébrale fonctionnelle

Sommaire

2.1	La cartographie fonctionnelle	12
2.1.1	Introduction	12
2.1.2	Bref historique	13
2.2	L'anatomie cérébrale	16
2.2.1	Architecture macroscopique	16
2.2.2	Architecture microscopique	18
2.3	Les techniques d'imagerie fonctionnelle	20
2.3.1	Les techniques électromagnétiques	21
2.3.2	Les techniques métaboliques et hémodynamiques	21
2.4	L'imagerie par résonance magnétique fonctionnelle (IRMf)	22
2.4.1	Principe physique de l'IRM	22
2.4.2	Bases physiologiques du signal en IRMf	26
2.4.3	Caractéristiques et artefacts des images IRMf-BOLD	28
2.4.4	Protocole expérimental	31

Ce premier chapitre est une introduction au domaine d'application qui a motivé tous nos travaux : les neurosciences, c'est-à-dire l'étude des relations entre le cerveau et les fonctions mentales. Au cours de l'histoire, différents courants de pensée ont proposé leurs modèles pour relier anatomie et fonction. En effet, si l'anatomie de l'homme a pu être étudiée précisément grâce à de nombreuses dissections, les processus mentaux sont beaucoup plus complexe à étudier, en l'absence d'outils de mesure adéquats *in vivo*. L'imagerie cérébrale a permis aux cognitivistes d'avoir à leur disposition de nouvelles techniques d'exploration de l'activité cérébrale *in vivo*, dont nous présentons ici quelques éléments. A noter que le lecteur trouvera d'excellentes introductions à ce domaine dans [Mazoyer, 2001; Houdé *et al.*, 2002].

2.1 La cartographie fonctionnelle

2.1.1 Introduction

L'étude des processus mentaux et de leur substrat biologique a intéressé au cours des siècles de nombreux corps de métier : philosophes, biologistes, anatomistes, psychologues, physiologues, neurologues, etc. Chacun a abordé cette question selon son point de vue, créant tout autant de domaines scientifiques. Deux grands courants se distinguent pourtant : d'un côté les neurosciences privilégiant la connaissance de l'anatomie cérébrale afin de comprendre, à terme, ses propriétés fonctionnelles, et de l'autre la psychologie qui cherche à comprendre les mécanismes de la pensée grâce à des techniques expérimentales comportementalistes (la psychophysique¹ par exemple) servant à confirmer ou infirmer des modèles théoriques du fonctionnement cérébral. D'autres disciplines se trouvent à l'interface de ces deux courants comme la neuropsychologie et la psychophysologie. La première cherche à étudier les relations entre un dysfonctionnement cognitif et une pathologie physique (lésion d'une structure cérébrale) tandis que la deuxième se fonde sur l'étude couplée des événements mentaux et cérébraux locaux (psychophysique, électrophysiologie).

On peut faire un parallèle entre ces différentes techniques d'exploration du cerveau et les techniques de *reverse engineering*. Supposons ainsi qu'un industriel ait à sa disposition un composant électronique fabriqué par un de ses concurrents dont il souhaiterait connaître le fonctionnement pour le fabriquer lui-aussi. Il peut soit démonter le composant et étudier le circuit intégré au microscope pour essayer d'en déduire le schéma électronique (neurosciences), soit affecter une équipe d'ingénieur à la conception d'un modèle (un organigramme) des différentes fonctions que doit nécessairement posséder le composant pour effectuer sa tâche et vérifier par des mesures la pertinence du modèle (psychologie expérimentale). Il est également possible de rechercher quel est l'élément d'un composant défaillant qui se distingue par sa différence avec ceux de composants en bon état et ainsi comprendre l'utilité de cet élément au sein du composant (neuropsychologie). Enfin, il est possible de mesurer grâce à un oscilloscope les variations de courant (tension et in-

1. Psychophysique: mesure physique de réponses comportementales comme les temps de réponse ou le taux d'erreur.

tensité) intervenant au cours du fonctionnement du composant électronique et en déduire des informations utiles (psychophysologie). On voit déjà ici les dangers de conclusion hâtive concernant une expérience comportementale : les problèmes de causalité. Prenons un exemple simple et considérons un ordinateur réduit ici à son unité centrale et dont on ne connaît pas les fonctions de ses éléments (processeur, carte vidéo, mémoire, etc). Si, à la recherche du composant effectuant les calculs numériques, on enlève le ventilateur situé sur le processeur pendant que l'ordinateur est en marche, celui-ci va immédiatement s'arrêter pour éviter toute surchauffe. Il serait cependant absurde de conclure de cette expérience que le composant "moteur" de l'ordinateur est le ventilateur car celui-ci est nécessaire à son fonctionnement...d'où le soin particulier qui doit être pris lors de la conception de protocoles expérimentaux pour que l'expérience puisse répondre effectivement à une question précise. À noter que l'on peut également faire un parallèle avec d'autres domaines telle la génétique, avec d'un côté le séquençage complet du génome humain et de l'autre l'étude des pathologies engendrées par un gène déficient, absent ou surnuméraire.

2.1.2 Bref historique

L'histoire des relations entre le cerveau (anatomie) et la pensée (fonctions cognitives) a connu de nombreuses errances depuis l'Antiquité, principalement en raison de l'absence de techniques permettant d'observer le cerveau au cours de son fonctionnement. C'est ainsi que les philosophes grecs de l'Antiquité (Platon, Aristote, Hérophile) proposèrent différents organes pour situer le siège des fonctions mentales : le cerveau, les ventricules ou même le cœur. Ce n'est que bien plus tard, pendant la Renaissance, que la question de la relation entre le corps et l'esprit connut une évolution grâce aux premières dissections. Ainsi, Descartes (1596-1660) proposa que le corps pouvait affecter l'esprit et réciproquement. Selon lui, la glande pinéale (l'épiphysse) était le siège de l'âme humaine. Bien que l'épiphysse joue un rôle important dans le fonctionnement humain, cette affirmation nous apparaît aujourd'hui erronée mais, par cette localisation, Descartes évoquait un lien fort entre l'esprit et l'entité cerveau-système nerveux.

C'est au XIXe siècle que l'hypothèse d'une corrélation entre anatomie et fonction vit le jour. La théorie de la localisation des fonctions cérébrales - la notion qui spécifie que les processus mentaux sont corrélés avec des régions précises du cerveau - et les tentatives d'établir des localisations grâce à des observations empiriques ont eu alors un essor considérable. Un pas important fut franchi par Franz Gall (1758-1828) avec l'élaboration de la phrénologie [Greenblatt, 1995]. Selon cette théorie, les caractéristiques psychiques de l'homme sont liées avec la conformation externe du crâne. Plus précisément, la phrénologie formule l'hypothèse d'une localisation précise des facultés et d'une projection apparente de ces localisations sur le crâne (voir figure 2.1). Ainsi, Gall proposait de déterminer les qualités morales et intellectuelles d'un individu en observant la morphologie de son crâne. Il s'agit là d'une dérive très contestable de cette théorie (on parle alors de pseudo-science), mais malgré tout, Gall fut le premier à proposer l'étude de l'activité cérébrale basée sur l'hypothèse d'une fragmentation de cette activité en fonctions élémentaires et est ainsi considéré comme le premier cartographe des fonctions cognitives.

C'est au cours de ce siècle que se développa la neuropsychologie, fondée sur l'obser-

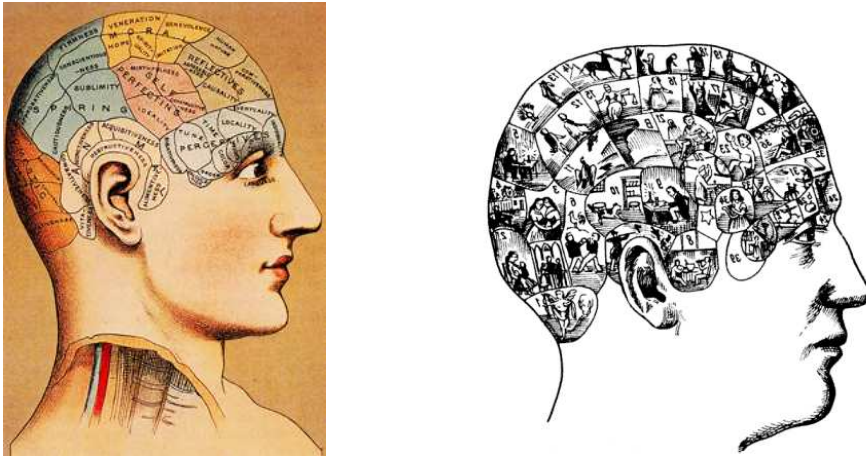


FIG. 2.1 – Cartes de la phrénologie localisant 35 facultés mentales et morales comme l'amour parental, l'estime de soi, l'amitié, l'espoir, la combativité... D'après <http://www.epub.org.br/cm/n01/frenolog/>

vation des lésions du cerveau et de son dysfonctionnement, dans le but d'effectuer la localisation cérébrale des fonctions mentales. Certains patients sont restés célèbres, tel Phineas Gage. Ce contremaître d'une équipe de construction de chemin de fer américaine fut gravement blessé en 1848 lors d'un accident au cours duquel une barre de fer traversa son crâne de part en part. La figure 2.2 présente un dessin et une reconstruction tridimensionnelle des dommages causés au crâne et au cerveau de Gage. Malgré leur importance au niveau du lobe frontal, celui-ci survécut à l'accident mais sa personnalité et son comportement se modifia irrémédiablement : il était dorénavant devenu changeant, irrévérencieux, capricieux, impatient, incapable de planifier une action à l'avance, etc. A sa mort, 12 ans plus tard, aucune autopsie ne fut pratiquée, ce qui limita à l'époque les interprétations et conclusions de son accident ; mais son cas mit cependant en évidence l'existence d'une relation entre le cerveau et le comportement.

C'est Paul Broca, neurochirurgien français, qui apporta en 1861 la preuve scientifique de la localisation d'une fonction mentale en mettant en évidence une zone particulière du cerveau nécessaire au langage située sur une circonvolution frontale gauche, désormais connue comme étant l'aire de Broca [Broca, 1861]. Il fit cette découverte en étudiant le cerveau *post mortem* de l'un de ses patients, M. Leborgne, qui souffrait d'aphémie (altération du langage portant sur l'articulation et pouvant se traduire par une perte de la parole) suite à un accident vasculaire cérébral. Ce patient fut surnommé "Tan" car c'était la seule syllabe qu'il pouvait prononcer ; il avait par contre conservé la compréhension. Lors de l'autopsie effectuée après son décès, Broca remarqua une lésion superficielle située sur la partie inférieure du lobe frontal gauche (voir figure 2.3) et établit ainsi le lien entre dysfonctionnement cognitif et lésion. Ici c'est la dissociation des processus cognitifs associés à la parole qui fut mis en évidence. D'autres études *post-mortem* de patients atteints du même symptôme révélèrent des lésions similaires et confirmèrent ainsi la théorie de la localisation des fonctions cérébrales.



FIG. 2.2 – Reconstitution de l'accident de Phineas Gage (voir à ce sujet ce site entièrement consacré à son histoire : <http://www.deakin.edu.au/hbs/GAGEPAGE/>) et reconstruction tridimensionnelle du crâne et cerveau lésés par Hanna Damasio et son équipe [Damasio et al., 1994].

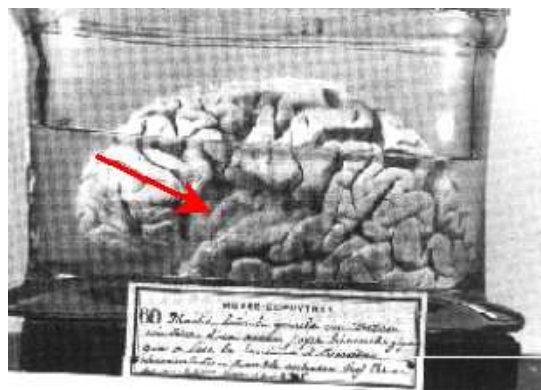


FIG. 2.3 – Cerveau du patient "Tan" de Paul Broca. On note une lésion du lobe frontal gauche responsable de la perte de la parole chez ce sujet.

A la même époque, d'autres découvertes anatomo-cliniques apportèrent de nouvelles preuves de la corrélation anatomo-fonctionnelle, telle celle du neurologue allemand Carl Wernicke qui mit en évidence une aire du lobe temporal dont la lésion entraînait un déficit sensoriel du langage [Wernicke, 1874] : le patient devenait incapable de reconnaître des mots alors qu'il pouvait très bien entendre des sons.

Malheureusement cette approche neuropsychologique n'est pas non plus exempte de critiques. Ses limites sont diverses :

- elle ne renseigne pas directement sur le fonctionnement du cerveau sain *in vivo*
- la lésion ne se restreint pas forcément à une aire précise mais touche généralement plusieurs aires fonctionnelles,
- la délimitation de la lésion cérébrale elle-même, nécessite une autopsie et ne peut

- être envisagée qu'à la mort du patient,
- l'existence d'une grande variabilité inter-individuelle sur le plan anatomique,
- l'existence d'une certaine plasticité cérébrale qui peut entraîner une récupération fonctionnelle alors même qu'il y a persistance de la lésion,
- elle donne lieu à des problèmes d'interprétation des troubles observés : s'agit-il de la perte d'informations, de la perte d'un processus ou d'une désorganisation globale ?

Il faudra attendre l'avènement des techniques d'imagerie fonctionnelle cérébrale au XXe siècle pour ouvrir la voie à de nouvelles expérimentations permettant de progresser dans une cartographie fonctionnelle du cerveau.

2.2 L'anatomie cérébrale

Avant d'aller plus loin, intéressons-nous quelques instants à l'anatomie du cerveau, siège de l'activité fonctionnelle. Cette description de l'anatomie cérébrale est d'autant plus importante que, comme nous le verrons par la suite, les localisations d'activation au cours de tâches cognitives sont toujours reportées sur des images anatomiques afin d'être interprétées. Il est donc essentiel, et nous y reviendrons, de définir un référentiel commun ainsi qu'un langage commun afin de pouvoir tirer partie de tous les résultats d'expériences cognitives et aboutir ainsi à une cartographie des fonctions cérébrales.

Le cerveau humain est un organe d'un poids moyen de 1400 grammes pour un volume moyen de 1500 cm³ (mais très variable selon les individus, de 1000 à 2000 cm³), situé dans la boîte crânienne et protégé par des membranes, appelées les méninges, qui incluent une couche externe dense (la dure-mère), une couche interne (la pie-mère) et une couche intermédiaire l'arachnoïde), constituée d'un réseau de fibres. Le cerveau baigne dans 150 millilitres environ de liquide céphalo-rachidien (LCR), qui a un rôle protecteur et nourricier. Exempt de nerfs sensitifs et donc insensible à la douleur, le cerveau est en revanche fortement vascularisé afin de répondre à ses besoins importants en oxygène.

L'architecture du cerveau peut être étudiée aux résolutions macroscopiques ou microscopiques selon que l'on se place du point de vue de l'anatomiste ou du neurobiologiste.

2.2.1 Architecture macroscopique

Le cerveau est constitué de deux hémisphères, reliés par une grande commissure, le corps calleux. Entre les deux hémisphères se trouve la fissure longitudinale, tandis que la fissure transverse sépare le cerveau et le cervelet (voir figure 2.4).

La surface du cerveau est constituée d'une couche de matière grise, appelée *cortex*, d'une épaisseur de 2 à 5 mm, et contenant le corps cellulaire des neurones (les cellules nerveuses). Il existe une très grande diversité de neurones (en forme et en taille) mais tous possèdent la même structure : le corps cellulaire (contenant le noyau), les dendrites et l'axone. Leur fonction est de participer au cheminement de l'influx nerveux. Pour cela, le corps cellulaire et les dendrites sont recouverts de synapses faisant office de système récepteur tandis que la partie émettrice est constituée de l'axone, la fibre nerveuse capable de conduire le potentiel d'action sur de longues distances. Ce sont les axones qui forment la matière blanche, située sous le cortex. Composée de faisceaux de fibres, sa

couleur est due à la myéline, gaine enveloppant les axones, et participant à la vitesse de transmission de l'influx nerveux. Il existe également de la matière grise n'appartenant pas au cortex et située au centre du cerveau : ce sont les noyaux gris centraux, constitués dans chaque hémisphère du noyau caudé, du noyau lenticulaire, du thalamus et du noyau subthalamique.

D'un point de vue macroscopique, on peut distinguer deux approches, intimement liées, pour décrire le cerveau : la parcellisation lobaire et l'anatomie sulco-gyrale.

Parcellisation lobaire

Chaque hémisphère est constitué de plusieurs lobes, constituant un découpage de la surface du cerveau. Pour des raisons anatomiques mais également historiques, on distingue cinq lobes, comme le montre la figure 2.4 :

- un *lobe frontal*,
- un *lobe pariétal*,
- un *lobe occipital*,
- un *lobe temporal*,
- un *lobe limbique*, également appelé *cortex cingulaire*, non visible sur la figure 2.4 car il se situe sur la face interne des hémisphères.

Il faut ajouter à ces cinq lobes, le lobe de l'*insula*, qui n'est pas visible sur la surface externe des hémisphères car entièrement recouvert par les lobes frontal, pariétal et occipital.

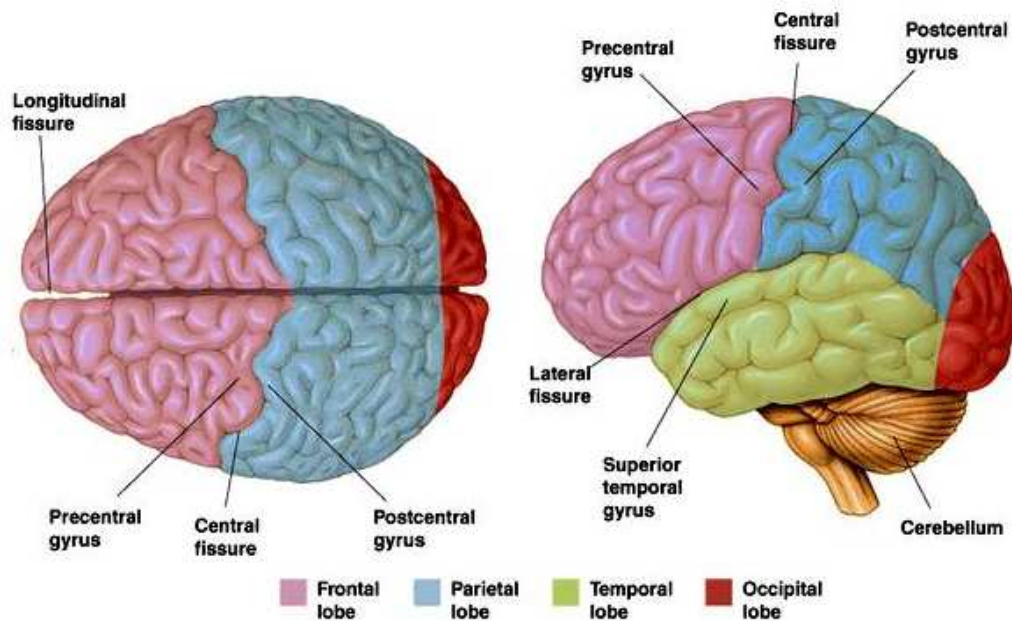


FIG. 2.4 – Anatomie cérébrale macroscopique : parcellisation lobaire et principaux sillons et gyri. A noter que le lobe limbique et le lobe de l'*insula* ne sont pas visibles sur les vues représentées. D'après <http://www.driesen.com/>

Certains lobes sont délimités par des scissures que l'on observe sur la face latérale de chaque hémisphère :

- le sillon central (ou scissure de Rolando) qui sépare le lobe frontal du lobe pariétal,
- le sillon latéral (ou scissure de Sylvius) qui sépare le lobe frontal du lobe temporal,
- le sillon occipital transverse (ou scissure pariéto-occipitale) qui sépare le lobe pariétal du lobe occipital.

Ces sillons sont dit primaires car ce sont ceux qui apparaissent le plus précocement lors du développement du cerveau chez le fœtus et qui sont présents chez tous les individus.

Anatomie sulco-gyrale

La surface des deux hémisphères présente un relief typique de circonvolutions ou gyri et de sillons ou sulci qui en augmentent considérablement la superficie. Ces plissements du cerveau apparaissent à différents moments du développement cérébral et l'homme est le primate possédant le cerveau le plus plissé car il contient le nombre le plus important de cellules nerveuses (de dix à cent milliards) et de connexions (chaque neurone pouvant établir de nombreuses relations avec d'autres neurones, jusqu'à 10000 par neurone). Comme nous venons de le voir, certains de ces sillons sont stables à travers les individus et servent à définir les frontières entre lobes. Pour les autres, on observe une grande variabilité inter-individuelle de forme, position, profondeur ou même d'existence [Ono *et al.*, 1990].

L'anatomie descriptive du cortex est fréquemment basée sur son organisation sulcale : on parle d'anatomie sulco-gyrale. Il n'y a pas cependant à ce jour de consensus concernant une nomenclature des sillons. On peut toutefois citer à ce sujet le site internet BrainInfo² qui utilise NeuroNames comme base de connaissance. Ce dernier se définit comme une nomenclature hiérarchique des structures du cerveau, référant plus de 850 entités neuroanatomiques regroupées entre elles par des relations de hiérarchies. Cette nomenclature permet d'attribuer à chacune des structures répertoriées un nom anatomique unique et d'accéder ensuite à une base de données d'informations associées à ces structures.

2.2.2 Architecture microscopique

Nous avons déjà vu que le cerveau est constitué de cellules nerveuses spécifiques, les neurones, formant la matière grise (corps cellulaire des neurones) et la matière blanche (les axones). La distribution des neurones au sein du cortex est loin d'être homogène ; ceci a permis d'établir des cartes de distribution architecturale et de distinguer différents territoires corticaux selon des critères cytoarchitectoniques, basés sur les distributions et tailles des corps cellulaires révélés par des marquages cellulaires (coloration de Nissl). La figure 2.5 présente les six couches cellulaires principales du cortex que l'on peut visualiser grâce à différents types de coloration. Il est à noter que, malgré la présence de différentes couches, le cortex n'a pas uniquement une structure laminaire car les neurones peuvent participer à différentes couches via leurs dendrites. On parle ainsi également de colonnes

2. <http://braininfo.rprc.washington.edu/>

corticales, chaque colonne correspondant à une unité fonctionnelle corticale organisée verticalement.

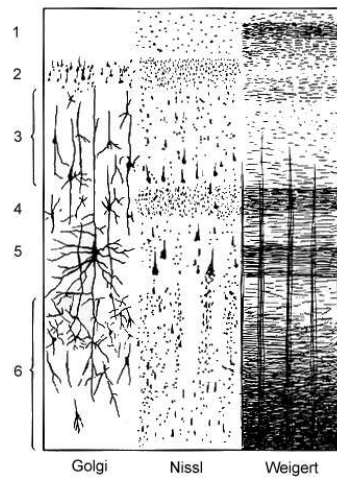


FIG. 2.5 – Les six couches cellulaires du cortex révélées selon trois types de coloration (Golgi, Nissl et Weigert). D'après [Vogt, 1991].

Les cartes cytoarchitectoniques les plus connues sont celles établies par [Brodmann, 1909] à partir d'études microscopiques de dissections. Comme nous le verrons par la suite, ces cartes, construites à partir d'un faible nombre de cerveaux, sont encore utilisées de nos jours pour décrire la localisation d'activations fonctionnelles. En effet, sous certaines réserves, ces cartes montrent une correspondance entre aires cytoarchitectoniques et aires fonctionnelles.

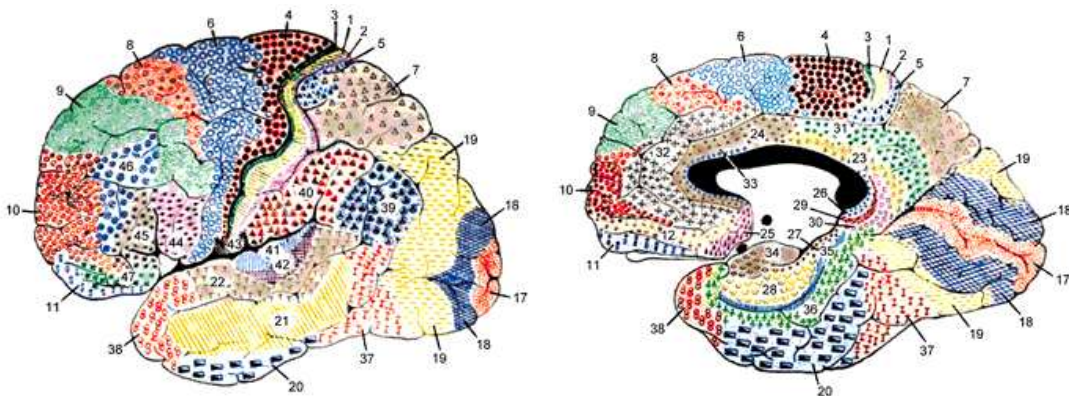


FIG. 2.6 – Cartes cytoarchitectoniques de Brodmann (1909) colorisées. Une cinquantaine d'aires ont été définies selon des critères portant sur la taille des corps cellulaires, sur leur densité, sur le nombre de couches et sur la densité des axones myélinisés. D'après <http://spot.colorado.edu/~dubin/talks/brodman/brodman.html>

2.3 Les techniques d'imagerie fonctionnelle

L'un des objectifs principaux des neurosciences est de localiser les zones associées à des fonctions cognitives et leurs différents modes d'interactions nécessaires à la réalisation de tâches fondamentales telles que la mémoire ou le langage. Le cerveau présentant une très grande variabilité anatomo-fonctionnelle inter-individuelle on cherche à décrire l'anatomie et à localiser précisément les zones fonctionnelles chez chaque sujet *in vivo*. Les grands progrès de la science et du traitement du signal ont permis de développer au XXe siècle des techniques d'exploration non invasives de l'activité cérébrale. Celles-ci se classent en deux grandes familles :

- les techniques d'observation de l'activité électromagnétique du cerveau :
 - l'électroencéphalographie (EEG).
 - la magnétoencéphalographie (MEG).
- les techniques d'imagerie métaboliques mesurant indirectement l'activité électrique du cerveau :
 - la tomographie par émission de positons (TEP), développée à la fin des années 1940 (abréviation anglaise PET).
 - la tomographie par émission monophotonique (TEMP), datant du début des années 1960 (abréviation anglaise SPECT).
 - l'imagerie fonctionnelle par résonance magnétique (IRMf), découverte en 1990.

Ces techniques permettent d'obtenir des cartes des différents types d'activité cérébrale : électrique, magnétique, métabolique et hémodynamique. Le type de signal mesuré va alors avoir des conséquences directes sur les résolutions temporelles et spatiales des acquisitions.

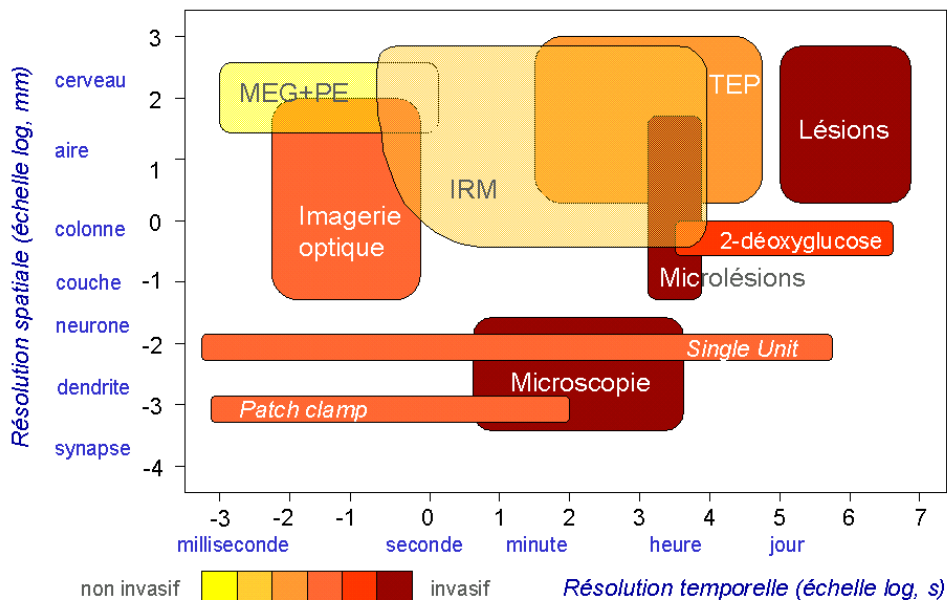


FIG. 2.7 – Classification de diverses techniques d'investigation cérébrale selon leur résolution spatiale, temporelle et leur caractère invasif.

2.3.1 Les techniques électromagnétiques

Lors de tâches cognitives, le cerveau génère des ondes électromagnétiques qui sont le reflet de l'activité électrique des neurones. Les potentiels d'action pré et post-synaptiques corticaux se propagent dans le cerveau à une vitesse proportionnelle à la vitesse de la lumière et sont donc observables en temps réel à la surface du scalp ou son voisinage à l'aide de macro-électrodes, positionnées à différents endroits de la tête du sujet. Electroencéphalographie et magnétoencéphalographie travaillent donc directement à l'échelle temporelle des événements synaptiques (résolution temporelle de l'ordre de la milliseconde). Les mesures effectuées sont donc un reflet direct et instantané des informations échangées entre neurones. Pour l'EEG, le matériel nécessaire à l'acquisition des signaux électriques est assez léger : on place quelques dizaines d'électrodes sur le scalp du sujet et on enregistre les variations de potentiel électrique (de l'ordre de quelques millivolts) toutes les millisecondes. En revanche, pour la MEG, l'appareillage nécessaire est beaucoup plus lourd. En effet, les variations du champ magnétique induites par l'activité cérébrale sont très faibles (de l'ordre du femto-Tesla), ce qui nécessite de placer le détecteur (SQUID - *Supraconducting Quantum Interference Device*) dans une cage de Faraday pour atténuer les perturbations magnétiques dues à l'environnement et de refroidir le bobinage à des températures permettant la supraconductivité.

Une fois obtenues les cartes de surface du potentiel électrique ou du champ magnétique, il reste à résoudre le difficile problème de trouver les positions des dipôles qui peuvent expliquer les cartes mesurées. On a donc ici affaire à un problème inverse, classiquement mal posé (il n'existe pas de solution unique) [Baillet and Garnero, 1997], qui restreint de ce fait la précision spatiale de ces techniques inverses. [Freeman *et al.*, 2003] avancent que la résolution intrinsèque en EEG serait assez limitée : une aire ne serait observable qu'à partir de plusieurs centimètres de diamètre.

2.3.2 Les techniques métaboliques et hémodynamiques

L'activité hémodynamique peut être indirectement mesurée à l'aide de techniques d'imagerie telles la tomographie par émission de positon (TEP) ou l'imagerie par résonance magnétique fonctionnelle (IRMf). En effet, l'activité neuronale étant corrélée à l'activité métabolique, elle-même couplée à la variation du débit sanguin cérébral, ces modalités permettent de cartographier l'activité engendrée par l'exécution d'une tâche cognitive au moyen de traceurs radioactifs (pour la TEP) ou paramagnétiques (pour l'IRMf).

La TEP permet d'étudier de façon quantitative et locale, le devenir de différentes molécules dans le corps humain. Des éléments radioactifs, émetteurs de positons produits à l'aide d'un cyclotron, sont incorporés à la place de leurs isotopes naturels dans des molécules biologiques qui sont injectées au patient. La présence de ces molécules ou de leurs métabolites dans une région du cerveau est repérée par les émissions de positons dont on détecte les photons d'annihilation à l'aide de photodétecteurs. Les images de concentration locale de la molécule injectée sont ensuite reconstruites grâce à des algorithmes de rétroprojection (on parle de reconstruction tomographique). Pour mesurer le débit sanguin régional cérébral, on utilise un traceur comme l'eau marquée à l' O^{15} . La réponse métabolique est elle-aussi mesurable grâce à l'utilisation d'un traceur du métabolisme du glucose

comme le déoxyglucose marqué au F^{18} . L'étude des variations temporelles des concentrations permet d'estimer localement la valeur de certains paramètres physiologiques reliés à l'activité cérébrale comme le débit sanguin régional. La résolution spatiale des images reconstruites est de l'ordre de 5 mm. La résolution temporelle est, elle, dépendante de la qualité de la caméra TEP : en raison de sa sensibilité limitée, une intégration temporelle sur un temps suffisamment long est nécessaire (typiquement plusieurs minutes). Il faut également attendre que la radioactivité décroisse avant de pratiquer une nouvelle acquisition chez le même patient.

Depuis les années 1990, la TEP est concurrencée par l'imagerie par résonance magnétique fonctionnelle dont les principaux avantages comme technique d'acquisition d'images de l'activité cérébrale sont le fait que :

1. aucune injection de substances radioactives n'est requise,
2. le temps d'acquisition d'un volume complet est très court (quelques secondes),
3. une possibilité de résolution très fine.

Etant la modalité utilisée dans nos travaux, l'IRMf est décrite plus en détail dans le paragraphe suivante. Notons tout de même dès à présent la tendance actuelle à combiner ces différentes modalités d'imagerie fonctionnelle afin de tirer partie des avantages de chacune d'elle : fusion MEG-EEG [Baillet *et al.*, 1999; Clerc *et al.*, 2002], EEG-IRMf [Goldman *et al.*, 2000].

2.4 L'imagerie par résonance magnétique fonctionnelle (IRMf)

L'IRM fonctionnelle a connu au début des années 1990 un essor spectaculaire car cette modalité permet d'obtenir de manière non-invasive et *in vivo* des images tridimensionnelles ayant à la fois une bonne résolution spatiale (quelques millimètres) et une bonne résolution temporelle (quelques secondes). Elle peut de plus facilement être couplée avec l'IRM anatomique (même imageur), permettant ainsi de détecter les aires activées dans les images fonctionnelles et de les reporter sur l'IRM anatomique pour la visualisation et l'interprétation des résultats. Après avoir succinctement décrits le principe physique de la résonance magnétique nucléaire (RMN) sous-jacent à l'acquisition de ces images, nous présentons l'hypothèse de couplage hémodynamique et décrivons les caractéristiques des images obtenues et de leurs éventuels artefacts. Enfin, nous décrivons le protocole d'acquisition de séries d'images.

2.4.1 Principe physique de l'IRM

Dans ce paragraphe, nous donnons les grandes lignes du principe physique de l'IRM, afin de pouvoir expliquer par la suite le contenu du signal fonctionnel mesuré. Le lecteur trouvera dans [Stark and Bradley, 1992; Van de Moortele, 1999; Houdé *et al.*, 2002] d'excellentes introductions aux techniques d'imagerie par résonance magnétique nucléaire.

L'imagerie par résonance magnétique (IRM) a été inventé au début des années 1970 par Paul Lauterbur [Lauterbur, 1973], co-lauréat du prix Nobel de Médecine 2003, et a connu depuis lors un essor considérable en permettant l'acquisition d'images tridimensionnelles de grande précision anatomique. Sous le terme IRM, on regroupe en fait diverses modalités qui ont en commun d'être toutes acquises avec le même imageur : images anatomiques (pondérées en T_1 ou T_2), images fonctionnelles (pondérées en T_2^*), images de diffusion (orientation des principaux faisceaux de fibres de matière blanche), angiographies IRM (visualisation des vaisseaux cérébraux), images de perfusion (distribution du sang) et spectroscopie RMN (distribution de certains métabolites énergétiques).

Les images qui nous intéressent ici sont des mesures du signal de résonance magnétique nucléaire des protons de l'hydrogène H de l'eau (les molécules d'eau représentant plus de 80 % du poids du cerveau). Pour acquérir une image du cerveau, on place la tête du sujet dans un champ magnétique intense (plusieurs Tesla) obtenu grâce à un électro-aimant constitué de matériaux supraconducteurs. Dans la suite, on considère l'espace muni d'un repère orthonormal (O_x, O_y, O_z) plongé dans un champ magnétique \vec{B}_0 dirigé selon l'axe O_z .

Fréquence de Larmor et précession

Placés dans un champ magnétique \vec{B}_0 , les protons de l'hydrogène de l'eau acquièrent une aimantation \vec{M}_0 (moment magnétique ou *spin*) proportionnelle à \vec{B}_0 et animée d'un mouvement de précession autour de l'axe O_z du champ magnétique. La relation de Larmor :

$$f_{\text{larmor}} = \frac{\gamma}{2\pi} \|\vec{B}_0\|$$

relie la fréquence de cette rotation avec l'amplitude du champ magnétique. γ dénote le rapport gyromagnétique du noyau. La fréquence de Larmor est également la fréquence de résonance à laquelle une onde radio peut exciter les protons.

Impulsion d'excitation

L'amplitude M_0 de cette aimantation est l'objet de la mesure RMN. Pour mesurer le phénomène de relaxation, il faut la basculer dans le plan O_{xy} , perpendiculaire à la direction de \vec{B}_0 : c'est l'effet macroscopique obtenu en émettant une impulsion radio-fréquence à la fréquence de Larmor, perpendiculairement à O_z (on parle de radio-fréquence puisque la fréquence du champ magnétique exciteur se trouve dans la gamme des 10/100 MHz) : c'est la situation dite de *résonance*. L'impulsion radio-fréquence, dont la durée va déterminer l'angle de bascule dont \vec{M}_0 s'écarte de l'axe O_z , est créée par une antenne placée autour de la tête du sujet.

Temps de relaxation

A la fin de l'impulsion, le vecteur aimantation \vec{M} retourne à sa valeur d'origine par un mouvement de précession, à la fréquence f_{larmor} , en émettant un champ électromagnétique qui permettra la mesure. On appelle *relaxation* le retour proprement dit à l'équilibre. On

peut le décomposer en relaxation longitudinale pour ce qui concerne le retour vers \vec{B}_0 de la composante longitudinale M_z de cette aimantation et relaxation transversale pour ce qui concerne le retour à 0 de la composante transversale M_{xy} . La cinétique de ces composantes s'écrit :

$$\begin{aligned} M_z(t) &= M_0 (1 - e^{-t/T_1}) \\ M_{xy}(t) &= M_0 e^{-t/T_2} \end{aligned}$$

On définit ainsi deux temps de relaxation, T_1 et T_2 , correspondant, respectivement, au temps de relaxation longitudinal et transversal. Ces constantes de temps dépendent de façon complexe des propriétés physico-chimiques du corps étudié. Le tableau 2.1 présente les valeurs mesurées pour différents tissus cérébraux.

Tissu	T_1 (s)	T_2 (ms)
Matière grise	1.09-2.15	61-109
Matière blanche	0.76-1.08	61-100
Liquide céphalo-rachidien	0.8-20	110-2000
Muscle	0.95-1.82	20-67
Graisse	0.2-0.75	53-94

TAB. 2.1 – Temps de relaxation T_1 et T_2 de divers tissus cérébraux. D'après <http://www.cis.rit.edu/htbooks/mri/inside.htm>

La constante de temps T_1 dépend plus spécifiquement de la mobilité des atomes d'hydrogène ou de celle des molécules où ils sont engagés. T_1 sera d'autant plus court que ces atomes seront liés à de grosses molécules. La constante de temps T_2 dépend elle aussi de la mobilité des atomes d'hydrogène et est toujours inférieure au temps de relaxation T_1 . En pratique, on observe une décroissance du signal plus rapide que celle prévue en raison des hétérogénéités du champ magnétique qui existent au sein du volume mesuré :

$$M_{xy} = M_0 e^{-t/T_2^*} \quad \text{avec} \quad T_2^* < T_2$$

Ces hétérogénéités proviennent des imperfections des bobines de champ magnétique employées et des différences de susceptibilité magnétique entre les tissus du corps étudié. Elles provoquent une variation des fréquences de rotation au sein d'un volume donné, ce qui entraîne un déphasage accéléré des aimantations microscopiques au sein de ce volume, et donc la disparition rapide du signal.

L'acquisition de ces valeurs, T_1 , T_2 , T_2^* ainsi la norme de M_0 , permet d'obtenir des images dites respectivement pondérées en T_1 , T_2 , T_2^* et densité de proton.

Construction des images

Le signal RMN que nous venons de décrire ne permet pas de localiser les sources, car les dimensions du cerveau sont petites devant les longueurs d'onde utilisées (quelques mètres). C'est Lauterbur qui montra en 1973 la possibilité de reconstruire une image en superposant des gradients de champ linéaire au champ magnétique principal. Etant donné que la fréquence de résonance est une fonction du champ magnétique appliqué,

l'application d'un champ variant spatialement permet de spécifier des zones précises dans le volume d'intérêt imagé. Typiquement, trois étapes d'encodage spatial sont utilisées, correspondant aux directions de l'espace 3D : sélection de coupe, encodage par la phase et encodage par la fréquence.

Pour acquérir un volume tridimensionnel, on sélectionne un plan de coupe en appliquant un premier gradient de champ linéaire, perpendiculairement au plan de la coupe souhaitée. La largeur de la bande de fréquence utilisée permet de déterminer l'épaisseur de la coupe. On sélectionne ensuite à l'intérieur de ce plan les différentes lignes en appliquant un deuxième gradient de champ, appelé gradient de codage de phase. Enfin, on sélectionne les différentes colonnes par un gradient de fréquence ou gradient de lecture. Ces gradients de champ linéaire sont obtenus grâce à des bobines de gradient, dont l'amplitude est de quelques dizaines de milli-Tesla. Les signaux reçus par une antenne tête contiennent alors la superposition des informations de résonance magnétique encodées spatialement par la phase et la fréquence. On a en fait réalisé un échantillonnage du plan de Fourier (*k-space*) et l'image d'une coupe est reconstruite par application de la transformée de Fourier inverse.

Séquences d'acquisition

Il existe différentes séquences d'acquisition permettant de mesurer les temps de relaxation, dont les deux plus grandes familles sont les séquences dites *spin echo* et *gradient echo*. Le principe de l'écho est de refaire croître le signal de résonance à une valeur maximale au bout d'un temps d'écho TE, afin de pouvoir mesurer le signal à une amplitude convenable et en se donnant le temps de pouvoir le faire. Ainsi la séquence *spin echo* permet l'acquisition d'images pondérées en T_2 . La séquence *gradient echo* ne permet pas, à la différence du *spin echo*, de compenser les déphasages liés à l'hétérogénéité du champ magnétique et permet ainsi de mesurer la décroissance exponentielle en T_2^* du signal. Un autre paramètre important est le temps de répétition noté TR, qui mesure le temps séparant l'application de deux séquences successives. C'est le choix des valeurs attribuées aux paramètres TE et TR qui va déterminer le contraste des images (T_1 , T_2 ou densité de proton). Ainsi avec un TR long afin de ne pas avoir à subir la pondération en T_1 et un temps d'écho TE de l'ordre de grandeur du T_2 moyen des tissus, on obtient une image pondérée en T_2 (si on a utilisé une séquence *spin echo*) ou T_2^* (si on a utilisé une séquence *gradient echo*).

Un des problèmes avec ces séquences est que l'acquisition d'une image est assez longue, car il faut effectuer une impulsion radio-fréquence pour chaque ligne composant une coupe. Or dans le cas de l'imagerie fonctionnelle, le temps d'acquisition est essentiel et doit être inférieur à la constante de temps associée au phénomène biologique que l'on souhaite observer. La séquence *echoplanar imaging* (EPI) mise au point par Mansfield, également co-lauréat du prix Nobel de médecine 2003 [Mansfield, 1977] permet d'acquérir une coupe entière avec une seule impulsion radio-fréquence, grâce à des inversions rapides de gradient (balayage en dents de scie de la totalité du plan de Fourier en un seul signal). Il est ainsi possible d'acquérir une coupe en une centaine de millisecondes. Par contre, le temps d'acquisition limité à un seul TR par coupe impose de fait une limite à la résolution spatiale atteignable en EPI ; la résolution sera donc moindre qu'avec une séquence plus

traditionnelle en *spin echo*. Nous reviendrons par la suite sur les caractéristiques des images acquises avec une séquence EPI.

Enfin, il est à noter que d'autres types de séquences ont fait leur apparition, comme par exemple les séquences SMACH [Sodickson and Manning, 1997] ou SENSE [Pruessmann *et al.*, 1999; Golay *et al.*, 2000] qui sont des séquences ultra-rapides effectuant des acquisitions en parallèle avec plusieurs antennes. Ces nouvelles séquences joueront certainement un rôle important dans les années à venir.

2.4.2 Bases physiologiques du signal en IRMf

L'IRMf ne permet pas de mesurer directement l'activité neuronale mais seulement une conséquence de cette activité sur l'hémodynamique.

Le couplage hémodynamique

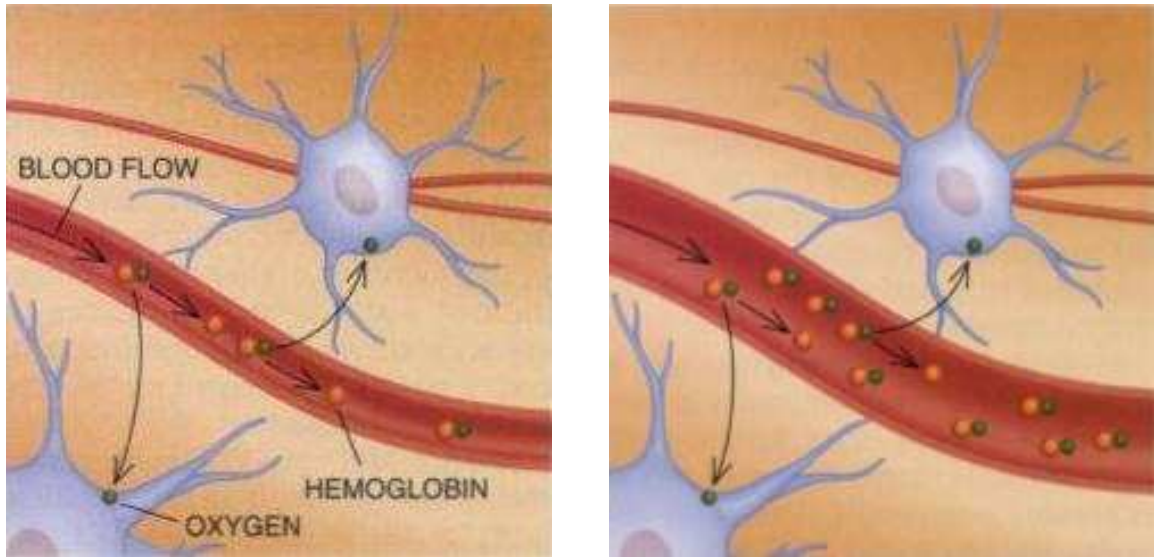
A la fin du siècle dernier, Roy et Sherrington [Roy and Sherrington, 1890] observèrent une modification locale de la couleur du cortex de sujets soumis à des stimulations. Ce changement de couleur s'explique par le fait que le sang artériel (saturé en oxygène) est rouge vif tandis que le sang veineux (appauvri en oxygène) est bleu violacé. Ces observations suggèrent un couplage entre l'activité électrique des neurones et le débit sanguin : c'est le couplage hémodynamique. Ainsi l'activité cognitive du cerveau engendrerait des variations du volume sanguin cérébral régional mais aussi de la concentration du sang en oxygène, à l'origine du signal BOLD. Cependant les conditions de ce couplage restent controversées et l'on s'interroge encore sur les relations liant activité neuronale, débit sanguin, volume sanguin, concentration en oxygène et métabolisme cérébral.

L'effet BOLD

Le glucose est la seule source d'énergie du cerveau. Comme il n'est quasiment pas stocké au sein du cerveau, son apport permanent par le système sanguin est nécessaire. Ainsi, en raison du couplage hémodynamique, dans des conditions de perfusion normale, le rCBF (débit sanguin cérébral régional ou *regional Cerebral Blood Flow*) reflète l'activité synaptique locale [Fox *et al.*, 1986]. De même pour le rCBV (volume sanguin cérébral régional ou *regional cerebral blood volume*) [Belliveau *et al.*, 1991].

Pour mesurer la perfusion régionale, on peut utiliser un produit de contraste (Gadolinium, Mion) dont les propriétés magnétiques vont modifier localement le champ magnétique et ainsi modifier la relaxation magnétique transverse des protons alentours. L'acquisition d'images pondérées en T_2^* permet justement d'observer les concentrations d'un produit paramagnétique dans le sang, en étant sensible aux différences de susceptibilité.

Or il s'avère que la désoxyhémoglobine présente dans le sang est diamagnétique alors que la molécule d'hémoglobine est elle paramagnétique. La molécule de désoxyhémoglobine peut alors être utilisée comme produit de contraste endogène. Une image pondérée en T_2^* sera sensible à la concentration d'oxygène dans le sang : les tissus contenant des vaisseaux riches en désoxyhémoglobine présenteront un signal moindre que les tissus contenant des vaisseaux riches en hémoglobine oxygénée. Le signal ainsi mesuré est le



(a) Etat de Repos

(b) Etat d'Activation

FIG. 2.8 – *Phénomène physique à la base de l'IRMf (extrait de [Raichle, 1994]). Une activation provoque une augmentation locale de la perfusion sanguine. L'activité synaptique étant quasi-anaérobie, on observe alors une diminution nette du contenu vasculaire en désoxyhémoglobine (en orange) ce qui entraîne une remontée du signal IRM T_2^* .*

signal BOLD (Blood Oxygen Level Dependent) [Kwong *et al.*, 1992; Ogawa *et al.*, 1992; Bandettini *et al.*, 1992], principalement utilisé en IRMf.

L'effet BOLD fait donc référence à une relation entre débit sanguin cérébral régional et consommation d'oxygène. Ce que l'on observe est une augmentation significative du débit sanguin à la suite d'une activation mais qui n'est pas suivie d'une augmentation significative du métabolisme de l'oxydation du glucose : l'augmentation du débit sanguin semble être totalement disproportionnée avec les besoins effectifs en oxygène des neurones irrigués [Kim and Ugurbil, 1997]. Une des hypothèses émise est que le métabolisme du glucose serait quasi-anaérobie [Fox *et al.*, 1988] expliquant la faible surconsommation d'oxygène des neurones en activité par rapport aux neurones au repos (cf figure 2.8), mais cela reste à démontrer. D'autres théories ont été proposées ; notons en particulier les travaux de Buxton [Buxton *et al.*, 1998] proposant un modèle biomécanique, dit du *ballon*, permettant d'expliquer les divers phénomènes observés en IRM par une différence d'élasticité entre secteurs veineux et artériels. Il s'agit en effet de pouvoir expliquer non seulement l'effet BOLD mais aussi les particularités du signal mesuré : un creux au début de la réponse (*initial dip*) et une chute temporaire du signal au dessous du niveau de base à la fin de l'activation (*final undershoot*).

Ceci étant, quelle que soit l'explication de la différence (découplage) entre débit sanguin et utilisation de l'oxygène, c'est le phénomène que mesure l'IRMf-BOLD. En effet, l'excès de volume sanguin associé avec une faible augmentation de l'extraction de l'oxy-

gène entraîne une diminution de la concentration en desoxyhémoglobine, et donc une modification locale du champ magnétique qui peut être mesurée par l'IRM T_2^* . En particulier, une diminution de la concentration du sang en desoxyhémoglobine va engendrer un temps de relaxation T_2^* plus long, puisque le champ magnétique local sera plus homogène et donc une augmentation du signal IRM.

Ainsi, l'origine du signal BOLD n'est que partiellement connue et il existe encore de nombreuses zones d'ombres quant à la description et l'explication des différents couplages mis en œuvre (électrique, hémodynamique, métabolique) [Heeger and Rees, 2002]. À noter toutefois les récents travaux de Logothetis [Logothetis *et al.*, 2001] concernant la base neurophysiologique du signal IRMf-BOLD. Grâce à l'acquisition simultanée du signal BOLD et de celui mesuré avec des électrodes implantées sur le singe éveillé, il a pu mettre en évidence que le signal BOLD est corrélé avec les *local field potentials* qui sont le produit des entrées des autres aires du cerveau et des processus locaux d'une aire du cortex. À n'en pas douter, ce sont les acquisitions simultanées de ce type ou les études IRMf/EEG, moins invasives, qui permettront dans le futur d'expliquer et de modéliser le signal BOLD.

2.4.3 Caractéristiques et artefacts des images IRMf-BOLD

Nous venons donc de voir que les images fonctionnelles IRMf-BOLD mettent en évidence des variations locales d'oxygénation du sang traduisant indirectement une activité neuronale. Ces images sont pondérées en T_2^* et classiquement acquises par une séquence d'acquisition rapide EPI. On peut dès lors en déduire un ensemble de caractéristiques que vont vérifier les images acquises en IRMf et expliquer l'apparition de plusieurs artefacts qu'il s'agira de prendre en compte.

Caractéristiques des images IRMf-BOLD

La résolution temporelle des images IRMf acquises avec une séquence EPI-BOLD est de 100 à 200 ms pour une coupe de 64×64 voxels (variable selon le champ de vue). Ainsi, si l'on se restreint à une acquisition monocoupe, on peut obtenir un échantillonnage temporel de cet ordre de grandeur. En pratique, on acquiert souvent un ensemble de coupes couvrant un plus grand domaine du cerveau. Un des jeux d'images que nous détaillerons plus avant est constitué de volumes constitués de 18 coupes de taille 64×64 pour un temps de répétition (TR) de 2 secondes. Il est en revanche important de noter que cette résolution temporelle est celle de l'imageur et non pas celle du signal mesuré. En effet, à la différence de la réponse neuronale à une activation, la réponse hémodynamique est tributaire de l'inertie vasculaire, entraînant un retard de plusieurs secondes pour le maximum d'activation. Le phénomène mesuré par effet BOLD a donc une résolution temporelle beaucoup moins fine que la résolution physique de la séquence EPI.

La résolution spatiale est elle directement conditionnée par le choix du temps de répétition et du champ de vue. En pratique, les images acquises ont une résolution de l'ordre de $3\text{mm} \times 3\text{mm} \times 3\text{mm}$, avec parfois une épaisseur de coupe plus grande que les dimensions intracoupes. À noter également que, ici aussi, la résolution spatiale est celle de l'imageur et non pas celle du phénomène physiologique observé (vasodilatation).

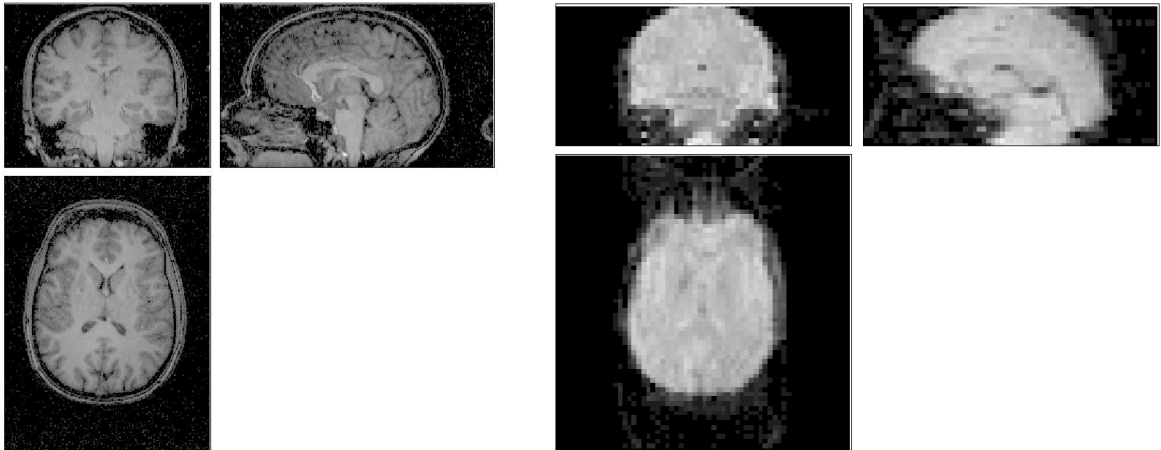


FIG. 2.9 – Coupes axiales, coronales et sagittales d'images IRM d'un même sujet, couvrant l'intégralité de l'encéphale, acquises sur une machine GE 1.5 T. L'image de gauche est pondérée en T_1 , avec une matrice $192 \times 256 \times 128$, pour une taille de voxels de $1 \times 1 \times 1.2 \text{ mm}^3$, acquise en une dizaine de minutes. L'image de droite est pondérée en T_2^* avec une séquence EPI, avec une matrice $64 \times 64 \times 26$, pour une taille de voxels de $3.75 \times 3.75 \times 4.5 \text{ mm}^3$ et avec un temps d'acquisition de 3 secondes.

Les différences de susceptibilité entre le sang entièrement oxygéné et désoxygéné est faible [Frackowiak *et al.*, 1997], ce qui implique que les variations du signal IRM sont très faibles : de 2 à 5 % avec un champ magnétique de 1.5 T. Par contre, à temps d'écho et type de séquence donnés, l'amplitude du signal est trois fois plus élevée à 4 T qu'à 1.5 T. Sachant également que le rapport signal sur bruit croît parallèlement avec la force du champ magnétique, on comprend dès lors l'intérêt de travailler à très haut champ en imagerie fonctionnelle (mais ceci se fait au détriment de la qualité de l'image avec l'augmentation d'un certain nombre d'artefacts). Avec un champ traditionnel de 1.5 T, on peut toutefois augmenter le rapport signal sur bruit en répétant la mesure de nombreuses fois chez le même sujet ($SNR' = SNR \times \sqrt{n}$ où n représente le nombre de répétitions). A noter que le signal évoqué ici est le signal de résonance magnétique : les améliorations du rapport signal sur bruit sont moindres sur le signal BOLD lui-même en raison de bruits physiologiques.

Au delà de la résolution spatiale imposée par l'imageur, se pose aussi le problème de la localisation des activations : en effet, par essence même, le signal BOLD provient des capillaires ou même des veines de drainage, qui peuvent se trouver à distance de la zone corticale activée qu'ils irriguent [Turner, 2002]. Ceci est particulièrement observable à haut champ (3 T) avec des zones activées mesurées qui semblent suivre des veinules. On voit ainsi une des limitations de l'IRMf-BOLD.

Artefacts des images IRMf-BOLD

Les images IRMf-BOLD peuvent être affectées par divers artefacts qu'il s'agit au mieux de corriger, ou simplement de prendre en compte.

Nous venons ainsi de préciser que d'un point de vue spatial, le signal BOLD n'est qu'une mesure indirecte de l'activité neuronale, et que sa localisation se situe sur les vaisseaux veineux et non pas sur les neurones d'où proviennent l'activation.

L'IRMf est également affectée par des artefacts de susceptibilité. En effet, cette modalité est sensible aux variations de susceptibilité magnétique (on mesure ainsi le signal BOLD) et donc en particulier, il peut y avoir des distorsions ou des pertes de signal. Ces variations sont particulièrement visibles aux interfaces tissu-air, comme les sinus frontaux.

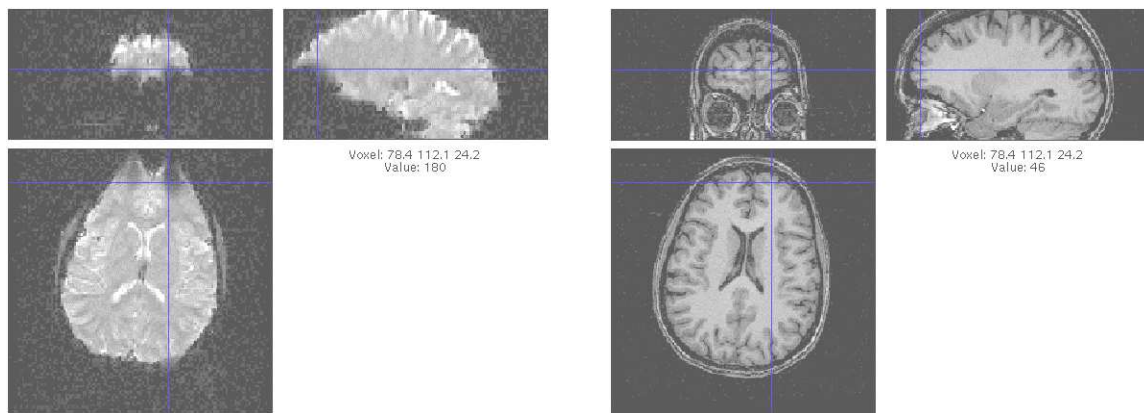


FIG. 2.10 – Exemple d'artefacts de susceptibilité observés pour les sinus frontaux, avec une IRM à haut champ (3 Tesla). L'image de gauche présente une IRM pondérée en T_2^* tandis que l'image de droite présente l'IRM anatomique correspondante, dans le même espace. La croix bleue indique la position des pertes de signal.

De même, toute hétérogénéité du champ magnétique peut créer des distorsions [Jezzard and Clare, 1999]. Pour les corriger, deux solutions sont envisageables : mettre en place des antennes "shim" permettant de rétablir un certain degré d'homogénéité du champ magnétique, ou alors acquérir une image de phase pendant le protocole pour corriger les images *a posteriori* [Jezzard and Balaban, 1995].

On note également l'apparition de fantômes (*ghost*), dus à l'échantillonnage dans le domaine de Fourier, se traduisant par l'apparition d'autres images du cerveau, d'intensité plus faible et décalés d'un demi-champ de vue par rapport à l'image principale.

Enfin, tous les mouvements pouvant survenir au cours de l'acquisition sont susceptibles de créer des artefacts : respiration, battements cardiaques, mouvements du sujet, etc. Pour les éviter, la tête du sujet est soigneusement calée dans l'aimant de l'IRM et si besoin est, le rythme respiratoire ainsi que les battements cardiaques peuvent être mesurés pour être ensuite filtrés dans les signaux mesurés et éviter ainsi de les attribuer à des activations.

Nous présenterons dans le chapitre suivant les différentes techniques informatiques permettant de corriger les artefacts que nous venons d'énumérer ici.

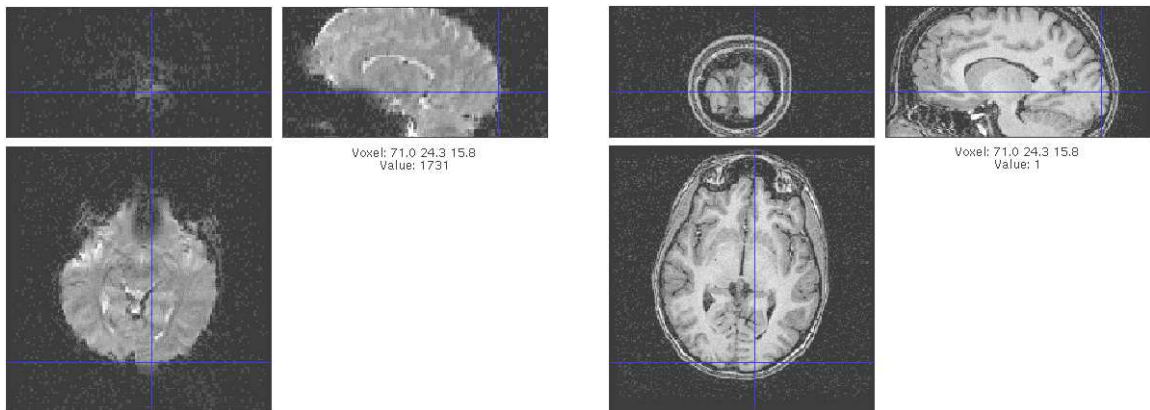


FIG. 2.11 – Exemple de distorsions géométriques dues aux hétérogénéités du champ magnétique dans la direction de l'encodage de la phase (antéro-postérieure), avec une IRM à haut champ (3 Tesla). La croix bleue, située dans le lobe occipital, met en évidence l'importance des déformations, de plusieurs millimètres ici.

2.4.4 Protocole expérimental

Le paradigme expérimental désigne la séquence de stimulations et actions à exécuter par le sujet. Par déviation sémantique, ce même terme sert à désigner la fonction modélisant le déroulement temporel attendu à cette séquence expérimentale.

On distingue classiquement deux sortes de protocoles:

- protocoles en bloc : ceux-ci ont pour caractéristique de faire répéter continûment la même tâche au sujet pendant une période de plusieurs dizaines de secondes avec également une répétition du stimulus, s'il a lieu.
- protocoles événementiels : dans ce cas, la tâche à effectuer n'est réalisée que pendant une durée très brève ou alors le stimulus est lui aussi très court. Comme pour un paradigme par bloc, ce processus est répété de nombreuses fois. Avec ce type de protocole, on met en évidence la réponse hémodynamique dont on peut alors estimer les caractéristiques : temps au pic, amplitude, etc. On peut alors obtenir des informations sur la chronométrie mentale.

La mise en œuvre d'une expérience en IRMf engendre un certain nombre de contraintes :

- travail en présence d'un fort champ magnétique : le matériel utilisé doit être soit diamagnétique soit placé dans une cage de Faraday,
- séquences d'acquisition très bruyantes, perturbant les expériences auditives,
- tunnel de l'aimant très étroit : le sujet peut difficilement se mouvoir et les systèmes de stimulus visuels doivent être gérés en conséquence.

Au cours d'une expérience fonctionnelle, les images classiquement acquises pour chacun des sujets y prenant part, sont :

- une image anatomique haute résolution (1 mm^3) sur laquelle seront reportées les activations,



FIG. 2.12 – *IRM Bruker 3 Tesla du centre IRMf du CHU de la Timone à Marseille et antenne tête (système radio-fréquence) de type "cage d'oiseau". Extrait de <http://irmfmrs.free.fr/>*

- une ou plusieurs séries d'images fonctionnelles EPI-BOLD,
- éventuellement une image de phase pour la correction des distorsions.

A noter que dans le cas de protocoles événementiels, il faut également conserver les instants de présentation des stimuli (*onsets*) car, pour des soucis d'efficacité, ceux-ci sont souvent aléatoires et donc spécifiques à chaque acquisition. Ils seront par la suite utilisés pour proposer un modèle de la réponse observée.

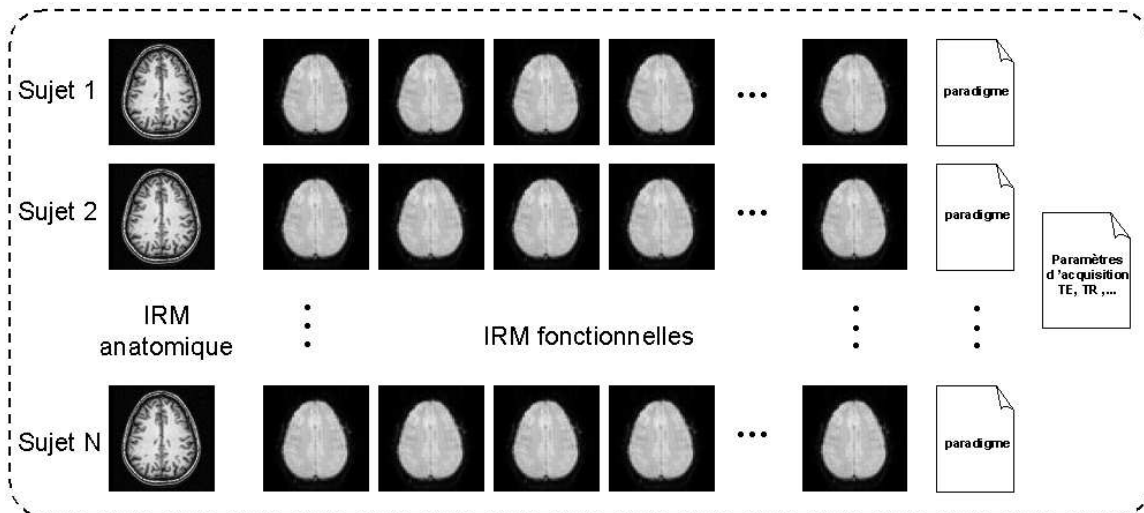


FIG. 2.13 – *Exemple typique des données acquises lors d'une expérience fonctionnelle : une image anatomique ainsi qu'une série d'images fonctionnelles sont acquises pour chaque sujet intervenant dans l'expérience. Les informations spécifiques au paradigme et aux paramètres d'acquisition sont également conservés et seront utilisés lors de l'analyse des données.*

Chapitre 3

Méthodes d'analyse des images fonctionnelles

Sommaire

3.0	Notations	34
3.1	Prétraitements	34
3.1.1	Correction des distorsions géométriques	35
3.1.2	Correction du délai temporel intercoupe (<i>slice timing</i>)	36
3.1.3	Correction du mouvement	37
3.1.4	Mise en correspondance des images anatomiques et fonctionnelles	40
3.1.5	Lissage spatial des données	40
3.2	Normalisation spatiale des images	41
3.2.1	Présentation du problème	41
3.2.2	Corrélation anatomo-fonctionnelle	42
3.2.3	Algorithmes de normalisation spatiale	45
3.2.4	Labellisation des activations fonctionnelles	53
3.3	Analyse statistique des données fonctionnelles	56
3.3.1	Problématique	56
3.3.2	Le modèle linéaire général	58
3.3.3	Inférence statistique	66
3.3.4	Méthodes exploratoires	67
3.3.5	Les analyses de groupe	69
3.4	Prise en compte d'une information spatiale	70
3.4.1	Analyse par région d'intérêt	70
3.4.2	Modèles régularisés spatialement	72
3.4.3	Analyse sur la surface corticale	72
3.5	Conclusion	76

L'analyse des images fonctionnelles fait classiquement intervenir plusieurs domaines : le traitement des images, le traitement du signal, les statistiques. . . Les données fonctionnelles se situent dans un espace $3D+t$ et on peut les décrire, soit comme une série de N images scalaires, soit comme une image vectorielle (un décours temporel par voxel). Mais dans tous les cas, il ne faut jamais oublier que l'on ne peut pas dissocier le spatial du temporel : une transformation spatiale entraîne des modifications des séries temporelles et réciproquement. L'analyse des données fonctionnelles par résonance magnétique consiste en la *détection* et la localisation dans le cerveau des modifications hémodynamiques traduisant une activité cérébrale. D'autres analyses peuvent également être effectuées selon le protocole et les questions de neurosciences auxquelles l'expérience essaye de répondre. Ainsi on peut également être intéressé par l'*estimation* de paramètres physiologiques de la réponse hémodynamique ou par l'étude de la connectivité fonctionnelle.

On voit déjà apparaître ici un dilemme entre estimation et détection : la forme du signal d'intérêt n'est pas connue précisément car la réponse hémodynamique varie selon les sujets et la localisation ; ainsi il apparaît difficile de pouvoir à la fois estimer la forme du signal d'intérêt et tester sa présence dans les données [Frackowiak *et al.*, 2004].

Dans ce chapitre, nous nous proposons de présenter l'ensemble des techniques mises en œuvre lors de l'analyse des images fonctionnelles par résonance magnétique. Nous verrons ainsi les différents prétraitements appliqués aux images avant leur analyse statistique. Nous avons étudié à part un des prétraitements, la normalisation spatiale, afin de mettre en évidence la difficulté de cette étape ainsi que les problèmes conceptuels qu'elle pose. Nous présenterons ensuite l'analyse statistique proprement dite et les différentes voies possibles pour obtenir des informations pertinentes à partir des données. Enfin, nous évoquerons les diverses approches proposées dans la littérature pour prendre en compte l'information spatiale lors de l'analyse des données.

3.0 Notations

Voici quelques notations utilisées au cours de ce chapitre :

n	nombre d'échantillons temporels du signal (nombre d'images).
$\{I_i\}_{i=1}^n$	ensemble des images fonctionnelles acquises.
Y	vecteur de dimension n contenant le décours temporel du voxel à étudier.
m	nombre de régresseurs.
X	matrice de plan expérimental $n \times m$.
β	vecteur de dimension m contenant les paramètres à estimer.
ϵ	vecteur résidu de dimension n .
R	projecteur $n \times n$ sur l'espace des résidus : $R = Id - XX^+$.
c	vecteur contraste, de dimension m .
TR	temps de répétition séparant les acquisitions de deux volumes successifs

3.1 Prétraitements

Le terme *prétraitement* regroupe l'ensemble des traitements appliqués aux données fonctionnelles pour corriger des effets variables, soit spatialement soit temporellement,

au sein des séries d'images qui ne sont pas liées à la tâche effectuée au cours de l'expérience. Leur mise en œuvre ne nécessite donc pas, a priori, la connaissance du protocole expérimental.

En particulier, les problèmes de compensation du mouvement et des distorsions ainsi que le recalage d'images inter-sujets ont fait l'objet d'une action de développement "IRMf" entre l'INRIA (projets Epidaure, Robotvis et Vista), le CEA/SHFJ et l'équipe U494 de l'INSERM. Le lecteur trouvera sur le site Epidaure³ les rapports d'activité ainsi qu'un des logiciels conçus au cours de cette action de recherche.

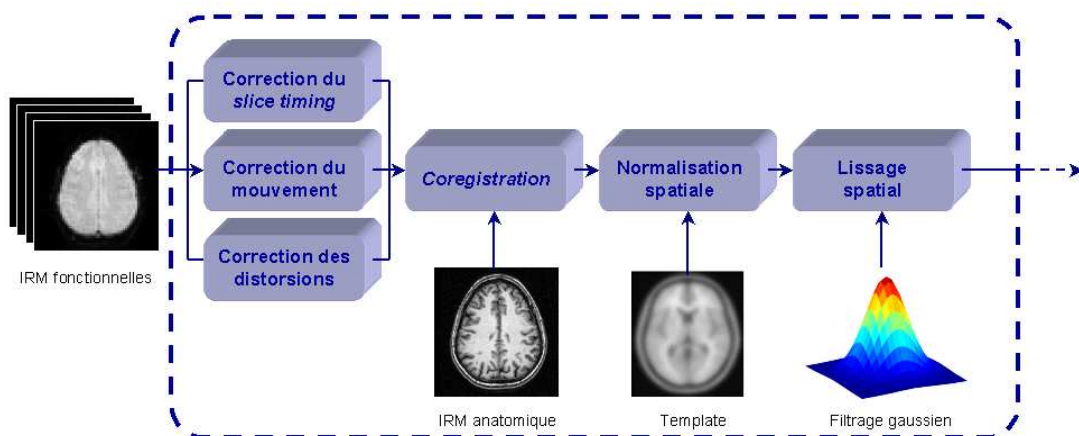


FIG. 3.1 – Les différentes étapes constituant les prétraitements fréquemment appliqués aux images acquises avant analyse statistique. Des corrections sont tout d'abord appliquées aux images fonctionnelles pour éliminer les différents artefacts (mouvements, distorsions, délai intercoupe (slice timing)) puis celles-ci sont recalées avec l'image anatomique du même sujet. Par la suite, les images sont souvent normalisées spatialement de manière à aligner les cerveaux des sujets ou à les placer dans un espace standard. Enfin, classiquement, un lissage spatial isotrope est appliqué pour, entre autre, augmenter le rapport signal sur bruit des images.

3.1.1 Correction des distorsions géométriques

Comme nous l'avons vu dans le chapitre précédent, les images acquises avec une séquence EPI sont sujettes à des distorsions géométriques, dues aux hétérogénéités du champ magnétique [Jezzard and Clare, 1999]. En effet, avec une séquence *echo planar*, une coupe entière est acquise avec une seule impulsion radio-fréquence suivie d'un parcours de l'image dans le plan de Fourier ligne par ligne. Il y a donc un délai de lecture différent entre deux échantillons voisins selon qu'ils sont sur la même ligne ou sur la même colonne : si une image compte N lignes, le délai temporel entre deux échantillons successifs sur une même colonne est N fois plus important que sur la même ligne. Cela entraîne une erreur sur la fréquence mesurée et puisque les images sont encodées spatialement en phase et en fréquence,

3. <http://www-sop.inria.fr/epidaure/Collaborations/IRMf/>

ces erreurs de lecture de la fréquence entraînent des distorsions géométriques principalement dans la direction de l'encodage de la phase (fréquemment antéro-postérieur). Un simple calcul [Bahrami, 2000] montre que les déformations engendrées sont proportionnelles à l'amplitude du champ magnétique B_0 (une IRM à 3 Tesla engendrera de plus grandes distorsions géométriques qu'une IRM à 1.5 Tesla, mais le rapport signal sur bruit sera meilleur) et à ses hétérogénéités. Ainsi, les déformations peuvent facilement atteindre plusieurs voxels dans la direction de l'encodage de la phase alors qu'elles sont inférieures au dixième de voxel dans la direction orthogonale (distorsions identiques aux séquences traditionnelles).

La solution proposée par [Jezzard and Balaban, 1995] consiste à acquérir une image de phase grâce à une séquence *double écho* possédant les mêmes distorsions que les images EPI. Il est alors possible, connaissant les retards de phase ligne à ligne, de corriger l'ensemble des images fonctionnelles avec ces cartes de phase dépliées (acquises modulo 2π).

Malheureusement, les distorsions peuvent changer au cours des acquisitions si le sujet bouge dans le scanner. En ce cas, il faudrait acquérir une image de phase avec chaque image EPI, ce qui est rédhibitoire. [Andersson *et al.*, 2001] propose de prendre en compte les paramètres de mouvement, comme nous le verrons par la suite, pour estimer les hétérogénéités du champ et ainsi corriger les distorsions géométriques.

Notons enfin que certains ont abordé le problème de la correction des distorsions géométriques comme un problème de recalage multimodal non rigide en recalant les images EPI sur une IRM T_1 , sur laquelle les distorsions sont négligeables [Hellier and Barillot, 2000; Kybic *et al.*, 2000; Studholme *et al.*, 2000].

3.1.2 Correction du délai temporel intercoupe (*slice timing*)

L'acquisition d'un volume fonctionnel peut prendre plusieurs secondes en fonction du nombre de coupes ; chacune étant acquise en une centaine de millisecondes. C'est le cas en particulier lorsque l'on souhaite faire l'acquisition de l'ensemble du cerveau, le temps de répétition étant alors de 3 à 6 secondes. Chaque volume étant acquis coupe par coupe, il y a donc un délai important entre l'instant de l'acquisition de la première coupe et celui de la dernière coupe d'un volume. Dans le cas d'acquisitions entrelacées, ce délai inter-coupe peut atteindre jusqu'à $\frac{1}{2}TR$ entre deux coupes adjacentes [Henson *et al.*, 1999].

Ainsi, si l'on souhaite comparer des signaux provenant de coupes différentes, il faut tenir compte du fait que ces décours temporels n'ont pas été échantillonnés aux mêmes instants. Il serait alors possible d'analyser les données en prenant en compte les instants d'acquisition des coupes d'où proviennent les signaux (et la séquence d'acquisition entrelacée ou séquentielle, ainsi que l'ordre d'acquisition ascendant ou descendant) [Worsley *et al.*, 2002]. Cette solution n'est pas très efficace d'un point de vue informatique car elle va nécessiter de définir différents modèles associés à chacune des coupes du volume. Une autre solution pour tenir compte du délai intercoupe consiste à rééchantillonner les signaux temporels mesurés afin qu'ils suivent tous le pas d'échantillonnage de la première coupe acquise dans chaque volume. Pour ce faire, il faut donc décaler temporellement les signaux de chaque coupe d'un certain Δt ; ceci est effectué en pratique grâce à un déphasage dans l'espace de Fourier :

$$TF[x(t - \Delta t)] = e^{-j2\pi f \Delta t} TF[x(t)]$$

Cette technique correspond à une interpolation sinus cardinal dans le domaine temporel et elle est donc valable en principe sur les signaux mesurés respectant la condition de Shannon pour l'échantillonnage. Malheureusement, si la tête du sujet a bougé au cours des acquisitions, les instants d'échantillonnages des différentes coupes correspondantes ne sont plus connus et ne sont plus réguliers, ce qui invalide la modélisation décrite ci-dessus. On trouvera dans [Paradis, 2001] une discussion sur le choix de l'ordre d'acquisition des coupes (séquentiel ou entrelacée) et ses conséquences lors de mouvements des sujets.

Enfin, notons que plus le TR est grand, plus la correction du délai temporel intercoupe est nécessaire ; malheureusement un TR long implique également un pas d'échantillonnage faible, rendant la réinterpolation du signal imprécise. Souvent négligée, la correction du délai intercoupe n'a pas de solution simple et pourtant elle peut avoir des conséquences importantes sur la sensibilité lors de la détection d'activations [Henson *et al.*, 1999].

3.1.3 Correction du mouvement

Lors de l'acquisition de séries d'images fonctionnelles, il est demandé aux sujets d'éviter tout mouvement de la tête ; ils y sont d'ailleurs "aidés" par divers moyens de contention. Malheureusement les séquences d'acquisition sont souvent longues et les mouvements de la tête sont inévitables, surtout dans le cas de protocoles faisant intervenir des tâches motrices. Le rythme cardiaque et la respiration y contribuent également. Il est donc fréquent d'observer un mouvement de la tête du sujet, soit au cours d'une séquence d'acquisition, soit entre deux sessions successives d'une même expérience. Les déplacements engendrés par ces mouvements peuvent avoir deux conséquences concomitantes : une perte de sensibilité lors de la détection d'activations et l'apparition d'activations artefactuelles à proximité de zones où les niveaux de gris varient rapidement (interface cerveau-LCR).

Le *réalignement* des volumes fonctionnels acquis se place dans le contexte du recalage rigide monomodal : il s'agit de trouver les transformations rigides à 6 degrés de liberté (3 pour la translation et 3 pour la rotation) permettant de réaligner chaque volume avec un volume de référence (classiquement le premier de la série). L'estimation de ces paramètres se fait par la minimisation d'un critère de similarité, le plus souvent fondé sur des comparaisons d'intensité entre les images (mesure iconique). [Roche, 2001] a proposé une classification des mesures de similarité iconiques en fonction des hypothèses formulées sur le type de dépendance entre les intensités des images à recalcr ; nous présentons ici cette taxonomie associée aux principales mesures de similarité correspondantes :

- **conservation de l'intensité** : la somme des différences au carré (SSD, *Sum of Square Differences*) par exemple utilisée dans [Friston *et al.*, 1995a],
- **dépendance affine** : le coefficient de corrélation,
- **dépendance fonctionnelle** : le critère de Woods [Woods *et al.*, 1993], le rapport de corrélation [Roche *et al.*, 1998],
- **dépendance statistique** : l'information mutuelle [Collignon *et al.*, 1995; Studholme *et al.*, 1996; Viola and Wells III, 1997].

Dans le cadre du réalignement de volumes fonctionnels correspondant à la même modalité, l'hypothèse d'une relation de conservation des intensités apparaît la plus naturelle. Ainsi la somme des différences au carré est une des mesures de similarité les plus utilisées dans ce contexte.

La méthode d'optimisation est également importante car le résultat du recalage dépend alors de sa sensibilité aux minima locaux du critère. Classiquement, les méthodes choisies font soit partie de la famille des algorithmes de descente de gradient comme la méthode de Newton [Friston *et al.*, 1995a], soit des méthodes génériques évitant le calcul de gradient du critère comme la méthode de Powell [Jenkinson *et al.*, 2002].

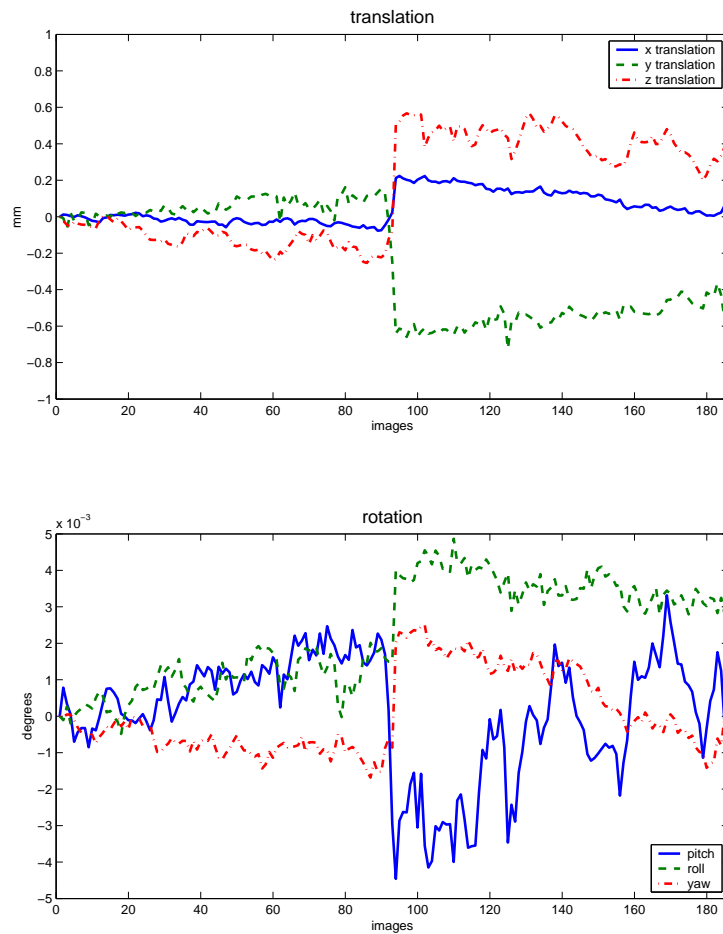


FIG. 3.2 – Paramètres des mouvements estimés (translation et rotation) avec l'algorithme présenté dans [Friston *et al.*, 1995a] sur un protocole constitué de 186 images acquises en deux sessions. On note que les mouvements de translation sont inférieurs au millimètre et que les paramètres de rotation restent inférieurs au degré. Par contre, on remarque que le sujet a bougé entre les deux acquisitions correspondant aux deux sessions. De plus, il semble que les mouvements mesurés soient corrélés avec le paradigme (périodique par bloc, pour une tâche de préhension).

[Freire and Mangin, 2001] démontrent cependant que les algorithmes reposant sur un

critère SSD ne sont pas robustes face aux données aberrantes (*outliers*) dont les voxels activés font partie. En effet, les voxels contenant des variations hémodynamiques violent l'hypothèse de conservation des intensités du critère SSD. L'estimation des paramètres de la transformation rigide est alors biaisée par les activations et ce biais est de plus corrélé avec le paradigme expérimental. Il en découle que si les images sont rééchantillonnées pour corriger les mouvements faussement détectés, il apparaît alors des activations artefactuelles, qui sont d'autant plus importantes que les activations "réelles" sont fortes et que le champ magnétique à l'intérieur de l'IRM est important. Pour corriger cela, Roche a proposé l'utilisation d'estimateurs robustes tel que l'estimateur de Geman-McClure pour prendre en compte la présence d'éventuels outliers. Une implémentation de cet algorithme, adaptée de [Friston *et al.*, 1995a] est disponible en ligne⁴.

[Freire *et al.*, 2002] ont ensuite étudié le choix de la mesure de similarité pour la correction des mouvements en IRMf. Il apparaît ainsi que les critères basés sur l'information mutuelle ou le rapport de corrélation apparaissent plus robustes, avec des estimations moins biaisées par les activations. Destinées à l'origine au recalage multimodal, ces mesures de similarité peuvent toutefois contribuer à améliorer la détection des mouvements intervenant dans une série d'images de même modalité.

Cependant, même après correction du mouvement, il s'avère que les mouvements estimés du sujet expliquent encore une grande part de la variance des signaux mesurés. Si ces mouvements ne sont pas corrélés avec le paradigme expérimental, une solution consiste à enlever la composante mouvement de chacun des décours temporels par filtrage, ou d'ajouter les estimations des six paramètres de mouvement rigide comme variable de non-intérêt au sein d'un modèle explicatif des données (nous reviendrons sur un tel modèle dans la suite de ce chapitre) mais cela se fera au détriment de la sensibilité de la détection d'activation. Le cas où le mouvement est réellement corrélé avec le paradigme expérimental reste enfin un problème délicat à gérer [Friston *et al.*, 1996]. En effet, en ce cas, il n'est plus possible d'ajouter les paramètres de mouvement dans le modèle. [Andersson *et al.*, 2001] ont pour cela proposé une explication à la présence de variance résiduelle après correction du mouvement : la position de la tête du sujet dans l'IRM influe sur le champ magnétique local et les artefacts associés vont ainsi dépendre des mouvements de la tête. Une correction du mouvement, même non biaisée, ne va ainsi pas suffire pour rendre les images comparables les unes avec les autres. Les auteurs parlent d'interaction de susceptibilité par le mouvement. La solution qu'ils proposent consiste à estimer les champs de déformation engendrés par les mouvements (eux-même préalablement estimés avec une des méthodes décrites précédemment) et de corriger les images pour supprimer ces artefacts.

Ainsi, il apparaît qu'il est souvent difficile de considérer indépendamment chacune des étapes de prétraitement des données (correction du *slice timing*, des artefacts et du mouvement) car elles interagissent : la présence de mouvement engendre des artefacts supplémentaires et rend caduque la correction classique des délais inter-coupes. Une solution consiste à modéliser l'ensemble de ces prétraitements comme un problème 3D+t non séparable [Bannister *et al.*, 2002] mais ceci se fait au détriment du temps de calcul.

4. INRIAalign: <http://www-sop.inria.fr/epidaure/software/INRIAalign/>

3.1.4 Mise en correspondance des images anatomiques et fonctionnelles

Une stratégie fréquente consiste à reporter les résultats des analyses fonctionnelles sur une image anatomique, acquise au cours de l'expérience. La mise en correspondance des images anatomiques et fonctionnelles d'un sujet, aussi appelée recalage, consiste à définir la transformation spatiale permettant de passer des images fonctionnelles à l'image anatomique. On a donc affaire ici à un problème de recalage multimodal que l'on peut considérer comme rigide si les distorsions intervenant sur les images EPI sont faibles ou ont été corrigées.

Comme nous l'avons vu pour la correction du mouvement, l'information mutuelle [Collignon *et al.*, 1995] ou le rapport de corrélation [Roche *et al.*, 1998] peuvent être utilisés comme mesure de similarité pour le recalage multimodal. On peut également essayer de se ramener artificiellement dans le cadre du recalage monomodal en synthétisant à partir des images d'une modalité une nouvelle image ressemblant en intensité à l'image de l'autre modalité. Ainsi, dans le cas de recalage TEP-IRM, [Friston *et al.*, 1995a; Ashburner and Friston, 1997] proposent de segmenter dans l'image anatomique la matière grise, la matière blanche et le LCR pour construire, par combinaison linéaire de ces segmentations, une image ressemblant à une image TEP qui peut alors être recalée avec une image de référence de la série TEP en utilisant un critère SSD comme dans le cas monomodal.

Enfin, dans le cas d'une expérience faisant intervenir plusieurs sujets ou pour comparer les résultats de différentes expériences, il faut effectuer la mise en correspondance d'images provenant de différents sujets, ce qui pose de nombreux problèmes conceptuels. Nous reviendrons plus longuement sur ce point, appelé *normalisation spatiale* dans la section suivante.

3.1.5 Lissage spatial des données

Le filtrage spatial passe-bas (lissage) des images est un prétraitement fréquemment utilisé lors de l'analyse des images fonctionnelles [Frackowiak *et al.*, 1997]. En pratique, ce filtrage est effectué par convolution avec un filtre gaussien isotrope de largeur à mi-hauteur donnée (FWHM, *Full Width at Half Maximum*). Les motivations sont multiples :

- améliorer le rapport signal sur bruit,
- renforcer l'hypothèse que les images sont des représentations discrètes de champs aléatoires gaussiens,
- atténuer la variabilité anatomique et fonctionnelle entre sujets, dans le cas d'analyses de groupe.

Classiquement, le lissage effectué a une largeur à mi-hauteur de une à deux fois la taille des voxels pour une analyse mono-sujet et le double pour une analyse multi-sujets.

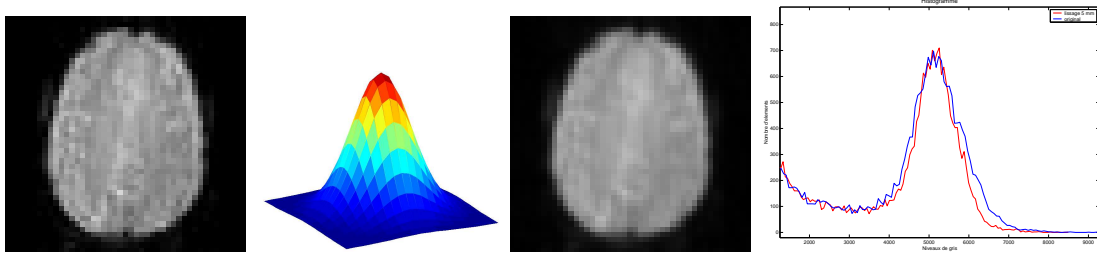


FIG. 3.3 – Coupe axiale d'une image fonctionnelle (voxels de $3.75 \times 3.75 \times 4.5 \text{ mm}^3$) avant et après lissage par un filtre gaussien isotrope de FWHM 5 mm. La courbe de droite représente l'histogramme de l'image initiale et de l'image filtrée : on note la réduction de la dynamique de l'image.

3.2 Normalisation spatiale des images

3.2.1 Présentation du problème

La normalisation spatiale consiste à placer dans le même espace spatial les données provenant de plusieurs sujets distincts [Fox, 1995]. En effet, afin de pouvoir tirer des conclusions à l'échelle d'une population et non pas d'un simple sujet, les expériences fonctionnelles font classiquement appels à de nombreux sujets. Pour pouvoir les comparer, il est alors nécessaire d'établir une transformation spatiale permettant d'obtenir une correspondance unique point à point entre chaque position dans un cerveau et sa position correspondant dans un autre cerveau : c'est le but du *recalage intersujet*.

Par ailleurs, un des principaux objectifs de la neuroimagerie fonctionnelle est de participer à la cartographie des fonctions cérébrales. Pour cela, les résultats d'expériences effectuées par divers laboratoires à travers le monde doivent pouvoir être comparés en étant placés dans un *espace standard*. On peut alors construire un *atlas* numérique sur lequel sont représentées l'ensemble des aires activées détectées et effectuer par la suite des *meta* analyses reposant sur l'étude à un deuxième niveau, de résultats d'expériences déjà effectuées. A l'espace standard utilisé est également associé un *système de coordonnées* permettant de définir une position "géographique" standard, universelle, de chacun des éléments du cerveau.

Les images anatomiques du cerveau étant de meilleure résolution et affichant plus précisément les détails anatomiques que les images fonctionnelles, ce sont ces premières que l'on recherche à mettre en correspondance. Les images fonctionnelles de chaque sujet ayant été au préalable recalées sur les images anatomiques correspondantes (*coregistration*), les transformations estimées sur les images anatomiques peuvent être appliquées à l'ensemble des images fonctionnelles. Ainsi, à la fin du processus, images anatomiques et fonctionnelles se retrouvent toutes dans le même espace, avec une correspondance point à point.

Malheureusement, on observe une incroyable variabilité interindividuelle des cerveaux humains. Si, de prime abord, deux cerveaux semblent être comparables, on observe rapide-

ment qu'il existe de grandes différences globales (volume, forme générale) mais également une grande diversité dans l'anatomie sulco-gyrale [Ono *et al.*, 1990]. Ainsi les différences topologiques sont nombreuses, même pour un sillon considéré comme stable à travers les individus tel que le sillon central : celui-ci présente parfois deux composantes connexes au lieu d'une. Et il en est de même pour l'architecture microscopique, la cytoarchitecture [Amunts *et al.*, 1999], avec d'importantes variabilités en taille et en localisation des différentes aires cytoarchitectoniques à travers les individus. Le recalage intersujet apparaît donc comme une tâche extrêmement difficile pour laquelle même la définition du problème reste floue.

De plus, dans le cadre des neurosciences, le problème de la normalisation spatiale ne devrait pas être défini comme la mise en correspondance de structures anatomiques (i.e. ce que l'on observe sur les IRM anatomiques) mais plutôt comme celui des aires corticales (définies fonctionnellement). Malheureusement cette information n'est pas disponible dans les images anatomiques et l'hypothèse de travail utilisée est celle d'une *corrélation anatomo-fonctionnelle* : le recalage de structures anatomiques (restant à déterminer lesquelles sont pertinentes) participerait ainsi à celui d'aires fonctionnelles sous-jacentes. Plus précisément, il est reconnu qu'il existe entre les individus de grandes variabilités anatomiques et fonctionnelles et on peut considérer qu'une partie de la variabilité fonctionnelle s'explique par la variabilité anatomique [Tzourio *et al.*, 1998]. L'objectif de la normalisation spatiale est alors de minimiser la variabilité anatomique (les différences interindividuelles des formes macroscopiques du cerveau) pour que seule la variabilité fonctionnelle perdure au cours de l'étape suivante (la détection d'activations).

Enfin, une fois les images fonctionnelles normalisées et analysées, un des points importants et de pouvoir décrire la localisation fonctionnelle des différentes aires activées [Brett *et al.*, 2002b]. Cette étape de *labellisation des activations* est fortement liée à celle de la normalisation spatiale. Ici encore il est fait appel à la relation entre anatomie et fonction : comment décrire d'un point de vue fonctionnel la localisation d'une aire activée dans le cerveau dont on dispose de sa position spatiale dans un référentiel standard ?

Nous nous proposons de présenter dans la suite de cette section quelques pistes pour l'ensemble des points clés évoqués ci-dessus concernant la normalisation spatiale : hypothèses de corrélation anatomo-fonctionnelle, définition d'un espace standard associé à un système de coordonnées, utilisation d'un atlas et labellisation des activations.

3.2.2 Corrélation anatomo-fonctionnelle

Nous avons déjà évoqué l'hypothèse d'une corrélation anatomo-fonctionnelle dans le chapitre précédent avec les travaux de Broca, Wernicke ou Brodmann.

En particulier, si l'on observe précisément les cartes cytoarchitectoniques de Brodmann de la figure 2.6, on peut noter une corrélation entre l'organisation cytoarchitectonique et la spécialisation régionale fonctionnelle [Roland and Zilles, 1998]. Ainsi si l'on reprend les cartes de Brodmann, on peut attribuer chacune des 51 aires de Brodmann à des aires fonctionnelles comme celles impliquées pour la vision, l'audition ou la parole. La figure 3.4 présente une proposition de l'attribution fonctionnelle des aires de Brodmann réalisée par le professeur Mark Dubin de l'université du Colorado, USA.

Ainsi la structure microanatomique du cerveau permettrait de faire un lien avec la

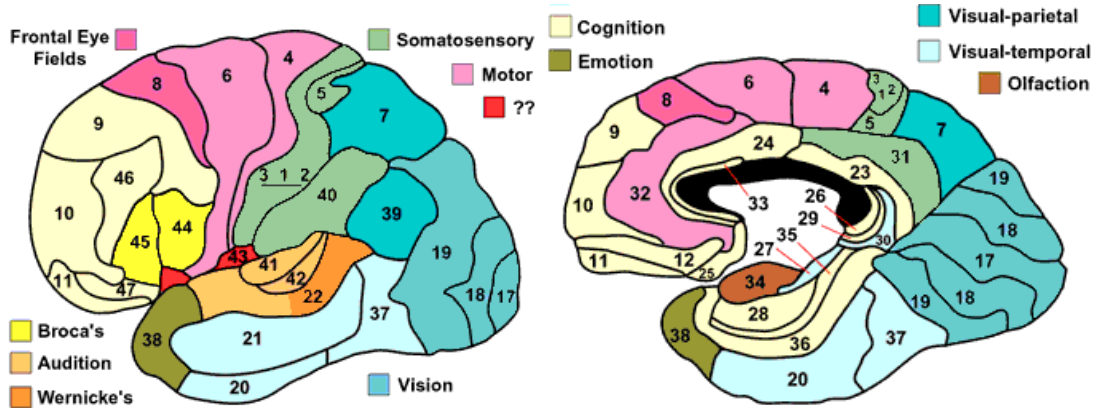


FIG. 3.4 – Attribution fonctionnelle des cartes cytoarchitectoniques de Brodmann. A comparer avec les cartes de Brodmann présentées sur la figure 2.6. D'après <http://spot.colorado.edu/~dubin/talks/brodmann/brodmann.html>

fonction cérébrale. Il faut cependant noter qu'il existe une importante variabilité en taille et en localisation des différentes aires cytoarchitectoniques à travers les individus [Amunts *et al.*, 1999].

On note également sur les figures 2.6 et 3.4 que certaines frontières d'aires cytoarchitectoniques correspondent à des sillons. C'est le cas par exemple pour le sillon central qui effectue la transition sensori-motrice. Ce sillon a été particulièrement étudié par Wilder Penfield et Théodore Rasmussen au centre de recherche en neurophysiologie du Royal Victoria Hospital de Montréal en 1950 [Penfield and Rasmussen, 1950]. Ils ont ainsi construit une topographie corticale sensitive et sensorielle du sillon central, présentée sur la figure 3.5. On observe une correspondance topologique entre les différentes parties du corps humain et leur représentation sur ce sillon : on parle de somatotopie. Cette carte topographique permet de générer un *homonculus*⁵ pour lequel chaque segment corporel représenté est proportionnel à la surface corticale de l'aire correspondante dans le sillon central.

On retrouve cette notion de somatotopie pour d'autres fonctions. Ainsi, pour les aires visuelles, les zones dédiées à une région du champ de vue se retrouvent à proximité sur la surface corticale. On parle en ce cas de rétinotopie. Le système visuel, invariablement situé au sein de la scissure calcarine est donc un autre exemple de corrélation forte entre anatomie et fonction.

Une hypothèse favorisant la corrélation anatomo-fonctionnelle serait que les sillons corticaux seraient des frontières entre les zones fonctionnelles [Welker, 1990]. Si cette hypothèse se vérifie pour les principaux sillons et les aires primaires, il en est différemment pour les autres sillons, qui sont par ailleurs beaucoup moins stables que ce soit en position ou en morphologie [Ono *et al.*, 1990]. Ainsi dans le lobe frontal, la morphologie sulcale du cortex préfrontal est très variable et on observe que très peu de relation entre anatomie sulcale et cytoarchitectonie ; les frontières des aires cytoarchitectoniques restant elle-mêmes incertaines.

5. *petit homme* en latin.

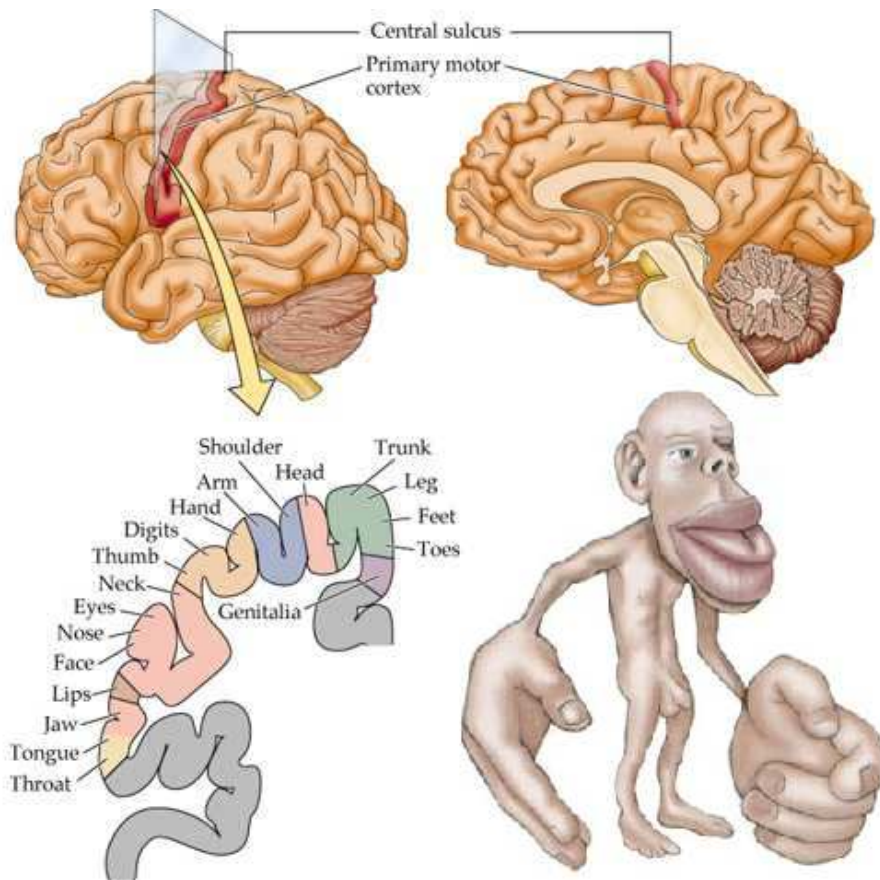


FIG. 3.5 – Carte topographique du sillon central mettant en évidence le cortex somatosensoriel primaire et Homunculus de Penfield. D'après http://www.utdallas.edu/~tres/integ/mot2/display2_08.html et http://www.vanderbilt.edu/neuro_majors/homunculus.htm

Une théorie plus récente avance que les racines sulcales (les fonds des premiers sillons apparaissant sur le cortex du fœtus) seraient des amers anatomiques plus stables que les sillons eux-mêmes [Régis *et al.*, 1995]. En effet, au début du développement fœtal, le cerveau ne possède que quelques plis faiblement marqués commun à tous les fœtus. Ce n'est que lors de son évolution ultérieure, lorsque le cortex gagne en superficie, que ces plis initiaux plongent dans le cerveau et que des plis additionnels se créent. Au demeurant, ce développement pourrait expliquer pourquoi certains de ces plis sont des limites de zones fonctionnelles. Le reste du sillon est moins stable anatomiquement puisque sa forme dépend de l'étendue du processus de plissement du cortex, qui varie sensiblement en fonction des individus, et explique la présence ou l'absence de plis ou de branches secondaires autour du sillon principal. En particulier, cette hypothèse que les racines sulcales seraient similaires à tous les individus est particulièrement étudiée au service hospitalier Frédéric Joliot par Jean-François Mangin et son équipe.

Pour conclure, il apparaît qu'il existe une certaine corrélation anatomo-fonctionnelle validant en partie le fait de recalcr les cerveaux de sujets différents guidé par l'anatomie.

En particulier, le recalage anatomique peut être guidé par la mise en correspondance d'amers géométriques pertinents tels que les sillons et plus précisément les racines sulcales. Toutefois, il existe de nombreuses aires cervicales dont la correspondance entre sujets est inconnue, ou tout du moins ne fait pas l'unanimité chez les experts. Certains s'interrogent même sur la faisabilité d'un recalage fonctionnel (ce que l'on désire en neuroimagerie) en utilisant uniquement de l'information anatomique.

3.2.3 Algorithmes de normalisation spatiale

Il existe dans la littérature de nombreux algorithmes de recalage en imagerie médicale. Nous invitons le lecteur à se reporter sur les états de l'art publiés par [Maintz and Viergever, 1998; Hill *et al.*, 2001; Modersitzki, 2004] pour une vision exhaustive du domaine.

Classifier l'ensemble des algorithmes de recalage est une tâche ardue car il en existe un nombre important, souvent adaptés à un contexte précis. [Maintz and Viergever, 1998] propose neuf paramètres permettant de les ordonner. La normalisation spatiale est un cas particulier de recalage monomodal 3D/3D entièrement automatique. Les principaux éléments constitutifs d'un algorithme de recalage sont au nombre de trois :

1. une classe de transformations : linéaire (rigide, affine), linéaire par morceaux, non linéaire (polynomiale, base de fonctions (Fourier, Cosinus, ondelettes), splines, modèle physique),
2. un critère à minimiser ou une mesure de similarité (voir ci-dessous),
3. et un algorithme de minimisation (descente de gradient, algorithme de Powell, recherche stochastique ou exhaustive, etc) avec d'éventuelles optimisations (approche multi-résolution et multi-échelle) afin d'accélérer la convergence et éviter les minima locaux.

Concernant le critère à minimiser, on définit classiquement deux classes de mesures de similarité selon la nature des paramètres utilisés pour le calcul de ladite mesure :

- les intensités des voxels : en ce cas, ce sont directement les niveaux de gris des images qui sont utilisés pour calculer le critère. On parle de recalage *iconique*.
- des structures géométriques : en ce cas, des particularités géométriques sont extraites des images et ce sont celles-ci qui sont injectées dans le calcul de la mesure de similarité. On parle de recalage *géométrique*.

Dans le cadre de l'exposé unidimensionnel que représente la lecture de ce mémoire, nous avons estimé que l'axe de présentation le plus pertinent pour la normalisation spatiale était le nombre de degré de liberté de la transformation recherchée. Les deux principaux sous-groupes sont les algorithmes rigides et non-rigides, ces derniers regroupant eux-mêmes diverses catégories.

Approche rigide

Nous avons déjà abordé le recalage rigide monomodal lors de la correction de mouvement dans les séries fonctionnelles. Ces méthodes reposent classiquement sur la maximisation de l'information mutuelle [Collignon *et al.*, 1995; Studholme *et al.*, 1996;

Viola and Wells III, 1997], du rapport de corrélation [Roche *et al.*, 1998] et de leurs différentes variantes. Pour le recalage intersujet, l'approche rigide est bien trop contrainte car elle ne permet pas de modéliser la variabilité anatomique observée sur les cerveaux de différents sujets. Cependant, un recalage rigide est fréquemment préconisé comme une première étape pour la normalisation car il permet par la suite de s'affranchir de l'estimation de ces paramètres pour se concentrer sur des déformations plus locales.

Approche proportionnelle

L'approche géométrique la plus ancienne mais encore très utilisée en neurosciences est la méthode proposée Talairach et Tournoux [Talairach and Tournoux, 1988; 1993], permettant de gérer la variabilité des formes du cerveau. Ceux-ci ont en fait publié un atlas stéréotaxique du cerveau, introduisant trois nouveaux concepts : un atlas papier macro- et micro-anatomique d'un cerveau, un système de coordonnées associé à cet atlas ainsi qu'une méthode de normalisation dite proportionnelle permettant de plonger un cerveau dans cet espace.

Le système de coordonnées de Talairach repose sur le repérage des commissures antérieures (CA) et postérieures (CP), supposées invariantes entre les individus, le long du plan inter-hémisphérique. L'axe CA-CP associé aux axes passant à la verticale de ces deux amers anatomiques dans le plan inter-hémisphérique permettent de définir un ensemble de boîtes dont les faces sont tangentes aux bords du cerveau (voir la figure 3.6). La construction d'une grille tridimensionnelle dans chacun des 12 parallélépipèdes ainsi obtenus permet de définir un système de coordonnées : un triplet (x,y,z) permet de repérer n'importe quel point de l'espace stéréotaxique de Talairach.

Talairach et Tournoux ont alors proposé une méthode de normalisation proportionnelle permettant de transformer un cerveau en un autre en adaptant globalement leurs formes et leurs volumes. Pour cela, les commissures antérieures et postérieures des deux sujets à recaler sont alignées en effectuant un changement d'échelle global des cerveaux. Ensuite, des facteurs de normalisation sont appliqués sur chacune des 12 boîtes englobantes du cerveau constituant l'espace de Talairach afin qu'elles soient de dimensions égales pour les deux cerveaux. Cette méthode correspond donc à un recalage affine par morceaux. Elle permet de recaler globalement des cerveaux en corrigeant pour les différences de volumes et de formes.

Enfin, Talairach et Tournoux ont publié un atlas papier d'un hémisphère d'un cerveau placé dans cet espace stéréotaxique [Talairach *et al.*, 1967]. Nous y reviendrons dans la section suivante car il concerne plus précisément l'étape de labellisation des activations.

Etant initialement destinée au repérage des structures cérébrales profondes, comme le thalamus, à des fins neurochirurgicales, la normalisation proportionnelle de Talairach donne de bons résultats pour le recalage des structures internes mais reste en revanche beaucoup plus approximative pour l'anatomie corticale. La figure 3.7 met en évidence la variabilité résiduelle de la position des sillons de nombreux cerveaux normalisés dans l'espace de Talairach.

En effet, une des principales critiques formulées envers les approches géométriques (dont l'approche proportionnelle fait partie – les commissures antérieures et postérieures étant les primitives, ici ponctuelles, à recaler) est qu'elle entraîne de grandes variabilités

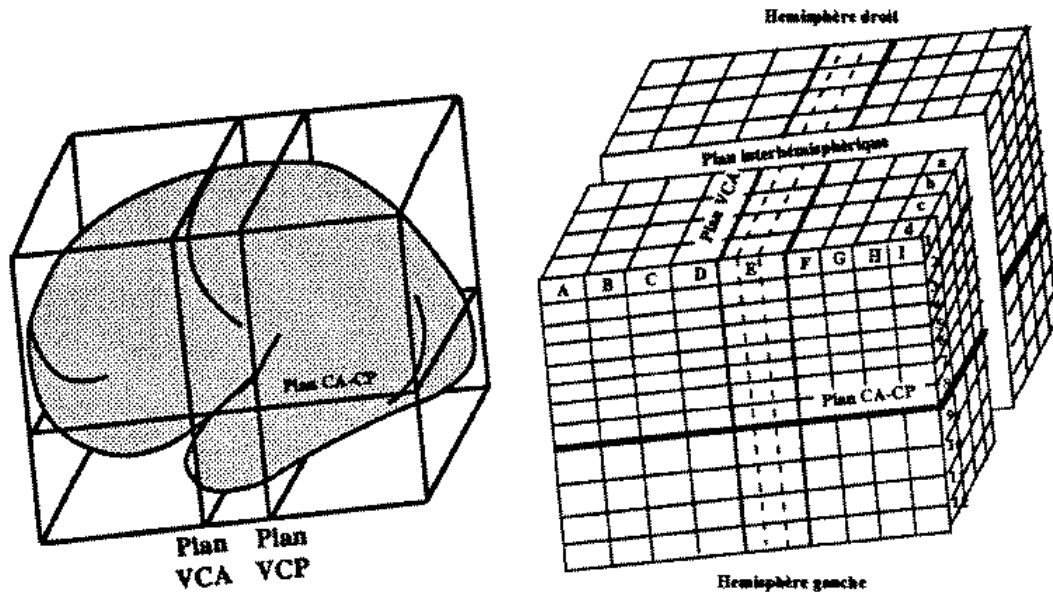


FIG. 3.6 – Le système de grille proportionnelle tridimensionnelle de Talairach [Talairach and Tournoux, 1988]. Chaque structure cérébrale a des coordonnées dans l'espace 3D, ce qui permet de la comparer directement aux structures correspondantes identifiées à la même position chez d'autres sujets.

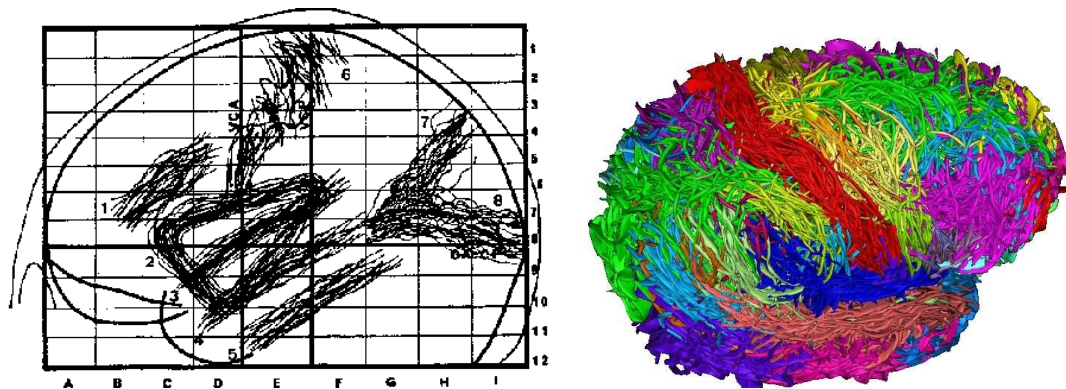


FIG. 3.7 – Variabilité en forme et en position des principaux sillons dans l'atlas de Talairach [Talairach et al., 1967]. L'image de droite est constituée de la superposition des sillons de 50 cerveaux normalisés dans l'espace de Talairach, obtenue par Jean-François Mangin.

sur la position de structures éloignées des repères anatomiques. De nombreuses études ont montré l'insuffisance de l'approche proportionnelle de Talairach pour la normalisation spatiale en mesurant des variations de plusieurs centimètres pour la localisation de sillons

corticaux [Van Essen *et al.*, 1998; Thompson and Toga, 1996].

Augmentation du nombre de degré de liberté des transformations admissibles

Nous venons de voir que malgré sa popularité, le système de normalisation proportionnelle de Talairach est vivement critiqué en raison de son imprécision dans la mise en correspondance de l'anatomie sulco-gyrale. Ainsi la plupart des méthodes utilisées aujourd'hui pour placer une image dans la grille proportionnelle définie par Talairach diffèrent de l'approche originale proposée par celui-ci. Elles reposent le plus souvent sur le recalage sur une image cible correspondant à un modèle (*template*) placé préalablement dans le repère proportionnel de Talairach. Ainsi une image recalée sur un de ces modèles est alors directement placée dans le repère de Talairach, permettant d'utiliser ses coordonnées pour les comparer avec les images d'autres sujets.

Le *Montreal Neurological Institute* (MNI) a, à cet effet, distribué dans la communauté des neurosciences plusieurs modèles qui sont depuis utilisés comme standard pour la normalisation spatiale [Evans *et al.*, 1992; 1993]. Le premier modèle, MNI305, a été obtenu à partir de la moyenne de 305 cerveaux normalisés linéairement [Collins *et al.*, 1994] sur un modèle placé dans Talairach lui-même constitué du moyennage de 241 sujets recalés manuellement dans l'espace de Talairach. Le deuxième modèle, utilisé comme référence internationale dans de nombreux projets, est lui basé sur le moyennage de 152 cerveaux recalés sur le MNI305 par une méthode linéaire à 9 paramètres. Ce dernier modèle est par ailleurs disponible pour les modalités T_1 , T_2 et densité de proton. La figure 3.8 présente quelques coupes extraites de ce modèle.

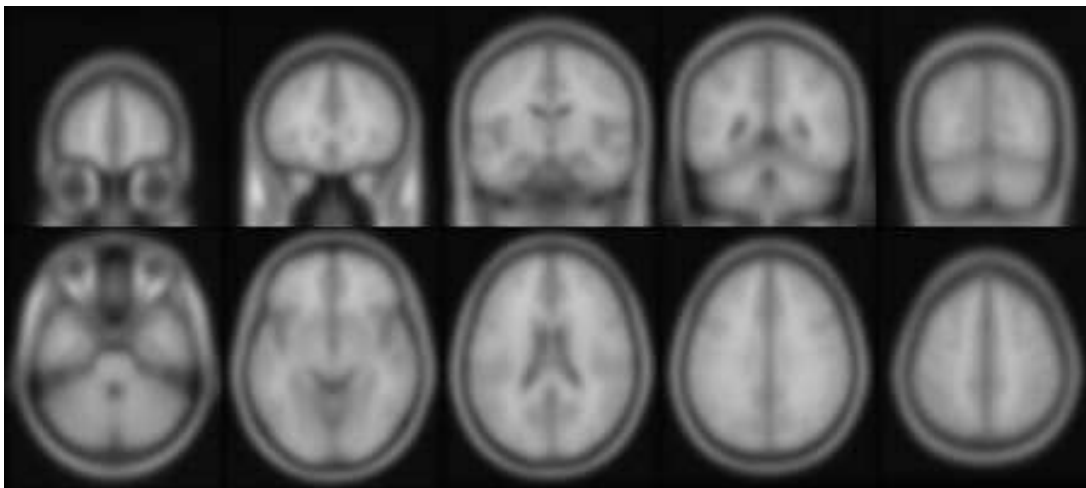


FIG. 3.8 – Coupes axiales et coronales du modèle MNI utilisé comme le standard international dans de nombreux projets de cartographie cérébrale. On note la perte de nombreux détails anatomiques en raison du flou introduit par le moyennage de sujets comportant une variabilité anatomique.

Malheureusement, on observe des différences de forme et de volume entre le cerveau utilisé dans l'atlas de Talairach et le cerveau moyen du MNI, en particulier dans les lobes

temporaux (voir la page internet de Matthew Brett⁶ à ce sujet). Cette différence peut s'expliquer par le fait que l'atlas de Talairach repose sur un seul cerveau, d'une femme âgée, alors que le template MNI est constitué d'un ensemble beaucoup plus important de sujets plus jeunes. Il est alors dangereux d'utiliser les coordonnées des images recalées sur le template MNI comme étant les coordonnées du système de Talairach. [Brett *et al.*, 2001] propose une transformation à appliquer aux coordonnées d'images normalisées sur le template MNI pour obtenir des coordonnées dans l'espace de Talairach.

Dans certaines études de groupe où les sujets présentent des anomalies cérébrales, il peut être délicat de recalcr les images sur un modèle crée à partir de sujets sains. Une solution consister alors à choisir un des sujets comme étant le sujet de référence sur lequel tous les sujets seront normalisés mais cela peut entraîner un biais dans les résultats si le sujet choisi n'est pas représentatif du groupe. Pour construire un modèle permettant de s'affranchir de ces biais, Guimond propose un algorithme itératif convergeant vers un modèle anatomique moyen (en intensité et en forme) [Guimond *et al.*, 1998; 2000].

Les algorithmes permettant de recalcr le cerveau d'un sujet sur un modèle constitué d'une image tridimensionnelle utilisent principalement les approches iconiques – liées aux intensités. En effet, celles-ci permettent de s'affranchir du choix des amers anatomiques à recalcr et offrent également la possibilité d'estimer des transformations incluant de nombreux degrés de liberté. Un état de l'art des différents algorithmes de recalcrage non linéaire en imagerie médicale a par ailleurs été réalisée en 1999 [Lester and Arridge, 1999] et plus récemment pour l'imagerie cérébrale dans [Crum *et al.*, 2003].

Les algorithmes de recalcrage non rigide recherchent à estimer un champ de déformation dense associant un vecteur déplacement à chacun des voxels constituant l'image à recalcr. La mesure de similarité utilisée est la somme des différences d'intensité au carré (SSD) [Bajcsy and Kovačič, 1989] ou plus récemment l'information mutuelle [Rueckert *et al.*, 1999]. On peut séparer ces algorithmes en deux groupes selon le type du champ de déformation recherché. Dans un premier cas, le champ est défini paramétriquement, typiquement par une combinaison linéaire de fonctions de bases, telles que des cosinus [Friston *et al.*, 1995a; Ashburner and Friston, 1997; Ashburner *et al.*, 1997; Ashburner and Friston, 1999], des polynômes [Woods *et al.*, 1998], des B-splines [Rueckert *et al.*, 1999], des fonctions radiales de base [Rohde *et al.*, 2003], etc. Une seconde approche consiste à modéliser la déformation d'une image en une autre comme un processus physique de déformation d'un matériau élastique ou fluide [Bro-Nielsen, 1996; Christensen *et al.*, 1997] ou des régularisations les approximant [Thirion, 1998; Cachier, 2002; Cachier *et al.*, 2003; Hermosillo, 2003]. Dans tous les cas, les critères ou énergies à minimiser sont souvent modifiés pour ajouter un terme de régularisation contraignant la transformation à être régulière.

Un comparatif des principales méthodes de recalcrage non rigide a été récemment effectué [Hellier *et al.*, 2001a; 2002; 2003]. Les critères permettant de mesurer la qualité des recalcrages effectués reposaient sur des mesures globales (volume moyen, recouvrement des tissus de matière blanche et de matière grise, etc) et des mesures locales liées à l'appa-

6. <http://www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html>

riement des sillons. Les résultats présentés montrent que la qualité des recalages pour les mesures globales est directement reliée aux nombres de degré de liberté des transformations associées à chacun des algorithmes. Par contre, avec les critères locaux vérifiant le recalage du sillon central, on ne remarque pas de différences notables de recalage entre les méthodes rigides et non-rigides. [Hellier *et al.*, 2002] étudie également l'impact des différentes méthodes de recalage sur celui de données fonctionnelles (dipôles MEG). En ce cas, il est montré que l'approche proportionnelle de Talairach ainsi qu'une méthode non rigide (base de cosinus) diminuent la variabilité fonctionnelle (position des dipôles MEG) comparées à un recalage rigide (validant ainsi qu'une partie de la variabilité fonctionnelle est régie par la variabilité anatomique). Par contre, ces deux méthodes donnent des résultats similaires (avec même un léger avantage pour l'approche de Talairach) alors que l'on pouvait espérer obtenir de meilleurs résultats pour l'approche non rigide, disposant de beaucoup plus de degrés de liberté.

Afin de pallier au manque de précision dans la mise en correspondance des structures anatomiques lors de la normalisation spatiale, une stratégie communément appliquée consiste à effectuer un filtrage spatial gaussien sur les images normalisées. Ce lissage permet ainsi de réduire la variabilité anatomique résiduelle entre les individus en favorisant le recouvrement des aires corticales fonctionnelles homologues. En revanche, cela diminue la résolution effective des images fonctionnelles et rend plus imprécise la localisation spatiale des activations détectées.

Une des critiques formulées envers les méthodes non rigides utilisées dans le comparatif est qu'elles ne prennent pas en compte les connaissances de plus haut niveau dont on dispose concernant le cortex et les sillons [Mangin *et al.*, 1998]. Ces algorithmes se rapprochent plutôt des algorithmes de type *morphing* en permettant une mise en correspondance de plus en plus précise de toutes les structures cérébrales mais cela sans réelle justification anatomique. Une voie ouverte récemment et mise en avant dans [Collins *et al.*, 1998; Hellier and Barillot, 2001; Cachier *et al.*, 2001] propose de contraindre les algorithmes iconiques non rigides avec la mise en correspondance d'amers géométriques tels que les sillons.

Approche hybride combinant géométrique et iconique

Les sillons corticaux, repères anatomiques et fonctionnels (supposés) du cerveau, peuvent être utilisés comme contrainte dans l'estimation de la transformation spatiale des algorithmes de recalage non rigide, afin d'en garantir la cohérence anatomique.

Pour ce faire, les principaux sillons sont segmentés et labellisés dans les deux images à recaler [Le Goualher *et al.*, 1997; Rivière *et al.*, 2002]. Une contrainte géométrique d'appariement des sillons homologues est alors ajoutée à l'énergie de recalage iconique à minimiser. Les résultats présentés montrent une plus grande précision obtenue pour le recalage des sillons que les méthodes "aveugles". En particulier, [Cachier *et al.*, 2001] observe que les algorithmes de recalage sans contraintes sulcales proposent parfois un recalage pour lequel deux sillons distincts se retrouvent superposés. L'application d'une contrainte géométrique supplémentaire permet alors d'éviter ce genre de résultat. [Barillot *et al.*, 2002] compare également les résultats obtenus entre une approche non rigide sans contraintes [Hellier *et al.*, 2001b] et une approche hybride recalant localement les données

autour d'un sillon par application des modes de variations obtenus par une étude statistique des formes des sillons [Corouge *et al.*, 2001]. Les résultats présentés pour le recalage de données fonctionnelles (dipôles MEG) donnent l'avantage aux approches mixtes, en diminuant la variabilité de la localisation de ces dipôles après recalage.

Dans ces méthodes proposées, ce sont les principaux sillons qui sont utilisés comme amers géométriques. Cependant, nous avons évoqué l'hypothèse selon laquelle les primitives géométriques stables au sein d'une population de cerveaux seraient plutôt les racines sulcales [Régis *et al.*, 1995; Mangin *et al.*, 1998], correspondant aux plis élémentaires apparaissant sur le cerveau du fœtus. Une étude récente [Cachia *et al.*, 2003a] tend à prouver que les racines sulcales pourraient être identifiées dans les cerveaux adultes en utilisant les variations de profondeur ou de courbure le long des lignes sulcales inférieures. Ce type d'approche plaide en faveur des approches structurales préconisées par [Mangin *et al.*, 2003].

Approche surfacique

Jusqu'à présent, nous avons présenté diverses techniques permettant de déformer un cerveau dans l'espace de Talairach, afin de pouvoir, entre autre, utiliser son système de coordonnées tridimensionnel. [Van Essen *et al.*, 1998] proposent une alternative à cette approche volumique en prenant en compte la nature surfacique du cortex. En effet, ce dernier est homotopique à une sphère et apparier les surfaces corticales apparaît ainsi pertinent. On se rapproche alors de la cartographie géographique terrestre.

Plus généralement, les approches surfaciques apportent trois améliorations aux approches conventionnelles [Van Essen and Drury, 1997]: une visualisation surfacique des résultats, la construction d'atlas surfaciques et la mise en place d'algorithmes de recalages surfaciques. En particulier, utiliser un système de coordonnées lié à la surface du cortex permet de préserver sa topologie: la distance mesurée dans l'espace 3D entre deux points de la surface corticale sous-estime drastiquement la vraie distance mesurée le long de la surface.

[Van Essen *et al.*, 1998] proposent ainsi d'extraire la surface corticale des images anatomiques puis de la déplier (*inflating, unfolding*). Des coupes sont alors effectuées sur la surface le long du fond de certains sillons pour être ensuite aplatie (*flattening*) sur un plan, en minimisant les distorsions introduites. Le cortex cérébral est alors représentée par une surface plane sur laquelle on peut définir un système de coordonnées bidimensionnel (voir figure 3.9).

Ces auteurs ont en particulier appliqué ce principe au cerveau du Visible Man⁷, considéré comme ayant une anatomie sulco-gyrale normale. La carte obtenue est utilisée comme référence pour déformer le cerveau d'un sujet sur celui-ci. Ils proposent ainsi un algorithme de recalage recherchant une déformation fluide des surfaces planes, guidée par un petit nombre d'amers anatomiques labellisés manuellement. De manière intéressante, cet approche surfacique permet d'effectuer des recalages inter-espèces, comme par exemple entre l'homme et le macaque.

D'autres équipes ont proposé des approches similaires, reposant sur le réalignement de

7. http://www.nlm.nih.gov/research/visible/visible_human.html

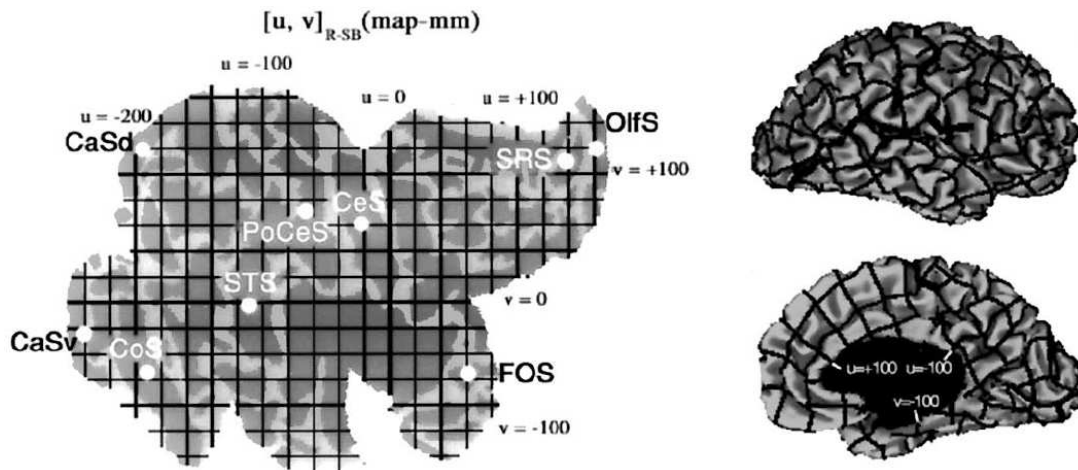


FIG. 3.9 – Définition d'un système de coordonnées bidimensionnel sur la surface du cortex [Van Essen and Drury, 1997]. L'image présente ici l'hémisphère gauche du Visible Man après dépliement. L'origine du système de coordonnées est définie relativement au sillon central et la grille représentée est espacée de 20 mm.

surfaces corticales [Thompson and Toga, 1996; Dale and Sereno, 1999; Fischl *et al.*, 1999a], permettant la construction d'atlas probabilistes surfaciques [Thompson *et al.*, 1996; 1997].

Les inconvénients de l'approche *applatissage* sont qu'elle entraîne de fortes distorsions géométriques pour permettre le dépliement complet et que ceci ne peut se faire qu'après la réalisation de plusieurs incisions manuelles sur la surface. Par ailleurs, le système de coordonnées cartésien défini n'est pas entièrement satisfaisant car deux points situés sur le bord de la surface aplanie peuvent se retrouver très éloignés alors qu'ils étaient proches sur la surface gonflée. Pour contourner cela, [Fischl *et al.*, 1999b] proposent de déformer les surfaces corticales sur une sphère en minimisant les distorsions. Un nouveau système de coordonnées (latitude et longitude) peut alors être défini sur cette surface (voir figure 3.10).

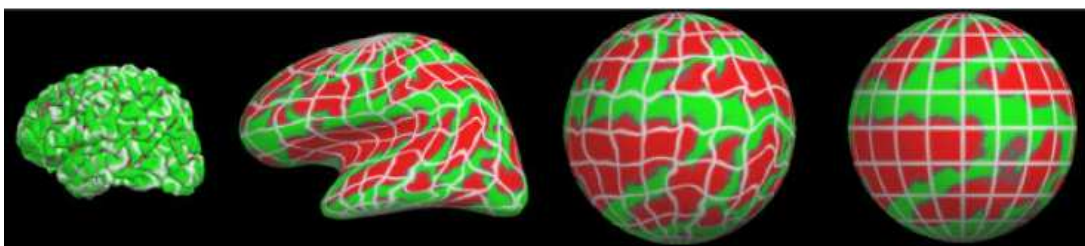


FIG. 3.10 – Différentes représentations de la surface corticale d'un individu. L'image de droite correspond au moyennage de 40 sujet dans ce repère sphérique [Fischl *et al.*, 1999b].

Fischl *et al.* proposent également un algorithme de recalage permettant de réaligner automatiquement la surface sphérique du cortex d'un individu sur celle obtenue par le moyennage de plusieurs sujets. Le critère à minimiser repose sur la différence quadratique

de courbure (représentative des sillons et gyri) des deux surfaces pénalisé par deux termes permettant de minimiser les distorsions.

Notons enfin les travaux récents de Roberto Toro [Toro and Burnod, 2003] définissant un modèle géométrique de l'anatomie corticale : le cerveau est modélisé par une sphère dont les principaux sillons sont les axes. Cette définition des axes est inspirée par l'arrangement des sillons observé sur une représentation stéréographique de la surface corticale. Grâce à ce modèle, on obtient une paramétrisation de la surface qui, à la différence des systèmes de coordonnées surfaciques angulaires, permet une localisation basée la distance du point considéré par rapport à des amers géométriques tels que les sillons.

3.2.4 Labellisation des activations fonctionnelles

Le problème de la labellisation des activations fonctionnelles reste un problème difficile en raison des réserves émises sur les hypothèses de corrélation anatomo-fonctionnelles [Brett *et al.*, 2002b; Bjaalie, 2002; Passingham *et al.*, 2002; Swallow *et al.*, 2003; Mangin *et al.*, 2003]. La normalisation spatiales des sujets sur un atlas, déterministe ou probabiliste, permet de proposer un label à chacune des aires fonctionnelles, relié à l'anatomie macroscopique ou microscopique ou à des aires cérébrales.

Systèmes de coordonnées

Une fois les données normalisées dans un espace standard, il est possible d'affecter des coordonnées, dites alors stéréotaxiques, à chacun des voxels de l'image. En particulier, nous avons déjà décrit le système de coordonnées tridimensionnel de Talairach et ceux bidimensionnels, sur la surface du cortex, de Van Essen (plan 2D) et Fischl (sphère). De nombreux articles de neuroscience présentent leurs résultats en indiquant les coordonnées de Talairach des maxima locaux d'activations des aires activées. Leur principal avantage est que les résultats d'expériences différentes peuvent ainsi être comparés aisément. Par contre, leur précision va dépendre de la qualité du recalage effectué.

Le projet *BrainMap Database*⁸ se propose de recenser l'ensemble des localisations présentées dans la littérature afin de favoriser les meta analyses. [Nielsen and Hansen, 2002] ont ainsi pu détecter la présence de données aberrantes dans certaines publications avec des erreurs de transcription, des confusions entre millimètres et centimètres, des inversions d'axes, etc. Par ailleurs, on note parfois des erreurs dans la labellisation des structures anatomiques correspondantes.

Labels macroanatomiques

Les labels macroanatomiques se réfèrent à l'anatomie sulco-gyrale. Ceux-ci sont particulièrement intéressants pour les régions où la corrélation anatomo-fonctionnelle est vérifiée. A partir de données normalisées dans l'espace de Talairach, il est possible d'utiliser l'atlas AAL (*Anatomical Automatic Labeling*)⁹ [Tzourio-Mazoyer *et al.*, 2002], construit à partir d'une parcellisation anatomique manuelle : 45 volumes d'intérêt sont définis pour

8. <http://www.brainmapdbj.org/>

9. <http://www.cyceron.fr/freeware/>

chaque hémisphère relativement aux sillons (aire motrice supplémentaire, gyrus supra-marginal, gyrus angulaire, etc). La figure 3.11 présente plusieurs coupes de cet atlas superposées sur l'IRM T_1 d'un sujet normalisé sur l'atlas du MNI.

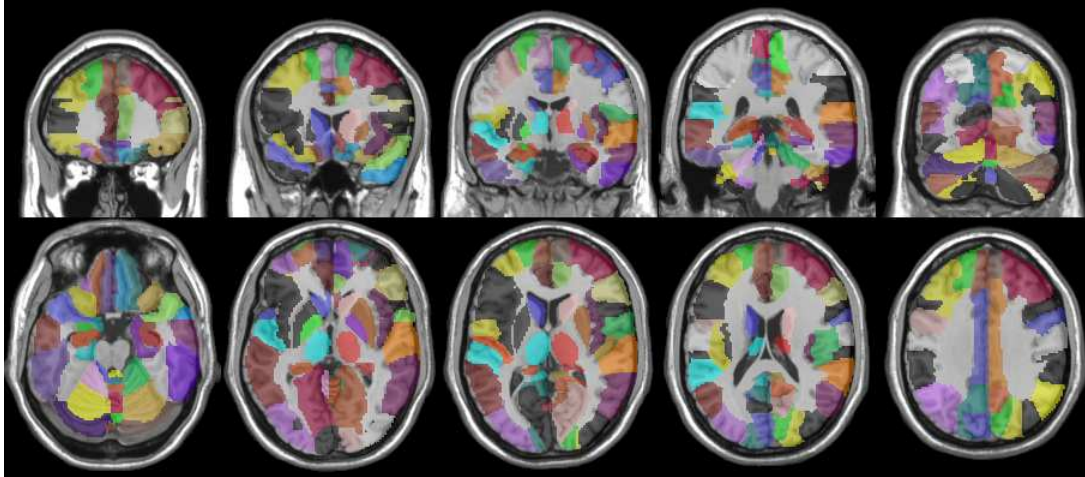


FIG. 3.11 – Différentes coupes axiales et coronales de l'atlas AAL [Tzourio-Mazoyer et al., 2002] définissant 45 volumes anatomiques d'intérêt par hémisphère sur le modèle MNI.

Ce type d'atlas a cependant l'inconvénient d'être déterministe et de ne pas prendre en compte la variabilité anatomique des cerveaux humains. Le consortium ICBM¹⁰ (*International Consortium for Human Brain Mapping*), regroupant plusieurs laboratoires de cartographie cérébrale nord-américains, a, parmi ses objectifs, la construction d'atlas probabilistes de la structure du cerveau, à la fois anatomique et fonctionnelle. Leurs travaux permettent ainsi, entres autres, d'étudier la variabilité de l'anatomie sulco-gyrale pour des sujets sains ainsi que pour des patients [Thompson et al., 1996; 1997]. A l'heure actuelle, plus de 7000 cerveaux ont été analysés, à la fois *in vivo* et *post mortem* pour la cytoarchitectonie.

Labels microanatomiques

Identifier l'aire de Brodmann associée à une zone cérébrale n'est pas chose aisée car il n'est pas possible à l'heure actuelle d'obtenir des cartes cytoarchitectoniques *in vivo* d'un sujet. Pour contourner ce problème, de nombreux chercheurs utilisent l'atlas papier de Talairach [Talairach and Tournoux, 1988]. Ce dernier a publié un atlas stéréotaxique du cerveau post-mortem d'une femme de 60 ans. Ce cerveau est placé dans l'espace de Talairach en utilisant la méthode proportionnelle relatives aux commissures antérieures et postérieures. Talairach et Tournoux ont alors étiquetés diverses coupes d'un hémisphère de ce cerveau avec des labels anatomiques relatifs aux principaux sillons et gyri. Puis ils ont estimé les localisations des 51 aires cytoarchitectoniques de Brodmann en comparant l'anatomie de ce cerveau avec les planches de Brodmann du début du XXème siècle. Il est alors possible d'associer un label anatomique et cytoarchitectonique à chaque point situé

10. <http://www.loni.ucla.edu/ICBM/>

dans l'espace de Talairach. Le *Talairach Daemon*¹¹ permet d'interroger en ligne une base de données avec des coordonnées tridimensionnelles et obtenir les labels correspondants [Lancaster *et al.*, 2000]. Il faut cependant prendre garde de la labellisation effectuée car de nombreux paramètres entrent en compte, chacun influençant sa validité : la précision de la normalisation spatiale dans l'espace de Talairach, l'imprécision du report des aires de Brodmann sur l'atlas de Talairach et la variabilité de la cytoarchitectonie à travers les individus.

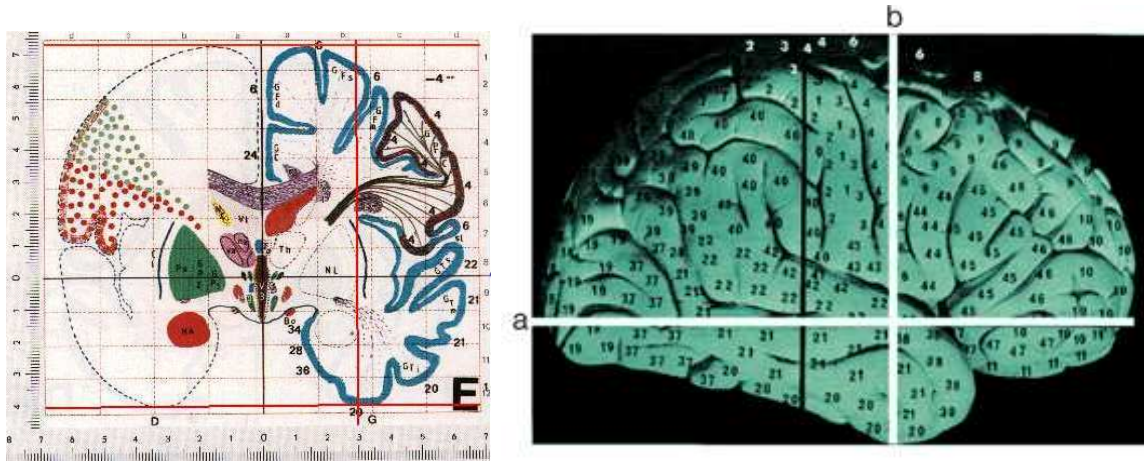


FIG. 3.12 – (à gauche) Planche de l'atlas de Talairach [Talairach and Tournoux, 1988] où un hémisphère est labellisé selon des critères anatomiques et cytoarchitectoniques. (à droite) Aires cytoarchitectoniques de Brodmann reportées sur le cerveau de l'atlas de Talairach.

Le programme international NeuroGenerator¹² se propose de regrouper dans une base de données les images IRMf normalisées, les résultats d'expériences ainsi que les cartes cytoarchitectoniques *post mortem* [Roland *et al.*, 1997; Roland and Zilles, 1994]. Ce projet vise à construire des atlas probabilistes de l'anatomie microscopique du cerveau et de les comparer avec les structures macroscopiques. NeuroGenerator et le consortium ICBM permettront ainsi à terme d'affiner les hypothèses de corrélation anatomo-fonctionnelles.

Labels fonctionnels

Enfin, une dernière possibilité pour labelliser les activations consiste à relier leur position par rapport à des aires fonctionnelles mises en évidence avec d'autres expériences. Une stratégie pouvant être mise en œuvre consiste alors à préalablement faire participer le sujet à une expérience simple, parfaitement calibrée (*localizers*) permettant de localiser des aires précises (*fusiform face area*, aires motrices, aires visuelles (V_1 , V_2 , MT)) dans l'anatomie du sujet. Il est alors possible de reporter les zones détectées par la nouvelle expérience sur la carte construite avec les aires

11. <http://ric.utahscsa.edu/projects/talairachdaemon.html>

12. <http://www.neurogenerator.org/>

définies fonctionnellement. Cette stratégie est fréquemment utilisée pour les expériences faisant intervenir les aires corticales sujettes à une organisation topologique (rétinotopie, somatotopie, tonotopie). Ainsi des labels fonctionnels ont été utilisés pour des expériences concernant l'imagerie mentale [O'Craven and Kanwisher, 2000; Klein *et al.*, 2003].

Notons enfin que de nouvelles solutions émergent grâce au progrès des techniques d'acquisition. [Kruggel *et al.*, 2003] ont ainsi réussi à caractériser de manière cytoarchitectonique les aires corticales d'un cerveau *post mortem* grâce à une acquisition IRM à très haute résolution (voxels de 0.035 mm^3 pour un temps d'acquisition de 12 heures). Les progrès en physique de l'IRM associés aux aimants à haut champ permettront peut-être à terme de pouvoir construire *in vivo* les cartes cytoarchitectoniques de différents sujets et ainsi ouvrir de nouvelles perspectives.

De même, des modalités récentes comme l'imagerie de diffusion permettent un accès *in vivo* à des informations sur la microstructure des tissus [Le Bihan *et al.*, 2001] et les différents faisceaux de fibres de la matière blanche connectant les aires cérébrales. [Wiegell *et al.*, 2003] ont ainsi pu labelliser les nucléi du thalamus en classifiant les voxels d'images pondérées en diffusion.

3.3 Analyse statistique des données fonctionnelles

3.3.1 Problématique

Le but de l'analyse des images fonctionnelles est d'extraire des données des informations pertinentes sur l'expérience au cours de laquelle les images ont été acquises. Généralement, l'objectif est de pouvoir tester des hypothèses reliées au protocole expérimental et construire des cartes statistiques d'activations.

L'analyse statistique des images fonctionnelles et plus précisément la détection d'activations est un problème délicat pour plusieurs raisons. Tout d'abord, la quantité de données peut être un facteur limitant dans les techniques d'analyses. En effet, on dispose de plusieurs centaines à plusieurs milliers de volumes tridimensionnels constitués de plusieurs dizaines de milliers de voxels; et ceci pour chacun des sujets participant à une expérience. Ensuite, le signal recherché dans les données n'est que très partiellement connu (couplage hémodynamique, caractéristiques de l'imageur, influence des différents artefacts) et fait l'objet de corrélations à la fois temporelles et spatiales. Enfin, le rapport signal sur bruit relié à l'effet BOLD est très faible.

Il existe un nombre assez important de logiciels permettant d'analyser les images fonctionnelles, certains se restreignent à l'analyse statistique (sous plusieurs formes) tandis que d'autres permettent aussi d'effectuer l'ensemble des prétraitements que nous avons décrit dans les sections précédentes. Le logiciel le plus utilisé dans la communauté des neurosciences est SPM¹³ (*Statistical Parametric Mapping*) [Frackowiak *et al.*, 1997], développé par le *Wellcome Department of Cognitive Neuroscience, Functional Imaging Laboratory*,

13. <http://www.fil.ion.ucl.ac.uk/spm/>

Londres. Les autres alternatives les plus connues sont les suivantes, certaines n'étant que des adaptations du logiciel SPM :

- AFNI [<http://afni.nimh.nih.gov/afni/>],
- BrainVoyager [<http://www.brainvoyager.de/>],
- FMRISTAT [<http://www.bic.mni.mcgill.ca/~keith/>],
- FreeSurfer [<http://surfer.nmr.mgh.harvard.edu/>],
- FSL [<http://www.fmrib.ox.ac.uk/fsl/>],
- Lipsia [<http://194.95.183.247/Lipsia/>],
- Lyngby [<http://hendrix.imm.dtu.dk/software/lyngby/>],
- MEDx [<http://medx.sensor.com/>].

En réalité, l'acronyme SPM désigne à la fois le logiciel présenté ci-dessus mais également un type d'images (les cartes statistiques paramétrées) sur lesquelles est prise la décision de détection des activations fonctionnelles. Par extension, on parle également de méthodologie SPM pour désigner l'ensemble de la chaîne de traitement que met en œuvre le logiciel SPM : transformation spatiale des images (réalignement, normalisation), construction de cartes statistiques paramétrées et inférence statistique.

Cependant, si la méthodologie SPM reste la plus usitée à l'heure actuelle, de nombreuses techniques d'analyse des images fonctionnelles ont été proposées et peuvent être classifiées selon différents critères : univariées ou multivariées, inférentielles ou non, paramétriques ou non, exploratoires ou guidées par des hypothèses, globales ou focales, etc. La figure 3.13 présente une taxonomie des principales méthodes d'analyse des données fonctionnelles.

La distinction qui est faite entre méthodes univariées et multivariées réside dans le choix de l'entité élémentaire prise en compte. Dans le cas de méthodes univariées, les signaux temporels sont analysés voxel par voxel, indépendamment. A l'opposé, les méthodes multivariées prennent en compte directement l'ensemble des décours temporels provenant de tous les voxels. Les données fonctionnelles étant par nature intrinsèquement multivariées, plongées dans un espace spatio-temporel, les méthodes multivariées apparaissent comme étant les plus pertinentes. Cependant, les méthodes univariées restent les plus usitées, car elles s'avèrent rapides et efficaces.

Une seconde dichotomie importante entre les techniques d'analyse est le caractère exploratoire ou non. Les méthodes exploratoires travaillent directement à partir des données fonctionnelles sans connaissance a priori du paradigme expérimental ; Ce sont essentiellement des méthodes multivariées. Les alternatives sont les méthodes testant des hypothèses reliées à une modélisation explicite du paradigme. Il ne s'agit pas ici de décider quelle serait le type de méthode le plus judicieux car il n'y a pas de réponse à une telle question. En effet, chacune des méthodes présente des avantages et des inconvénients, les rendant plus ou moins attrayantes selon l'application précise. Ainsi, les méthodes multivariées peuvent permettre de détecter des signaux aberrants dans les données (dûs par exemple à une mauvaise correction des artefacts de mouvements) tandis que les méthodes univariées s'avèrent donner de bons résultats quand le paradigme expérimental est bien connu.

Dans la suite de ce chapitre, nous avons choisi de présenter de manière privilégiée la méthode univariée basée sur le modèle linéaire général. Nous évoquerons dans le paragraphe suivant les différentes stratégies pour prendre en compte l'information spatiale.

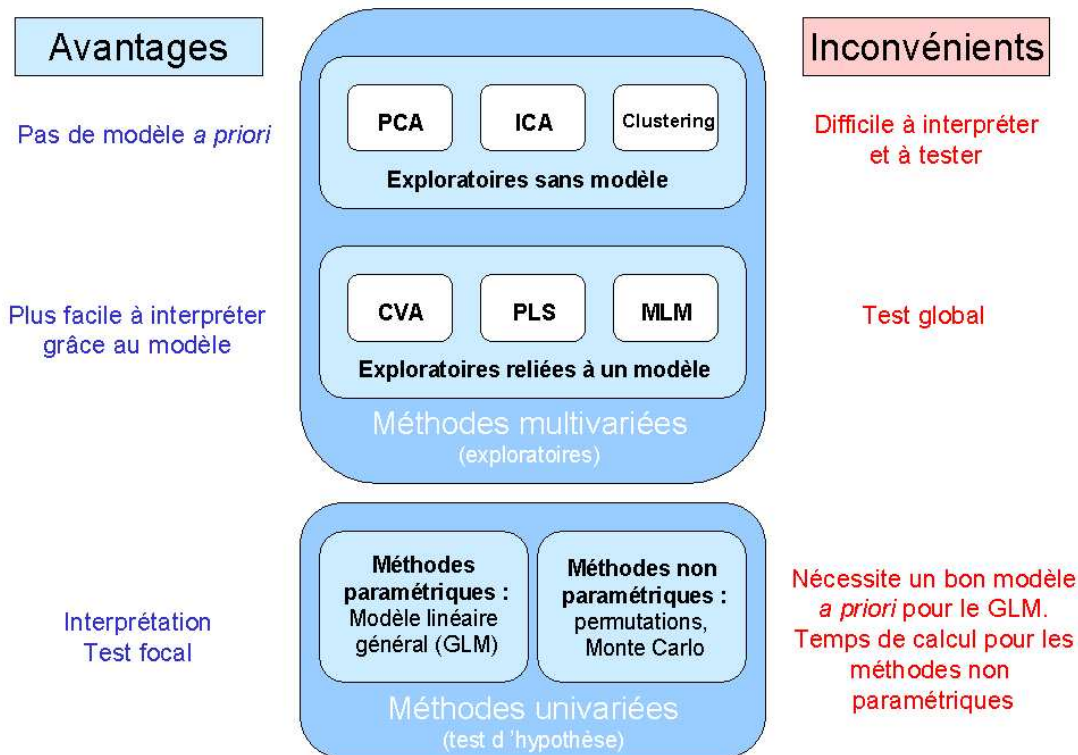


FIG. 3.13 – Les principales méthodes d'analyse de données d'imagerie par résonance magnétique fonctionnelle (les acronymes sont indiqués en langue anglaise). D'après [Kherif, 2003].

Enfin dans le chapitre suivant, nous reviendrons sur une sous-catégorie des approches multivariées exploratoires : les approches de type classification.

3.3.2 Le modèle linéaire général

Le modèle linéaire général est un cadre statistique standard de l'analyse de données, que l'on retrouve pour les analyses de variance et les modèles de multiples régressions. Il a été popularisé par le logiciel SPM pour les analyses univariées de données fonctionnelles [Friston *et al.*, 1995b; 1994; 1995c; Worsley and Friston, 1995]. Il établit une relation linéaire entre le signal observé et un ensemble de décours temporels modélisant ce que l'on s'attend à observer (le paradigme expérimental). Plus précisément, il modélise le décours temporel observé $y(t)$ comme étant issu de la combinaison linéaire de m variables explicatrices $\{x_i(t)\}$, appelés les régresseurs, et d'une erreur résiduelle $\epsilon(t)$:

$$y(t) = \beta_1 x_1(t) + \beta_2 x_2(t) + \dots + \beta_m x_m(t) + \epsilon(t) = \sum_{i=1}^m \beta_i x_i(t)$$

ce qui peut s'écrire pour tous les instants d'acquisition :

$$\begin{bmatrix} y(t_1) \\ \vdots \\ y(t_n) \end{bmatrix} = \begin{bmatrix} x_1(t_1) & \cdots & x_m(t_1) \\ \vdots & \vdots & \vdots \\ x_1(t_n) & \cdots & x_m(t_n) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \epsilon(t_1) \\ \vdots \\ \epsilon(t_n) \end{bmatrix}$$

et donc sous forme matricielle, pour tout voxel :

$$Y = X\beta + \epsilon \quad \text{avec} \quad E[\epsilon] = 0 \quad \text{et} \quad \text{Var}(\epsilon) = \sigma^2 I$$

La matrice X est appelée matrice de plan expérimental¹⁴ et contient sur ses colonnes les régresseurs constituant le modèle. Typiquement, les régresseurs sont obtenus en convoluant le ou les stimuli P_i par un modèle de réponse hémodynamique canonique $h(t)$:

$$x_i(t) = h(t) \otimes p_i(t)$$

La figure 3.14 présente un modèle de réponse hémodynamique $h(t)$, correspondant à la différence de deux fonctions Gamma, utilisé par le logiciel SPM. D'autres modélisations ont été proposées, paramétriques ou non, éventuellement définies selon des considérations physiologiques, mais le modèle de HRF du logiciel SPM reste le plus usité, puisqu'il permet de modéliser simplement l'activation et l'*undershoot* (légère dépression post-activatoire).

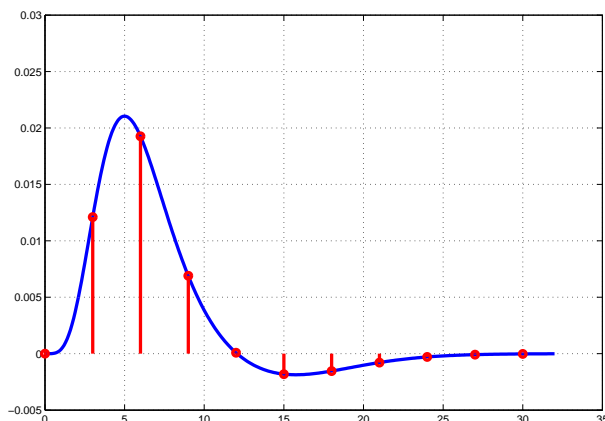


FIG. 3.14 – Modèle de réponse hémodynamique utilisé par SPM, obtenu par différence de deux fonctions Gamma (en bleu). Un exemple d'échantillonnage avec un TR de 3 secondes est également présenté (en rouge).

Cette modélisation – le signal observé est la sortie d'un filtre linéaire – correspond à plusieurs hypothèses sous-jacentes :

- linéarité : le principe de superposition s'applique et stipule que si une stimulation entraîne une réponse BOLD dans une région, alors la même stimulation répétée plusieurs fois dans le temps aura pour réponse la somme des réponses à chaque stimulation,

14. aussi appelée *design matrix*.

- invariance temporelle : la relation d'entrée-sortie est indépendante du temps. Ainsi, un délai dans la stimulation entraîne le même délai dans la réponse,
- causalité : la sortie ne dépend pas d'instants futurs.

Si l'hypothèse de la linéarité de la réponse BOLD semble vérifiée lorsque les stimuli sont suffisamment distants dans le temps [Boynton *et al.*, 1996; Glover, 1999] et en particulier pour les protocoles par blocs, on observe des effets non-linéaires lorsque les événements sont trop rapprochés (comme ce peut être le cas pour les paradigmes événementiels rapides). [Friston *et al.*, 1998] ont proposé l'utilisation de séries de Volterra pour modéliser ces effets non-linéaires, tout en restant dans le cadre du modèle linéaire général, grâce à l'ajout de régresseurs supplémentaires (nous y reviendrons dans le chapitre 6).

Nous présentons sur la figure 3.15 deux exemples de matrices de plan expérimental pour un paradigme par bloc et pour un paradigme événementiel rapide. En plus des régresseurs modélisant les stimulations, des régresseurs représentant les dérivées temporelles premières et secondes de ces premiers régresseurs ont été ajoutés pour modéliser la variabilité de la réponse hémodynamique (en particulier le délai). D'autres régresseurs sont parfois ajoutés, tels les estimations des paramètres de mouvement (cf le paragraphe 3.1.3) ou des signaux basses fréquences qui se comportent alors comme un filtre passe-haut permettant de modéliser les variations basses fréquences (effets respiratoires, effets cardiaques, dérives lentes lors de l'acquisition).

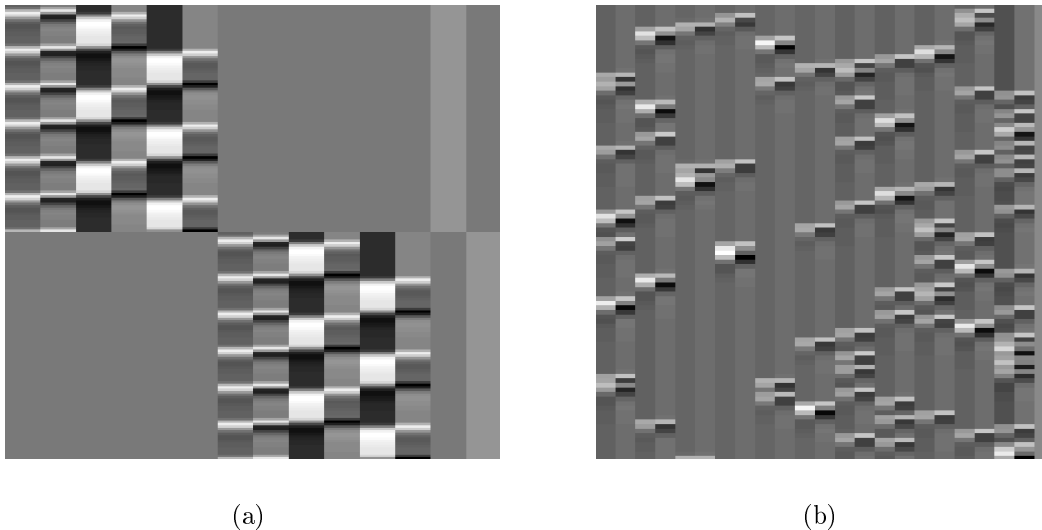


FIG. 3.15 – Exemples de matrices de plan expérimental pour un paradigme par bloc (a) et pour un paradigme événementiel rapide (b). Chacune des colonnes de ces matrices représente un régresseur modélisant une composante du signal.

Estimation au sens des moindres carrés

Le nombre de paramètres du modèle étant inférieur au nombre d'observations, l'estimation des coefficients de régression $\hat{\beta}$ est effectuée au sens des moindres carrés :

$$\begin{aligned}
 S &= \widehat{\epsilon}^T \widehat{\epsilon} \\
 &= (Y - X\widehat{\beta})^T (Y - X\widehat{\beta}) \\
 &= Y^T Y - 2\widehat{\beta}^T X^T Y + \widehat{\beta} X^T X \widehat{\beta}
 \end{aligned}$$

La minimisation du critère S fait apparaître les équations dites normales :

$$\frac{\partial S}{\partial \widehat{\beta}} = 0 \iff X^T Y = (X^T X) \widehat{\beta}$$

Si la matrice X est de rang plein, la matrice $X^T X$ est définie positive et donc inversible. L'estimation des paramètres de régression au sens des moindres carrés est donc :

$$\widehat{\beta} = (X^T X)^{-1} X^T Y$$

Dans le cas contraire, il existe une infinité de solutions et il faut faire appel aux inverses généralisés pour définir une solution unique, comme la pseudo-inverse de Moore-Penrose notée $(X^T X)^-$. Dans le cas de matrices pleines, inverse et pseudo-inverse coïncident.

On peut effectuer une interprétation géométrique de la solution aux moindres carrés, comme le montre la figure 3.16. En effet, on a :

$$\begin{aligned}
 \widehat{Y} &= X\widehat{\beta} = X(X^T X)^{-1} X^T Y \\
 \widehat{\epsilon} &= Y - \widehat{Y} = (I - X(X^T X)^{-1} X^T) Y = RY
 \end{aligned}$$

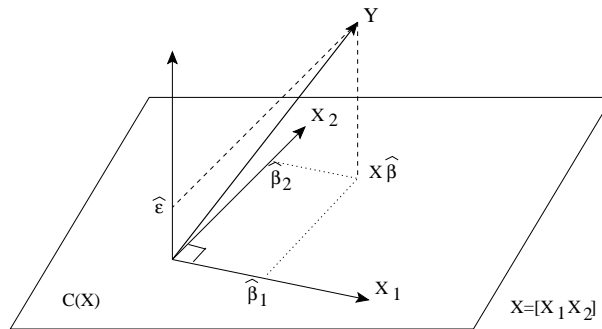


FIG. 3.16 – Explication géométrique de l'estimation au sens des moindres carrés comme étant une projection orthogonale de l'espace vectoriel engendré par les colonnes de X .

Ainsi, $P = X(X^T X)^{-1} X^T$ est le projecteur orthogonal sur l'espace engendré par les colonnes de la matrice du plan expérimental X tandis que $R = I - P$ est le projecteur orthogonal sur l'espace des résidus. L'estimation au sens des moindres carrés des paramètres de régression revient donc à projeter les données dans le sous-espace engendré par les régresseurs contenus dans la matrice X . Par ailleurs, les résidus sont orthogonaux au modèle.

On peut étudier quelques-unes des propriétés de l'estimateur au sens des moindres carrés : on note tout d'abord que c'est un estimateur linéaire, de la forme $\hat{\beta} = AY$. Par ailleurs, on a :

$$E[\hat{\beta}] = E[(X^T X)^{-1} X^T Y] = (X^T X)^{-1} X^T X \beta = \beta$$

$$Var(\hat{\beta}) = Var((X^T X)^{-1} X^T Y) = (X^T X)^{-1} X^T Var(Y) X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1}$$

L'estimateur au sens des moindres carrés est donc un estimateur linéaire et non biaisé. Le théorème de Gauss-Markov [Pelat, 1998] assure qu'il est l'estimateur possédant la plus petite variance parmi tous les estimateurs linéaires non biaisés.

Estimation du maximum de vraisemblance

Par la suite, il est intéressant de se placer dans un cadre statistique et d'étudier l'estimateur du maximum de vraisemblance, permettant de construire des tests d'hypothèses. Pour cela, faisons l'hypothèse que les résidus suivent une loi normale centrée :

$$\epsilon \sim \mathcal{N}(0, \sigma^2 I)$$

On en déduit, d'après le modèle linéaire, que les observations Y suivent également une loi normale dont les paramètres sont :

$$Y \sim \mathcal{N}(X\beta, \sigma^2 I)$$

La densité de probabilité s'écrit donc :

$$p(Y) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{(Y - X\beta)^T (Y - X\beta)}{2\sigma^2}\right)$$

La log-vraisemblance s'écrit alors :

$$\begin{aligned} L(Y) &= -\log(p(Y)) \\ &= \frac{n}{2} \log(2\pi\sigma^2) + \frac{(Y - X\beta)^T (Y - X\beta)}{2\sigma^2} \end{aligned}$$

L'estimateur $\hat{\beta}$ selon le maximum de vraisemblance est donné par :

$$\begin{aligned} \frac{\partial L}{\partial \beta} &= -\frac{X^T (Y - X\beta)^T}{\sigma^2} \\ \frac{\partial L}{\partial \beta} = 0 &\iff X^T Y = X^T X \hat{\beta} \end{aligned}$$

On retrouve les équations normales intervenant pour l'estimateur au sens des moindres carrés. L'estimateur selon le maximum de vraisemblance est donc le même et on peut en déduire sa distribution statistique :

$$\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X^T X)^{-1})$$

Prise en compte de la corrélation temporelle

Dans le cas de signaux fonctionnels, on observe une corrélation temporelle qui n'est pas prise en compte dans le modèle décrit ci-dessus [Zarahn *et al.*, 1997; Friston *et al.*, 2000; Woolrich *et al.*, 2001; Worsley *et al.*, 2002]. Les origines de cette autocorrélation ne sont que partiellement connues et des auteurs se réfèrent au principe d'acquisition, à des bruits physiologiques, etc. Jusqu'à présent, la matrice de covariance des résidus était supposée être égale à $\sigma^2 I$. Nous présentons maintenant le cas de l'estimation au sens des moindres carrés lorsque le bruit est supposé coloré :

$$\epsilon \sim \mathcal{N}(0, \Sigma) \quad \text{avec} \quad \Sigma = \sigma^2 V$$

Avec le même type de calculs que précédemment, on obtient [Pelat, 1998] :

$$\begin{aligned} \hat{\beta} &= (X^T V^{-1} X)^{-1} X^T V^{-1} y \\ \text{Var}(\hat{\beta}) &= \sigma^2 (X^T X)^{-1} X^T V X (X^T X)^{-1} \end{aligned}$$

L'estimateur obtenu est non biaisé et est même le meilleur estimateur linéaire non biaisé. Par contre, il nécessite de connaître la matrice de covariance du bruit, ce qui n'est généralement pas le cas. Sans hypothèses a priori, son estimation est délicate car on dispose alors de plus d'inconnues que de données. Ainsi les techniques d'estimation directe de la structure de covariance du bruit ne sont vraisemblablement pas suffisamment précises et sont susceptibles de biaiser l'estimation des paramètres de régression [Zarahn *et al.*, 1997]. Deux principales méthodes ont été proposées pour obtenir des estimations non biaisées des paramètres : le blanchiment et l'imposition d'une structure de covariance connue. Elles peuvent toutes deux s'écrire comme correspondant à un filtrage temporel des données, par une matrice de convolution K (de forme Toeplitz) :

$$KY = KX\beta + K\epsilon$$

En ce cas, les estimations des paramètres de régression sont :

$$\begin{aligned} \hat{\beta} &= (X^T K^T K X)^{-1} X^T K^T K Y \\ \text{Var}(\hat{\beta}) &= \sigma^2 (X^T K^T K X)^{-1} X^T K^T K V K^T K X (X^T K^T K X)^{-1} \end{aligned}$$

et en particulier, la distribution des résidus devient :

$$K\epsilon \sim \mathcal{N}(0, \sigma^2 K V K^T)$$

La première approche consiste à choisir K tel que $K = V^{-1/2}$: ceci correspond à un blanchiment des données (*whitening*). En effet, en ce cas, la structure de covariance du bruit devient :

$$\text{Var}(K\epsilon) = \sigma^2 K V K^T \approx \sigma^2 I$$

Bien sûr, il faut ici aussi disposer d'une estimation de la matrice de covariance V . Ceci est réalisé en pratique en imposant des contraintes sur sa forme. En particulier, un modèle auto-régressif d'ordre 1 est souvent utilisé. Nous avons représenté sur la figure 3.17 les quatre premiers coefficients d'un modèle autorégressif estimé en chaque voxel des images fonctionnelles. Pour cela, nous avons utilisé la technique décrite dans [Worsley *et al.*, 2002], qui prend en compte le fait que les signaux sur lesquels sont estimés les paramètres sont en fait les résidus d'un modèle linéaire (source d'un biais). On observe que le premier coefficient semble prépondérant (validant les approches se limitant à un modèle AR(1)) mais aussi qu'il existe de fortes disparités spatiales. Ainsi, la matière grise présente une autocorrélation temporelle bien plus élevée que la matière blanche. Cette constatation plaide en la faveur d'analyses prenant en compte des informations anatomiques comme l'appartenance à différents tissus. Nous y reviendrons à la fin de ce chapitre.

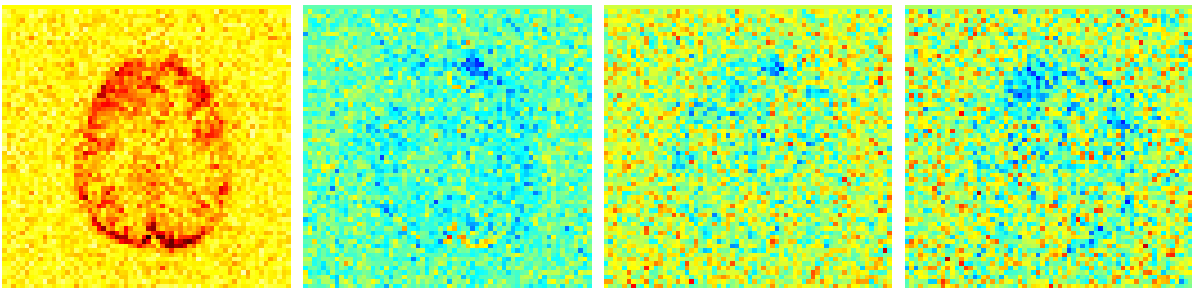


FIG. 3.17 – Coupes axiales représentant les quatre premiers coefficients d'un modèle auto-régressif (de gauche à droite) estimés sur les décours temporels associés à un paradigme par bloc avec un TR de 2 secondes.

Une autre solution consiste à imposer une structure de corrélation connue en effectuant un filtrage passe bas K suffisamment "fort" pour négliger la corrélation intrinsèque (*pre-colouring*) :

$$\text{Var}(K\epsilon) = \sigma^2 K V K^T \approx \sigma^2 K K^T$$

[Friston *et al.*, 2000] montrent que la technique de blanchiment est l'approche la plus efficace pour prendre en compte la corrélation temporelle mais qu'une estimation approximative des autocorrélations peut engendrer un biais qui restreindra l'efficacité des tests statistiques. Le filtrage passe bas temporel des données s'avère en ce cas plus robuste pour une perte d'efficacité modérée.

La figure 3.18 présente deux exemples de régression d'un signal pour un paradigme par bloc : les points bleus indiquent les échantillons temporels mesurés tandis que les courbes rouges permettent de visualiser le modèle correspondant aux paramètres estimés.

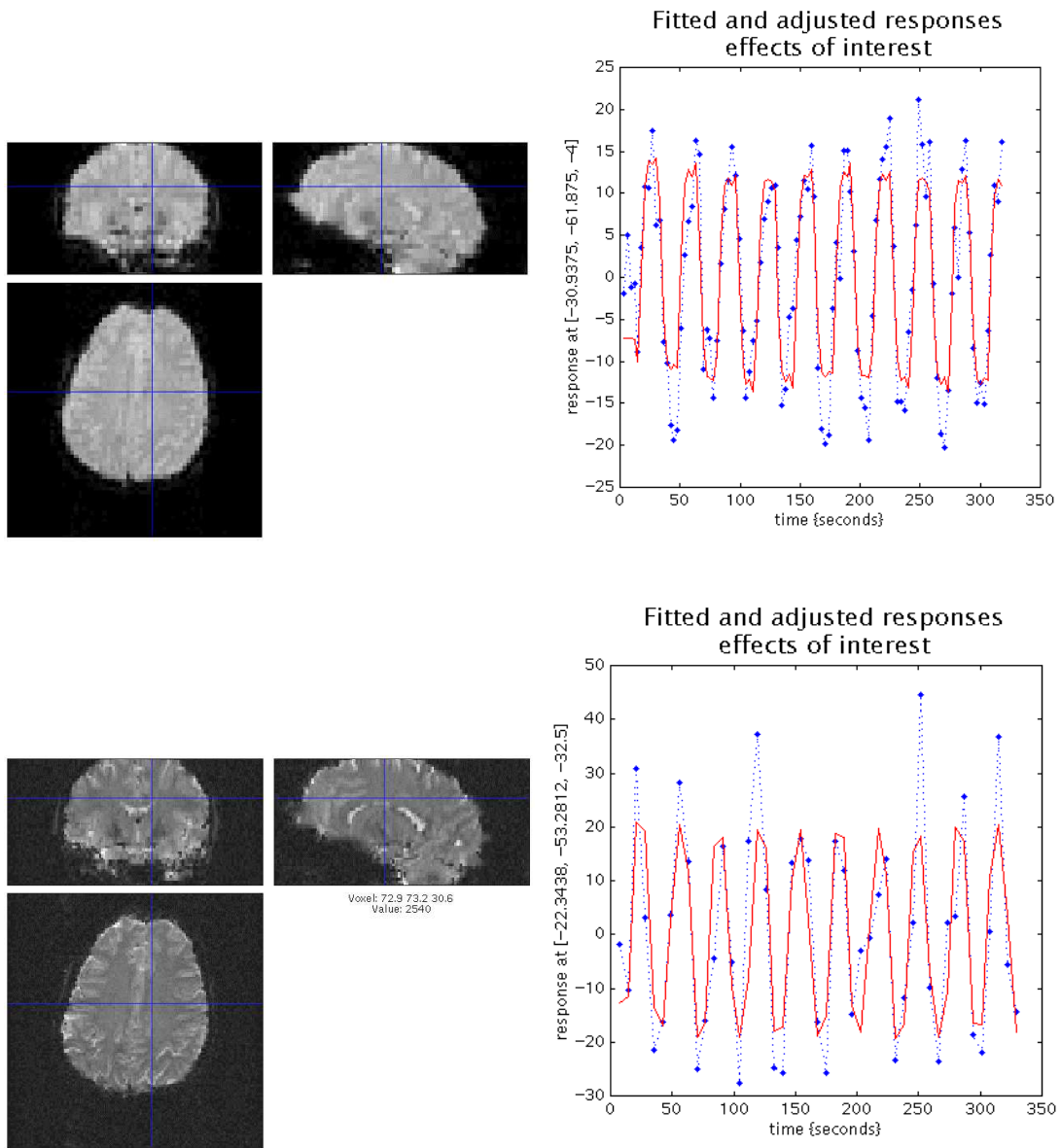


FIG. 3.18 – Exemples d'application du modèle linéaire général pour le voxel présentant le maximum d'activation dans un protocole par bloc. Dans ces deux exemples, seul les paramètres d'acquisition changent. Pour le premier, le TR est de 3 secondes pour 108 images de taille $64 \times 64 \times 29$ et de résolution $3.44 \times 3.44 \times 4 \text{ mm}^3$. Pour le second, le TR est de 7 secondes pour 48 images de taille $128 \times 128 \times 44$ et de résolution $1.72 \times 1.72 \times 2.5 \text{ mm}^3$. Ceci met en évidence le dilemme entre résolution temporelle et résolution spatiale: travailler avec un TR faible permet d'échantillonner plus finement la réponse BOLD mais ceci se fait au détriment de la résolution spatiale.

3.3.3 Inférence statistique

Il s'agit maintenant d'interroger le modèle afin de construire des cartes statistiques paramétriques (SPM) [Petersson *et al.*, 1999b]. On définit pour cela un contraste c correspondant aux coefficients de pondération d'une combinaison linéaire des régresseurs et permettant de définir une hypothèse nulle à tester :

$$H_0 : c^T \beta = 0$$

On cherche ainsi à tester si l'effet associé à une condition donnée (un régresseur de la matrice du plan expérimental) est plus grand (ou plus petit) que celui associé à une autre condition, ou, plus généralement, si l'effet associé à un groupe de conditions est plus grand (ou plus petit) que celui associé à un autre groupe. Une fois calculés les paramètres du modèle, on peut calculer le test statistique suivant :

$$t = \frac{c^T \hat{\beta}}{\text{Var}(c^T \hat{\beta})} = \frac{c^T \hat{\beta}}{\sqrt{\hat{\sigma}^2 c^T (X^T X)^{-1} X^T V X (X^T X)^{-1} c}} \sim t_\nu$$

où t_ν est une distribution de Student avec ν degrés de liberté. Ceci se déduit des résultats que nous avons présentés concernant les distributions statistiques de $\hat{\beta}$. L'estimation de σ^2 et le nombre de degré de liberté du test sont donnés par [Friston *et al.*, 1995b] :

$$\hat{\sigma}^2 = \frac{\hat{\epsilon}^T \hat{\epsilon}}{\text{trace}(RV)} \quad \text{et} \quad \nu = \frac{\text{trace}(RV)^2}{\text{trace}(RV RV)}$$

A noter que le test obtenu est signé et permet de tester si un effet est plus grand ou plus petit qu'un autre effet mais non pas qu'il est différent. Il faut pour cela utiliser un test F et définir un ensemble de contrastes $\Gamma = \{c_1, \dots, c_f\}$ de sorte que le test statistique est :

$$F = \frac{(\Gamma \hat{\beta})^T (\Gamma \hat{\beta})}{\Gamma^T \text{Var}(\hat{\beta}) \Gamma} \frac{d_2}{d_1} \sim F_{d_1, d_2}$$

où F_{d_1, d_2} est une distribution de Fisher avec d_1 et d_2 degrés de liberté. Cette statistique permet en particulier de tester si un sous-ensemble de régresseurs donnés est "utile" pour expliquer les variations du signal.

Notons que le dénominateur des tests t et F fait intervenir la matrice de covariance des résidus, ce qui explique l'importance de son estimation, compte tenu des corrélations temporelles.

Le calcul de ces tests statistiques, t ou F , en chaque voxel permet alors de construire une carte statistique, sur laquelle on peut décider d'accepter ou de rejeter l'hypothèse nulle. Une valeur élevée de t ou de F permet ainsi de rejeter l'hypothèse nulle. La connaissance de la distribution des tests sous l'hypothèse H_0 permet de déterminer analytiquement la probabilité que le test obtienne un score supérieur à t_0 . L'inférence statistique consiste alors à rejeter ou non l'hypothèse nulle en fonction de la probabilité calculée. Le choix d'un risque d'erreur α permet de définir un seuil. Tous les voxels situés au dessus de ce seuil sont déclaré comme répondant au protocole expérimental avec un risque d'erreur α .

Problèmes de multiples comparaisons

Les indications précédentes concernant le seuil des cartes statistiques ne sont valables que pour effectuer un test en un voxel. Ici, on souhaite tester l'ensemble des voxels constituant l'image et contrôler le nombre de faux positifs à l'échelle de l'image. Il faut donc prendre en compte le nombre de tests effectués : c'est résoudre le *problème de multiples comparaisons*. Dans le cadre des images fonctionnelles, les tests à effectuer ne sont pas indépendants car il existe une corrélation spatiale. Le nombre effectif de tests indépendants est donc moindre que le nombre de voxels.

Une solution élégante est apportée par la théorie des champs aléatoires gaussiens [Worsley *et al.*, 1992; 1998]. Pour cela, transformons la carte obtenue en convertissant voxel à voxel les probabilités t en Z . L'idée est que sous l'hypothèse nulle, la carte en z est un champ aléatoire gaussien de dimension D , dans un volume Vol , avec un certain degré de lissage. La caractéristique d'Euler est une mesure indiquant le nombre de composantes connexes d'un ensemble moins le nombre de trous dans ces composantes connexes. On peut estimer cette caractéristique sur le champ gaussien une fois seuillée par la valeur t_α :

$$E[\chi_t] = Vol|\Lambda|^{1/2}(2\pi)^{-\frac{D+1}{2}} H_D(t)e^{-\frac{t^2}{2}}$$

où Λ est la matrice de covariance du processus et H_D le polynôme de Hermite de degré D . Pour une valeur élevée de t_{alpha} , on peut faire l'approximation suivante :

$$p(Z_{max} > t_\alpha) \approx p(\chi_t \geq 1) \approx 1 - e^{-E[\chi_t]} \approx E[\chi_t]$$

Les deux équations précédentes permettent ainsi de définir un seuil sur un champ aléatoire gaussien, en contrôlant le risque de faux positifs. Worsley a également introduit la notion d'éléments de résolution (RESELS), définis par :

$$RESELS = \frac{Vol}{\prod_{i=1}^D FWHM_i}$$

où $FWHM_i$ représente la largeur spatiale à mi-hauteur de la réponse spatiale dans la direction i . La covariance du processus D -dimensionnel testé qui décrit la quantité de lissage de l'image peut s'exprimer ainsi :

$$Vol|\Lambda|^{1/2} = RESELS(4\log(2))^{\frac{D}{2}}$$

L'intensité maximale n'est pas le seul paramètre permettant de caractériser les objets dans la carte, d'autres méthodes ont été proposées, prenant en compte l'étendue spatiale des aires activées [Poline and Mazoyer, 1993].

3.3.4 Méthodes exploratoires

Nous présentons ici très succinctement les principales méthodes multivariées utilisées pour l'analyse exploratoire des données. Le lecteur pourra se référer à [Kherif, 2003] pour une description détaillée de ces techniques.

Les méthodes exploratoires caractérisent le signal sans utiliser de modèle spécifique des données. Bien qu'elles ne permettent pas d'effectuer d'inférence statistique (c'est-à-dire transformer des scores de tests concernant une hypothèse en valeurs probabilistes), ces méthodes peuvent s'avérer très utiles dans l'analyse des données fonctionnelles. Elles peuvent ainsi permettre de sélectionner un modèle qui sera utilisé pour une analyse inférentielle ultérieure ou vérifier que le modèle choisi est en adéquation avec les effets observables dans les données [Pettersson *et al.*, 1999a]. Les méthodes exploratoires peuvent également être utilisées pour les études de connectivité fonctionnelles, en raison de leur sensibilité aux corrélations temporelles dans le cerveau.

L'analyse en composantes principales (ACP) est la méthode multivariée la plus utilisée [Thirion and Faugeras, 2003a; Thirion, 2003]. Elle repose sur la décomposition en valeurs singulières (SVD) de la matrice des données (un décours temporel par colonne). On obtient alors une série d'images propres et de décours temporels propres. Des valeurs propres sont associées à chacune de ces composantes principales quantifiant leur importance relative. La première composante indique la principale variabilité observable sur les images tandis que les composantes suivantes modélisent la variabilité que l'on peut observer orthogonalement à toutes les composantes précédentes. Ainsi les composantes principales constituent une modélisation spatio-temporelle décrivant la structure de covariance d'un jeu d'observations multivariées, hiérarchisées par leur valeur singulière respective.

Appliquée sur les données Y , l'analyse en composantes principales permet d'observer, sans hypothèse préalable sur le paradigme, les modes de variation des données, avec par exemple un décours temporel propre corrélé avec le paradigme expérimental. Appliqué sur les résidus, l'ACP permet de détecter éventuellement des modes de variation ne respectant pas l'hypothèse de bruit gaussien, indiquant que des effets n'ont pas été correctement modélisés dans la matrice de plan expérimental ou que, par exemple, la correction du mouvement n'a pas été satisfaisante. D'autres travaux proposent d'utiliser l'analyse en composantes indépendantes (ACI) pour ne pas contraindre les décours temporels propres à être orthogonaux entre eux, et donc éventuellement pouvoir discriminer des effets reliés à une tâche d'autres effets corrélés comme des artefacts dûs aux battements cardiaques ou à des mouvements de la tête.

Plusieurs variantes ont été proposées pour introduire la matrice de plan expérimental X dans l'analyse, par exemple en effectuant la SVD de $X^T Y$. Les distinctions entre les variantes reposent sur des normalisations spécifiques des matrices X et Y . On ne peut plus alors parler de méthodes purement exploratoires mais elles restent utiles pour observer de manière non inférentielles les modes de variation des données projetée sur le sous-espace engendré par le modèle. Nous reviendrons sur l'application d'une de ces techniques dans le chapitre 6.

Enfin, des approches de type classification (*clustering*) ont été utilisées comme techniques d'analyses exploratoires des données, en regroupant les signaux temporels présentant des similarités selon une métrique à définir. Nous reviendrons dans le chapitre suivant sur ces approches et leur utilisation pour l'analyse de données fonctionnelles.

3.3.5 Les analyses de groupe

Jusqu'à présent, nous avons essentiellement envisagé l'analyse de données correspondant à un seul individu. Cependant, en pratique, la question de neurosciences est posée au niveau d'une population. Ainsi, les expériences sont classiquement menées sur un groupe composé de plusieurs sujets (en prenant en compte leurs spécificités : âge, sexe, dextrie). Selon la solution utilisée pour regrouper les informations issues des différents sujets, les résultats obtenus correspondent, soit à une étude de cas, soit peuvent être généralisés à l'échelle d'une population.

Les analyses à effets fixes

Une première analyse, permettant de moyenniser les résultats individuels sur un groupe, consiste à agréger les données provenant des sujets et à en faire l'analyse, comme s'il s'agissait d'un sujet unique. On parle d'analyse à effets fixes. En effet, cette approche revient à supposer que la variance est nulle entre les sujets : l'*effet sujet* est considéré comme étant un effet fixe.

Pour mettre en œuvre une analyse à effets fixes, le modèle linéaire général est utilisé en concaténant les données de tous les sujets (au nombre de n) en un seul vecteur :

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} X_1 & & 0 \\ & \ddots & \\ 0 & & X_n \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} E_1 \\ \vdots \\ E_n \end{bmatrix}$$

En particulier, on remarque que la matrice de plan expérimental est diagonale par blocs, ce qui fait que les estimations des paramètres pour un sujet est indépendante des données des autres sujets.

Ce modèle permet de tester individuellement chaque sujet par un test T, avec l'hypothèse que la variance du bruit est identique entre sujets. On peut également déterminer si un effet est significatif pour l'ensemble des individus mais les résultats ne peuvent être extrapolés en dehors de la population échantillonnée. Ce type d'analyses permet seulement de reporter des études de cas : il n'est en effet pas garanti que l'échantillonnage d'un autre groupe de sujets donnerait des résultats similaires.

Les analyses à effets aléatoires

Les analyses à effets aléatoires visent à prendre en compte la variance entre les sujets, modélisée par une distribution gaussienne, afin de pouvoir effectuer des inférences statistiques au niveau d'une population. Ce type d'analyse prend à la fois en compte la variance intra-sujet et inter-sujet.

Une méthode à deux niveaux permet d'effectuer ce type d'analyses, en résumant l'effet à tester pour chaque sujet. Tout d'abord, des cartes de contraste sont construites individuellement pour chaque sujet (c'est le premier niveau). Un test T peut alors être utilisé sur l'ensemble de ces images de contraste pour tester leur moyenne par rapport à leur variance (c'est le deuxième niveau). La variance par rapport à laquelle est effectuée le test

est alors entre sujets et les conclusions peuvent être généralisées à une certaine population parente qui dépend de l'échantillonnage des sujets.

L'inconvénient des analyses à effets aléatoires est qu'elles sont moins sensibles que les analyses à effets fixes. En effet, le nombre de degrés de liberté associé est très faible (égal au nombre de sujets moins un). C'est le prix à payer pour prendre en compte la variabilité inter-sujets qui est beaucoup plus grande que la variabilité intra-sujets.

3.4 Prise en compte d'une information spatiale

Ce paragraphe présente les principales méthodologies pour ajouter une information spatiale lors de l'analyse des données fonctionnelles. En effet, dans le cadre de l'analyse univariée utilisant le modèle linéaire général, les voxels sont traités indépendamment sans distinction du tissu dans lequel ils appartiennent. Le seul traitement faisant intervenir une information anatomique est l'opération de lissage spatial par un filtre gaussien, mais celle-ci est "aveugle" en ne prenant pas en compte les spécificités des différents tissus du cerveau. De nombreux travaux visent à améliorer cette opération et les trois principales approches sont : les analyses par région d'intérêt, la définition d'a priori spatiaux au sein de nouveaux modèles et les analyses restreintes à la surface corticale.

3.4.1 Analyse par région d'intérêt

La technique la plus utilisée pour prendre en compte une information spatiale est l'approche par région d'intérêt (ROI). Celle-ci consiste à définir une ou plusieurs régions anatomiques puis de s'intéresser aux variations hémodynamiques de ces seules zones. La définition des zones – étape cruciale pour la pertinence des résultats – peut être effectuée de différentes manières, que l'on peut mettre en parallèle avec les différentes techniques de labellisation des activations fonctionnelles évoquées dans le paragraphe 3.2.4.

La première méthode pour définir des régions d'intérêt repose sur l'utilisation d'un atlas. Il est en effet possible de normaliser les données dans un espace standard tel que l'espace de Talairach et sélectionner des régions à partir d'atlas placés eux-aussi dans le même espace : atlas AAL [Tzourio-Mazoyer *et al.*, 2002] avec des labels macroanatomiques ou atlas de Talairach (et l'aide du *Talairach Daemon*) avec des labels microanatomiques. [Maldjian *et al.*, 2003]¹⁵ proposent ainsi un outil logiciel permettant de définir automatiquement des régions d'intérêt correspondant aux aires de Brodmann, grâce à un atlas construit à partir du *Talairach Daemon*.

Une seconde méthode repose sur la délimitation manuelle de régions d'intérêt sur une IRM anatomique selon des considérations sur l'anatomie sulco-gyrale puis le report de ces régions sur les images fonctionnelles. Cette solution évite de devoir normaliser les images dans un même espace pour utiliser un atlas et permet ainsi de s'affranchir de la variabilité anatomique. Le repérage dans les images se fait en utilisant des amers géométriques (des invariants anatomiques) tels que les sillons corticaux. La boîte à outil Marsbar¹⁶ [Brett *et*

15. WFU_PickAtlas : <http://www.fmri.wfubmc.edu/download.htm>

16. MarsBar : <http://marsbar.sourceforge.net/>

al., 2002a] permet de définir ce type de région d'intérêt et par la suite d'analyser statistiquement les signaux moyens en utilisant SPM. [Nieto-Castanon *et al.*, 2003] proposent également de construire des régions d'intérêt à partir de l'anatomie intrinsèque à chaque sujet – et non pas sur un sujet moyen – en délimitant manuellement les sillons qui servent alors de frontières aux régions d'intérêt.

Une troisième méthode de définition de régions d'intérêt recherche à construire des aires fonctionnelles à l'aide de tâche *localizer* [Swallow *et al.*, 2003]. Pour cela, des expériences ont été mises en {oeuvre pour permettre d'identifier précisément pour un sujet la localisation fonctionnelle d'une aire donnée : *fusiform face area*, aires visuelles, complexe occipital latéral, etc. Les cartes paramétriques seuillées obtenues par l'analyse de ces expériences permettent de définir des masques sur lesquels les données du protocole suivant sont analysées. Le logiciel fROI¹⁷ est spécifiquement dédié à la construction de telles régions d'intérêt fonctionnelles.

Chacune de ces techniques de construction de région d'intérêt présente cependant des inconvénients. Ainsi, l'utilisation d'un atlas nécessite de normaliser les images et donc de d'effectuer une hypothèse forte de corrélation anatomo-fonctionnelle. Les régions obtenues ne reposent ensuite plus sur les anatomies spécifiques de chaque sujet. Leur construction manuelle s'avère alors plus robuste, en n'utilisant pas de sujet moyen mais on perd le caractère automatique de la méthode, qui devient alors très coûteuse en temps et relativement imprécise car le dessin des régions est sujet à caution. Ensuite, la définition de régions selon des critères fonctionnels nécessite l'acquisition et l'analyse de données supplémentaires permettant de localiser une aire précise : en dehors du temps supplémentaire que cela demande, il n'est pas évident que toutes les aires fonctionnelles cérébrales puissent être aisément localisées par un protocole simple préalable. Enfin, de manière générale, il faut garder à l'esprit que les analyses par régions d'intérêt ne testent des effets que sur les régions sélectionnées et n'autorisent donc pas l'exploration des données sans a priori sur la localisation des activations. Par ailleurs, les régions d'intérêt restent souvent grossières, de la taille des gyri, ce qui ne permet pas des détections focales précises.

En revanche, les analyses par région d'intérêt ont deux avantages précieux par rapport à la méthodologie classique décrite dans les paragraphes précédents (approche univariée dans un espace standard). Tout d'abord, on obtient une puissance statistique accrue car le nombre de tests effectués est faible (égal au nombre de régions d'intérêt définies) permettant d'éviter le caractère conservateur de la correction du problème de comparaisons multiples. Par ailleurs, si la région d'intérêt a été correctement dessinée, le signal moyen correspondant est alors équivalent à celui obtenu en utilisant le meilleur filtrage spatial pour lisser les données. Les analyses par région d'intérêt permettent ainsi d'obtenir une meilleure sensibilité en détection. Le deuxième avantage de ce type d'analyse concerne les études de groupe. En effet, cela permet de définir des homologies entre aires fonctionnelles sans reposer sur la qualité de la mise en correspondance des images anatomiques. Les régions d'intérêt sont définies sur chaque sujet selon des considérations anatomiques et fonctionnelles et ensuite analysées simultanément, même si les régions d'intérêt ne se

17. fROI: <http://froi.sourceforge.net/>

superposent pas exactement une fois normalisées dans l'espace de Talairach , en raison de la variabilité anatomo-fonctionnelle.

3.4.2 Modèles régularisés spatialement

Plusieurs travaux ont été présentés pour remplacer l'approche standard lissage spatial suivie d'une analyse univariée par un modèle intrinsèquement spatio-temporel. Le principal objectif de ces travaux est de se libérer du lissage spatial gaussien stationnaire, favorisant la détection d'activations de la taille et de la forme du noyau de lissage utilisé et mélangeant des signaux fonctionnels éventuellement distincts.

[Descombes *et al.*, 1998a] ont ainsi proposé de remplacer l'étape de lissage spatial gaussien par un algorithme de restauration de données utilisant une modélisation de type champ de Markov, permettant de préserver les transitions des signaux. Ils ont par la suite proposé un modèle spatio-temporel complet permettant l'analyse des données [Descombes *et al.*, 1998b] mais les inférences statistiques dans ce cadre devient ardue, nécessitant des simulations de Monte Carlo pour déterminer la distribution a posteriori des *tâches* d'activation spatio-temporelles. La dimension de l'espace-état spatio-temporel est alors source de temps de calcul important. Par ailleurs, les méthodes markoviennes sont limitées par le coté arbitraire du modèle local d'interaction.

D'autres auteurs ont proposé des modèles faisant intervenir le voisinage des voxels lors de l'estimation de leurs paramètres, dans un cadre bayésien [Everitt and Bullmore, 1999; Hartvig and Jensen, 2000; Genovese, 1999; Gossel *et al.*, 2001] mais dans la plupart des cas, ces modèles spatiaux restent inestimables numériquement en raison de la dimensionalité des données, et les inférences statistiques reposent sur des méthodes de simulation très coûteuses en temps, les rendant inutilisables en routine. [Penny and Friston, 2003] a toutefois présenté un modèle possédant un nombre de paramètre restreint mais dont les estimations restent délicates. Nous reviendrons sur ce modèle dans le chapitre 7, en en présentant une autre formulation possédant une résolution simplifiée.

Notons également les travaux de [Poline and Mazoyer, 1994a; 1994b] utilisant une analyse multi-échelle à divers degrés de lissage pour détecter les activations. Cette idée a par la suite été utilisée par [Coulon *et al.*, 2000] pour mettre en place une description structurale multi-échelle des cartes d'activation, qui peuvent alors être mises en correspondance dans le cadre d'analyses multi-sujets sans reposer fortement sur l'étape de normalisation spatiale. L'inconvénient de ces approches est qu'avec de forts lissages, il devient difficile détecter les différents amas de voxels (*blobs*).

3.4.3 Analyse sur la surface corticale

Une autre voie proposée dans la littérature pour ajouter une information spatiale lors de l'analyse des données fonctionnelles consiste à restreindre ladite analyse sur la surface du cortex. En effet, il s'agit de prendre en compte non seulement la distribution spatiale des différents tissus cérébraux (matière grise et blanche) mais également la particularité topologique du cortex d'être homotopique à une sphère et donc de pouvoir être modélisé simplement de manière surfacique. En particulier, travailler sur la surface corticale permet mettre en œuvre un lissage spatial anisotrope, ne moyennant que les signaux appartenant

à la matière grise et évitant ainsi de mélanger les signaux provenant de tissus différents, comme c'est le cas dans le cadre d'un lissage spatial tridimensionnel. Ce principe est illustré sur la figure 3.19.

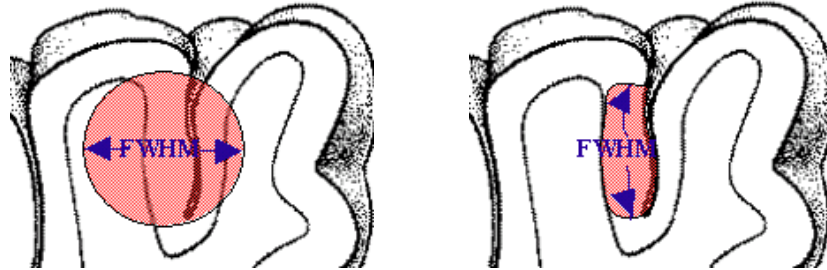


FIG. 3.19 – *Lissage sur la surface corticale. Ces deux schémas illustrent le fait qu'un lissage effectué sur la surface respecte mieux l'anatomie sulco-gyrale en comparaison avec un filtre gaussien isotrope. Dans le cas surfacique, seul les voxels appartenant au même tissu sont moyennés. D'après [Andrade et al., 2001].*

[Andrade *et al.*, 2001] ont proposé une méthodologie complète, appelée CSM (*Cortical Surface Mapping*), permettant l'analyse de données fonctionnelles sur la surface corticale. Les différentes étapes sont :

- Mise en correspondance des images fonctionnelles et anatomiques, en utilisant les techniques de recalage décrites dans le paragraphe 3.1.
- Segmentation de l'image anatomique en matière grise et matière blanche et construction d'une surface triangulée située au milieu de la matière grise. Le maillage obtenu, homotopique à une sphère, contient environ 20 000 nœuds par hémisphère.
- Attribution de valeurs fonctionnelles en chaque nœuds du maillage. Pour cela, les images fonctionnelles sont interpolées de manière à obtenir un décours temporel par nœuds. Les analyses à venir ne travaillent désormais que sur ces données interpolées et restreintes au cortex.
- Lissage spatial des données effectué selon le voisinage inhérent à la surface et implémenté comme étant la résolution de l'équation de la chaleur sur le maillage du cortex.
- Estimation des paramètres d'un modèle linéaire général : identique à l'analyse univariée tridimensionnelle.
- Correction pour les problèmes de multiples comparaisons grâce à une estimation de la quantité effective de lissage sur les données.
- Tests d'hypothèse à partir des paramètres estimés du modèle linéaire.

La figure 3.20 présente un résultat de détection d'activations utilisant CSM sur une version dépliée d'un hémisphère gauche. La comparaison avec les résultats fournis par SPM interpolés a posteriori sur la surface met en évidence une légère augmentation de la sensibilité, que l'on peut attribuer au lissage "intelligent" des données.

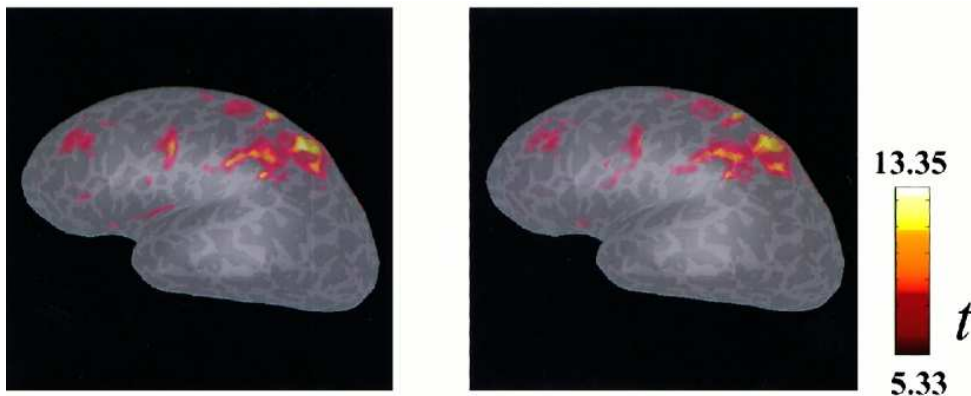


FIG. 3.20 – Cartes d'activations obtenues par CSM (à gauche) et SPM (à droite) pour une tâche motrice (probabilité corrigée à $5 \cdot 10^{-2}$). On observe une région insulaire détectée par CSM et pas par une approche tridimensionnelle classique. Par ailleurs, les valeurs en t obtenues sont légèrement supérieures dans le cas de CSM.

L'utilisation d'information géométrique, apparaissant sous forme de lissage surfacique ici, permet donc d'augmenter la sensibilité des analyses statistiques. Cependant, une des étapes critique de cette technique est la projection des données fonctionnelles tridimensionnelles sur la surface corticale. En effet, cette étape nécessite une très bonne mise en correspondance des images anatomiques et fonctionnelles, puisque le maillage cortical est construit à partir de l'image anatomique puis plongé dans l'espace des données fonctionnelles par interpolation. La figure 3.21 présente le résultat de l'interpolation d'un volume fonctionnel sur un maillage surfacique : on observe des pertes de signal sur les fonds des sillons, symptomatique d'une projection défailante. Nous avons en effet mis en évidence la présence de distorsions géométriques dans les images fonctionnelles obtenues par une acquisition rapide EPI. Ces pertes de signal entraînent la non-détection d'activations dans les régions dans lesquelles elles apparaissent.

D'autres travaux [Kiebel *et al.*, 2000; Kiebel and Friston, 2002] se proposent d'ajouter une information surfacique par le biais de fonctions de base spatiales contraintes anatomiquement. La méthode proposée est conceptuellement proche des techniques de déconvolution en cherchant à estimer les paramètres d'un ensemble de fonctions réparties spatialement sur la surface corticale et modélisant les données fonctionnelles. Le principal inconvénient de ces travaux est qu'ils utilisent une version aplatie de la surface corticale, source de distorsions géométriques supplémentaires comparé à [Andrade *et al.*, 2001] qui travaille directement sur le maillage original. Par ailleurs, les fonctions de base sont réparties de manière régulière sur cette surface corticale aplatie sans que la position de leur centre puisse être remise en question en fonction des données à modéliser.

Enfin, les travaux de [Fischl *et al.*, 1999b] tendent à valider l'approche surfacique en comparant les résultats obtenus en détection d'activations pour une analyse de groupe selon deux types de normalisation spatiale des sujets (cf la figure 3.22). Dans un cas, la mise en correspondance des anatomies est effectuée par l'approche proportionnelle

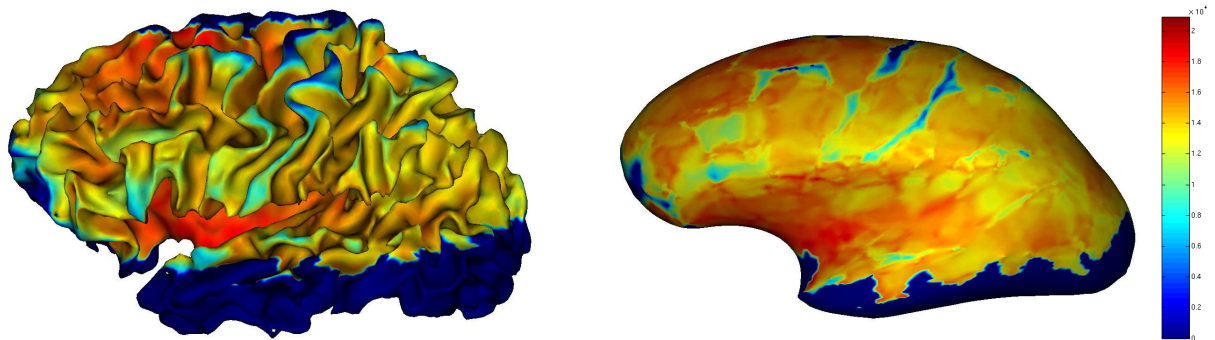


FIG. 3.21 – *Interpolation d'une image fonctionnelle sur la surface corticale. Ces deux figures représentant le même cerveau avant et après gonflement. On observe ici que, malgré un recalage jugé satisfaisant entre les images anatomiques et fonctionnelles, il y a des pertes de signal importantes (en bleu), principalement au niveau du sillon central.*

de Talairach tandis que dans l'autre cas, ce sont les surfaces corticales qui sont recalées grâce à l'algorithme présenté dans [Fischl *et al.*, 1999a]. On observe une amélioration spectaculaire de la sensibilité par l'approche surfacique, en raison d'une meilleure mise en correspondance des sujets lors de l'analyse de groupe. Ces résultats montrent principalement l'importance de la qualité de la normalisation pour les analyses multi-sujets et son lien direct avec la sensibilité des analyses statistiques.

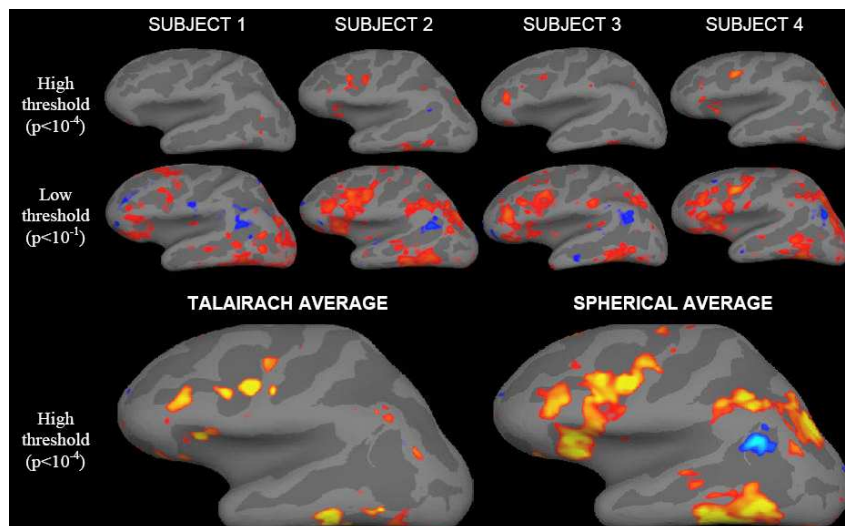


FIG. 3.22 – *Comparaison de la sensibilité en détection d'activations pour une analyse multi-sujets selon deux normalisations: l'une reposant sur la normalisation proportionnelle de Talairach et l'autre utilisant un recalage des surfaces corticales. La seconde technique montre sur cet exemple une sensibilité accrue comparée à l'approche volumique de Talairach. D'après [Fischl *et al.*, 1999b].*

3.5 Conclusion

Au cours de ce chapitre, nous avons décrit l'ensemble des étapes de la chaîne de traitement des images fonctionnelles, de leur acquisition à la détection des activations liées au protocole expérimental. Nous avons ainsi détaillé les prétraitements appliqués sur les images pour corriger les différents artefacts : correction des distorsions géométriques, du délai temporel intercoupe et du mouvement. L'analyse statistique la plus employée est une approche univariée reposant sur le modèle linéaire général.

Un certain nombre d'approches ont été proposées pour prendre en compte une information spatiale lors des analyses. En effet, les activations potentielles se trouvent dans la matière grise (cortex et noyaux gris centraux), là où se trouvent les corps cellulaires des neurones. En particulier, plusieurs travaux ont montré de meilleurs résultats de sensibilité lorsque l'analyse est effectuée sur la surface corticale, avec un lissage spatial des données ne mélangeant pas les décours temporels associés à des tissus différents. Nous avons cependant mis en évidence les difficultés d'interpolation des images fonctionnelles sur un maillage obtenu à partir des images anatomiques.

Par ailleurs, dans le cadre multisujet, se pose le délicat problème de la mise en correspondance des images des différents sujets. En effet, l'essentiel des algorithmes de recalage reposent sur une hypothèse de corrélation anatomo-fonctionnelle, qui n'est que partiellement justifiée. Les conséquences d'un recalage imparfait sont importantes car la méthodologie univariée fait l'hypothèse forte d'une correspondance voxel à voxel des images de chaque sujet : on observe alors des pertes de sensibilité importantes dans les analyses multisujets. Pour amoindrir ce problème, les images sont fortement lissées spatialement mais ceci se fait au détriment de la localisation précise des activations. Une alternative consiste à effectuer les analyses par région d'intérêt, en détournant manuellement des aires homologues entre sujets, mais ce type d'approche reste limité à quelques aires et s'avère très coûteux en temps.

Il ressort de ce chapitre que l'introduction d'information spatiale dans l'analyse des données peut permettre d'améliorer la sensibilité en détection. Les approches proposées reposant sur l'utilisation d'un maillage de la surface corticales vont en ce sens mais elles restent difficiles à mettre en œuvre. Par ailleurs, la mise en correspondance fonctionnelle des images dans les analyses multisujets apparaît comme étant un problème difficile : une qualité intéressante des méthodes d'analyses est de s'avérer robuste face aux imperfections des recalages anatomiques.

Chapitre 4

Classification de données spatiales

Sommaire

4.1 Cluster Analysis	78
4.1.1 Unsupervised learning	78
4.1.2 Parametric Models	79
4.1.3 Self-Organizing Maps	80
4.1.4 Hierarchical Clustering	80
4.2 Mixture models	81
4.2.1 Maximum Likelihood and EM algorithm	81
4.2.2 Gaussian Mixture Models	84
4.2.3 Model Order Selection	96
4.2.4 Classification	96
4.2.5 Incorporating Spatial Priors	98
4.3 Clustering methods in fMRI data analysis	100
4.3.1 Definition of the feature space	100
4.3.2 Introduction of spatial information	101
4.4 Conclusion	101

Ce chapitre est rédigé en anglais.

This chapter is divided into two complementary parts. Firstly, we present an overview of the main techniques dedicated to cluster analysis. The aim is not to make an exhaustive survey of the field but rather to indicate the most common methods and then specifically focus on parametric methods (mixture model) and some adaptations that have been proposed to handle spatial data. We carefully detail the derivation steps of the Expectation-Maximization algorithm. As this is not strictly necessary to the understanding of mixture models, this part may be skipped in a first reading.

Secondly, we describe several works linked with functional magnetic resonance imaging that rely on the unsupervised learning framework. Again, this section is not an exhaustive survey but rather a selection of reference papers to better place our work in the following of this dissertation.

4.1 Cluster Analysis

4.1.1 Unsupervised learning

When some data x are given labels y that define the membership to a certain number of categories, we call classifier every application C giving the correspondence between y and x :

$$y = C(x)$$

The training of the classifier is performed on a certain number of samples (x,y) and its performances can be measured in terms of generalisation, that is its capacity to correctly label data x that were not part of the learning dataset. We are here in the framework of supervised learning : the classifier is built using labeled data

Unsupervised learning refers to every learning method that does not use the label information to build a classifier. Motivations are various, here are a few of them :

- missing or expensive labels: this is the case, for example, when the labelling is performed, sample by sample, by an expert. This appears to be cumbersome and expensive in manpower.
- exploratory analysis: when the amount of data is important, it appears that it is often useful to look beforehand for structures, or hidden relations, within them. The discovery of distinct classes (clusters containing similar samples) may suggest to take this structure into account if it had not been detected during the acquisition.
- dimension reduction: if acquired data span a high dimension space, it is difficult to visualize them. Unsupervised learning algorithms can help determining a subspace (of dimension 2 or 3, adapted to human) allowing for an optimal representation of the data. It may also happen that some algorithms may appear intractable from a computer science point of view due to the dimensions of the problem. In that case, unsupervised clustering can contribute to lower the dimensionality of the data and can be considered as a preprocessing to be applied to the data before their real analysis.

Applications of unsupervised learning are numerous. To cite only a few examples: compression of data into a small set of representatives; preprocessing to summarize data in order to apply computer intensive algorithms; structure discovery, by separating data in consistent subgroups, that can be potentially instructive; determination of interesting projecting axes like highest variance directions (visualisation of multidimensional data), etc.

Cluster analysis is part of unsupervised learning and has a variety of goals. All relate to grouping a collection of objects into subsets (*clusters*) such that those within each cluster are more closely to one another than objects assigned to different clusters. Central to all of the goals of cluster analysis is the notion of degree of similarity (or dissimilarity) between the individual objects being clustered. A clustering method attempts to group the objects based on the definition of similarity supplied to it.

The main clustering algorithms are: parametric models (associated with probability density), self-organizing maps and hierarchical methods.

4.1.2 Parametric Models

Clustering algorithms for classification of unlabeled data can be formulated in the context of probability density estimation. Following Bishop [Bishop, 1995], we consider three approaches to clustering: parametric, non parametric and semi-parametric methods. In a semi-parametric approach, the probability density is modeled as a mixture of parametric functions where the number of elements is also an unknown parameter of the model, thus allowing a more flexible form for the densities. There are then several procedures to estimate “the best” model order according to the training data.

The underlying idea is to consider data x as a mixture of observations generated by a certain number of different probability densities. These densities are often chosen as being multivariate gaussian, because this kind of distribution is well-known, easy to deal with, and has been proven able to model a wide number of phenomenon.

$$p(X|\Theta) = \sum_{k=1}^K \pi_k p(X|k; \theta_k)$$

Thus the parametric probability density $p(X|\Theta)$ used as a generative model of the data to be classified is a linear combination of K simple densities $\{p_k\}$ (usually gaussian) with parameters $\{\theta_k\}$ and the mixing proportions $\{\pi_k\}$.

The estimation of the parameters is classically performed according to the maximum likelihood principle. This brings several algorithms that are standard tools for the unsupervised classification of data:

- Expectation-Maximization (EM) algorithm,
- Fuzzy c-means,
- K-Means algorithm.

We will describe in details this model, its variants and the algorithms that allow to estimate its parameters in the second part of this chapter.

4.1.3 Self-Organizing Maps

Self-organizing maps, also called Kohonen maps, are popular tools for visualization and interpretation of large high-dimensional datasets. They have similarities with both mixture models and neural networks approaches, but they add a notion of topology : The objective of a Kohonen network is to map input vectors (patterns) of arbitrary dimension N onto a discrete map with 1, 2 or 3 dimensions. Patterns close to one another in the input space should be close to one another in the map: they should be topologically ordered.

The online algorithm [Kohonen, 2001] is quite simple : we consider a low dimensional grid (for example a two-dimensional rectangular grid) of K prototypes $\{m_k\}$. The aim is to adjust this grid (associated with the prototypes) to the dataset x , which is in a higher dimensional space. To do so, the observations x_i are processed one at a time. We find the closest prototype m_j to x_i and we update all of the prototypes in the neighborhood of m_j , using the formula :

$$m_j \leftarrow m_j + \alpha(x_i - m_k)$$

The effect of the update is to move the prototypes close to the data, but also to maintain a smooth low dimensional spatial relationship between the prototypes. At the end, the algorithm provides K prototypes representative of the data x distributed over a grid, that can be displayed for visualization purpose.

[Heskes, 2001] shows that self-organizing maps can be interpreted as kind of mixture models with additional regularization. However, their underlying model is not well known and their implementation requires the use of several parameters that are difficult to set.

4.1.4 Hierarchical Clustering

The results of K-Means algorithm or related methods depend on the choice for the number of clusters to be searched and a starting configuration assignment. In contrast, hierarchical clustering methods do not require such specifications. Instead, they require the user to specify a measure of dissimilarity between disjoint groups of observations. They provide hierarchical representations in which the clusters at each level of the hierarchy are created by merging clusters at the next lower level. At the lowest level, each cluster contains a single observation. At the highest level there is only one cluster containing all of the data.

Strategies for hierarchical clustering divide into two basic paradigms : agglomerative (bottom-up) and divisive (top-down). Agglomerative strategies are the most used ; That's the one we will describe now.

Agglomerative strategies start at the bottom and at each level recursively merge a selected pair of clusters into a single cluster. This produces a grouping at the next higher level with one less cluster. The pair chosen for merging consist of the two groups with the smallest intergroup dissimilarity. A summary of the different steps are the following:

1. initialization : each observation represent a singleton cluster,
2. compute dissimilarity measure between all the clusters,
3. merge the two closest clusters into a single one,

4. return to step 2 until there is only one cluster left.

A measure of dissimilarity between two clusters must be defined. Several definitions have been proposed :

single linkage	$d_s = \min_{i,j} \ x_i - x_j\ $
complete linkage	$d_{co} = \max_{i,j} \ x_i - x_j\ $
average linkage	$d_a = \frac{\sum_{i,j} \ x_i - x_j\ }{N_k N_l}$
centroid linkage	$d_{ce} = \ c_k - c_l\ $

The choice of this measure has an important influence on the output of the algorithm. If the data dissimilarities exhibit a strong clustering tendency, with each of the clusters being compact and well separated from others, then all three methods produce similar results. However in the other cases, we can observe discrepancies : single linkage will, for example, have a tendency to combine at relatively low thresholds, observations linked by a series of close intermediate observations. Indeed, the main difficulty with hierarchical clustering algorithms is that they do not rely on a generative model and they may appear tricky to use.

4.2 Mixture models

In this section, we focus our description of unsupervised learning on mixture models, and more specifically on gaussian mixture models (GMM).

As stated by Fisher (1922), inferences can be split into three aspects: model specification (mixture models), estimation of model parameters (Maximum Likelihood principle and EM algorithm) and estimation of precision (model order selection). We will discuss these three aspects below.

We begin here by describing the EM algorithm in a generic way because it is a very general algorithm that has achieved considerable notoriety in the last decade especially in the computer vision, neural networks and medical imaging research communities. We will then derive this algorithm to the estimation of model parameters of a GMM and show its relations with other clustering algorithms such that fuzzy C-means and K-means. At last, we will present some variants to regularize the estimations and incorporating spatial priors when dealing with spatial data.

4.2.1 Maximum Likelihood and EM algorithm

The Expectation-Maximization (EM) algorithm can be derived of several ways. In this paragraph we present the EM as a *lower bound maximization*. In the Appendix B, we show that the EM algorithm can also be seen as an *alternate optimization* of a fuzzy criterion penalized by an entropy term.

Let X denote the observable variables and let Z denote some hidden unknown variables associated to data X . We note Θ the parameters of the parametric model for the probability density function from which X are realizations.

The parameters can be estimated using a criterion based on the maximization of the likelihood of the parameters for the given data set. The log-likelihood is as follows:

$$L(\Theta) \triangleq \log p(X|\Theta) \quad (4.1)$$

The maximum-likelihood estimation of Θ is then:

$$\hat{\Theta} \triangleq \arg \max_{\Theta} L(\Theta) = \arg \max_{\Theta} \log p(X|\Theta) \quad (4.2)$$

Since Z is not observed, the probability of the data X is indeed a marginal probability and the log-likelihood is said to be the *incomplete log-likelihood*:

$$L(\Theta) = \log \sum_Z p(X, Z|\Theta) \quad (4.3)$$

As you can see, the logarithm is separated from the joint probabilities $p(X, Z|\Theta)$ by a summation sign, preventing the problem decoupling. To get rid of that, let's introduce the variable Z in the log-likelihood formula using the definition of joint probabilities:

$$p(X|\Theta) = \frac{p(X, Z|\Theta)}{p(Z|X, \Theta)} \quad (4.4)$$

We also introduce $q(Z)$, the probability density function of the random variable Z , using a small trick:

$$\begin{aligned} L(\Theta) &= \log \frac{p(X, Z|\Theta) q(Z)}{p(Z|X, \Theta) q(Z)} \\ &= \log \frac{p(X, Z|\Theta)}{q(Z)} + \log \frac{q(Z)}{p(Z|X, \Theta)} \end{aligned}$$

Z being a random variable, $L(\Theta)$ is also a random quantity: to remove this randomness, we take the expectation¹⁸ of the previous equation over the variable Z :

$$L(\Theta) = \underbrace{\sum_Z q(Z) \log \frac{p(X, Z|\Theta)}{q(Z)}}_{\mathcal{L}(q, \Theta)} + \underbrace{\sum_Z q(Z) \log \frac{q(Z)}{p(Z|X, \Theta)}}_{KL(q(Z)||p(Z|X, \Theta))} \quad (4.5)$$

$KL(.||.)$ denotes the Kullback-Leibler divergence and $\mathcal{L}(q, \Theta)$ is an auxiliary function from which we will find a concrete interpretation.

A simple use of the Jensen's inequality (see Appendix A) shows that this divergence is non-negative:

$$KL(q(Z)||p(Z|X, \Theta)) \geq 0 \quad (4.6)$$

Furthermore the divergence is uniquely minimized and equals 0 when the two distributions are equal, which gives here:

$$q(Z) = p(Z|X, \Theta) \quad (4.7)$$

18. We use the discrete formulation for the expectation but equivalently we could have used the continuous expression replacing the discrete sum by an integral. Same remark with the marginalization in equation 4.3.

So we can conclude that the auxiliary function is a **lower bound** of the incomplete log-likelihood, whatever the distribution of Z , $q(Z)$, and there is equality when $q(Z) = p(Z|X, \Theta)$.

Another interesting thing to point out is the fact that, in the definition of the auxiliary function $\mathcal{L}(q, \Theta)$, the summation sign is outside of the logarithm, giving the hope that this formulation may be more easily computable.

The Expectation-Maximization algorithm [Dempster *et al.*, 1977; Jordan and Bishop, 2002; Bilmes, 1998] takes advantage of the lower bound to maximize the log-likelihood. In an *E-step*, we compute a lower bound on the log-likelihood using Jensen's inequality and in an *M-step* we maximize that lower bound. By iterating these two procedures, we are bounding and maximizing the objective function which is therefore guaranteed to increase monotonically.

0. **Initialization** of the parameters $\Theta^{(0)}$ (randomly, or using more subtle strategies)
1. **Expectation step:** $q^{(t+1)} = \arg \max_q \mathcal{L}(q, \Theta^{(t)})$
2. **Maximization step:** $\Theta^{(t+1)} = \arg \max_{\Theta} \mathcal{L}(q^{(t+1)}, \Theta)$

The algorithm iterates over these two steps until a convergence criterion is reached. Usually, the criterion is based on variations of the likelihood:

$$|L(X|\Theta^{(t+1)}) - L(X|\Theta^{(t)})| \leq \epsilon \quad (4.8)$$

Now we need to explain why the maximization of a lower bound of the log-likelihood also maximizes the log-likelihood itself [Minka, 1998; Dellaert, 2002].

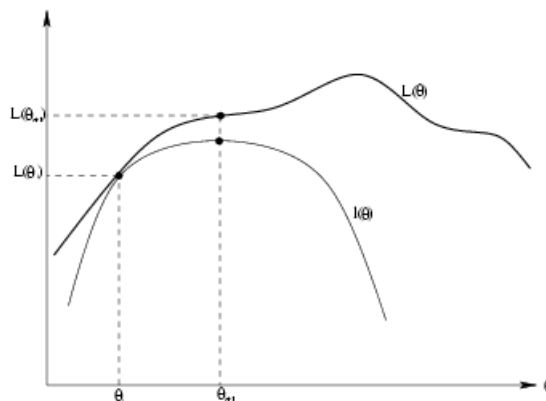


FIG. 4.1 – *Expectation-Maximization as Iterated Bound Maximization (similar to [Russell, 1998])*

In the **E-step**, we try to find an optimal bound for the incomplete log-likelihood. Thus we wish to maximize $\mathcal{L}(q, \Theta^{(t)})$ with respect to q . Indeed we try to find the best bound, defined as the bound $\mathcal{L}(q, \Theta^{(t)})$ that touches the objective function $L(\Theta)$ at the current guess Θ^t . We have already solved this problem when recognizing a Kullback-Leibler divergence as being the gap between the auxiliary function and the incomplete

log-likelihood: the lower bound is maximized and reaches $L(\Theta)$ when:

$$L(\Theta) = \mathcal{L}(q, \Theta^{(t)}) \quad \text{when} \quad q(Z) = p(Z|X, \Theta) \quad (4.9)$$

Thus the E-step consists of computing the distribution of the hidden variables conditioned by the data using the current estimation of the parameters. This is called the Estimation step because the bound is expressed as an expectation.

In the **M-step**, we are *maximizing the bound*, that is we wish to maximize $\mathcal{L}(q^{(t+1)}, \Theta)$ with respect to Θ . Using equation 4.7, the auxiliary function $\mathcal{L}(q, \Theta)$ becomes:

$$\begin{aligned} \mathcal{L}(q, \Theta) &= \sum_Z p(Z|X, \Theta) \log \frac{p(X, Z|\Theta)}{p(Z|X, \Theta)} \\ &= \sum_Z p(Z|X, \Theta) \log p(X, Z|\Theta) - \sum_Z p(Z|X, \Theta) \log p(Z|X, \Theta) \\ &= \sum_Z p(Z|X) \log p(X, Z|\Theta) - \sum_Z p(Z|X) \log p(Z|X) \end{aligned} \quad (4.10)$$

because Z is independent from Θ . The maximization of $\mathcal{L}(\Theta)$ over Θ can then be rephrased as the maximization of the so-called Q -function:

$$Q(\Theta) = \sum_Z p(Z|X) \log p(X, Z|\Theta) \quad (4.11)$$

$$Q(\Theta) = E_{Z|X} [\log p(X, Z|\Theta)] \quad (4.12)$$

Q is the *expected complete log-likelihood* and can be more easily maximized with respect to Θ than the initial incomplete log-likelihood.

However the M-step is problem dependent and it may be difficult to find the maximum over Θ . Fortunately, it is not strictly necessary to maximize the bound over Θ : indeed any improvement of the Q is enough to converge to a zero-gradient point. This is called as **Generalized EM** [Dempster *et al.*, 1977].

In the same way, it is also not necessary to fully maximize $\mathcal{L}(q, \Theta^{(t)})$ over q during the E-step: the bound need to touch the objective function. This allows to justify incremental, sparse and other variants of the EM algorithm [Neal and Hinton, 1998].

Other variants of the EM algorithm such as the Conditional EM (ECM) [Jebara and Pentland, 1998] or the Stochastic EM (SEM) are described in the literature.

4.2.2 Gaussian Mixture Models

Finite Mixture Model

In a finite mixture model [McLachlan and Peel, 2000], the probabilistic density function of $X|\Theta$ is defined as a weighted sum of K parametric functions (parameterized by θ_k):

$$p(X|\Theta) = \sum_{k=1}^K \pi_k p(X|k; \theta_k) \quad (4.13)$$

$\{\pi_k\}$ are the mixing coefficients (proportions) of each density function such that $\sum_{k=1}^K \pi_k = 1$

The unknown parameters are $\Theta = \{\theta_k, \pi_k\}$ and K , the number of mixture components, is an hyper-parameter.

Gaussian Mixture Model

In a Gaussian mixture model, the distribution of the data is modeled as the combination of K Gaussian distributions with parameters $\theta_k = \{\mu_k, \Sigma_k\}$:

$$\begin{aligned} p(X = x|k; \theta_k) &= (2\pi)^{-d/2} |\Sigma_k|^{-1/2} \exp\left\{-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k)\right\} \\ &= G_{\mu_k, \Sigma_k}(x) \end{aligned} \quad (4.14)$$

p denotes the dimension of the space from which the data X are observations. With this model, the parameters are then:

$$\Theta = \{\pi_1, \dots, \pi_{K-1}, \mu_1, \dots, \mu_K, \Sigma_1, \dots, \Sigma_K\} \quad (4.15)$$

The EM algorithm provides an elegant solution for the estimation of these parameters, given data x . Indeed, in the case of a mixture model, there is an obvious hidden variable, corresponding to the belonging of each sample to a particular class: if this information was known (that is, if the data were labelled), then the estimation would be easy, using the Bayesian framework. We will show that, in this case, the EM algorithm leads to simple closed form equations, that appear to be efficient compared to gradient descent algorithms.

In the following, the unknown parameters $Z_i \in 1 \dots K$ indicate that the i^{th} sample x_i is generated by the Z_i^{th} mixture component. We also define γ_i^k by:

$$\gamma_i^k = p(Z_i = k | X = x_i; \theta_k) \quad (4.16)$$

1. **Expectation step:** Using Baye's rule, we obtain:

$$\gamma_i^k = \frac{p(Z_i = k)p(X = x_i | Z_i = k; \theta_k)}{p(X = x_i; \theta_k)}$$

$$\gamma_i^k = \frac{\pi_k G_{\mu_k, \Sigma_k}(x_i)}{\sum_{l=1}^K \pi_l G_{\mu_l, \Sigma_l}(x_i)} \quad (4.17)$$

2. **Maximization step:** With a Gaussian mixture model and the assumption that the x_i are independent, the Q -function becomes:

$$\begin{aligned} Q(\Theta) &= \sum_Z p(Z|X) \log p(X, Z | \Theta) \\ &= \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log p(Z_i = k, X = x_i | \theta_k) \end{aligned} \quad (4.18)$$

Using Baye's rule once again, we get:

$$\begin{aligned}
 p(Z_i = k, X = x_i | \theta_k) &= p(X = x_i | Z_i = k, \theta_k) p(Z_i = k | \theta_k) \\
 Q(\Theta) &= \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log p(X = x_i | Z_i = k, \theta_k) + \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log p(Z_i = k | \theta_k) \\
 &= \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \left(-\frac{1}{2} \log |\Sigma_k| - \frac{1}{2} (x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) \right) \\
 &\quad + \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log \pi_k + cst \tag{4.19}
 \end{aligned}$$

The maximization of Q is obtained by setting the derivatives of Q over the parameters equal to 0.

- for the mixing coefficients, there is the constraint that their sum equals 1 so we use the Lagrange Multipliers technique to incorporate this constraint. We thus define:

$$\begin{aligned}
 Q'(\Theta) &= Q(\Theta) + \lambda \left(\sum_{k=1}^K \pi_k - 1 \right) \\
 \frac{\partial Q'(\Theta)}{\partial \pi_k} &= \sum_{i=1}^N \frac{\gamma_i^k}{\pi_k} + \lambda = 0
 \end{aligned}$$

We obtain the following system:

$$\begin{cases} \lambda \pi_k = - \sum_{i=1}^N \gamma_i^k \\ \sum_{k=1}^K \pi_k = 1 \end{cases}$$

from which we find that $\lambda = -N$ and then:

$$\boxed{\pi_k = \frac{1}{N} \sum_{i=1}^N \gamma_i^k} \tag{4.20}$$

- for the mean of the Gaussian components, we have:

$$\frac{\partial Q(\Theta)}{\partial \mu_k} = \sum_{i=1}^N \gamma_i^k \Sigma_k^{-1} (x_i - \mu_k)$$

using results on matrix derivations given in Appendix C. Setting this derivative equal to 0 gives:

$$\boxed{\mu_k = \frac{\sum_{i=1}^N \gamma_i^k x_i}{\sum_{i=1}^N \gamma_i^k}} \tag{4.21}$$

- at last, for the covariance matrices, we can consider several cases according the shape of the covariances we expect: full, diagonal or spherical.

Considering a *full covariance matrix*, we have:

$$\frac{\partial Q(\Theta)}{\partial \Sigma_k^{-1}} = \sum_{i=1}^N \frac{\gamma_i^k}{2} \Sigma_k - \frac{1}{2} \sum_{i=1}^N \gamma_i^k (x_i - \mu_k)(x_i - \mu_k)^T = 0$$

And then:

$$\Sigma_k = \frac{\sum_{i=1}^N \gamma_i^k (x_i - \mu_k)(x_i - \mu_k)^T}{\sum_{i=1}^N \gamma_i^k} \quad (4.22)$$

If we consider a *spherical covariance matrix* $\Sigma_k = \sigma_k^2 I_d$, we have a new expression for Q :

$$Q(\Theta) = \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \left(-d \log \sigma_k - \frac{1}{2\sigma_k^2} (x_i - \mu_k)^T (x_i - \mu_k) \right) + \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log \pi_k + cst$$

The derivation of Q over σ_k^{-1} yields:

$$\frac{\partial Q(\Theta)}{\partial \sigma_k^{-1}} = \sum_{i=1}^N \gamma_i^k (d\sigma_k - \sigma_k^{-1} (x_i - \mu_k)(x_i - \mu_k)^T) = 0$$

And then:

$$\sigma_k^2 = \frac{\sum_{i=1}^N \gamma_i^k (x_i - \mu_k)^T (x_i - \mu_k)}{d \sum_{i=1}^N \gamma_i^k} \quad (4.23)$$

Finally, for a *diagonal covariance matrix* $\Sigma_k = \text{diag} [\sigma_k^{2[1]} \dots \sigma_k^{2[d]}]$, we find:

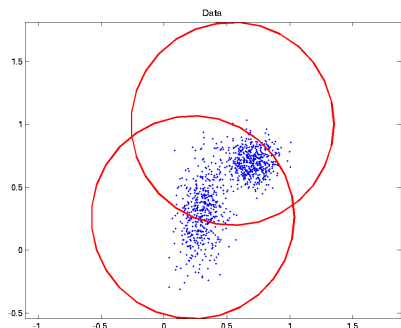
$$\sigma_k^{2[l]} = \frac{\sum_{i=1}^N \gamma_i^k (x_i^{[l]} - \mu_k^{[l]})^2}{\sum_{i=1}^N \gamma_i^k} \quad (4.24)$$

It is also possible to add constraints on covariance matrices to prevent correlation from some variables by setting to zero the corresponding elements in the matrices. For example, we can easily constraint covariance matrices to be block-diagonal, by just computing covariance terms within each block.

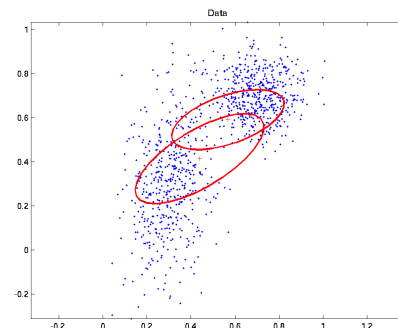
Fuzzy C-Means

If we add constraints to the parameters we can find a relationship between the EM algorithm and **fuzzy c-means** [Bezdek, 1981].

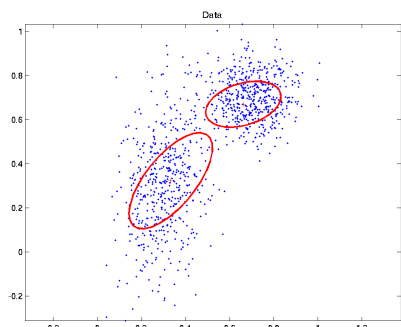
Let's consider a Gaussian mixture model with identical spherical covariance matrices and equal mixture proportions:



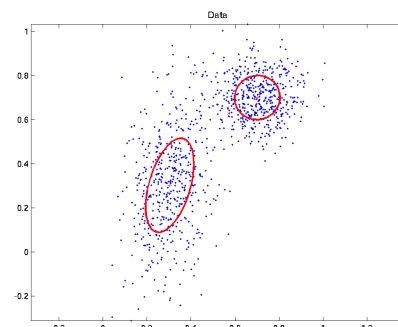
(a) Initialization



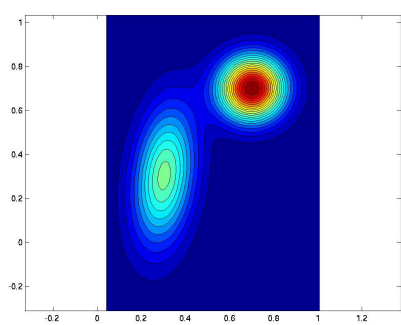
(b) After 5 iterations



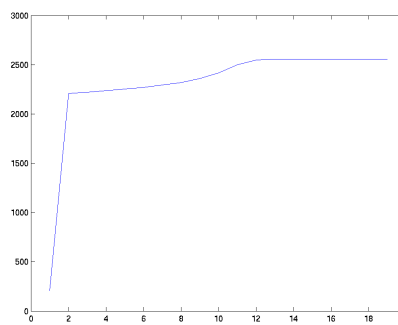
(c) After 10 iterations



(d) After 15 iterations

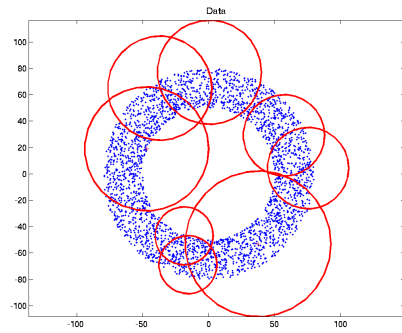


(e) View of the estimated PDF

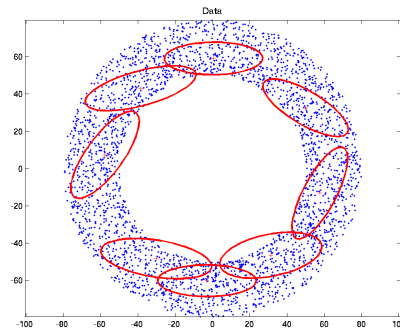


(f) Log-likelihood over time

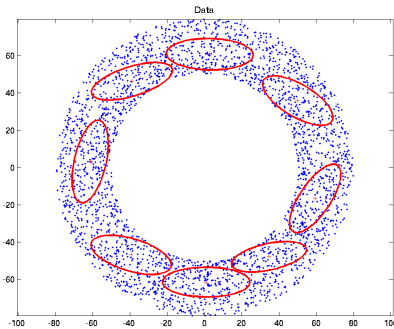
FIG. 4.2 – EM algorithm on a 2D synthetic data set (two Gaussian).



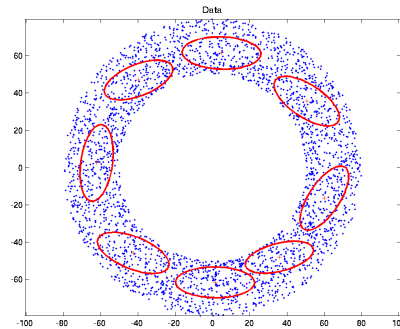
(a) Initialization



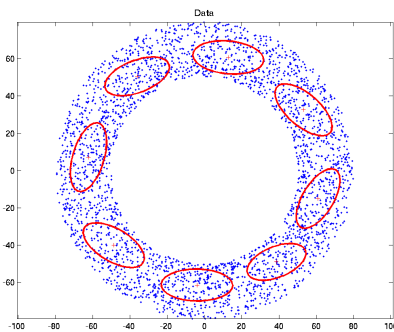
(b) After 5 iterations



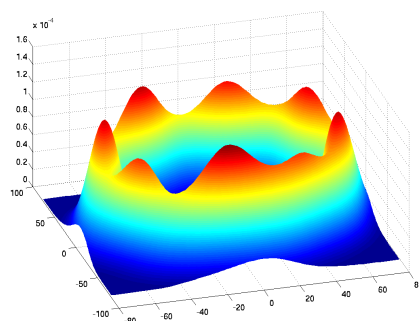
(c) After 10 iterations



(d) After 15 iterations



(e) After 20 iterations



(f) 3D representation of the estimated PDF

FIG. 4.3 – *EM algorithm applied on another 2D synthetic data set, using 8 classes. A similar example is shown in [Bishop, 1995].*

$$\begin{cases} \Sigma_k = \sigma^2 I_d, \forall k \\ \pi_k = \frac{1}{K}, \forall k \end{cases}$$

The Q -function (equation 4.19) becomes:

$$\begin{aligned} Q(\Theta) &= \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \left(-d \log \sigma - \frac{1}{2\sigma^2} (x_i - \mu_k)^T (x_i - \mu_k) \right) - \log K \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k + cst \\ &= -\frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k (x_i - \mu_k)^T (x_i - \mu_k) + cst \end{aligned}$$

Thus the maximization of $Q(\Theta)$ is equivalent to the minimization of the criterion:

$$\boxed{\widehat{Q}(\Theta) = \sum_{k=1}^K \sum_{i=1}^N \gamma_i^k (x_i - \mu_k)^T (x_i - \mu_k)} \quad (4.25)$$

This can be compared with the *fuzzy c-means* objective function:

$$J = \sum_{k=1}^K \sum_{i=1}^N (\gamma_i^k)^p (x_i - \mu_k)^T (x_i - \mu_k) \quad (4.26)$$

The parameter p is a weighting exponent on each fuzzy membership and determines the amount of fuzziness of the resulting classification: the larger p , the fuzzier the partition.

We can see that the objective function of an EM algorithm with Gaussian mixtures (spherical covariances and same proportions) is the *fuzzy c-means* criterion with $p = 1$.

The fuzzy partition is carried out through an iterative optimization with the update of membership γ_i^k and the cluster centres μ_k by:

$$\gamma_i^k = \frac{1}{\sum_{l=1}^K \left(\frac{(x_i - \mu_k)^T (x_i - \mu_k)}{(x_i - \mu_l)^T (x_i - \mu_l)} \right)^{\frac{1}{p-1}}} \quad \text{and} \quad \mu_k = \frac{\sum_{i=1}^N (\gamma_i^k)^p x_i}{\sum_{i=1}^N (\gamma_i^k)^p}$$

We can notice that with $p = 1$, we find back the EM update equation 4.21 for μ_k but the equations for the membership γ_i^k are different (compared with 4.17). [Ichihashi *et al.*, 2001] study the similarities between Gaussian mixture models and fuzzy C-means: they show that the same algorithm as the EM for Gaussian mixtures can be derived from a modified objective function of the fuzzy C-means using a regularization term by Kullback-Leibler information.

Gaussian Mixture Model and Uniform Noise

If data is corrupted by random noise, we can add a uniform noise component in the mixture model.

$$p(X|\Theta) = \sum_{k=1}^K \pi_k p(X|k; \theta_k) + \pi_{K+1} U_{\mathbb{D}}(X) \quad (4.27)$$

where $U_{\mathbb{D}}(X) = \frac{1}{\lambda(\mathbb{D})}1_{\mathbb{D}}(X)$ and $\mathbb{D} \subset \mathbb{R}^d$.

In this case, the M-step remains the same, only the E-step is modified. In this formulation, the difficulty is that the domain \mathbb{D} is unknown and should be estimated.

See figure 4.4 for an example on a synthetic dataset. Two 2D gaussian distributions are corrupted by uniform noise. In that case, estimating the parameters of the mixture model using two classes fails because one class tries to model outliers data from the uniform distribution. Adding a third class is a good compromise because the uniform noise will be modeled by one gaussian. The latest case uses explicitly a uniform class in the model to be fitted to the data: results are correct but the domain \mathbb{D} was considered as known and results might be poorer if it had to be estimated.

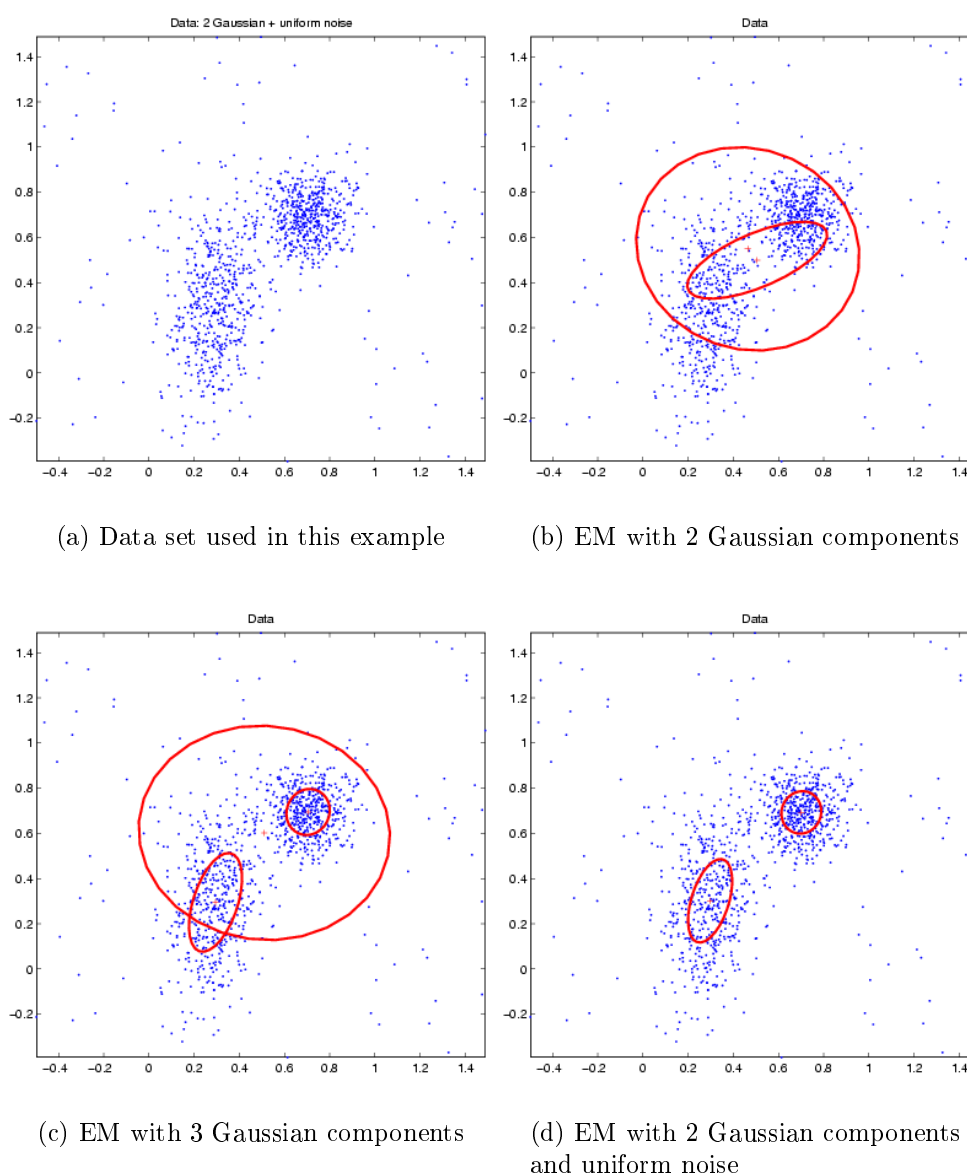


FIG. 4.4 – *Gaussian mixture model corrupted by uniform noise.*

Other Mixture Models

The EM algorithm can be derived for other mixture components than Gaussian distributions: there are references for mixture models of Poisson, logistic, Bernoulli or Weibull distributions (see the Mixture Modeling Page from David Dawes¹⁹) but the Gaussian (normal) one is the most often used due to its simplicity and the analytical formulas for the M-step.

Bayesian Regularization using Conjugate Priors

As noted by Bishop [Bishop, 1995], maximizing the likelihood function is non-trivial and may lead to local maxima or degeneracy. For instance, we can easily demonstrate with a simple example that the likelihood of a mixture density approaches infinity if one of the kernels is placed on a data point and the variance is let to zero, i.e. the likelihood is an unbounded function. Furthermore, local maxima can be reached when small groups of points are close together, leading to a poor estimation of the underlying distribution.

Several strategies have been proposed to overcome singular solutions and local maxima. The first one is to constrain all the components of the mixture to have equal covariance matrices (homogeneous process). Another strategy often implemented in EM algorithms consists in checking the singularity of covariance matrices using a Singular Value Decomposition, at each iteration. If one component is found to have a covariance matrix with a small value on one of its dimensions, then it is reinitialized to a larger value (e.g. its initial value). However this often leads to a decrease of the loglikelihood function at the next step, and, more importantly, to an oscillatory behavior of the algorithm.

A proper way to deal with this ill-posed problem of the maximum likelihood estimation is to use *Bayesian inference*. Although the results obtained with *maximum likelihood* and *Bayesian inference* are frequently nearly identical, the approaches are conceptually quite different. Maximum likelihood view the parameters as quantities whose values are fixed but unknown; the best estimates of their value is defined to be the one that maximizes the probability of obtaining the samples actually observed. In contrast, Bayesian methods view the parameters as random variables having some prior distribution. Observation of the samples converts this to a posteriori density, thereby revising our opinion about the true values of the parameters.

We can see the close link between *Maximum a Posteriori* learning and *Maximum Likelihood* inference using Bayes' theorem:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} \quad \text{i.e.} \quad \text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}} \quad (4.28)$$

Thus, maximizing the likelihood is equivalent to posterior maximization if no priors are affected to the parameters (uniform law). Otherwise, *Maximum a Posteriori* (MAP) principle can be seen as a *penalized likelihood maximization*:

$$\max_{\theta} \log \text{posterior} = \max_{\theta} \log \text{likelihood} + \log \text{prior} \quad (4.29)$$

19. <http://www.csse.monash.edu.au/~dld/mixture.modelling.page.html>

In penalized maximum likelihood, the likelihood is augmented by an extra term that can be interpreted either as a prior on the parameters [Ormoneit and Tresp, 1997; 1998; Ridolfi and Idier, 2000] or as a regularization that punishes for complexity of the model [Green, 1990]. In the following, we will focus on Ormoneit et al. work [Ormoneit and Tresp, 1997; 1998] on MAP learning. Using priors on the parameters eliminates the need for ugly hacks like resetting singular covariance matrices or using equal covariance matrices (strong hypothesis).

One of the key issue is the choice of the prior distribution on the parameters. If the prior is chosen to be from a *conjugate family*, then we can obtain closed-form update rules in the EM derivation. A conjugate family assures that for any distribution on this family, the posterior also belongs to the same family, i.e.,

A family F of probability distributions on θ is said to be conjugate if, for every $p \in F$, the posterior $p(\theta|x)$ also belongs to F .

We give in the following tabular the conjugate prior distribution for a few probability distributions:

Probability distribution generating the sample	Corresponding family of conjugate prior distribution
Normal	Normal \times Wishart
Multinomial	Dirichlet
Poisson	Gamma

Thus the conjugate prior of a multivariate normal density $G_{\mu_k, \Sigma_k}(x)$ is the product of a normal density $G_{\hat{\mu}_k, \eta_k^{-1} \Sigma_k}(\mu_k)$ and a Wishart density $W_{\alpha_k, \beta_k}(\Sigma_k^{-1})$ defined as:

$$W_{\alpha_k, \beta_k}(\Sigma_k^{-1}) = c |\Sigma_k^{-1}|^{\alpha_k - (d+1)/2} \exp(-tr(\beta_k \Sigma_k^{-1})) \quad \text{with } \alpha_k > (d+1)/2 \quad (4.30)$$

The conjugate prior for the multinomially distributed mixing proportions π_k is a Dirichlet density $D_\nu(\pi)$ defined by:

$$D_\nu(\pi) = b \prod_{k=1}^K \pi_k^{\nu_k - 1} \quad \text{with constraints } \pi_k \geq 0 \quad \text{and} \quad \sum_{k=1}^K \pi_k = 1 \quad (4.31)$$

b and c are normalizing constants and $\hat{\mu}, \eta, \alpha, \beta, \nu$ are parameters of the prior distributions (*hyperparameters*).

The full prior is then given by:

$$p(\Theta) = D_\nu(\pi) \prod_{k=1}^K G_{\hat{\mu}_k, \eta_k^{-1} \Sigma_k}(\mu_k) W_{\alpha_k, \beta_k}(\Sigma_k^{-1}) \quad (4.32)$$

The penalized loglikelihood is:

$$\begin{aligned} L_p(\Theta) &= L(\Theta) + \log p(\Theta) \\ &= \sum_{i=1}^N \log \left(\sum_{k=1}^K \pi_k G_{\mu_k, \Sigma_k}(x_i) \right) + \log D_\nu(\pi) + \sum_{k=1}^K \left[\log G_{\hat{\mu}_k, \eta_k^{-1} \Sigma_k}(\mu_k) + \log W_{\alpha_k, \beta_k}(\Sigma_k^{-1}) \right] \end{aligned}$$

Let us derive the EM algorithm for this penalized loglikelihood. The E-step remains the same than in the unpenalized case (equation 4.17). For the M-step, the expected complete loglikelihood Q becomes:

$$\begin{aligned}
 Q(\Theta) &= \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \left(-\frac{1}{2} \log |\Sigma_k| - \frac{1}{2} (x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) \right) + \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log \pi_k \\
 &+ \sum_{k=1}^K \left[\left(-\frac{1}{2} \log |\eta_k^{-1} \Sigma_k| - \frac{1}{2} (\mu_k - \hat{\mu}_k)^T \eta_k \Sigma_k^{-1} (\mu_k - \hat{\mu}_k) \right) + \log W_{\alpha_k, \beta_k}(\Sigma_k^{-1}) \right] \\
 &+ \log D_\nu(\pi) + cst
 \end{aligned} \tag{4.33}$$

For the mixing proportions, using the Lagrange Multipliers technique, we obtain the following system:

$$\begin{cases} \lambda \pi_k = 1 - \nu_k - \sum_{i=1}^N \gamma_i^k \\ \sum_{k=1}^K \pi_k = 1 \end{cases} \quad \text{which gives} \quad \lambda = K - \sum_{k=1}^K \nu_k - N$$

yielding:

$$\boxed{\pi_k = \frac{\sum_{i=1}^N \gamma_i^k + \nu_k - 1}{\sum_{k=1}^K \nu_k + N - K}} \tag{4.34}$$

For the mean of the Gaussian components, we have:

$$\frac{\partial Q(\Theta)}{\partial \mu_k} = \frac{1}{2} \sum_{i=1}^N \gamma_i^k 2 \Sigma_k^{-1} (x_i - \mu_k) - \frac{1}{2} 2 \eta_k \Sigma_k^{-1} (\mu_k - \hat{\mu}_k) = 0$$

which gives:

$$\boxed{\mu_k = \frac{\sum_{i=1}^N \gamma_i^k x_i + \eta_k \hat{\mu}_k}{\sum_{i=1}^N \gamma_i^k + \eta_k}} \tag{4.35}$$

At last, for the covariance matrices, we obtain:

$$\frac{\partial Q(\Theta)}{\partial \Sigma_k} = \sum_{i=1}^N \frac{\gamma_i^k}{2} \Sigma_k - \frac{1}{2} \sum_{i=1}^N \gamma_i^k (x_i - \mu_k)(x_i - \mu_k)^T + \frac{1}{2} \Sigma_k \tag{4.36}$$

$$-\frac{\eta_k}{2} (\mu_k - \hat{\mu}_k)(\mu_k - \hat{\mu}_k) + (\alpha_k - (d+1)/2) \Sigma_k - \beta_k = 0 \tag{4.37}$$

and then:

$$\boxed{\Sigma_k = \frac{\sum_{i=1}^N \gamma_i^k (x_i - \mu_k)(x_i - \mu_k)^T + \eta_k (\mu_k - \hat{\mu}_k)(\mu_k - \hat{\mu}_k)^T + 2\beta_k}{\sum_{i=1}^N \gamma_i^k + 1 + 2(\alpha_k - (d+1)/2)}}$$

We can notice that this Bayesian approach requires only negligible additional computations compared to a standard EM algorithm.

Furthermore if we compare the M-steps of the unregularized and the regularized cases, we note that the regularization effect is formulated in the form of hyperparameters that correspond to an imaginary data set with known sufficient statistics. In the case of mixing proportions, for example, the hyperparameters ν_k correspond to the proportion of the imaginary data counts from the kernel indexed with k . The hyperparameters can be formulated with prior knowledge or simply, by experimenting with data.

If we take $\nu_k = 1$, $\eta_k = 0$, $\beta_k = 0$ and $\alpha_k = \frac{d+1}{2} \forall k$, then we find back the classical M-step updates with no prior.

In figure 4.2.2.0, we show the mixture model resulting of an EM algorithm applied on a circular bidimensional dataset, with 30 classes. Priors were just defined on covariance matrices with $\beta \gg 0$

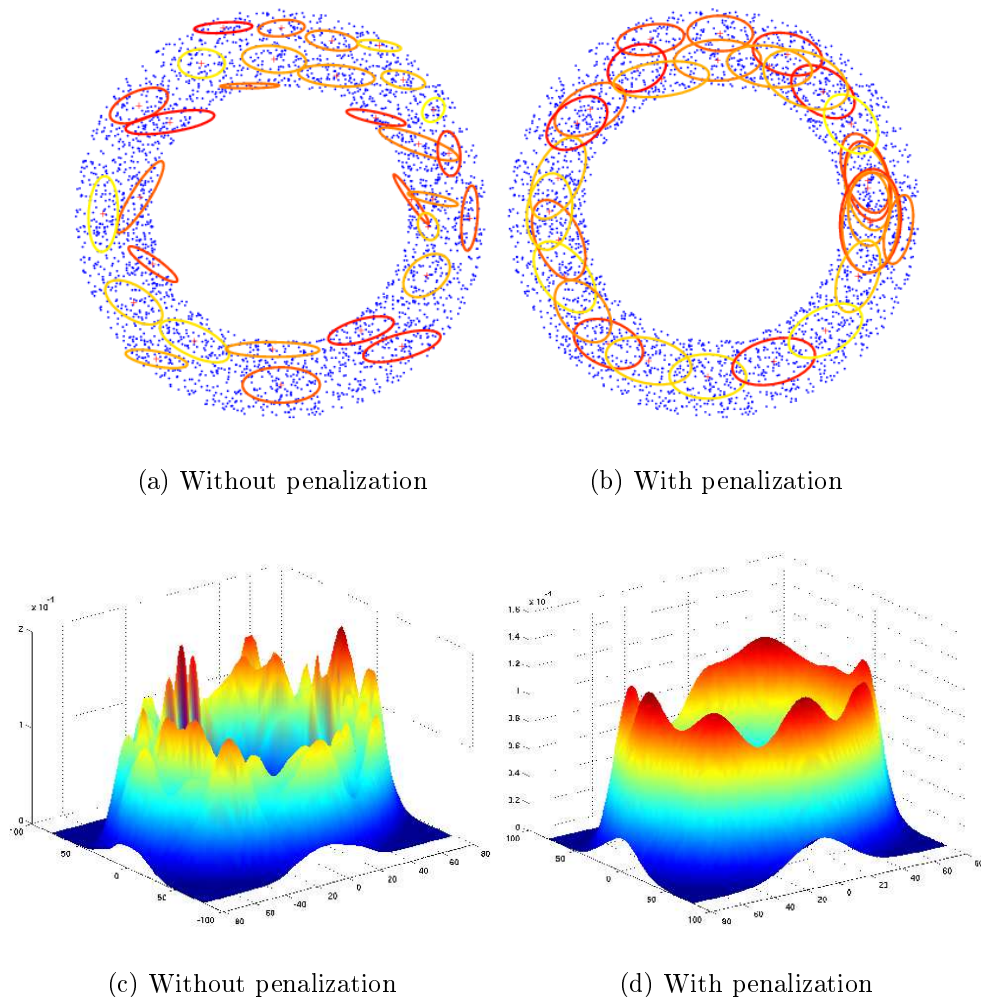


FIG. 4.5 – *Synthetic dataset (same as figure 4.3) with a high number of classes : we see that without constraints, some of the gaussian components converge to singularities. This kind of event is penalized using conjugate priors on the covariance matrices.*

4.2.3 Model Order Selection

The number K of mixture components is a hyper parameter that can be estimated through an information criterion. The strategy is to compute the ML-estimator with the EM algorithm for several values of K , $K = 1, \dots, K_{max}$ and then choose K that maximizes the criterion.

The maximized log-likelihood can be bias-corrected to penalize models with too many parameters: the aim is to find a tradeoff between under-fitting and over-fitting. The model with the smallest criterion value is estimated to be “closest” to the truth and is the best approximation for the information on the data, relative to the other models considered (in the sense of the principle of parsimony).

The Akaike Information Criterion (AIC) [Akaike, 1973; 1974] is a parsimonious approach to the estimation of relative expected Kullback-Leibler divergence:

$$\boxed{AIC(K) = -2 \log p(X|\Theta_K) + 2\nu_K} \quad (4.38)$$

where Θ_K and ν_K are respectively the maximum likelihood estimator and the number of degrees of freedom of the model.

This criterion is asymptotically equivalent to the cross-validation criterion (*leave one out*) but it has a tendency to choose too complex models.

A second-order AIC has also been derived with an additional bias correction term to handle cases better when there are too many parameters in relation to the size of the sample [Burnham and Anderson, 1998]:

$$AIC_c = -2 \log p(X|\Theta_K) + 2\nu_K \left(\frac{N}{N - \nu_K - 1} \right) = AIC + \frac{2\nu_K(\nu_K + 1)}{N - \nu_K - 1} \quad (4.39)$$

However the Bayesian Information Criterion (BIC) from Schwarz [Schwarz, 1978] is often preferred:

$$\boxed{BIC(K) = -2 \log p(X|\Theta_K) + \nu_K \log N} \quad (4.40)$$

This criterion is of bayesian inspiration but is equivalent, in that framework, to the Minimum Description Length (MDL) criterion, coming from the information theory [Rissanen, 1978; Bishop, 1995].

For a Gaussian mixture model with full covariance matrices, the number of degrees of freedom are:

$$\nu_K = K - 1 + Kd + K(d(d + 1)/2)$$

Other criteria have been defined such as the Integrated Classification Likelihood (ICL) [Biernacki *et al.*, 1998] or the Entropy Criterion (NEC) [Biernacki *et al.*, 1999] (more appealing in a classification framework) but the most common used is the BIC.

4.2.4 Classification

In many situations, the goal of a mixture modeling estimation is to classify each data value to one of the K classes: this is a hard partition of the data. Indeed, this information

is stored in the hidden variable Z ($Z_i \in 1 \dots K, \forall i = 1 \dots N$). There are two main ways to use a mixture model for hard classification:

- After convergence of the EM algorithm described below, we can compute the classification \hat{z}_i which maximizes the posterior probability (MAP classification). We simply assign each x_i to the most probable class according to the posterior probabilities γ_i^k :

$$\hat{z}_i \triangleq \arg \max_{1 \leq k \leq K} \gamma_i^k \quad (4.41)$$

- Celeux and Govaert proposed in [Celeux and Govaert, 1992] the **Classification EM** (CEM) in which they add an intermediate classification step between the E and the M step.

1. **E-step** Expectation: as EM
2. **C-step** Classification: update the partition by assigning each x_i to the cluster which provides the maximum current conditional probability γ_i^k (MAP principle).
3. **M-step** Maximization: as EM

CEM provides a biased estimator of Θ if clusters are not well separated, however it is much faster than EM, so results of a CEM algorithm are often used as initialization of the EM algorithm.

Let us add constraints to the mixture parameters to show the relationship between EM algorithm and the famous **K-means algorithm** [Duda and Hart., 1973].

As for the fuzzy c-means, let us consider a Gaussian mixture model with identical spherical covariance matrices and equal mixture proportions:

$$\begin{cases} \Sigma_k = \sigma^2 I_d, \forall k \\ \pi_k = \frac{1}{K}, \forall k \end{cases}$$

We remind ourselves that, in such a case, the Q -function (equation 4.19) becomes:

$$Q(\Theta) = -\frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k (x_i - \mu_k)^T (x_i - \mu_k) + cst$$

Furthermore, in a classification context (hard partition), we have:

$$\forall i, \gamma_i^k = \begin{cases} 1 & \text{if } Z_i = k \\ 0 & \text{otherwise} \end{cases}$$

Thus, the maximization of $Q(\Theta)$ corresponds at this point to the minimization of:

$$\hat{Q}(\Theta) = \sum_{k=1}^K \sum_{i \in C_k} (x_i - \mu_k)^T (x_i - \mu_k) \quad (4.42)$$

We recognize the *intra-class* variance sum, which is the criterion of the K-means algorithm [Hartigan and Wong, 1979]. Indeed the K-means algorithm is a special case of a Classification EM.

Therefore the EC-step is to assign each data point to its nearest cluster mean:

$$\hat{z}_i = \arg \max_{1 \leq k \leq K} \frac{G_{\mu_k}(x_i)}{\sum_{l=1}^K G_{\mu_l}(x_i)} = \arg \max_{1 \leq k \leq K} G_{\mu_k}(x_i) = \arg \min_{1 \leq k \leq K} (x_i - \mu_k)^T (x_i - \mu_k) \quad (4.43)$$

And the M-step consists of updating the positions of the cluster means (or centroids):

$$\forall k, \mu_k = \frac{\sum_{i=1}^N \gamma_i^k x_i}{\sum_{i=1}^N \gamma_i^k} = \frac{1}{\text{Card}\{C_k\}} \sum_{i \in C_k} x_i \quad (4.44)$$

We will present some specificities of the K-Means algorithm in the next chapter.

4.2.5 Incorporating Spatial Priors

In many cases people are dealing with spatial data (Cressie [Cressie, 1993] classifies them in three categories: geostatistical data, lattice data and point patterns): data collected from different spatial locations (geology, epidemiology, ecology, astronomy, etc.) or image processing (sensor images, medical or biological images, etc.). If spatial information is not used in the clustering algorithm, the resulting classes may appear spatially messy. Thus the objective when clustering spatial data [Dang and Govaert, 1998] is the following:

1. the clusters should be as homogeneous as possible, i.e. observations should be as similar to each other as possible within a group
2. it is assumed that the partition changes slowly in the geographic space, i.e. two neighboring observations are more likely to belong to the same group than two observations lying far apart (hypothesis of spatial smoothness).

In the following we mainly consider data sampled on a regular \mathbb{R}^d lattice, eg. multidimensional images.

A first approach to add spatial information in a clustering algorithm is to add data **spatial coordinates as additional variables** of each sample.

$$p(X, V | \Theta) \triangleq \sum_{k=1}^K \pi_k p(X, V | k; \theta_k) \quad (4.45)$$

We will discuss and interpret this last equation in chapter 7.

Another approach to take spatial constraints into account consists in **penalizing the likelihood criterion** by adding a “neighborhood term”. To do so, let’s define v_{ij} by:

$$v_{ij} = \begin{cases} \alpha \geq 0 & \text{if } x_i \text{ and } x_j \text{ are neighbors} \\ 0 & \text{otherwise} \end{cases} \quad (4.46)$$

The regularizing term proposed by Ambroise et al. [Ambroise *et al.*, 1997] is:

$$G(\gamma) \triangleq \frac{1}{2} \sum_{k=1}^K \sum_{i=1}^N \sum_{j=1}^N \gamma_i^k \gamma_j^k v_{ij} \quad (4.47)$$

Indeed the more the clusters contain adjacent samples the greater this term is. The overall criterion is then:

$$\widehat{\mathcal{L}}(\Theta) \triangleq \mathcal{L}(\Theta) + \beta G(\gamma) \quad (4.48)$$

where β is a positive weighting coefficient allowing to adjust the influence of the spatial regularization. Ambroise et al. propose a modified version of the EM algorithm called **Neighborhood EM** (NEM) [Ambroise *et al.*, 1997; Dang and Govaert, 1998] to optimize this criterion.

We present the derivation of the NEM in the framework of alternate minimization described in Appendix B. Thus the modified criterion is:

$$\widehat{\mathcal{L}}(\Theta, \gamma) = \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log [\pi_k G_{\mu_k, \Sigma_k}(x_i)] - \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log \gamma_i^k + \beta \frac{1}{2} \sum_{k=1}^K \sum_{i=1}^N \sum_{j=1}^N \gamma_i^k \gamma_j^k v_{ij} \quad (4.49)$$

The derivation of this criterion over γ_i^k (E-step) using the Lagrangian yields the following system:

$$\begin{cases} \gamma_i^k = \exp \left\{ -1 + \lambda_i + \log [\pi_k G_{\mu_k, \Sigma_k}(x_i)] + \beta \sum_{j=1}^N \gamma_j^k v_{ij} \right\} \\ \sum_{k=1}^K \gamma_i^k = 1 \end{cases}$$

The Lagrangian multipliers are then found to be:

$$\lambda_i = 1 - \log \left[\sum_{k=1}^K \pi_k G_{\mu_k, \Sigma_k}(x_i) \exp \left\{ \beta \sum_{j=1}^N \gamma_j^k v_{ij} \right\} \right]$$

And finally, the γ_i^k 's are:

$$\boxed{\gamma_i^k = \frac{\pi_k G_{\mu_k, \Sigma_k}(x_i) \exp \left\{ \beta \sum_{j=1}^N \gamma_j^k v_{ij} \right\}}{\sum_{l=1}^K \pi_l G_{\mu_l, \Sigma_l}(x_i) \exp \left\{ \beta \sum_{j=1}^N \gamma_j^l v_{ij} \right\}}} \quad (4.50)$$

The variable γ_i^k appears in both sides of this equation: thus γ_i^k is a fixed point of this equation. A way to estimate these variables is to iteratively compute new values using the previous ones up to convergence. The comparison with the equation obtained in a standard E-step of an EM algorithm (equation 4.17) is interesting: we can see that each posterior probability of X is weighted by this term:

$$\exp \left\{ \beta \sum_{j=1}^N \gamma_j^l v_{ij} \right\}$$

which smoothes the partitioning with a tuning parameter provided by β .

The derivation of the new criterion over the mixture parameters (M-step) is the same as in the EM algorithm since $G(\gamma)$ is just a function of the partitioning parameters.

Interestingly, [Dang and Govaert, 1998] show that the NEM algorithm can be related to an EM algorithm procedure for hidden Markov random fields (MRF) with mean field

approximation.

Indeed, **Markov random field models** are another tool to model spatial context when clustering spatial data. Previously, we assumed that each sample was labelled independently from the other samples (i.e. the $\{Z_i\}$'s are independent). Here, we can use a more complex model for Z : a MRF model. Such a model assumes the the probability that the sample x_i belongs to the k^{th} mixture component depends on the labeling of its neighbors. The Hammersley-Clifford theorem states that such a random field is a Gibbs random field [Besag, 1974], thus defining a new distribution for Z . This modifies the computation of the γ_i^k in the Estimation step of an EM algorithm, and the parameters of the Gibbs distribution need also to be estimated in the Maximization step. More details about the implementation of such an algorithm can be found in [Leemput, 2001]. This model has been used in the following papers [Kapur, 1999; Kapur *et al.*, 1998; Zhang *et al.*, 2000; 2001; Leemput, 2001; Leemput *et al.*, 2001] for medical imaging segmentation purpose.

4.3 Clustering methods in fMRI data analysis

Clustering algorithms are used for fMRI data analysis as an exploratory tool. The aim is to cluster fMRI time series in groups of voxel with similar activations. The underlying assumption is that the pattern of activation has a structure that can be divided into a few types of similar activations.

This kind of analysis can be useful to extract directly from the data what are the main archetypal times series that can be observed, without having to specify any a priori model. Clustering techniques can also be a way to detect fonctionnally connected areas in the brain, that share similar temporal properties.

Different algorithms have been proposed in the literature, we recall the principal ones in the next sections.

4.3.1 Definition of the feature space

The definition of the feature space in which the clustering algorithm will take place is of primary importance, because it is linked with the choice of the metric that is used to quantify the similarity between time courses [Thirion and Faugeras, 2003b]. The most classical one is the Euclidean distance in the original signal space on the raw time series [Baune *et al.*, 1999; Baumgartner *et al.*, 2000], however this provides a potentially high dimensional space, leading to computer issues. To reduce the dimensionality of the data space, [Goutte *et al.*, 1999] proposed to use the first cross-correlation coefficients between the time courses and a box-car function modeling the paradigm. This allows to use features whose dimension remain low. Furthermore, this adds some knowledge about the paradigm, which allows an easier interpretation of the “typical” time serie found for each cluster. At last, [Goutte *et al.*, 2001] explore the possibility of clustering features extracted from the fMRI time series,

such as the results of several standard single-voxel analyses: Student t test, Kolmogorov-Smirnov test, correlation with the paradigm, finite impulse response filter model, etc. Using these features in a clustering algorithm corresponds to perform a meta-analysis that is capable to separate clusters that are both activated but have different activation delays.

The main clustering algorithm have been used with these metrics yielding similar results: K-Means [Goutte *et al.*, 1999], fuzzy C-means [Baumgartner *et al.*, 2000; Fadili *et al.*, 2000], hierarchical clustering [Goutte *et al.*, 1999], Kohonen maps [Chuang *et al.*, 1999]. One of the key issue with clustering algorithms is the number of clusters: a commonly used criterion is the BIC for methods based on mixture models. [Baune *et al.*, 1999] proposed a dynamical cluster algorithm which does not need a prior initial specification of the number of clusters. Most of the clustering algorithms are also known to be very sensitive to the initialization: a solution is to perform several clusterings with different starting points and choosing the one maximizing the criterion.

4.3.2 Introduction of spatial information

Some attempts have been proposed to add spatial information in clustering algorithms. A solution is to favour the construction of connected clusters. Indeed, standard clustering algorithm do not rely on the fact that the data are spatially distributed. Thus, [Almeida and Ledberg, 2001] proposed to modify a hierarchical clustering algorithm to only merge clusters that are neighbors, using the correlation coefficient as similarity measure. The output clusters are then necessarily connected.

[Penny and Friston, 2003] specified a spatio-temporal framework, linked with functional clustering analyses because it aims to identify clusters of voxels with similar time courses that are in similar anatomical locations. A generative model is presented where the activation clusters are depicted with a Gaussian spatial shape. This model is however difficult to estimate on the dataset. We will come back on this model in chapter 7.

Clustering algorithms can also be used as an alternative to voxel-wise thresholding of statistical parametric maps. [Salli *et al.*, 2001; Salli, 2002] proposed to label voxels active or non-active using a contextual clustering algorithm: a voxel whose neighbors are labeled active is more likely to be also considered active. This segmentation is performed through an iterative conditional modes (ICM) algorithm.

4.4 Conclusion

In this chapter, we introduced the main algorithms used in unsupervised learning for cluster analysis. The aim of these methods is to model the distribution of the data in the feature space as a multimodal distribution and to characterize each mode as a data cluster. We pointed out that the well-known K-Means and fuzzy C-means algorithms are special cases of Gaussian mixture models and have been widely used in the machine learning community because of their simplicity and robustness.

Clustering algorithms are an exploratory, multivariate, model-free alternative to the univariate general linear model in fMRI data analysis. Such a tool allows to automati-

cally classify data in activated and non activated clusters, without making any prior on the shape of the activations or on temporal properties of the activations (note that no inference can be easily performed to assess the results). Clustering can also be performed on previously detected voxels to look for different temporal properties on the pattern of activations: time delay, negative activation, etc. However these techniques remain computationnally intensive because they deal with the whole dataset at the same time. For example, hierarchical clustering needs to build a $N \times N$ symmetric distance matrix (where N denotes the number of voxels), which takes a huge amount of memory. For these reasons, a lot of results presented in the litterature are performed on bidimensional slices or on a reduced dataset (activated voxels for example). The choice of the feature space and the corresponding metric is also an important point and there is no consensus at that time. At last, little work has be done to add spatial information in these clustering algorithms: this appears to be a way to investigate.

Chapitre 5

Parcellisation anatomique du cortex

Sommaire

5.1	Motivations	104
5.2	Segmentation des structures cérébrales	105
5.2.1	Segmentation du néocortex à partir d'IRM T_1	106
5.2.2	Segmentation du cerveau à partir d'IRM T_2^*	107
5.2.3	Obtention du domaine à parcelliser	111
5.3	Parcellisation d'un domaine convexe	112
5.3.1	Parcellisation et critère à optimiser	112
5.3.2	L'algorithme K-Means	114
5.3.3	Transformations de distance	116
5.3.4	Mise à jour des graines	117
5.3.5	Résultats obtenus	118
5.4	Prise en compte de distances géodesiques	121
5.4.1	Diagramme de Voronoï géodésique	122
5.4.2	Calcul du centre de masse géodésique	125
5.4.3	Exemples de parcellisations obtenues	127
5.5	Introduction de contraintes spatiales	130
5.5.1	Introduction de l'information <i>sillons</i>	130
5.5.2	Imposition de la position de certaines graines	133
5.6	Conclusion	135

Dans ce chapitre, nous présentons un algorithme permettant d'obtenir une parcellisation anatomique du cortex grâce à l'utilisation de diagrammes de Voronoï géodésiques centroïdaux. Nous montrons également comment introduire des contraintes à cette parcellisation comme la séparation des aires par les sillons corticaux, ceci grâce à l'utilisation de diagrammes de Voronoï pondérés.

L'intérêt d'une telle parcellisation est multiple: elle permet tout d'abord d'obtenir un cadre dans lequel les données fonctionnelles peuvent être plongées, permettant de réduire la quantité d'information à traiter. De plus, ce changement d'espace est informé par l'anatomie permettant ainsi d'effectuer une restauration des données (réduction du bruit).

Après avoir explicité plus amplement les motivations sous-jacentes à une parcellisation, nous précisons la définition du volume d'intérêt à parcelliser. Ensuite, nous présentons l'algorithme utilisé dans le cas de domaine convexe, que nous modifions pour prendre en compte des distances géodésiques. Enfin, nous proposons une autre modification de cet algorithme pour prendre en compte d'éventuelles contraintes spatiales.

5.1 Motivations

Nous avons vu dans le chapitre 3 qu'il existait principalement deux approches pour analyser les données fonctionnelles: soit une approche univariée, voxel à voxel, soit une approche par région d'intérêt. Ces deux approches ne sont pas exemptes de limitations. Pour les approches "voxels", les limitations sont:

- Le nombre de voxels (20000 environ) est beaucoup plus élevé que le nombre de mesures temporelles, ce qui empêche l'application directe de méthodes multivariées nécessitant l'inversion de matrices temps \times temps,
- L'information présente au niveau voxel est partiellement redondante en raison de la corrélation spatiale: il y a moins de degré de liberté que de voxels. Une représentation plus concise des données est alors possible sans pour autant perdre d'information.

Le même constat s'applique sur les représentations surfaciques des données fonctionnelles, avec de 20000 à 30000 nœuds par hémisphère sur le maillage (voir figure 5.1).

D'un autre côté, les approches reposant sur des régions d'intérêt ont les inconvénients d'être souvent trop grossières (une cinquantaine d'aires sur le cerveau), longues et difficiles à définir. En effet, les régions d'intérêt sont souvent délinées manuellement sur l'image anatomique puis reportées sur les images fonctionnelles.

La parcellisation anatomique que nous présentons dans ce chapitre permet de définir automatiquement une représentation intermédiaire quelconque des données entre le voxel et la région d'intérêt. La résolution effective de la représentation est alors un paramètre réglé par l'utilisateur. La figure 5.2 présente le diagramme de la construction d'une parcellisation du cortex, dont nous allons expliciter dans la suite de ce chapitre les différentes étapes.

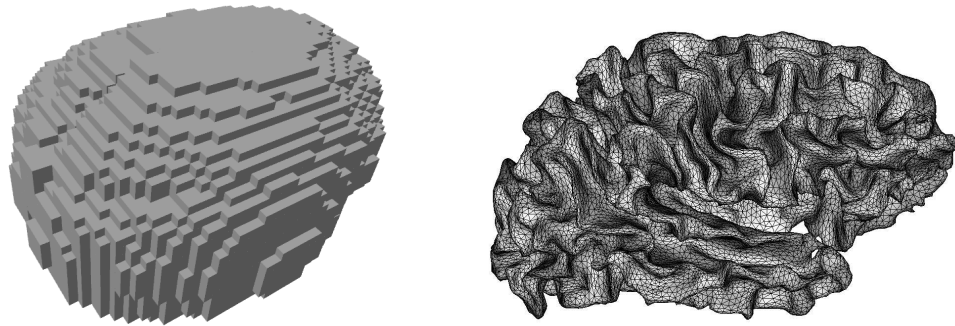


FIG. 5.1 – Ensemble de voxels fonctionnels modélisant le cerveau (environ 20000 pour une résolution standard) et maillage surfacique de l'interface entre matière grise et matière blanche (de 20000 à 30000 nœuds par hémisphère).

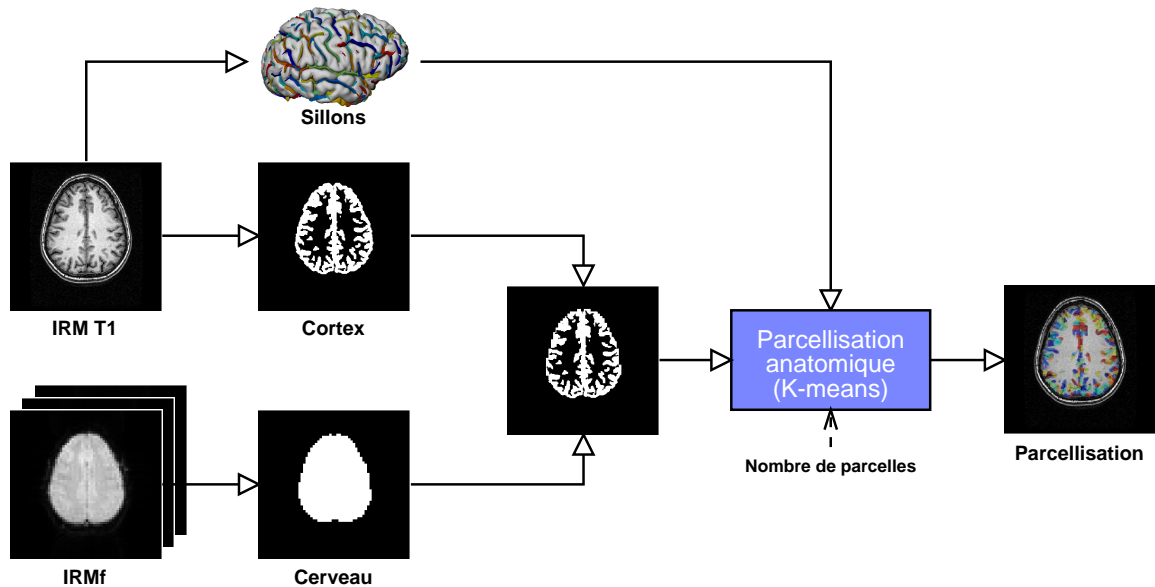


FIG. 5.2 – Diagramme de la construction d'une parcellisation du cortex à partir d'informations anatomiques.

5.2 Segmentation des structures cérébrales

Dans cette section, nous nous proposons de définir le volume d'intérêt à parcelliser. Comme indiqué précédemment, nous nous intéressons spécifiquement au cortex cérébral, contenant les potentielles activations fonctionnelles. Un traitement spécifique des noyaux sous-corticaux reste à envisager.

Nous disposons de deux modalités pour segmenter le cortex: IRM T_1 et T_2^* . Chacune de ces modalités a sa propre résolution spatiale et met en évidence des structures différentes. L'IRM T_1 va nous permettre d'obtenir une segmentation fine, de l'ordre du millimètre

des structures corticales. Il faudra toutefois segmenter le cerveau à partir des IRM fonctionnelles et ne conserver que les voxels anatomiques corticaux appartenant également au masque du cerveau issus des images fonctionnelles. En effet, la présence d'artefacts et de distorsions géométriques dans les IRM T_2^* acquises avec une séquence EPI entraîne l'absence de signal dans des voxels pourtant considérés comme appartenant au cortex.

5.2.1 Segmentation du néocortex à partir d'IRM T_1

La segmentation du cortex dans les IRM T_1 est effectuée grâce à l'utilisation des outils développés par Jean-François Mangin [Mangin *et al.*, 1995] et son équipe, et disponibles dans la plateforme de traitement d'images BrainVISA²⁰ [Cointepas *et al.*, 2001; Rivière *et al.*, 2003]. Les différentes étapes de la chaîne de traitements sont synthétisées dans l'annexe B de la thèse de [Cachia, 2003]. Nous résumons ici les principales étapes.

Correction du biais spatial en intensité

La plupart des images IRM présentent un biais spatial en intensité résultant des hétérogénéités du champ B_0 . Ainsi, selon la position dans l'image, le même tissu (matière grise, matière blanche) peut avoir des intensités différentes. Il est important de corriger ce biais dans les images si les algorithmes appliqués sur celles-ci font l'hypothèse d'une intensité constante par tissu. Les différents algorithmes proposés dans la littérature pour corriger le biais spatial diffèrent par le type de modélisation du champ (paramétrique ou non) et par l'algorithme de minimisation [Prima *et al.*, 2001]. La solution utilisée [Mangin, 2000] recherche à minimiser l'entropie de la distribution des intensités de l'image.

Segmentation du cortex par déformations homotopiques

La segmentation du cerveau repose sur une analyse de l'histogramme de l'image par une méthode espace-échelle pour déterminer les deux pics correspondant à la matière grise et la matière blanche. Ensuite, diverses opérations morphologiques sont effectuées pour séparer les hémisphères et le tronc cérébral.

Pour chaque hémisphère, l'ensemble formé par la matière grise et le liquide céphalo-rachidien est segmenté avec une contrainte topologique préservant l'homotopie du cortex à une sphère [Mangin *et al.*, 1995].

Détection et étiquetage automatique des plis corticaux

Les plis corticaux sont alors extraits du complexe gris/LCR grâce à une squelettisation homotomique. Le squelette obtenu est alors segmenté en surfaces simples, permettant ainsi de distinguer les sillons.

Les surfaces obtenues par la méthode précédente sont ensuite étiquetées automatiquement par l'algorithme de Denis Rivière reposant sur une modélisation markovienne du graphe des étiquettes des sillons corticaux [Rivière, 2000; Rivière *et al.*, 2002]. Ces

20. <http://www.brainvisa.info/>

informations prises en compte pour calculer ces probabilités sont de deux types : internes ou relationnelles.

1. Les informations internes comportent la position du sillon, sa longueur, sa profondeur maximale, ...
2. Les informations relationnelles comportent le nombre de voisins que possède un sillon, la distance envers ces voisins, leurs labels supposés, ...

Un réseau de neurones est d'abord entraîné sur un ensemble de cerveaux étiquetés manuellement. Puis, lorsque un nouveau cerveau est donné en entrée, une technique de recuit simulé est utilisée sur un critère de probabilité *a posteriori* donné par le réseau de neurones. La figure 5.3 présente un résultat d'étiquetage automatique de sillons corticaux, où chaque couleur représente une étiquette.

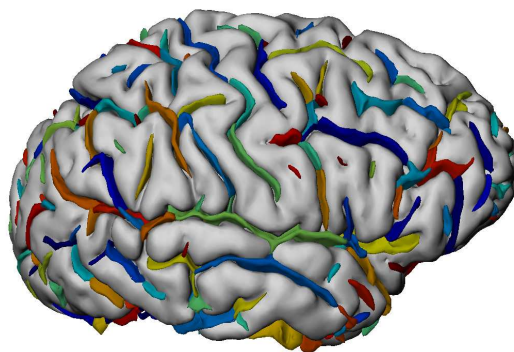


FIG. 5.3 – Labellisation des sillons corticaux par l'algorithme de [Rivière, 2000; Rivière et al., 2002].

5.2.2 Segmentation du cerveau à partir d'IRM T_2^*

Dans une optique de segmentation du cerveau, les images fonctionnelles peuvent être biaisées par les variations hémodynamiques temporelles. Pour s'en affranchir, on se propose d'effectuer la segmentation sur l'image moyenne des images fonctionnelles (le même principe est utilisé pour le recalage T_1 - T_2^* [Ashburner and Friston, 1997], où l'on met en correspondance la moyenne des images T_2^* avec l'IRM T_1).

L'approche proposée dans le logiciel SPM consiste à seuiller cette image moyenne afin de ne conserver que les voxels dont l'intensité est supérieure à 80% de la moyenne globale de l'image.

Nous utilisons également une segmentation par seuillage global mais nous proposons de déterminer le seuil automatiquement à partir de l'histogramme de l'image. Un histogramme typique d'une IRMf est présenté sur la figure 5.4.

La distribution de cet histogramme est clairement bimodale : le modèle choisi pour la distribution des intensités des voxels appartenant au cerveau est de type gaussien. Par contre, la distribution des voxels du fond est clairement non gaussienne. En effet, le bruit dans les images acquises par résonance magnétique peut être modélisé par une distribution

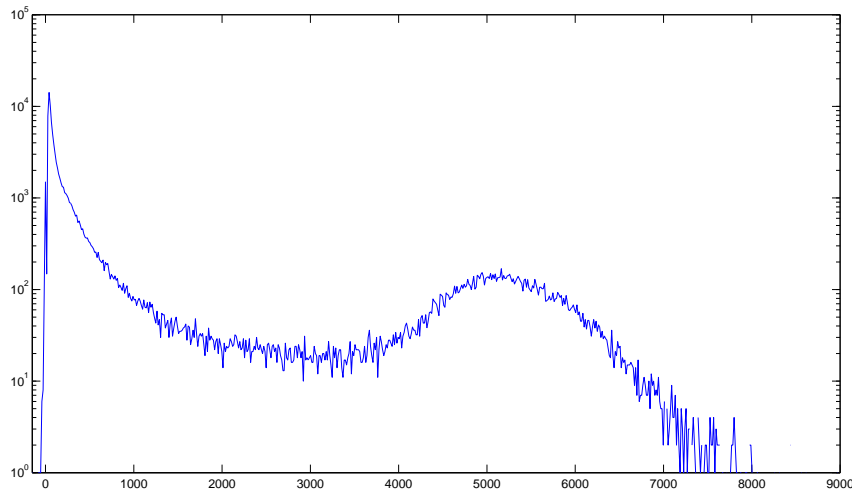


FIG. 5.4 – Histogramme d'une image fonctionnelle, en échelle logarithmique.

ricienne (distribution de l'amplitude de données complexes où chaque composante suit une distribution gaussienne) [Sijbers *et al.*, 1998], en raison de l'acquisition des données dans le plan de Fourier suivi d'une reconstruction. Afin de ne pas avoir à estimer les paramètres d'une telle distribution sur l'histogramme, nous proposons un algorithme en deux étapes, évitant de devoir définir explicitement une distribution paramétrique pour les voxels du fond de l'image:

- séparation des deux modes de l'histogramme,
- estimation des paramètres d'une loi gaussienne avec les données correspondant au mode ayant les intensités les plus élevées.

Considérons l'histogramme défini par les intervalles $X_{0\dots L}$ et leur contenu $N_{0\dots L}$:

La recherche de l'indice du seuil $S \in 0 \dots L$ se fait par le calcul itératif des centroïdes de chaque mode et de la mise à jour du seuil intermodal. L'algorithme est présenté dans l'encadré 1.

algorithme 1 Séparation des modes d'un histogramme

1. Initialisation : $S = 1$
2. Calcul des centroïdes C_1 et C_2 de chaque mode:

$$C_1 = \frac{\sum_{i=0}^S N[i]X[i]}{\sum_{i=0}^S N[i]} \quad \text{et} \quad C_2 = \frac{\sum_{i=S+1}^L N[i]X[i]}{\sum_{i=S+1}^L N[i]}$$

3. Mise à jour de l'indice du seuil :

$$S = \arg \min_i \left| N[i] - \frac{C_1 + C_2}{2} \right|$$

4. Répéter les étapes 2 et 3 tant que S évolue
-

En pratique, la convergence est rapide : quelques itérations suffisent malgré l'initialisation très éloignée de l'optimum. La figure 5.5 présente le résultat de la séparation en 2 modes de l'histogramme présenté à la figure 5.4.

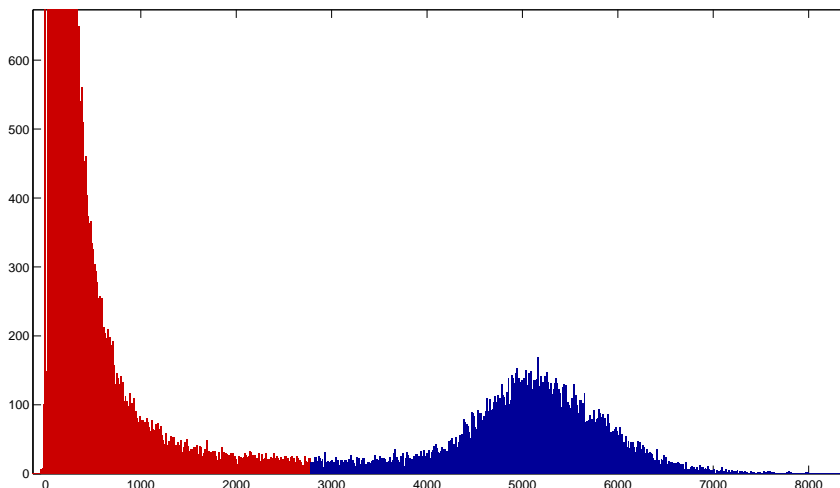


FIG. 5.5 – Séparation des deux modes de l'histogramme d'une image T_2^* .

L'étape suivante consiste à ajuster les paramètres d'une fonction gaussienne sur la partie de l'histogramme correspondant au second mode. Le modèle que nous proposons tend à prendre en compte l'effet de volume partiel intervenant dans l'image (du à sa faible résolution) et se manifestant dans l'histogramme par une distribution uniforme entre les deux modes.

Nous proposons deux fonctions d'ajustement f_i , la première modélisant le volume partiel par une composante constante uniforme sur tout l'histogramme et une seconde le modélisant par une autre gaussienne dont la covariance est contrainte à être élevée.

$$f_1(X) = x_0 + x_1 \times \exp\left(-\frac{(X - x_2)^2}{2x_3^2}\right)$$

$$f_2(X) = x_0 \times \exp\left(-\frac{(X - x_1)^2}{2x_2^2}\right) + x_3 \times \exp\left(-\frac{(X - x_4)^2}{2x_5^2}\right)$$

L'estimation des paramètres x_i est faite au sens des moindres carrés en utilisant l'algorithme de Levenberg et Marquardt [Marquardt, 1963], bien adapté à ce type de minimisation et correspondant à une descente de gradient à pas adaptatif. Les résultats d'ajustement avec f_1 et f_2 sur le deuxième mode de la figure 5.5 est présenté sur la figure 5.6.

La modélisation du volume partiel par une gaussienne apparaît plus pertinente ici. En effet, ajouter une distribution uniforme sur tout l'histogramme permet bien de modéliser l'effet de volume partiel mais biaise également l'estimation des paramètres de la gaussienne pour les voxels d'intensité élevée.

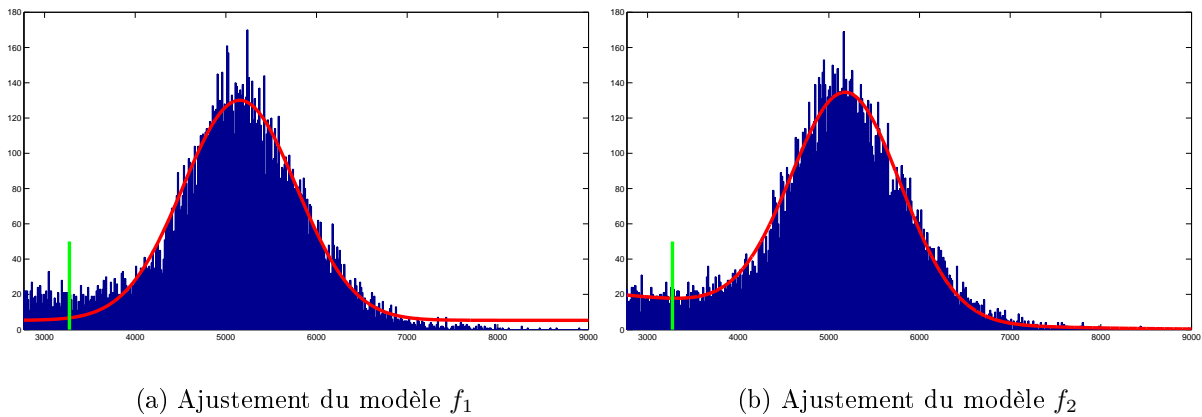


FIG. 5.6 – Ajustement des fonctions f_1 et f_2 sur l'histogramme.

Une fois la gaussienne ajustée, le seuil à appliquer à l'image est classiquement défini comme étant situé à 3σ de la moyenne de la gaussienne. Ce seuil est indiqué sur la figure 5.6 par une segment de droite verte. On voit ainsi que le seuil est sensiblement le même quel que soit le modèle choisi.

Un exemple de segmentation obtenue par la méthode décrite ci-dessus est disponible sur la figure 5.7. Il est à noter qu'aucune opération morphologique n'est appliquée en sus pour combler les trous par exemple, car ce sont des voxels ne contenant pas d'information pertinente sur le signal, qui biaiserait les estimations à venir s'ils étaient pris en compte.

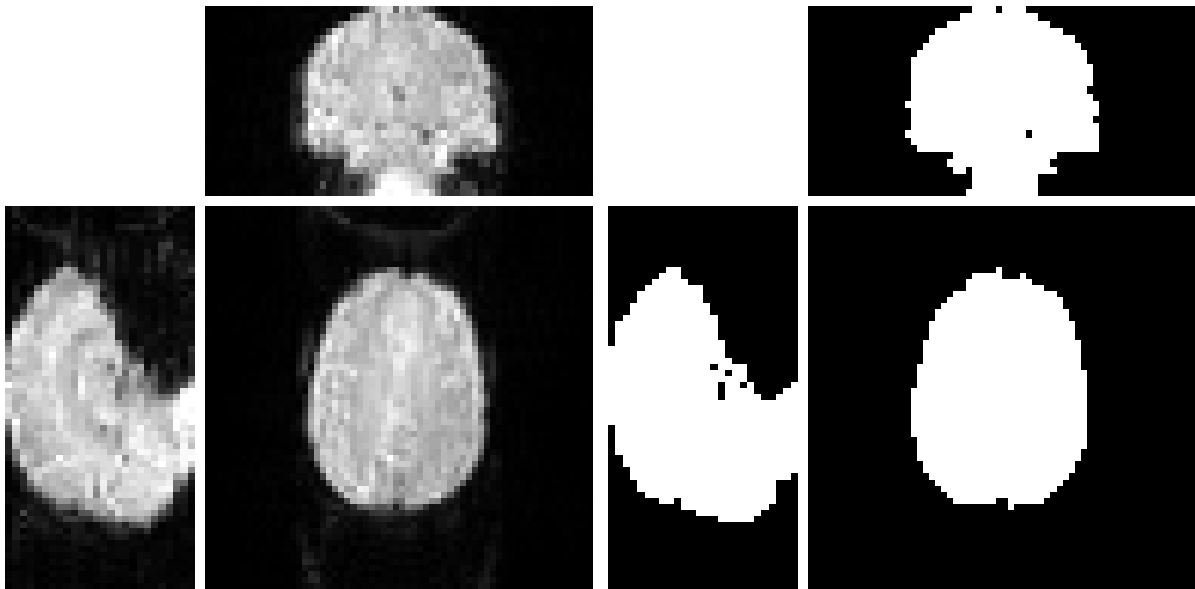


FIG. 5.7 – Image fonctionnelle moyenne (à gauche) et le masque du cerveau obtenu par seuillage automatique (à droite).

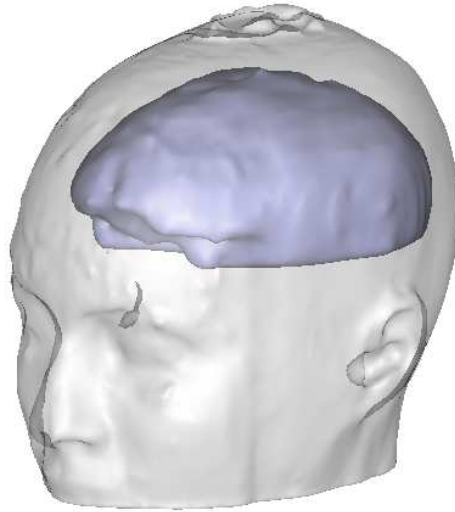


FIG. 5.8 – Vue tridimensionnelle d'un cerveau segmenté à partir d'images fonctionnelles.

5.2.3 Obtention du domaine à parcelliser

Enfin, le domaine à parcelliser est obtenu en effectuant un ET logique entre les deux masques obtenus précédemment. En particulier, le masque fonctionnel est rééchantillonné à la résolution anatomique avec un interpolateur plus-proche-voisin.

La figure 5.9 présente un exemple de domaine obtenu. On note en particulier l'ensemble de voxels anatomiques supprimés dans les lobes frontaux et temporaux car non présents dans le masque fonctionnel, vraisemblablement dûs à des artefacts dans les IRM T_2^*

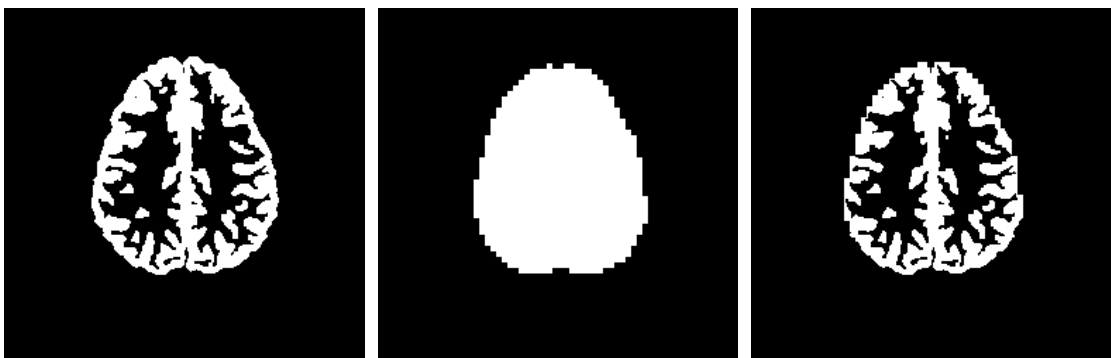


FIG. 5.9 – Coupe axiale de la segmentation du neocortex dans l'IRM anatomique (gauche), segmentation du cerveau dans la moyenne des images fonctionnelles (milieu) et domaine résultant (gauche).

5.3 Parcellisation d'un domaine convexe

Dans un premier temps, considérons le domaine à parcelliser comme étant convexe ; cette hypothèse est bien sûr non vérifiée dans notre cas comme en témoigne la figure 5.9.c, cependant cette hypothèse nous permet de mettre en place le cadre théorique et algorithmique que nous adapterons par la suite dans le cas de domaines non-convexes.

Notre objectif dans cette section est donc d'effectuer une parcellisation (un découpage) d'un domaine discret convexe en un nombre K , connu, de cellules. Afin d'obtenir une représentation *homogène* du domaine, les cellules doivent avoir des volumes comparables. Cependant, le critère de volume n'est pas suffisant pour caractériser une parcellisation car il y a de nombreuses configurations non souhaitées qui peuvent apparaître. La figure 5.10 présente trois parcellisations différentes d'un domaine carré bidimensionnel en 16 cellules. Les trois fournissent des cellules ayant la même aire ($\frac{1}{16}$ de l'aire du domaine) et la même géométrie : ce qui les distingue est leur compacité et ici c'est la parcellisation la plus compacte (à gauche) qui fournit la représentation la plus *homogène*, de notre point de vue.

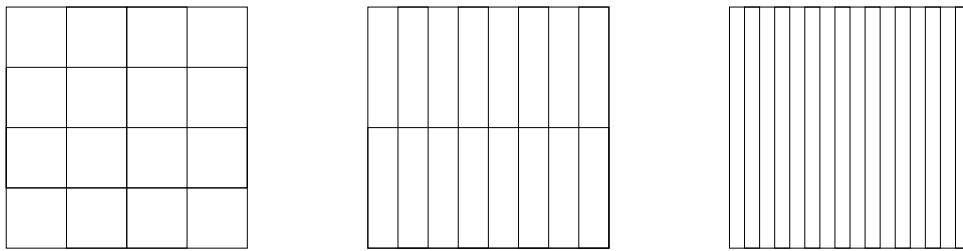


FIG. 5.10 – Exemple de différentes parcellisations d'un domaine carré en 16 éléments de même volume.

Un meilleur critère permettant de définir cette notion d'homogénéité est ainsi celui de la compacité : l'objectif est d'obtenir des cellules les plus compactes possibles au sein du domaine. On montrera par la suite que pour un domaine convexe, ce critère tend également à obtenir des cellules de même volume.

5.3.1 Parcellisation et critère à optimiser

Il existe différentes catégories de parcellisation, emboîtées les unes dans les autres constituant une hiérarchie : parcellisation quelconque, parcellisation de Voronoï et parcellisation de Voronoï Centroïdale.

Parcellisation

Tout d'abord, précisons le terme *parcellisation*²¹ dans un cadre général :

Définition 5.1 *Etant donné un ensemble fini $D = \{x_i\}_{i=1}^N$, on appelle parcellisation de*

21. Nous utilisons le terme *parcellisation* pour définir une *partition* de l'espace. Les termes anglophones correspondant sont *parcellation* ou *tesselation*.

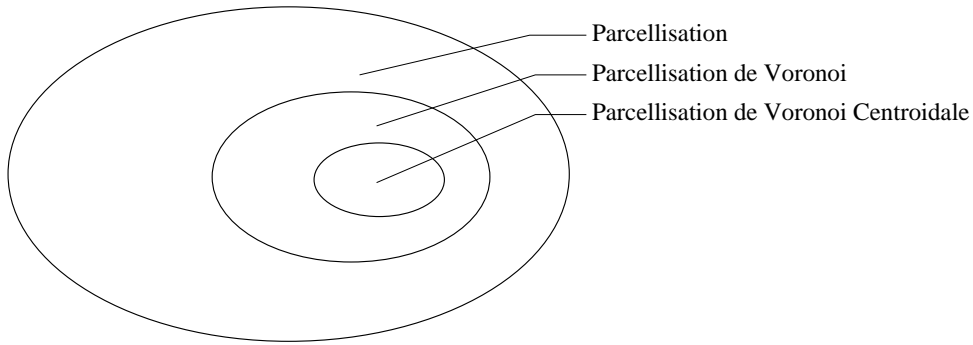


FIG. 5.11 – *Les différentes catégories de parcellisation: les parcellisations de Voronoï Centroidales sont un sous-ensemble des parcellisations de Voronoï, elles-mêmes étant un cas particulier de parcellisation.*

D en K éléments, tout jeu de sous-ensembles D_1, D_2, \dots, D_K tels que:

- aucun membre d'un sous-ensemble D_i est membre d'un sous-ensemble D_j ,
- tous les membres de D appartiennent à un des sous-ensembles D_i .

On a ainsi:

$$\{D_1, D_2, \dots, D_K\} \text{ parcellisation de } D \iff D = \bigcup_{i=1}^K D_i \text{ et } D_i \cap D_j = \emptyset, \forall i \neq j$$

Il est à noter que, d'après cette définition, les parcelles D_i sont quelconques et ne sont pas contraintes à être connexes par exemple.

Parcellisation de Voronoï

Une parcellisation de Voronoï de D est un cas particulier d'une parcellisation définie précédemment où chaque parcelle V_k peut être définie par un générateur (une graine) \overline{x}_k de telle sorte que les éléments de V_k sont plus proche de \overline{x}_k que de toute autre graine $\overline{x}_{k'}$.

Définition 5.2 *Etant donné un ensemble fini $D = \{x_i\}_{i=1}^N$, K graines $\{\overline{x}_k\}_{k=1}^K \in D$ et une fonction de distance d , on appelle parcellisation de Voronoï de D en K éléments, la parcellisation définie par les sous-ensembles $\{V_k\}_{k=1}^K$ tels que:*

$$V_k = \{x \in D \mid d(x, \overline{x}_k) < d(x, \overline{x}_l), \forall k \neq l\} \quad (5.1)$$

Effectuer une parcellisation d'un domaine D associé à une distance d en K parcelles revient donc à définir K graines \overline{x}_k , associées à K régions V_k selon l'équation 5.1.

On montre alors que les parcelles associées à un telle parcellisation sont contraintes à être connexes et convexes, par construction avec une distance d euclidienne.

Parcellisation de Voronoï Centroïdale

Il existe enfin un cas particulier de parcellisation de Voronoï correspondant au cas où les graines \bar{x}_k se trouvent au barycentre de chaque parcelle ; on parle alors de parcellisation de Voronoï centroïdale (PVC).

Définition 5.3 *Etant donné un ensemble fini $D = \{x_i\}_{i=1}^N$, K graines $\{\bar{x}_k\}_{k=1}^K \in D$ et une fonction de distance d , $\{V_k\}_{k=1}^K$ est une parcellisation de Voronoï centroïdale si :*

- les régions $\{V_k\}_{k=1}^K$ sont des régions de Voronoï pour les graines $\{\bar{x}_k\}_{k=1}^K$,
- les graines $\{\bar{x}_k\}_{k=1}^K$ sont les barycentres des régions $\{V_k\}_{k=1}^K$

La figure 5.12 présente deux exemples de parcellisation de Voronoï d'un domaine bidimensionnel carré en 500 éléments, l'une quelconque et l'autre centroïdale.

En pratique, une parcellisation de Voronoï centroïdale est obtenue par construction en utilisant une propriété de minimisation.

Proposition 5.1 *Etant donné un domaine discret $D = \{x_i\}_{i=1}^N$ et un entier $K > 1$, soit $\{\bar{x}_k\}_{k=1}^K$ un ensemble de K points de D et $\{V_k\}_{k=1}^K$ une parcellisation du domaine D en K éléments. Soit le critère suivant :*

$$C((\bar{x}_k, V_k), k = 1 \dots K) = \sum_{i=1}^K \sum_{x \in V_i} d(x, \bar{x}_i)^2$$

Une condition nécessaire pour que C soit minimal est que $\{\bar{x}_k\}_{k=1}^K$ et $\{V_k\}_{k=1}^K$ forment une parcellisation de Voronoï centroïdale du domaine D .

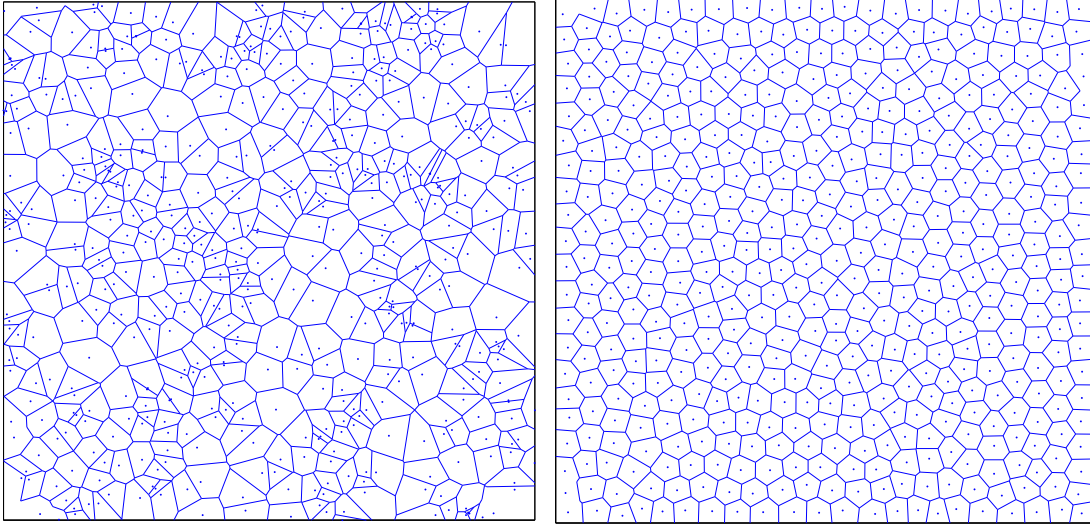
La démonstration de cette propriété est disponible dans [Du et al., 1999]. C'est la minimisation de ce critère qui va être utilisée pour construire une parcellisation de Voronoï centroïdale d'un domaine discret. L'algorithme permettant de minimiser ce critère est l'algorithme K-Means, que nous avons présenté au chapitre 4, parmi les algorithmes de *clustering*.

5.3.2 L'algorithme K-Means

Selon le domaine d'application, l'algorithme K-Means [Hartigan and Wong, 1979] peut être appelé de différentes manières : méthode de Lloyd, algorithme LBG, moyennes mobiles, nuées dynamiques, algorithme ISODATA, ... On rappelle que cet algorithme est en fait un CEM (Classification EM) pour un mélange de gaussiennes isotropes et de même covariance et avec des pondérations identiques (cf paragraphe 4.2.4)

La démonstration de la convergence de cet algorithme se fait en montrant qu'à chaque itération de l'algorithme, le critère C (inertie intraclasse, égale à la somme des variances de chaque classe) diminue, et puisqu'il y a un nombre fini de classifications possibles, on atteint forcément un point de convergence.

De plus, comme indiqué dans le chapitre précédent, le K-Means est en fait un cas particulier d'un Classification-EM, pour lequel la convergence est aussi assurée.



(a)

(b)

FIG. 5.12 – Exemples de parcellisation de Voronoï, (a) quelconque, (b) centroïdale

algorithme 2 Algorithme K-Means

 Initialisation du nombre de classe K et de leur centre associé $\{\bar{x}_k\}_{k=1}^K$
répéterClassifier les échantillons selon le centre \bar{x}_k le plus procheMettre à jour les $\{\bar{x}_k\}_{k=1}^K$ selon la classification**jusqu'à** aucun changement pour les $\{\bar{x}_k\}_{k=1}^K$

 Enfin, on peut également voir l'algorithme K-Means comme une minimisation alternée des deux paramètres \bar{x}_k et V_k du critère C . En effet, étant donné une parcellisation $\{V_k\}$, la mise à jour des paramètres $\{\bar{x}_k\}$ minimisant le critère C est donnée par :

$$\frac{\partial C}{\partial \bar{x}_k} = -2 \sum_{x \in V_k} (x - \bar{x}_k) = 0$$

ce qui donne

$$\bar{x}_k = \frac{1}{\text{Card}(V_k)} \sum_{x \in V_k} x, \quad \forall k$$

De même, étant donné la position des graines $\{\bar{x}_k\}$ constante, soit $y \in D$. Supposons que le critère soit minimisé lorsque y appartient à la cellule V_k . On en déduit immédiatement que le critère est plus faible que celui obtenu pour y appartenant à $V_{k'}$ quel que soit

$k' \neq k$:

$$\sum_{i=1}^K \sum_{x \in V_i} d^2(x, \bar{x}_i) \leq \sum_{i=1}^K \sum_{x \in V_i} d^2(x, \bar{x}_i)$$

avec $y \in V_k$ avec $y \in V_{k'}, \forall k' \neq k$

d'où on obtient, en éliminant les termes identiques de part et d'autre de l'inégalité :

$$d^2(y, \bar{x}_k) \leq d^2(y, \bar{x}_{k'}), \quad \forall k' \neq k$$

Ainsi, y est attribué à la cellule dont la graine est la plus proche. Ceci est valable pour tout y donc la minimisation du critère par rapport à $\{V_k\}$ revient à effectuer un diagramme de Voronoï.

Il est également intéressant de noter que dans le cas d'un domaine convexe (utilisation d'une distance d euclidienne), il y a équivalence entre minimiser l'inertie intraclasse C et maximiser l'inertie interclasse, grâce au théorème de Huygens.

$$\begin{aligned} I_{tot} &= \sum_{i=1}^N d^2(x_i, \bar{x}) \\ &= \sum_{K=1}^K \sum_{i \in V_k} d^2(x_i, \bar{x}) \\ &= \sum_{k=1}^K \left[\sum_{i \in V_k} d^2(x_i, \bar{x}_k) + d^2(\bar{x}, \bar{x}_k) \right] \quad \text{par le théorème de Huygens} \\ &= \underbrace{\sum_{k=1}^K \sum_{i \in V_k} d^2(x_i, \bar{x}_k)}_{\text{inertie intraclasse}} + \underbrace{\sum_{k=1}^K d^2(\bar{x}, \bar{x}_k)}_{\text{inertie interclasse}} \end{aligned}$$

L'inertie totale, ne dépendant que des données, est constante (indépendante de la partition). Il apparait donc que l'algorithme K-Means tend à obtenir des cellules les plus compactes possibles (minimisation de l'inertie intraclasse) mais également à maximiser les distances entre les centres des cellules (maximisation de l'inertie interclasse).

Dans le cas d'un domaine discret dont les éléments sont régulièrement répartis sur une grille, chacune des étapes de l'algorithme K-Means peut être reformulée dans le cadre de la géométrie discrète, bénéficiant ainsi d'algorithmes adaptés performants.

5.3.3 Transformations de distance

Les transformations de distance permettent de calculer des cartes de distance : à partir d'une image binaire (fond et objet), ces algorithmes permettent de calculer une image flottante où la valeur de chaque voxel de l'objet est égale à sa distance la plus courte du fond. Ces transformations ont de nombreuses applications en traitement d'images (construction

de diagrammes de Voronoï, calcul de squelettes, etc) et ont donc suscité l'attention de la communauté scientifique à partir des années 1980 pour concevoir des algorithmes, exacts ou approchés, de calcul de cartes de distance [Danielsson, 1980; Borgefors, 1984; 1986; Thiel, 1994; Cuisenaire, 1999]. En effet, l'algorithme de type "force brute" est rapidement très coûteux en temps de calcul.

Définition 5.4 *Soit une image binaire I contenant un objet O (label 1). Pour une distance d donnée, la transformée en distance D de I est définie en tout point par:*

$$D(p) = \min \{d(p,q), q \in O\} \quad (5.2)$$

En général, d est la distance euclidienne

$$d_E(p,q) = \left(\sum_i (p_i - q_i)^2 \right)^{1/2}$$

cependant toute application $d : \mathbb{Z}^n \rightarrow \mathbb{R}^+$ vérifiant les axiomes de symétrie et de définition ainsi que l'inégalité triangulaire peuvent être utilisées pour définir une transformation de distance.

Un bon compromis entre précision et coût calculatoire est obtenu grâce aux masques de chanfrein, popularisés par Borgefors. L'algorithme consiste à propager localement des distances entières à partir de masques dits de *chanfrein*. En fait, les masques définissent un ensemble de déplacements élémentaires associés à des coûts de déplacement. Une implémentation naïve, parallèle, fonctionne ainsi : le masque de chanfrein est appliqué itérativement sur chaque voxel de l'image et la nouvelle valeur du voxel est obtenue en calculant la somme des poids du masque avec la valeur des voxels associés et en en conservant la valeur minimale. Il suffit ensuite d'appliquer successivement cette opération sur l'image créée jusqu'à ce qu'aucune valeur n'évolue plus.

Heureusement, il existe une version séquentielle qui ne nécessite que deux balayages de l'image, en séparant le masque en deux et en appliquant le premier dans le sens où les coordonnées augmentent et le second en sens inverse, et ceci quelle que soit l'image et la complexité de l'objet qu'elle contient. On trouvera une démonstration de l'équivalence entre la version séquentielle et parallèle dans [Mangin, 1995].

La construction du diagramme de Voronoï peut se faire en même temps que le calcul de la carte de distance [Bertin, 1994; Cuisenaire, 1999]. En effet, il suffit de propager les labels associés aux voxels d'où proviennent les distance les plus faibles.

5.3.4 Mise à jour des graines

Dans le cas euclidien (domaine convexe), la mise à jour de la position des graines connaissant le diagramme de Voronoï associé revient simplement à calculer le barycentre de chaque cellule:

$$\bar{x}_k = \frac{1}{\text{Card}(V_k)} \sum_{i \in V_k} x_i \quad \forall k \quad (5.3)$$

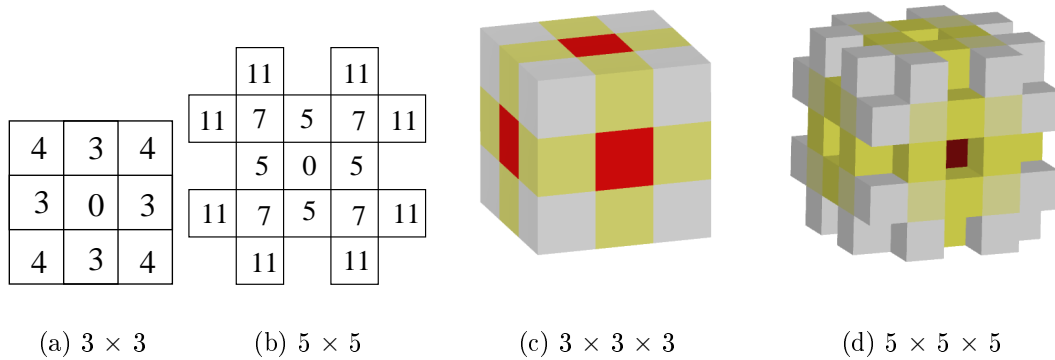


FIG. 5.13 – (a, b) masques de chanfrein 2D, (c, d) masques de chanfrein 3D

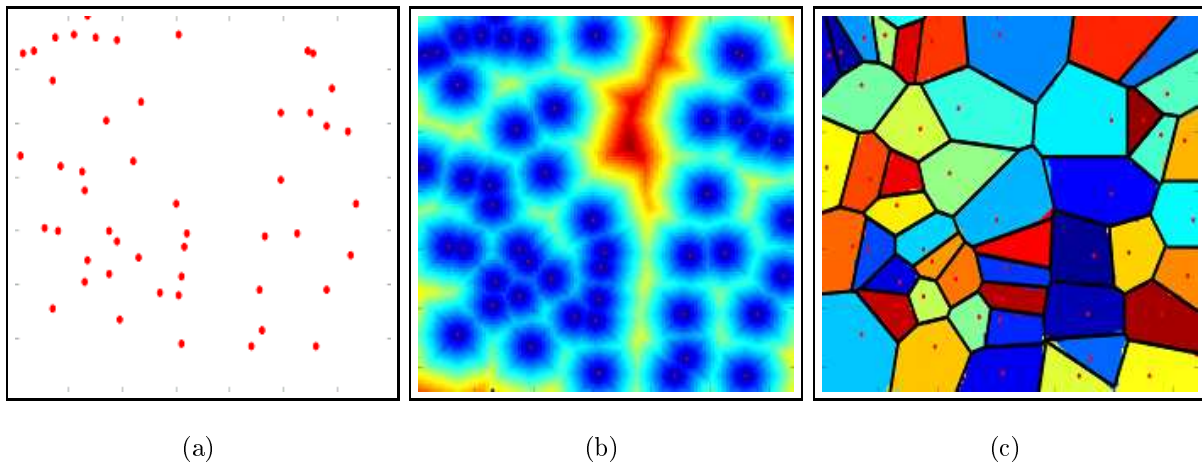


FIG. 5.14 – Exemple de carte de distance bidimensionnelle (b) calculée à partir de 50 pixels (a). Le diagramme de Voronoï associé est représenté en (c).

En pratique, les positions des graines sont définies de manière à se trouver sur un des noeuds de la grille afin de pouvoir calculer le diagramme de Voronoï discret à l'itération suivante. Les centres de classe \bar{x}_k sont donc affectés à leur position sur la grille la plus proche.

5.3.5 Résultats obtenus

Afin de montrer le comportement de l'algorithme, nous présentons sur la figure 5.15 la parcellisation d'un domaine convexe 512×512 en 250 éléments. L'image (a) montre l'initialisation avec le diagramme de Voronoï construit à partir de graines tirées aléatoirement uniformément sur le domaine (tirage uniforme sans remise). L'image (b) montre le diagramme de Voronoï obtenu après convergence de l'algorithme K-Means (40 itérations).

Les couleurs attribuées à chaque classe sont aléatoires et ne servent que dans un but de visualisation des classe²².

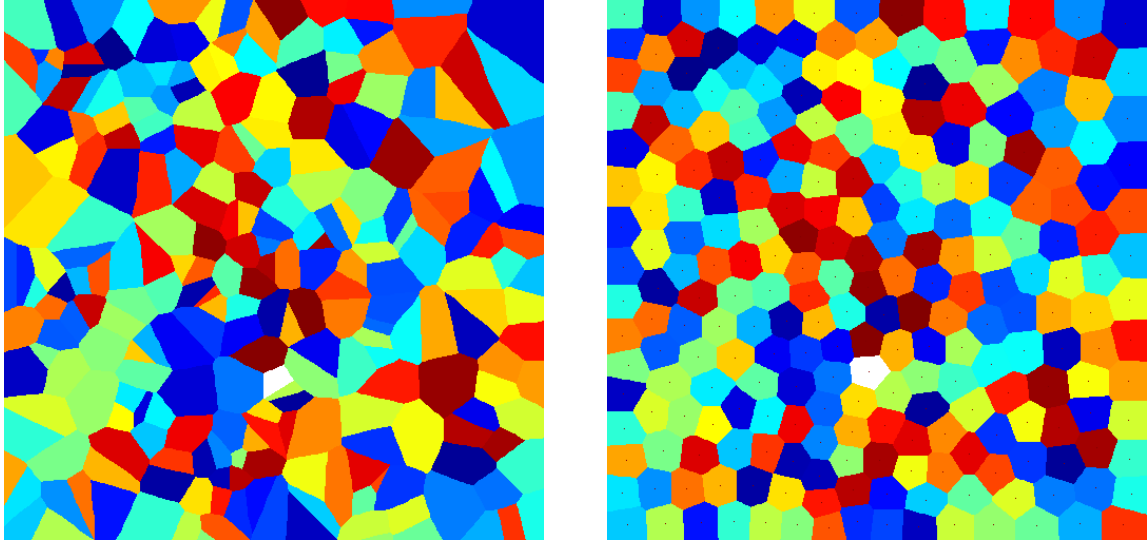


FIG. 5.15 – *Parcellisation de Voronoï centroïdale d'un domaine convexe 512×512 en 250 classes. L'image de gauche présente la parcellisation de Voronoï calculé à partir de graines initialisées aléatoirement dans le domaine tandis que l'image de droite présente la parcellisation de Voronoï centroïdale obtenue après convergence de l'algorithme K-Means.*

On observe que la parcellisation obtenue est beaucoup plus régulière que celle obtenue à l'initialisation. En particulier, on note bien que chaque cellule est beaucoup plus compacte qu'initialement et que les graines se situent bien au barycentre de chaque cellule. Enfin, on vérifie que les centres de chaque cellule sont bien les plus éloignés possible des centres des cellules voisines.

La figure 5.16 présente l'évolution des volumes des parcelles avant et après application de l'algorithme K-Means. On note que les parcelles ont des volumes équivalents ($\frac{512 \times 512}{250} \approx 1050$) après convergence. La variance résiduelle sur les volumes s'explique par le fait, d'une part que les cartes de distances intervenant dans la construction des diagrammes de Voronoï sont approchées et non exactes et d'autre part que la géométrie du domaine (ici de forme carrée) va conditionner la forme des cellules situées aux bords de l'image.

Enfin, la figure 5.17 montre l'évolution de la parcellisation au cours des itérations de l'algorithme K-Means pour la partition d'un domaine carré 128×128 en 9 éléments.

De la parcellisation du cortex à la structure "nids d'abeille"

Il est intéressant de noter la forme particulière des cellules obtenues en 2D avec les diagrammes de Voronoï centroïdaux (cf figure 5.15). Celles-ci ont en effet des formes

²². Dans le cas bidimensionnel, le théorème dit des quatre couleurs garantit que quatre couleurs distinctes suffisent pour représenter un pavage de l'espace, tel qu'une mappemonde.

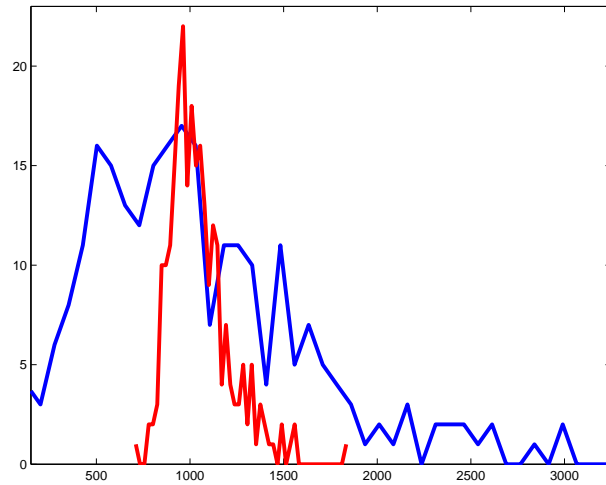


FIG. 5.16 – Histogramme des volumes des 250 cellules de la parcellisation présentée sur la figure 5.15. La courbe bleue correspond aux volumes calculés à l’initialisation tandis que la courbe rouge correspond aux volumes des cellules après convergence.

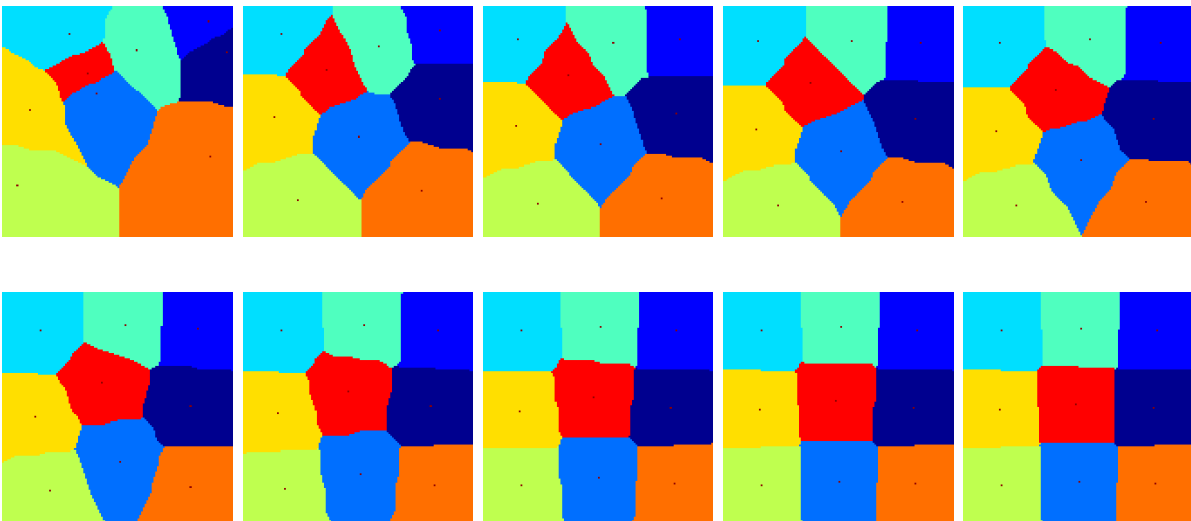


FIG. 5.17 – Evolution de l’algorithme *K-Means* pour la parcellisation d’un domaine 2D carré en 9 éléments. D’en haut à gauche à en bas à droite : itérations 1, 4, 7, 10, 13, 16, 19, 22, 25, 29 (convergence).

voisines à celle d’un hexagone régulier. Ceci n’est en aucun cas lié au hasard et fait référence à la *conjecture des nids d’abeille* qui prétend que toute partition d’un plan en régions d’aires égales a un périmètre au moins égal à celui obtenu par un pavage constitué d’hexagones réguliers disposés en nids d’abeille. Cette conjecture a été récemment démontrée par Thomas Hales [Hales, 2000; 2001], après avoir intéressé les mathématiciens pendant des siècles : Pappus d’Alexandrie évoquait le problème en parlant d’isopérimètre

dans son cinquième livre en 180 avant J.C.

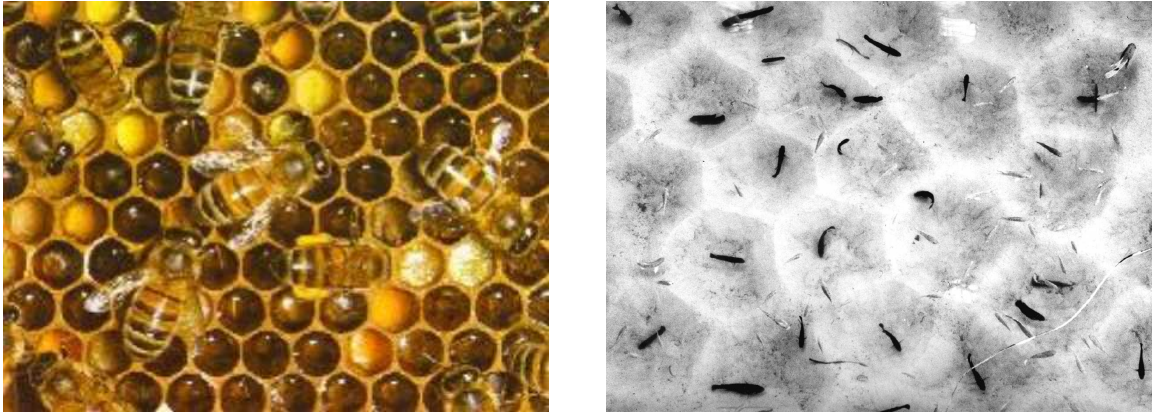


FIG. 5.18 – (a) Structure en nid d'abeille. (b) Extrait de G. W. Barlow, Hexagonal Territories, *Animal Behavior*, Volume 22, 1974, Academic Press.

Au delà de l'inspiration divine des abeilles longtemps évoquée, cette structure permet aux abeilles d'utiliser le moins de cire possible pour construire le logement des larves, tout en gardant un volume égal pour chaque larve.

La figure 5.18.b présente une autre configuration naturelle faisant intervenir des diagrammes de Voronoï centroïdaux: le comportement territorial des animaux. C'est une photographie vue de dessus des territoires du poisson *Tilapia mossambica*: les mâles de cette espèce creusent des trous dans le sable et rejettent le sable du centre de leur territoire vers ses voisins. S'il y a une densité suffisante de poissons, ce comportement fait apparaître des parapets de sable qui sont alors des délimitations visibles de leur territoire. Des ethnologues ont montré que ces territoires pouvaient être approximés par un diagramme de Voronoï. Un modèle comportemental décrivant comment les poissons établissaient leur territoire a même été proposé. Quand les poissons entrent dans une région, ils sélectionnent aléatoirement le centre de leur territoire. Leur désir de constamment centrer leur espace le plus loin de leurs voisins les pousse à modifier leur position pour rester en permanence au centre, ce qui redéfinit en temps réel les frontières de leur domaine. Les poissons pouvant être tous considérés de force égale, ce processus d'ajustement aboutit à la construction d'un diagramme de Voronoï centroïdal.

De nombreux autres exemples peuvent être trouvés dans la nature: carapace d'une tortue, cou d'une girafe réticulée, etc.

5.4 Prise en compte de distances géodesiques

Considérons maintenant un domaine pas forcément convexe, tout en restant connexe, ce qui correspond au cas réel que nous avons à traiter. Contrairement au cas précédent, on ne peut plus utiliser la distance euclidienne d_E dans tous les termes faisant intervenir des

distances entre éléments de l'espace à classifier. La figure 5.19 met en évidence ce principe sur le maillage triangulé de l'interface gris/blanc d'un hémisphère du cerveau. Étant donné un nœud du maillage, nous avons calculé la distance entre ce nœud et les nœuds voisins en utilisant une distance euclidienne (figure 5.19.a). Sur le même maillage déplié, on observe que le domaine associé est constitué de deux morceaux. En effet, la distance euclidienne ne respecte pas la topologie du maillage: deux nœuds du maillage proches de manière euclidienne (distance associée à l'espace dans lequel se trouve le maillage) peuvent se retrouver en réalité éloignés dès lorsque l'on raisonne avec la distance intrinsèque associée au maillage, i.e. la distance géodésique. En utilisant une telle distance (calculée par l'algorithme de Dijkstra, issu de la théorie des graphes et utilisé pour déterminer le plus court chemin entre deux points sur un graphe valué), on obtient la figure 5.19.c et sa version dépliée 5.19.d. On vérifie qu'ainsi les distances sont calculées à partir de la longueur de chemins restreints à rester sur la surface: le domaine obtenu se rapproche beaucoup plus du disque théorique, les fluctuations résiduelles de l'isocontour sont dues à l'approximation des calculs de distances géodésiques par l'algorithme de Dijkstra et aux déformations géométriques créées par le dépliement de la surface.

Dans cette section, nous considérons l'utilisation de distances géodésiques dans le critère à minimiser:

$$C((\bar{x}_k, V_k), k = 1 \dots K) = \sum_{i=1}^K \sum_{x \in V_i} d_{geo}(x, \bar{x}_i)^2$$

Regardons comment adapter les deux étapes de l'algorithme K-Means présenté précédemment pour minimiser ce critère.

5.4.1 Diagramme de Voronoï géodésique

La première étape consiste à attribuer chacun des éléments du domaine à la classe dont le représentant est le plus proche: ceci correspond à construire un diagramme de Voronoï géodésique. Pour cela, définissons la notion de transformée en distance géodésique:

Définition 5.5 Soit une image binaire I contenant un objet O (label 1). La transformée en distance géodésique D_X de I restreinte au domaine X est définie en tout point par:

$$D_X(p) = \min \{d_X(p, q), q \in O\} \quad (5.4)$$

où $d_X(p, q)$ est l'infimum de la longueur des chemins, s'il en existe, allant de p à q en restant inclus dans X et $+\text{Inf}$ sinon.

Piper et Granum [Piper and Granum, 1987] ont étudié le calcul de cartes de distance géodésique dans un domaine non-convexe et ont tout d'abord proposé d'appliquer itérativement l'algorithme en deux passes utilisant les masques de chanfrein jusqu'à ce que plus aucune valeur de voxel n'évolue. Malheureusement le nombre de passes à appliquer est inconnu au début du processus et dépend de la configuration du domaine à considérer. Cet algorithme perd ainsi l'essentiel des avantages qu'il pouvait apporter dans le cas de distances euclidiennes.

Ces auteurs ont alors proposé une autre méthode reposant sur la notion de propagation: l'algorithme est initialisé sur les voxels constituant l'objet et les distances sont

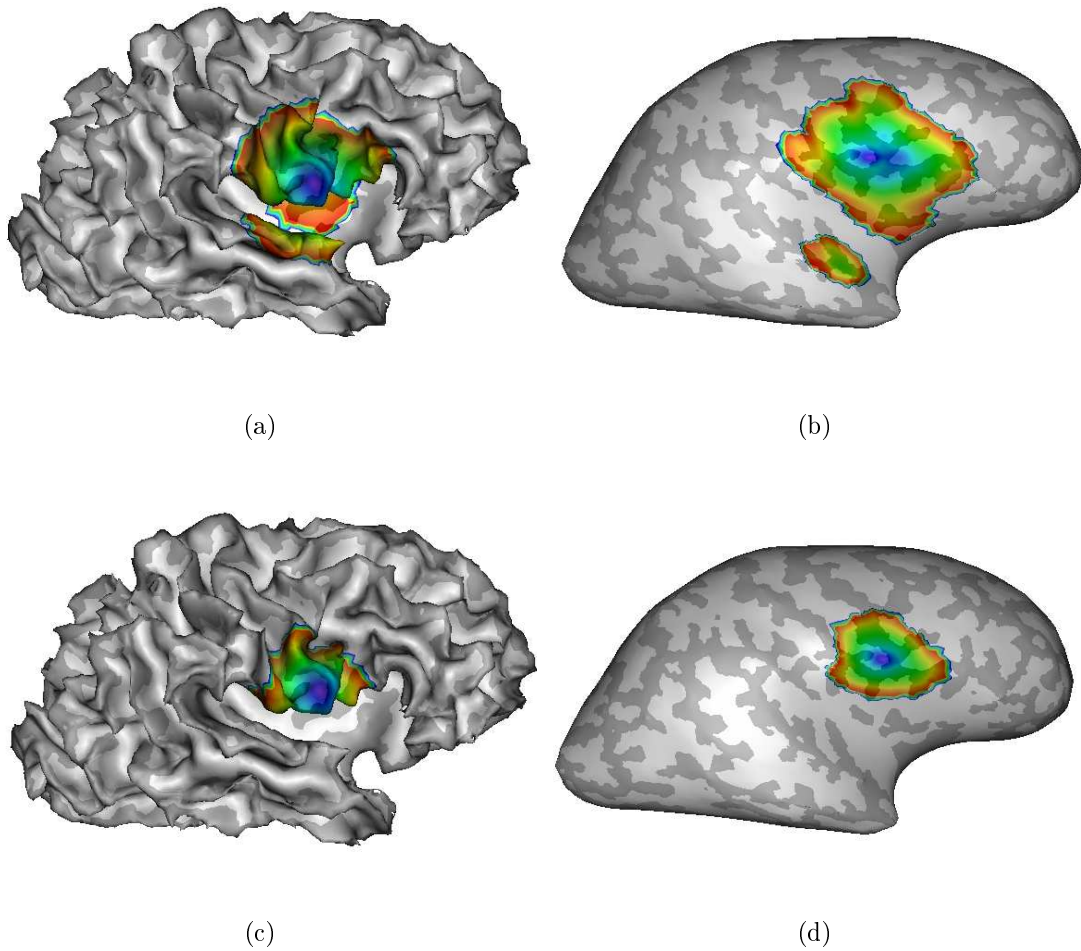


FIG. 5.19 – Comparaison entre distance euclidienne (a et b) et distance géodésique (c et d) sur un maillage triangulé de la surface corticale (a et c) dépliée (b et d).

propagées de proche en proche en considérant les voisins de ces voxels, puis les voisins des voisins, etc. Toutefois, l'algorithme proposé n'est optimal que dans le cas de l'utilisation d'une distance de type 4-connexe. Cette idée de propagation a toutefois été reprise par la suite dans [Verwer *et al.*, 1989] pour proposer un algorithme efficace par propagation ordonnée (reposant sur l'algorithme A^*): à chaque itération de l'algorithme, les voxels sont naturellement triés selon les valeurs croissantes de distance et seuls les voxels ayant pour valeur la distance la plus faible sont pris en compte à une itération donnée. Il s'agit en fait d'un algorithme par propagation de front. Grâce à cette subtilité, la complexité calculatoire reste faible et est proportionnelle au nombre de voxels à traiter (indépendamment de la géométrie du domaine considéré).

Cuisenaire [Cuisenaire, 1997] a repris et détaillé cet algorithme, dont nous présentons ici l'implémentation. Afin de ne propager à chaque itération que les voxels constituant le front, il est nécessaire pour cela de trier tous les voxels selon la métrique utilisée. Pour ce faire, les voxels sont ordonnés dans une liste de piles représentée sur la figure 5.20.

Comme dans le cas convexe, les distance géodésiques utilisées reposent sur la définition d'un ensemble de distances associées à des déplacements élémentaires, c'est-à-dire un masque de chanfrein. Il y a donc un nombre fini de distances possibles à considérer pour l'ensemble des voxels et on peut associer une pile à chacune de ces distances.

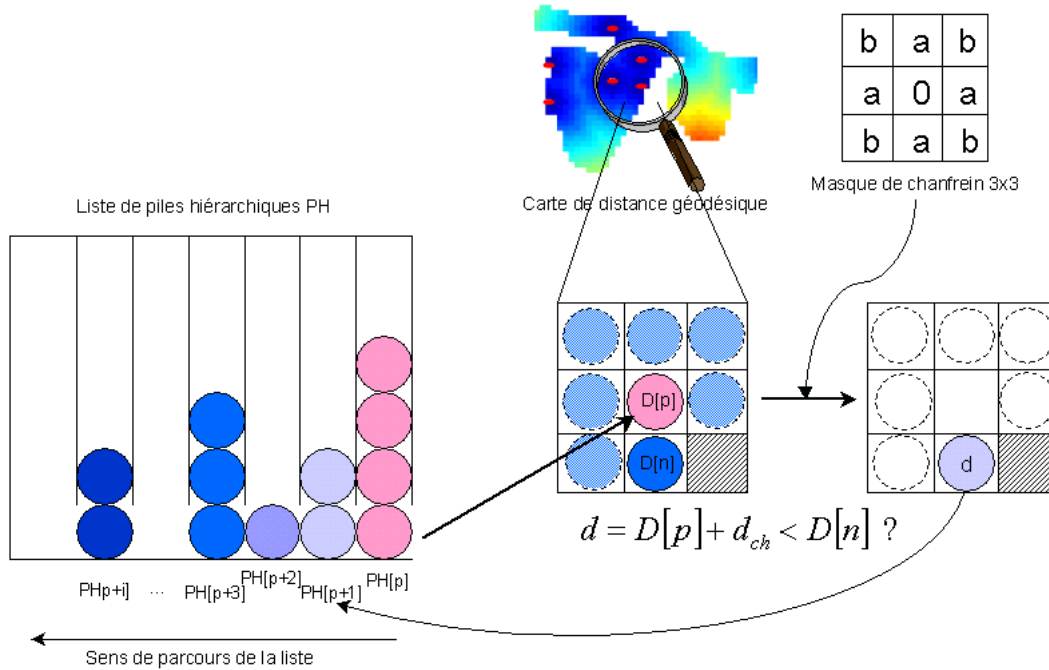


FIG. 5.20 – Calcul d'une carte de distance géodésique par propagation de proche en proche de distances de chanfrein. Les détails de l'algorithme sont donnés dans le texte et repris dans l'algorithme 3.

A l'initialisation de l'algorithme, on remplit la première pile, correspondant à la distance 0, avec les indices des voxels constituant l'objet. La carte de distance géodésique est initialisée avec une distance nulle pour tout les voxels de l'objet et infinie sinon. Ensuite, comme le montre la figure 5.20, à chaque itération de l'algorithme de propagation, on considère les voxels dont les indices sont situés dans la première pile non vide (le front). Pour chacun d'entre eux, on calcule la distance géodésique de chacun de ses voisins q en considérant que le chemin le plus court passe par le voxel courant p :

$$D[q] = D[p] + d_{chanfrein}(p,q)$$

Si la distance calculée $D[q]$ est inférieure à la distance stockée dans la carte de distance alors on met à jour cette valeur avec celle juste calculée et le voisin q du voxel courant considéré p est ajouté dans la pile correspondant à la distance $D[q]$. Cette stratégie est appliquée jusqu'à ce que toutes les piles soient vides.

Cet algorithme est résumé dans l'encart 3. Avec une implémentation soignée (allocation mémoire dynamique efficace des piles), la complexité calculatoire est proportionnelle au nombre de voxels constituant le domaine, ce qui en fait un algorithme rapide. Il est à

noter qu'il fait intervenir des masques de chanfrein et donc les distance calculées ne sont pas exactes mais permettent toutefois d'obtenir de bonnes approximations.

algorithme 3 Calcul d'une carte de distance géodésique D utilisant une liste de piles hiérarchiques PH

Initialisation:

$PH[0]$ contient la liste des voxels du fond

$PH[i]$ est vide pour tout $i \neq 0$

$D[p] = 0$ pour les voxels du fond, et ∞ sinon

Algorithme:

tant que PH est non vide **faire**

pour chaque voisin n de p **faire**

$d = D[p] + d_{chanfrein}(p,n)$

si $d < D[n]$ **alors**

$D[n] = d$

 ajouter n à $PH[d]$

fin si

fin pour

fin tant que

De manière anecdotique, cet algorithme permet de résoudre tous les problèmes de labyrinthe. La figure 5.21 présente un exemple de labyrinthe dont on connaît l'entrée et la sortie. Connaissant ces deux positions, pour trouver le chemin le plus court permettant d'aller de l'un à l'autre, il suffit de calculer une carte de distance géodésique en considérant que les murs correspondent au fond et que l'objet est situé au départ. On obtient l'image de gauche. Ensuite, il suffit de partir du point d'arrivée et de se déplacer de voisinage en voisinage vers le point dont la valeur dans la carte de distance est minimale. On obtient le chemin présenté en rouge sur l'image de droite. Cet algorithme (A^*) est fréquemment utilisé dans le calcul des déplacements de personnages intervenant dans les jeux vidéo, modulo certaines astuces de calcul.

Enfin, de manière similaire au cas euclidien, on peut aisément conserver au cours de l'exécution de l'algorithme le label correspondant à l'objet dont le voxel considéré est le plus proche. On obtient ainsi le diagramme de Voronoï géodésique pour un coup calculatoire supplémentaire négligeable.

5.4.2 Calcul du centre de masse géodésique

La deuxième étape de l'algorithme K-Means consiste à calculer le centre de masse de chaque classe. En effet, étant donné une parcellisation V_k , la mise à jour des centres \bar{x}_k se fait en minimisant:

$$\bar{x}_k = \arg \min_y \sum_{x \in V_k} d_{geom}^2(x,y) \quad \forall k \quad (5.5)$$

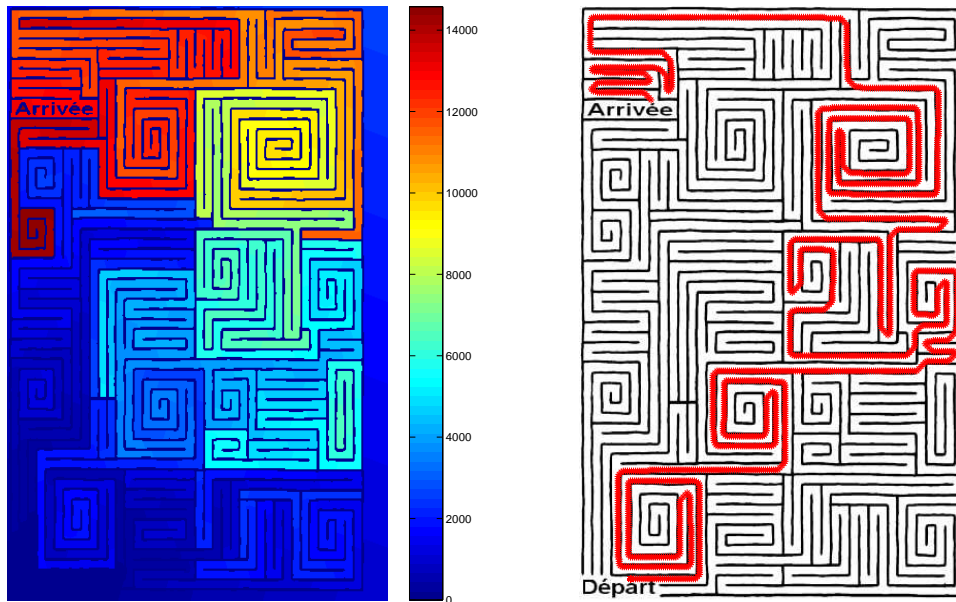


FIG. 5.21 – Résolution des problèmes de labyrinthe grâce aux cartes de distance géodésique. L'image de gauche présente la carte de distance géodésique obtenue avec une graine placée à l'entrée du labyrinthe. Trouver le chemin le plus court séparant le départ de l'arrivée s'effectue itérativement sur cette carte en partant de l'arrivée et en se déplaçant de proche en proche en minimisant la distance par rapport au départ.

Dans le cas d'une distance euclidienne, on avait une formulation explicite de ces centres de masse (le barycentre). Ici, il faut utiliser un algorithme itératif pour déterminer le minimum de ce critère. La figure 5.22.a présente un exemple de domaine non convexe où l'on a affiché en tout point la valeur du critère calculée en considérant ce point comme le centre de masse. Le point bleu correspond au barycentre euclidien calculé par simple moyenne des coordonnées des points constituant le domaine, tandis que le point rouge correspond au point où le critère calculé est minimal : c'est le centre de masse géodésique. On note que dans ce cas, le barycentre n'est même pas un point du domaine. La figure 5.22.b présente une carte d'élévation du critère. Il est intéressant de noter que par construction, le critère à minimiser est *lisse* et monotone (en raison des propriétés de l'application distance). Ainsi, la stratégie que nous avons utilisée pour déterminer le point du domaine minimisant le critère consiste à partir d'un des points du domaine et à calculer le critère associé à chacun de ses voisins (ceci nécessite de calculer une carte de distance géodésique par voisin). Il suffit ensuite d'itérer ce processus en considérant dorénavant les voisins du point pour lequel le critère était minimal, ce qui correspond à une simple descente de gradient. La convergence est obtenue lorsque tous les voisins du point considéré ont une valeur de critère associée plus élevée que celle du point lui-même.

Nous avons présenté sur la figure 5.23 deux exemple de calculs de centre de masse sur des objets non-convexes et comportant éventuellement des trous. Dans les deux cas présentés, l'initialisation de l'algorithme correspond à un cas extrême : les graines sont

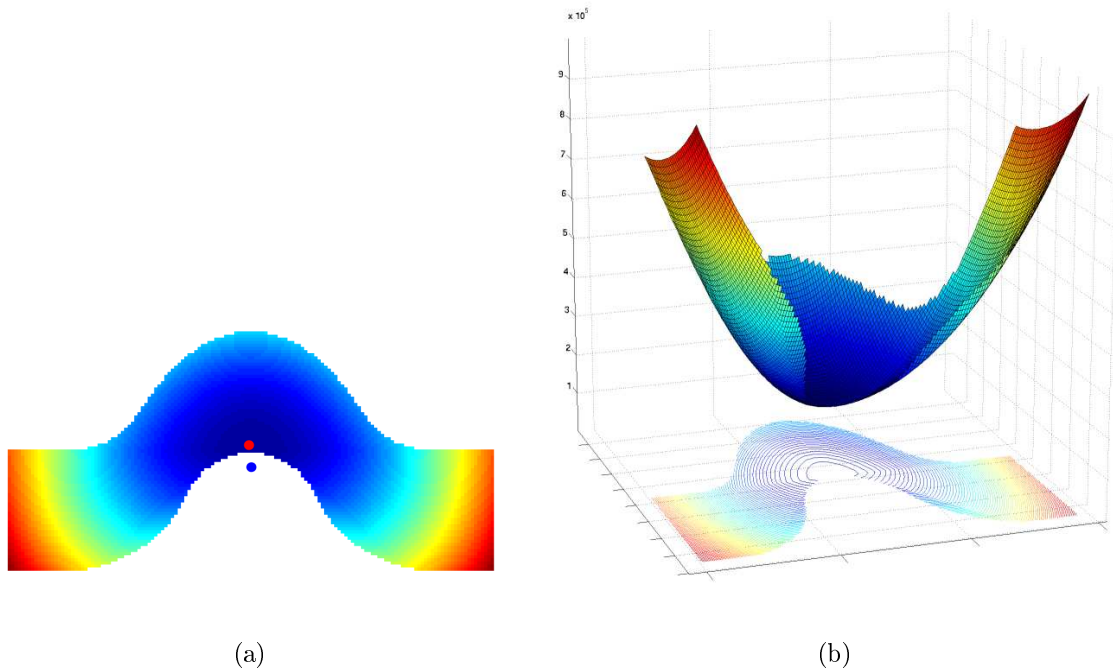


FIG. 5.22 – Critère associé à la détermination du centre de masse d'un objet non-connecté : pour chaque voxel situé à l'intérieur du domaine, le critère présenté par l'équation 5.5 est calculé. Le point en rouge représente alors la position minimisant le critère (le centre de masse géodésique) tandis que le point en bleu désigne le barycentre euclidien.

initialisées parmi les points les plus éloignés de la solution. Les courbes rouges représentent le parcours effectué pour atteindre la solution. Le critère a été calculé en chacun des points constituant ce chemin et en chacun de leurs voisins (8-connectivité en 2D).

A la différence du cas euclidien, le calcul itératif pour déterminer les nouvelles positions des graines est potentiellement lourd (nombreux calculs de cartes de distance géodésique mais les domaines considérés sont alors plus petits : on considère une parcelle et non le domaine tout entier). C'est pour cette raison que nous avons choisi de ne mettre en œuvre cet algorithme que dans les cas où le calcul du barycentre euclidien fournit un point n'appartenant pas au domaine.

5.4.3 Exemples de parcellisations obtenues

Les deux étapes de l'algorithme K-Means ont été décrites dans les sections précédentes. Pour effectuer la parcellisation d'un domaine tel que celui présenté dans la section 5.2, la méthodologie est donc la suivante :

1. choix du nombre de parcelles K par l'utilisateur et tirage aléatoire sans remise de K voxels parmi l'ensemble de voxels constituant le domaine à parcelliser.
2. calcul du diagramme de Voronoï géodésique correspondant aux K graines, grâce à l'algorithme décrit dans la section 5.4.1.

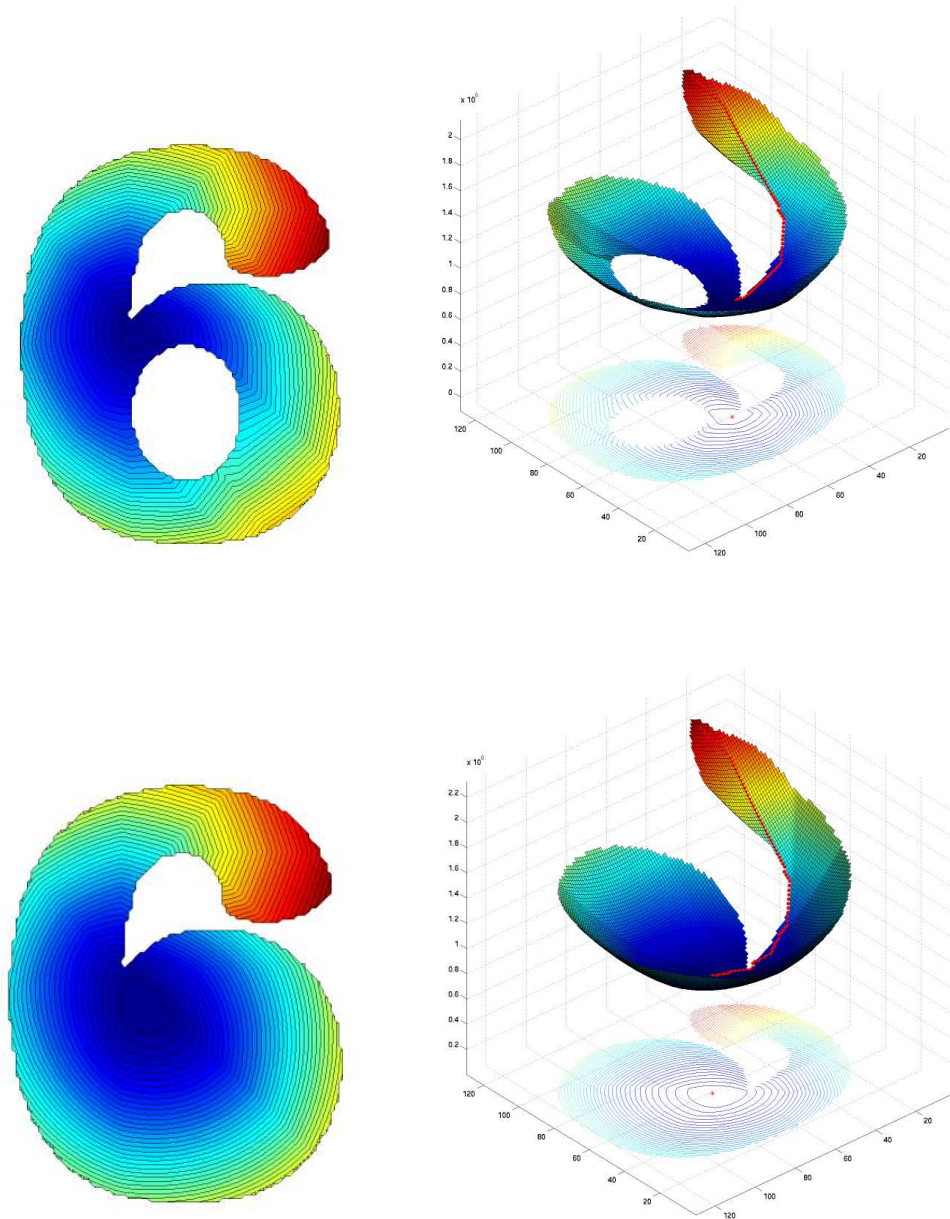


FIG. 5.23 – Détermination du centre de masse d'objets non convexes : le critère a été calculé en chaque point du domaine et la courbe rouge présente la progression de l'algorithme au cours des itérations pour trouver le centre de masse.

3. mise à jour de la position des graines en calculant le centre de masse géodésique de chaque parcelle. Comme indiqué précédemment, afin d'alléger les calculs, si le barycentre d'une parcelle se situe à l'intérieur de celle-ci alors la graine est mise à jour avec cette nouvelle position. Sinon l'algorithme itératif décrit dans la section précédente correspondant à une descente de gradient sur le critère, est utilisé.

4. itération des étapes 2 et 3 jusqu'à ce que la position des graines n'évolue plus.

Nous présentons sur la figure 5.24 un exemple de domaine 2D correspondant à la segmentation d'un hémisphère du cortex sur une coupe axiale T_1 . L'image de gauche montre la parcellisation initiale en 50 éléments répartis aléatoirement. L'image de droite montre elle la parcellisation obtenue après convergence de l'algorithme : ici, une quinzaine d'itérations ont été effectuées, en quelques secondes.

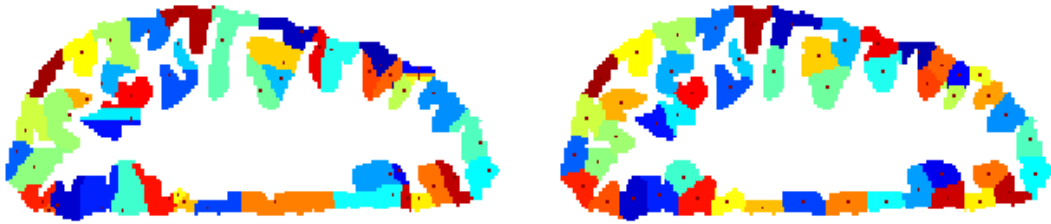


FIG. 5.24 – Parcellisation en 50 éléments d'une coupe axiale d'un hémisphère du cortex : initialisation à gauche et après convergence à droite.

Appliqué sur des domaines tridimensionnels, l'algorithme fournit les parcellisations représentées sur la figure 5.25. L'image de gauche correspond à une parcellisation en 300 éléments tandis que celle de droite est à une résolution beaucoup plus fine avec 1000 parcelles.

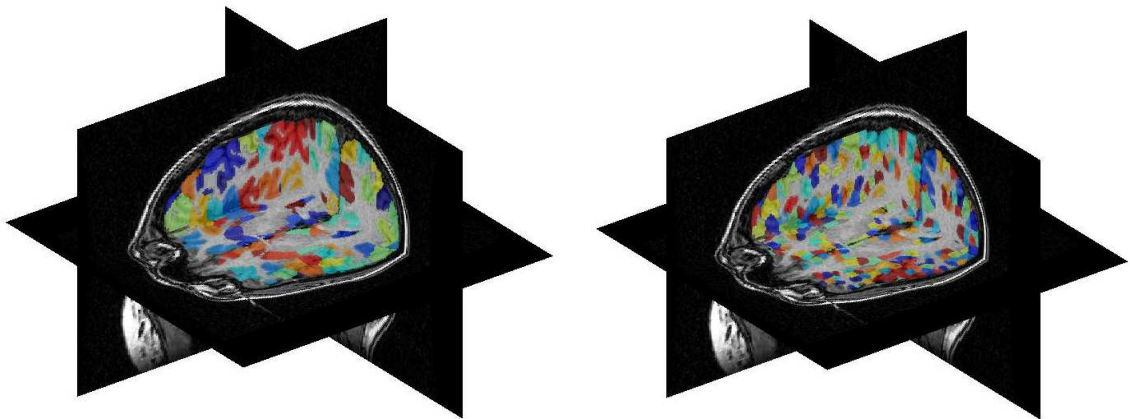


FIG. 5.25 – Parcellisations tridimensionnelles du cortex superposées à l'IRM T_1 en 300 éléments à gauche et 1000 à droite.

Nous avons également appliqué cet algorithme sur la segmentation du ruban cortical d'un singe (images provenant du projet européen MAPAWAMO²³). Dans ce cas, la taille

²³. MAPAWAMO : Mapping visual cortical regions in awake, behaving monkey using functional MRI, <http://www-sop.inria.fr/epidaure/Collaborations/MAPAWAMO/>

des voxels était similaire à celle fréquemment utilisée chez l'homme (1 mm^3), par contre le cerveau de singe est beaucoup plus petit qu'un cerveau humain (d'un facteur 4 en moyenne) ce qui implique la résolution effective de ces images est beaucoup plus faible.

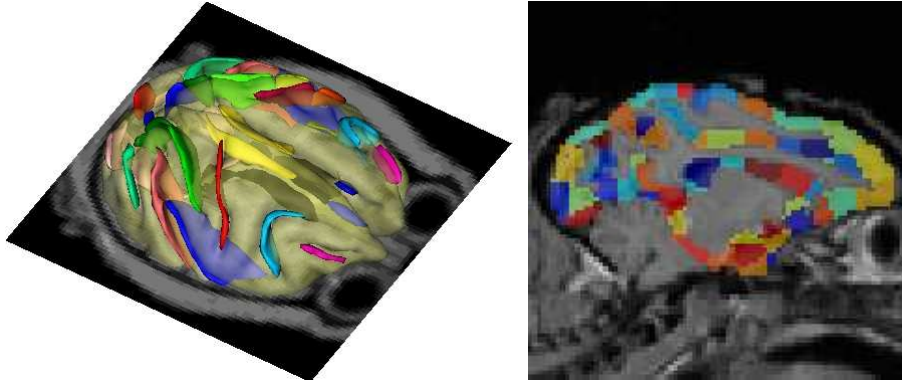


FIG. 5.26 – Parcellisation d'un hémisphère du cortex d'un singe du projet MAPAWAMO en 200 parcelles.

A partir d'une parcellisation donnée, il est possible de construire le graphe d'adjacence associé, indiquant les parcelles voisines. La figure 5.27 présente un exemple de ce type de graphe calculé à partir d'une parcellisation du cortex en 600 éléments. Ceci met en valeur le fait qu'une parcellisation peut être assimilée à un nouveau cadre dans lequel représenter des données. Au lieu d'utiliser comme unité de volume le voxel et comme relation de voisinage les connexités naturelles, l'élément élémentaire est la parcelle et les relations de voisinage ne sont plus immédiates (comme c'est le cas sur une grille régulière) mais se déduisent de la parcellisation après définition du type de connexité à considérer.

5.5 Introduction de contraintes spatiales

Dans cette section, nous proposons de modifier l'algorithme de parcellisation afin de prendre en compte de nouvelles contraintes. Les considérations en faveur de l'ajout de ces contraintes sont de l'ordre fonctionnel mais ont des répercussions sur la définition spatiale des parcelles. C'est pour cela que nous qualifions les contraintes d'anatomo-fonctionnelles. Deux types de contraintes sont décrites dans la suite : l'une correspondant à l'introduction de l'information de la position des sillons corticaux lors de la définition des parcelles et l'autre liée à l'imposition de la position de certaines graines.

5.5.1 Introduction de l'information *sillons*

Une des hypothèses concernant la relation anatomie-fonction pour le cerveau est que les sillons corticaux feraient office de séparateur entre les différentes aires fonctionnelles. Si cette hypothèse n'a pour l'instant jamais pu être entièrement confirmée dans le cas général, elle est vérifiée tout au moins pour les grands sillons. Ainsi le sillon central sépare

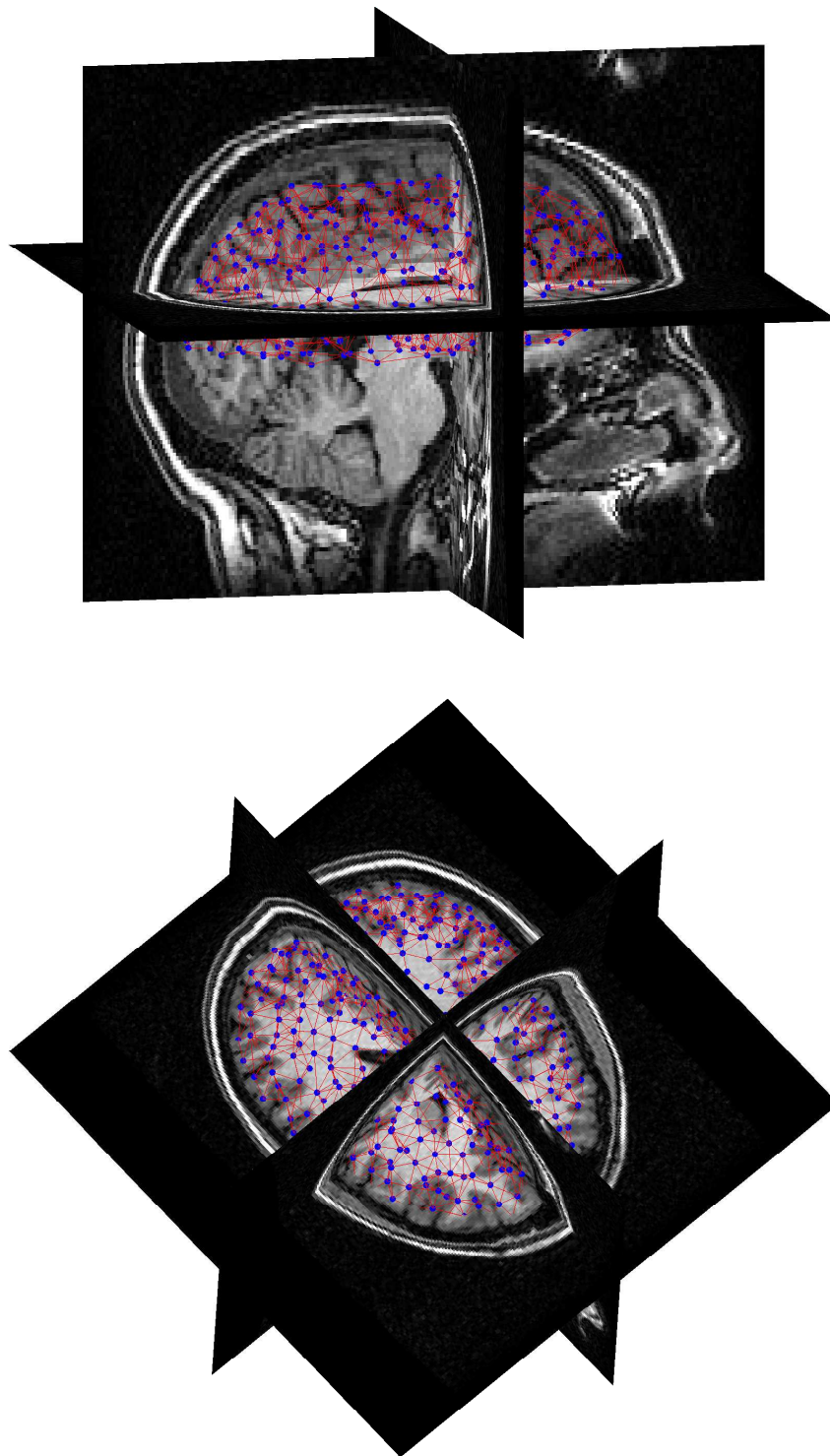


FIG. 5.27 – Graphe d'adjacence calculé à partir d'une parcellisation du cortex cérébral en 600 classes. Les centres de chaque classes sont modélisées par un point bleu.

les lobes frontaux et pariétal et par là-même les aires motrices et sensorielles. Connaissant la position de ces sillons, une des caractéristique intéressante d'une parcellisation du cortex serait donc de contraindre la définition des parcelles afin qu'une parcelle ne puisse se retrouver de part et d'autre d'un sillon. Il faut ainsi pouvoir imposer le fait que certains sillons puissent être les frontières physiques entre parcelles. Pour ce faire, la définition de la distance entre deux voxels du domaine est modifiée grâce à la définition pour chaque voxel de *coût de traversée* α_i du dit voxel :

$$D(p,q) = \sum_{i=1}^N \alpha_i d_{chanfrein}(p_i, q_i) \quad \text{avec} \quad \alpha_i \geq 1 \quad (5.6)$$

La distance entre les points p et q est décomposée en une somme de distances élémentaires, correspondant au masque de chanfrein pour lequel une pondération dépendant de la position courante du masque est appliquée sur le coefficient de chanfrein considéré. Le diagramme de Voronoï découlant de cette distance est dit pondéré. Il faut donc maintenant construire une image des α_i de telle manière que les distances faisant intervenir des chemins traversant des sillons soient pénalisées. Pour ce faire, nous avons utilisé la labellisation des sillons fournie par l'algorithme de [Rivière *et al.*, 2002] dont on crée une image binaire mettant en évidence les sillons pour lesquels on veut imposer une contrainte (sillon central, pariétal, etc). Pour des raisons de topologie et pour modéliser le volume partiel autour des sillons, cette image est lissée par un noyau gaussien isotrope de quelques millimètres :

$$\alpha_{\text{pondération}} = \beta G_{\sigma}(I_{\text{sillons}})$$

où β est un terme multiplicatif constant, paramétrant l'influence de la contrainte spatiale. Plus ce terme est élevé, plus la traversée d'un sillon sera difficile.

Une analogie peut être faite avec le déplacement à pied d'une personne sur un sentier escarpé : pour rejoindre un point B à partir d'un point A le plus rapidement possible, le randonneur doit choisir le chemin le plus court parmi les chemins disponibles (distance géodésique) tout en prenant en compte que plus l'altitude est élevée, moins il y a d'oxygène dans l'air et donc moins il marchera rapidement. Il faut donc trouver un compromis entre chemin le plus court et chemin le plus près du niveau de la mer. Dans notre cas, l'altitude est reliée à la proximité d'un sillon.

La figure 5.28 présente un exemple synthétique d'ajout de contrainte de type *sillon* à une parcellisation. L'image (a) représente une parcellisation quelconque d'un domaine convexe bidimensionnel en 200 classes, qui sert d'initialisation. La contrainte ajoutée est modélisée par l'image $\alpha_{\text{pondération}}$ représentée sur la figure (b) : c'est une courbe sinusoïdale partageant le domaine en deux. L'application de l'algorithme K-Means utilisant les distances pondérées par l'image $\alpha_{\text{pondération}}$ et initialisé par la parcellisation (a) fournit une nouvelle parcellisation représentée sur la figure (d). La parcellisation sans contrainte obtenue avec la même initialisation est présentée sur la figure (c) par comparaison. On observe que les parcelles qui étaient initialement de part et d'autre de la contrainte, se sont déplacées de telle sorte que leurs frontières sont dorénavant données par la contrainte.

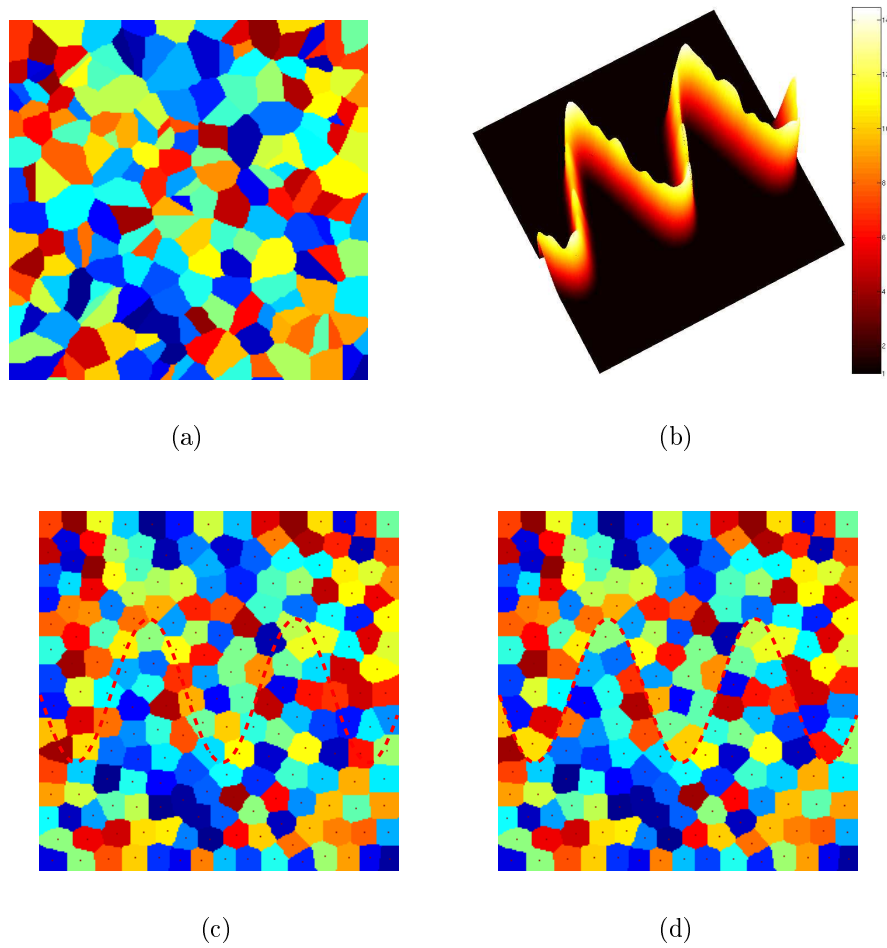


FIG. 5.28 – Exemple d'introduction de contrainte spatiale sur un domaine rectangulaire. L'image (a) présente la parcellisation aléatoire du domaine en 200 classes. La figure (b) représente une carte d'élévation de l'image $\alpha_{pondération}$ pénalisant un tracé sinusoïdal. Les images (c) et (d) présentent des résultats de parcellisation centroïdale, respectivement sans et avec la contrainte (b) (affichée en pointillé sur les images).

Bien qu'innovante, nous n'avons pas, dans la suite de nos travaux, utilisé cette pénalisation reposant sur les sillons lors de la construction de parcellisations anatomiques, car elle s'avère difficile à utiliser inter-sujets et nous sommes plutôt dirigés vers l'introduction directe de l'information fonctionnelle (cf chapitre 7) dans le processus de construction de parcellisations.

5.5.2 Imposition de la position de certaines graines

Une autre contrainte potentielle consiste à imposer la position de certaines graines au sein du domaine. En effet, lors de l'analyse de données fonctionnelles, on a parfois une information a priori de la position de certaines activations, par exemple par la

description dans la littérature de résultats obtenus avec des protocoles similaires ou par les résultats d'expériences effectuées au préalable (*localizers*). Ce type de contrainte s'ajoute facilement dans l'algorithme K-Means de parcellisation. Il suffit d'initialiser certaines graines aux positions prédéfinies et de ne pas mettre à jour leur position au cours de l'évolution de l'algorithme. En effet, le critère n'est en ce cas pas modifié, il n'y a que quelques paramètres en moins à estimer.

Une analogie intéressante peut être faite entre la parcellisation du cortex avec contraintes et le découpage du territoire français comme le montre la figure 5.29. La France a été au cours de l'histoire découpée administrativement en régions, départements, cantons, etc. Si nous considérons le domaine à parcelliser comme étant la France métropolitaine et la Corse, on peut effectuer une parcellisation automatique en environ 100 départements ou 22 régions, comme le montrent les figures b et c. Cependant si l'on compare ces découpages du territoire français avec ce qu'il est réellement (figure d), on observe de nombreuses différences.

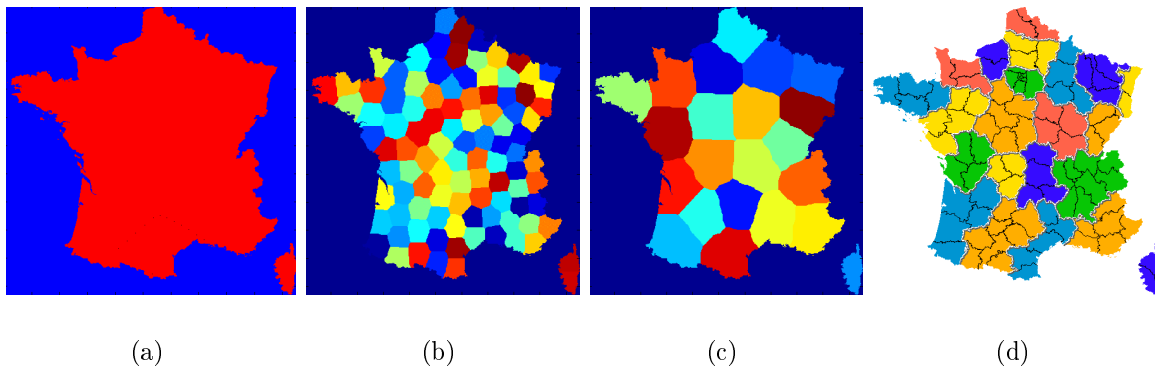


FIG. 5.29 – Analogie entre le découpage du territoire français et la parcellisation du cortex : l'image (a) présente la carte de France à découper en régions et départements. L'application de l'algorithme de parcellisation présenté dans ce chapitre fournit les résultats (b) et (c) pour respectivement 100 et 22 parcelles. La figure (d) présente le découpage réel du territoire français.

De manière simplifiée, on peut effectivement noter que fleuves et rivières sont souvent des frontières entre régions et départements ainsi que les montagnes (équivalence avec les sillons). De même, certaines villes importantes se trouvent au centre de leur département (équivalence avec le choix prédéfini de la position de certaines graines). Bien sûr, même si l'on appliquait ces contraintes aux parcellisations obtenues sur la figure 5.29, on n'obtiendrait pas le découpage réel du territoire français car celui-ci s'est modifié au cours du temps et il existerait de nombreux autres paramètres historiques à prendre en compte pour contraindre la parcellisation. Il en est probablement de même pour les cartes anatomo-fonctionnelles où les informations anatomiques sulco-gyrales et fonctionnelles BOLD ne sont que des éléments incomplets permettant l'obtention de telles cartes et la définition de frontières d'aires.

5.6 Conclusion

Dans ce chapitre, nous avons présenté un algorithme permettant d'obtenir une parcellisation d'un domaine non convexe tel que le cortex en un nombre quelconque d'éléments. La méthode repose sur la construction de diagrammes de Voronoï centroïdaux, obtenus par l'algorithme K-Means, bien connu en classification et modifié pour prendre en compte des distances géodésiques.

Les parcellisations de Voronoï centroïdales ont de très nombreuses applications [Du *et al.*, 1999], dans des domaines variés : analyse numérique, positionnement optimal des ressources, croissance de cellules, quantification vectorielle, etc. Leurs intéressantes propriétés ont récemment été utilisées pour obtenir des maillages isotropes de surfaces [Surazhsky *et al.*, 2003] (voir figure 5.30). D'autres références font appel à l'algorithme K-Means pour parcelliser un domaine, avec des applications en infographie pour simuler des mosaïques à partir d'images [Hausner, 2001] (voir par exemple la figure 5.31).

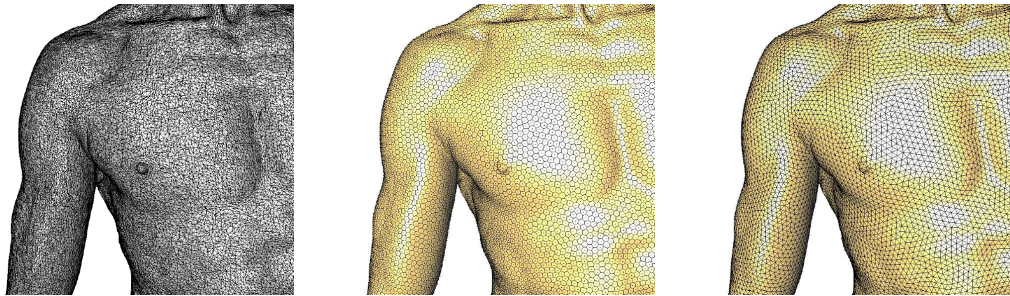


FIG. 5.30 – Maillage d'une numérisation du David de Michel-Ange : *original, partition uniforme et remaillage triangulaire. D'après [Surazhsky et al., 2003]*

Un point intéressant est la technique de calcul des diagrammes de Voronoï. En effet, en deux dimensions, il est possible d'utiliser la puissance de calcul des cartes graphiques : il suffit pour cela de placer en chaque graine l'apex d'un cône avec une large base. Calculer l'image en *z-buffer* issue d'une caméra placée au dessus des cônes permet d'obtenir directement un diagramme de Voronoï discret (voir figure 5.32).

Notre contribution concernant les diagrammes de Voronoï centroïdaux est la généralisation de l'algorithme K-Means en trois dimensions dans un domaine non convexe, utilisant des distances géodésiques. L'utilisation d'algorithmes de géométrie discrète permet d'effectuer les calculs de manière efficace. Nous proposons également une technique pour prendre en compte certaines contraintes, comme l'imposition de certaines frontières, grâce à l'utilisation de distances pondérées.

Dans le chapitre suivant, nous présentons plusieurs applications potentielles de parcellisations du cortex pour l'analyse de données fonctionnelles.



FIG. 5.31 – Utilisation de parcellisations de Voronoï centroïdales en infographie pour simuler des mosaïques décoratives à partir de dessins ou de photos. La "Sibylle libyenne" de Michel-Ange (chapelle sixtine) est ainsi transformée en une mosaïque de 2000 éléments avec un algorithme de type K-Means. D'après [Hausner, 2001].

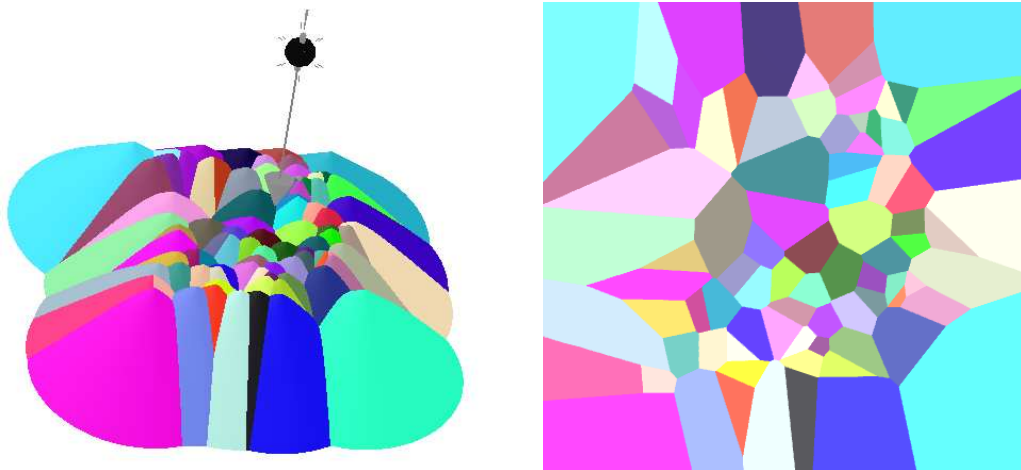


FIG. 5.32 – Calcul rapide de diagrammes de Voronoï 2D en utilisant la carte graphique de l'ordinateur: des cônes sont placés au dessous de chaque graine et une vue de dessus de cette scène tridimensionnelle obtenue par la technique du z-buffer permet d'obtenir directement le diagramme de Voronoï associé à ces graines.

Chapitre 6

Applications des parcellisations anatomiques

Sommaire

6.1	Introduction	138
6.2	Détection d'activations sur une parcellisation	138
6.2.1	Préambule à propos de la résolution spatiale des images fonctionnelles	138
6.2.2	Jeux de données, paradigme et modèle d'analyse	139
6.2.3	Méthode de parcellisation, affectation du signal fonctionnel et détection	140
6.2.4	Comparaison entre la détection à la résolution voxel et celle à la résolution parcelle	143
6.2.5	Discussion	147
6.3	Sélection de modèle par parcelle	149
6.3.1	Motivations	149
6.3.2	Une modélisation régionale des données	150
6.3.3	Application sur un jeu de données	151
6.4	Mesures robustes de la connectivité fonctionnelle	152
6.4.1	La connectivité fonctionnelle	153
6.4.2	Problème de dimensionnalité	153
6.4.3	Etude de la nature des interactions entre parcelles	154
6.4.4	Une nouvelle mesure de connectivité	155
6.4.5	Discussion	155
6.5	Conclusion	156

6.1 Introduction

Nous avons vu dans les chapitres d'introduction certaines limitations intrinsèques aux méthodes fondées sur l'analyse voxel à voxel et les avantages potentiels de l'utilisation d'une méthode fondée sur une parcellisation du volume cérébral. Le chapitre précédent a décrit une technique permettant de construire une telle parcellisation à partir de l'information anatomique. Cette technique peut être considérée comme une alternative à l'opération de lissage spatial communément appliquée sur les données fonctionnelles.

Dans ce chapitre, nous développons trois applications potentielles de cette méthode. Dans un premier temps, la parcellisation sera utilisée dans un but de détection du signal fonctionnel et évaluée par rapport à la méthode standard d'analyse de données. Ensuite, nous verrons comment on peut l'utiliser pour obtenir des modèles du signal temporel adaptés à chaque région cérébrale. Enfin, nous présenterons une application de la méthode à la mesure de la connectivité fonctionnelle entre régions.

6.2 Détection d'activations sur une parcellisation

La première application que nous présentons concerne la détection d'activations sur une parcellisation anatomique. En effet, une parcellisation permet d'obtenir une représentation plus concise des données fonctionnelles. La résolution étant amoindrie, on diminue le nombre d'observations permettant éventuellement d'obtenir des résultats plus sensibles.

Par ailleurs, en considérant une parcellisation comme une technique de lissage spatial des données, il est intéressant de comparer les résultats de détection d'activations obtenues avec une approche voxel (lissage spatial gaussien) et une approche parcelle.

Nous détaillons dans la suite de ce paragraphe la méthode mise en œuvre pour analyser des données fonctionnelles à une résolution parcelle puis nous proposons de comparer les résultats avec ceux obtenus par l'approche standard.

6.2.1 Préambule à propos de la résolution spatiale des images fonctionnelles

Il est bien connu que les images fonctionnelles sont le plus souvent analysées avec une résolution spatiale dégradée par rapport à la résolution d'acquisition. Cette dégradation apparaît bien souvent comme paradoxale, contradictoire avec les efforts mis en œuvre pour obtenir et conserver une bonne résolution spatiale.

La raison la plus souvent avancée pour justifier cette dégradation est le mauvais rapport signal à bruit des images fonctionnelles. Le signal est considéré comme étant plutôt basse fréquence et le bruit plutôt haute fréquence, d'où un gain obtenu par filtrage des hautes fréquences spatiales.

La seconde raison invoquée fait référence aux études de groupe. On a rappelé précédemment (cf chapitre 3) les problèmes de mise en correspondance des données de plusieurs sujets, du fait de la nature très variable de l'anatomie cérébrale d'une part mais aussi du fait de la variabilité fonctionnelle qui peut être observée entre les sujets. Le filtrage spatial passe bas, en étalant le signal sur des zones plus vastes, permet d'augmenter les

chances de recouvrement du signal entre sujets, et d'observer les résultats de groupe à une moindre résolution mais avec une meilleure sensibilité.

Le troisième argument n'est pas le moindre, mais est rarement aussi bien identifié que les deux premiers. Il concerne l'aspect test statistique du signal "d'activation". Dans le paradigme classique du test d'hypothèse, le risque α de faux positifs est contrôlé en déterminant un seuil tel que la probabilité de déclarer une région activée alors que l'on se trouve sous l'hypothèse nulle (bruit seulement) vaut α .

Une difficulté survient lorsque l'on réalise plusieurs tests (par exemple N). Si ceux-ci sont indépendants, on peut adapter le seuil α de telle sorte, qu'en moyenne, on ait une probabilité α d'obtenir un faux positif pour l'ensemble des N tests. Le seuil ainsi adapté sera dit "*corrigé pour les comparaisons multiples*" et est obtenu pour des tests indépendants par la formule de Bonferroni :

$$\alpha_c = 1 - (1 - \alpha)^{1/N} \simeq \frac{\alpha}{N}$$

En effet, on a, d'après les définitions indiquées ci-dessus :

$$\begin{aligned} p(\text{un test passe} | H_0) &= \alpha_c \\ p(\text{aucun test ne passe} | H_0) &= 1 - \alpha_c \\ p(\text{aucun test ne passe sur } N \text{ tests indépendants} | H_0) &= (1 - \alpha_c)^N = 1 - \alpha \end{aligned}$$

Dans le cas des images fonctionnelles du cerveau, si le test est réalisé en chaque voxel, N vaut environ 25000. Bien que l'indépendance des tests ne soit pas vérifiée, on imagine aisément la faible sensibilité d'une telle procédure. En utilisant les techniques permettant la prise en compte de la corrélation spatiale des données, la correction à apporter est bien moins sévère et donc la sensibilité de l'analyse meilleure (voir le paragraphe 3.3.3.0).

Quelques travaux ont vu le jour dans la littérature cherchant à préserver la résolution spatiale tout en améliorant le rapport signal à bruit. Citons en particulier [Descombes *et al.*, 1998b; Andrade *et al.*, 2001]. Ces travaux restent cependant encore marginaux et peu d'applications de neuroimagerie les utilisent. Notons aussi que ce problème de résolution versus sensibilité a été identifié dès le début des années 1990 et que des méthodes adoptant des stratégies multi-échelles ont été suggérées [Poline and Mazoyer, 1994a; 1994b].

6.2.2 Jeux de données, paradigme et modèle d'analyse

Le jeu de données que nous avons utilisé ici a pour but d'effectuer une cartographie du sillon pariétal grâce à plusieurs tâches mobilisant cette aire corticale [Simon *et al.*, 2002]. Le protocole expérimental consiste en trois blocs de conditions (mouvement de préhension ou "grasping") séparés par trois périodes de contrôle de 26 secondes chacune, précédée par 4 secondes d'instruction. Les sujets (au nombre de 10) étaient soumis à deux séquences de ce type. Le temps de répétition est de 2 secondes et correspond à la durée nécessaire à l'acquisition d'un volume fonctionnel, pour un total de 186 scans sur l'ensemble des deux sessions. La dimension spatiale des données est de $64 \times 64 \times 18$ avec $3.75 \times 3.75 \times 3.8 \text{ mm}^3$ pour taille de voxel. Ces données ont été acquises sur un scanner GE 1.5T et les distorsions entre volumes anatomiques et fonctionnels sont peu perceptibles.

L'image anatomique (pondérée en T_1) acquise peu après les images fonctionnelles a une résolution de $0.94 \times 0.94 \times 1.5 \text{ mm}^3$ (pour $256 \times 256 \times 124$ voxels).

Les données comportent d'autres sessions et d'autres conditions expérimentales qui seront décrites lors de la présentation des résultats sur la parcellisation à partir des données fonctionnelles.

6.2.3 Méthode de parcellisation, affectation du signal fonctionnel et détection

La méthode de traitement des données est rapidement résumée ici. Les aspects théoriques ayant déjà été présentés dans le chapitre 3, nous mettons l'accent sur des considérations pratiques qui font une part importante du succès de l'analyse de données.

Parcellisation

En premier lieu, une parcellisation est obtenue selon la technique décrite dans le chapitre précédent. Notons l'importance de la vérification de la correspondance des images anatomiques et fonctionnelles et la nécessité de vérifier très attentivement d'éventuelles distorsions ou non alignement des images. Les bords du cerveau sont bien sûr des repères essentiels, mais les structures internes tels que les ventricules, le sillon pariéto-occipital, une certaine partie de l'insula sont aussi de bon repères pour cette validation. Bien qu'une bonne correspondance soit essentielle, notons que la parcellisation est par essence une technique relativement robuste dès que la résolution choisie n'est pas trop précise. D'autre part, il n'est en général pas possible de recouvrer l'ensemble des distorsions et des pertes de signal dues aux inhomogénéités du champs magnétique dès qu'il s'agit de données acquise à 3 Tesla ou plus. Cet état des choses est appelé à évoluer avec la généralisation des systèmes d'acquisition parallèle (SENSE, SMASH, etc). Le choix du nombre de parcelles est discuté plus avant.

Affectation du signal et détection d'activation

On peut distinguer deux problèmes lors de l'affectation d'un signal fonctionnel à une parcelle (voir figure 6.1). D'une part, il faut résoudre un problème d'interpolation, la résolution des images anatomiques et fonctionnelles étant différente. D'autre part, il faut résumer une série de décours temporels et de nombreux choix sont possibles. Dans l'exemple que nous présentons ici, les voxels fonctionnels sont interpolés à la résolution des voxels anatomiques par fonctions splines. Le choix de la fonction d'interpolation n'est pas crucial dans ce contexte. L'ensemble des décours est ensuite résumé par leur valeur moyenne. D'autres stratégies sont envisageables mais sous l'hypothèse qu'à chaque parcelle est associée une activité fonctionnelle assez homogène, ce choix est raisonnable en première approximation. Dans le cas où il faudrait considérer que plusieurs signaux peuvent être présents dans la parcelle, il faudrait recourir à des techniques de type PCA (ou plus exactement CCA [Friman *et al.*, 2002] pour tenir compte des corrélations nécessairement présentes entre les voxels) voire ICA. Cependant, rappelons qu'un des buts de ce travail

est de proposer des parcelles dont on peut raisonnablement supposer qu'elles sont homogènes (du point de vue anatomique à ce stade du document, mais aussi du point de vue fonctionnel par la suite).

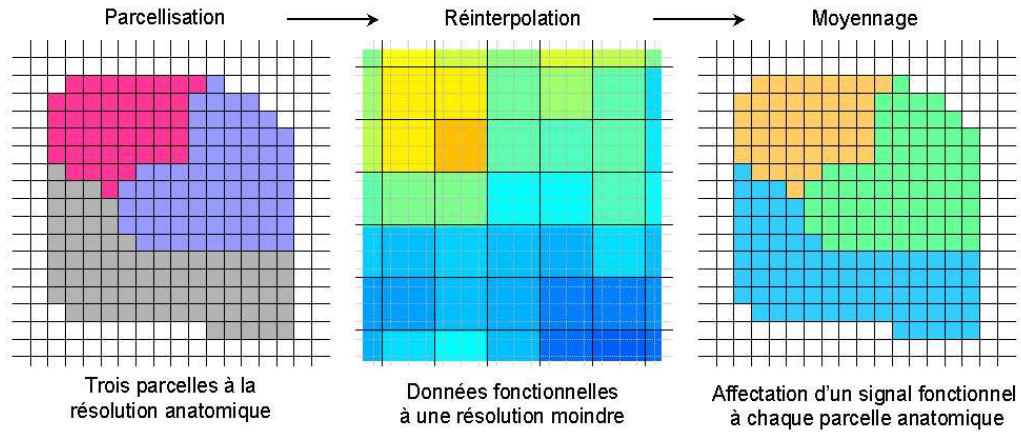


FIG. 6.1 – Affectation d'un signal en chaque parcelle en rééchantillonnant les images fonctionnelles et en moyennant les signaux au sein de chaque parcelle. Ceci est présenté schématiquement ici pour un instant temporel (une seule image) et doit être répété sur chacune des images du protocole afin d'obtenir des signaux temporels.

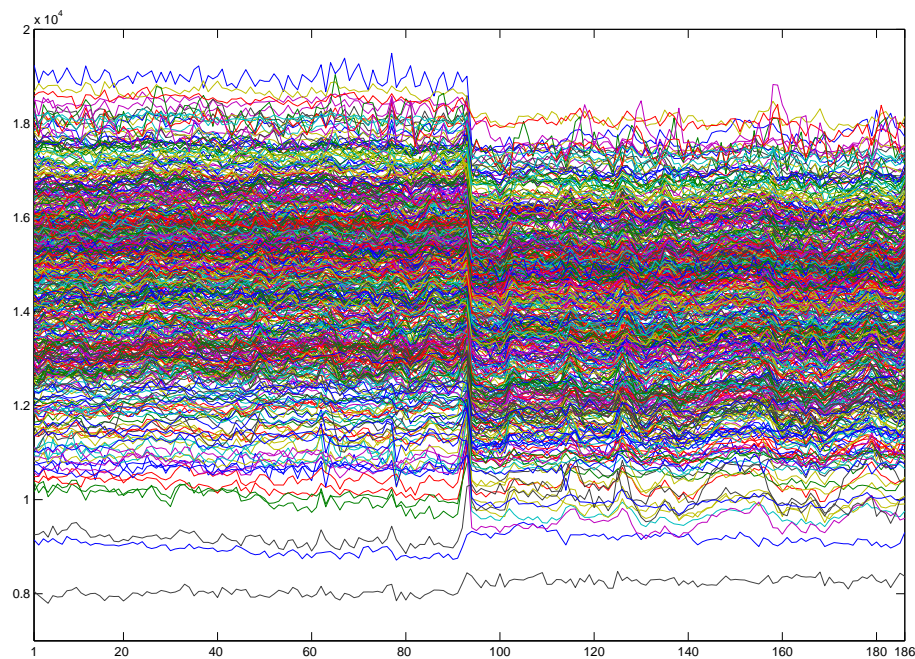


FIG. 6.2 – Décours temporel des signaux fonctionnels BOLD moyennés sur 500 parcelles couvrant les deux hémisphères du cerveau.

Le signal ainsi résumé dans chaque parcelle est analysé à partir d'un modèle linéaire de

type "SPM" décrit précédemment. Ici encore, de nombreuses variantes sont possibles. En particulier, l'estimation du bruit résiduel pourrait (devrait?) tenir compte de l'ensemble des décours temporels mais ces modèles supposent que la structure de covariance spatiale soit bien connue ou bien estimée, ce qui représente une difficulté que nous n'aborderons pas ici.

Si l'on note Y_k le signal fonctionnel correspondant à la parcelle k , le modèle général linéaire [Friston *et al.*, 1995b] s'écrit :

$$Y_k = X\beta_k + \epsilon_k$$

où X est la matrice codant pour l'information du paradigme expérimental, ϵ_k l'erreur résiduelle (supposée normale), et β_k les paramètres du modèle (cf chapitre 3).

Un contraste c des paramètres estimés est testé avec :

$$T = \frac{c^T \hat{\beta}_k}{\left(Var \left[c^T \hat{\beta}_k \right] \right)^{1/2}} = \frac{c^T \hat{\beta}_k}{\left(c^T \hat{\sigma}_k^2 (X^T X)^{-1} X^T V X (X^T X)^{-1} c \right)^{1/2}}$$

où $\hat{\sigma}_k^2$ est la variance de l'erreur estimée et V la structure de corrélation temporelle.

La distribution de T sous l'hypothèse nulle peut être approximée par une distribution de Student (degrés de liberté: $\nu = \frac{trace(RV)^2}{trace(RV RV)}$) avec R le projecteur sur les résidus tel que $\hat{\epsilon}_k = RY_k$.

Un modèle simple du décours temporel (X) est choisi correspondant à la convolution des alternances des conditions expérimentales avec une fonction hémodynamique standard et sa dérivée dans le temps. Bien que ce modèle soit aisément perfectible, il est répandu dans la pratique et suffit pour le but de comparaison que nous nous sommes donné.

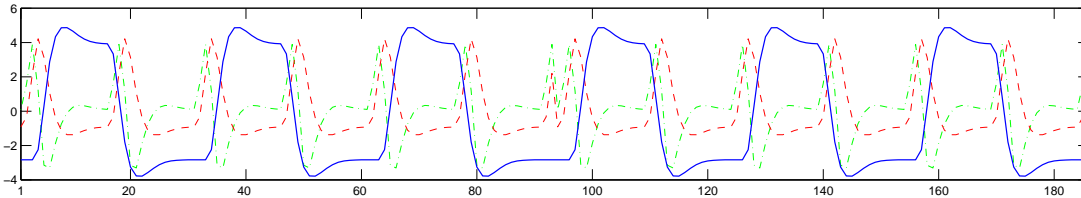


FIG. 6.3 – Décours temporel modélisant la tâche, obtenu en convoluant un modèle carré par une réponse hémodynamique canonique. Les courbes en pointillé correspondent aux dérivées premières et secondes, permettant de modéliser une partie de l'incertitude sur la forme de la réponse hémodynamique.

Correction pour le problème de multiples comparaisons

On peut raisonnablement émettre l'hypothèse que les signaux temporels issus de chaque parcelle sont statistiquement indépendants, en raison du processus même de leur construction (moyennage des signaux fonctionnels interpolés sur chaque parcelle). Ainsi, en ce cas, la correction de Bonferroni s'applique et le risque de faux positifs corrigé α_c est donné par $\alpha_c = \frac{\alpha}{N}$ où N est le nombre de parcelles.

Nous présentons sur la figure 6.4 les cartes d'activations obtenues sur la tâche de *grasping* avec 4000 parcelles. On note une activation dans le sillon central, correspondant au fait que le sujet a une activité motrice avec sa main. De plus, on observe que cette tâche active des régions frontales et pariétales. Les mêmes conclusions sont reportées dans [Simon *et al.*, 2002], détaillant les résultats obtenus sur ce jeu de données avec une analyse effectuée grâce à SPM.

A noter que la tâche de *grasping* entraîne une activité neuronale intense, avec des valeurs statistiques en t très élevées, expliquant les larges aires détectées.

6.2.4 Comparaison entre la détection à la résolution voxel et celle à la résolution parcelle

Nous avons étudié les performances de détection obtenues à partir des données traitées avec un filtrage gaussien passe-bas et celles obtenues par parcellisation de la matière grise. La figure 6.5 résume le diagramme des expériences menées pour cette comparaison. Clairement, cette comparaison est difficile du fait de plusieurs facteurs.

En premier lieu, les effets d'un mauvais recalage entre les données anatomiques et fonctionnelles devrait affecter la méthode de parcellisation, tandis que la prise en compte de l'information anatomique devrait sensibiliser l'analyse du fait de la réduction du nombre de tests statistiques à opérer.

En second lieu, une telle comparaison nécessite que l'on puisse comparer la "résolution spatiale" des deux types d'analyse. La manière de rendre cette comparaison interprétable n'est pas immédiate.

Enfin, les critères de détection statistique ne sont pas les seuls à prendre en compte : La localisation du signal est une composante importante qu'il est difficile de quantifier sur des données réelles. Des jeux de simulations ne seraient pas beaucoup plus intéressants car il serait difficile d'en valider la pertinence biologique.

Pour répondre à ces interrogations, nous proposons de conserver comparable la résolution spatiale des données et d'évaluer les résultats de manière semi-empirique étant donné la nature très différente des deux techniques. Nous présentons trois points de vue rendant possible la comparaison : géométrique, géométrique-statistique et statistique.

C₁ : Point de vue géométrique :

Une première solution consiste à faire se correspondre une mesure de la résolution spatiale des données. Typiquement, après filtrage spatial, on peut négliger la corrélation spatiale originelle des images. La largeur totale à mi-hauteur (FWHM *Full Width at Half Maximum*) du noyau gaussien est une mesure qui correspond à la distance avec laquelle il faut séparer deux sources ponctuelles pour les distinguer après filtrage. On peut donc raisonnablement calculer un nombre de parcelles correspondant au volume cérébral divisé par cette mesure de résolution (FWHM³ pour un volume). Pour un volume de FWHM³ mm³, on obtient une parcelle, ce qui permet

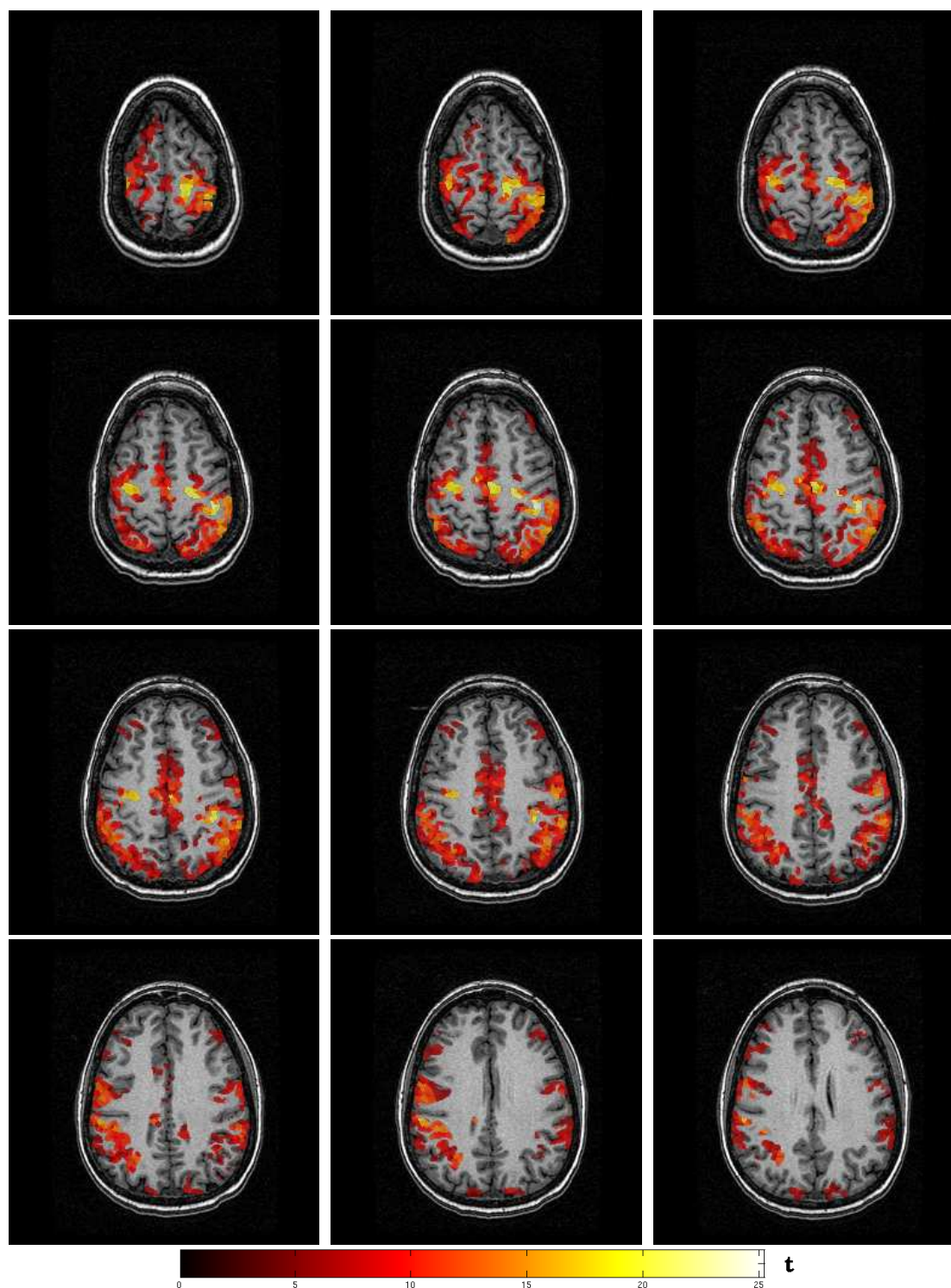


FIG. 6.4 – Coupes IRM T_1 axiales avec les activations détectées superposées ($p_c < 0.05$) pour la tâche de grasping en utilisant une parcellisation en 4000 parcelles. Les coupes sont représentées de la partie dorsale vers la partie ventrale toutes les 3 mm (de gauche à droite et de haut en bas).

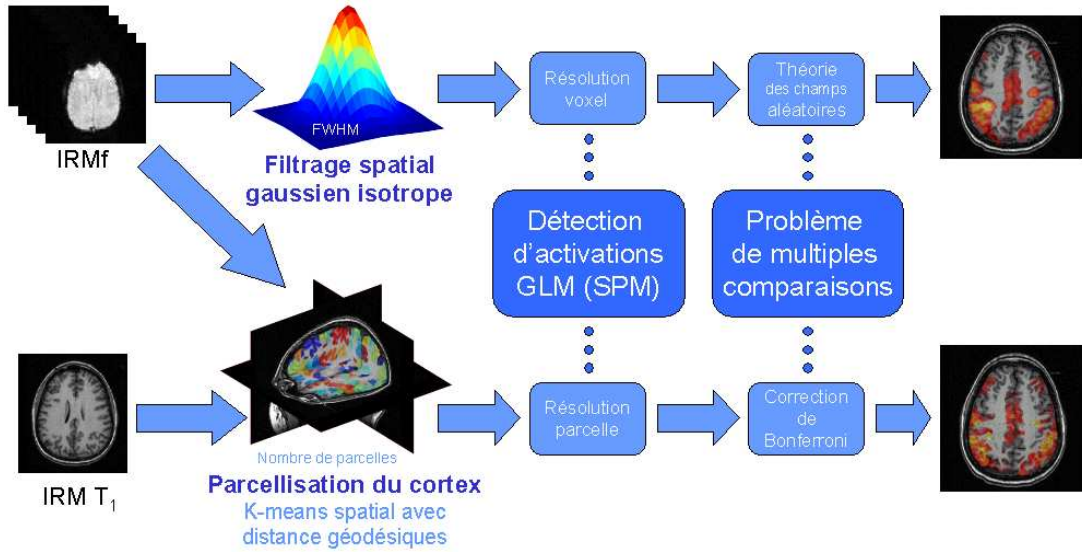


FIG. 6.5 – Schéma récapitulatif de l'application de la parcellisation du cortex pour la détection d'activation et la comparaison avec l'approche voxelique.

une interprétation aisée du nombre de parcelles.

$$N = \frac{V}{\prod_i FWHM_i}$$

C₂ : Point de vue géométrie-statistique :

Dans l'analyse des séries fonctionnelles à partir de la théorie des champs aléatoires, la résolution spatiale n'est pas calculée à partir des valeurs du filtre imposé aux données mais elle est mesurée sur les images "résiduelles" d'où le signal modélisé a été soustrait d'une part, et qui ont été normalisées par la variance estimée localement. Cette résolution aurait donc une interprétation géométrique mais ayant pris en compte le signal et le bruit estimé. Cette mesure correspond exactement à la notion d'éléments de résolution (*RESELS RESolution ELeMents*) introduite par Keith Worsley dans [Worsley *et al.*, 1992] (volume cérébral divisé par le produit des FWHM estimées sur les champs résiduels normalisés). On peut à partir de cette valeur définir le nombre de degrés de liberté spatial effectif ν [Worsley *et al.*, 1998].

$$\nu = RESELS \left(4 \log_e \frac{2}{\pi} \right)^{\frac{D}{2}}$$

Une solution pour comparer voxel et parcelle consiste à prendre comme nombre de parcelles N le nombre de degré de liberté spatial effectif estimé sur les résidus des images lissées avec un filtre de taille FWHM. Une des difficultés de cette méthode est que du fait de la mauvaise estimation de la variance du bruit localement (voxel à voxel), ces images (volumes) résiduelles peuvent apparaître moins lisses qu'en

réalité. Ceci donnerait lieu à un nombre de parcelles important.

C₃ : Point de vue statistique :

Une troisième approche consiste à utiliser la méthode de correction multiple pour fixer le nombre de parcelles de manière à ce qu'il corresponde à un nombre de tests indépendants. Typiquement, un seuil de 5% corrigé (α_c) pour les comparaisons multiples est choisi par la théorie des champs aléatoires et on observe la valeur de Z qui lui correspond. A cette valeur on peut associer une probabilité non corrigée (noté α). Le nombre de parcelles N est fixé de manière que $\alpha_c = 1 - (1 - \alpha)^{1/N}$, c'est-à-dire :

$$N = \frac{\log(1 - \alpha)}{\log(1 - \alpha_c)}$$

Dans la section suivante, nous présentons les résultats expérimentaux obtenus avec ces trois types de comparaison pour plusieurs valeurs de filtres gaussiens.

Résultats

Les images fonctionnelles ont été analysées avec des filtres gaussiens de tailles 8, 12 et 16 mm de FWHM. Les résultats se sont avérés similaires du point de vue de la comparaison lissage versus parcellisation pour ces différentes valeurs : on présente donc ceux obtenus pour une FWHM de 8mm (moins de perte de résolution). On trouve pour un filtre de 8mm une résolution équivalente de 1700 parcelles avec C_1 , 340 avec C_2 et 4900 avec C_3 . On note donc que le critère C_3 donne lieu à la parcellisation la plus fine (le plus grand nombre de parcelles), ce qui reflète la nature assez conservatrice des tests obtenus par la théorie des champs aléatoires. Le critère C_2 montre que les résidus normalisés sont (dans ce cas) bien plus dépendants que les données originelles. Cette résolution n'est pas très simple à interpréter. Enfin, C_1 fournit une valeur intermédiaire facilement interprétable vis-à-vis du filtre gaussien appliqué aux données.

La figure 6.6 montre les résultats avec ces trois valeurs du nombre de parcelles.

Elle montre que dans les trois cas la sensibilité est égale ou supérieure à celle obtenue par la méthode classique. Celle-ci est d'autant meilleure que le nombre de parcelles est petit et la correction pour les comparaisons multiples peu sévère. En particulier, on note l'augmentation de sensibilité dans les parties frontales antérieures et dans les parties plus latérales du cortex. Pour C_1 et C_2 les maxima globaux des cartes en t sont respectivement de 28.18 et 30.19, à comparer avec la valeur de 23.25 en utilisant SPM.

D'autre part, même s'il est impossible de comparer les localisations de façon certaine sur les données réelles, il est clair que la localisation des maxima locaux sur les données filtrées est influencée par le volume partiel et par le recouvrement des signaux dans les structures sous-jacentes différentes. La localisation obtenue par la méthode de parcellisation montre des signaux mieux placés vis à vis des résultats escomptés dans ce paradigme (e.g. le sillon précentral).

En se fondant sur ces premiers résultats, il semble que le critère C_1 soit un compromis acceptable pour trouver le nombre équivalent de parcelles étant donné un lissage spatial.

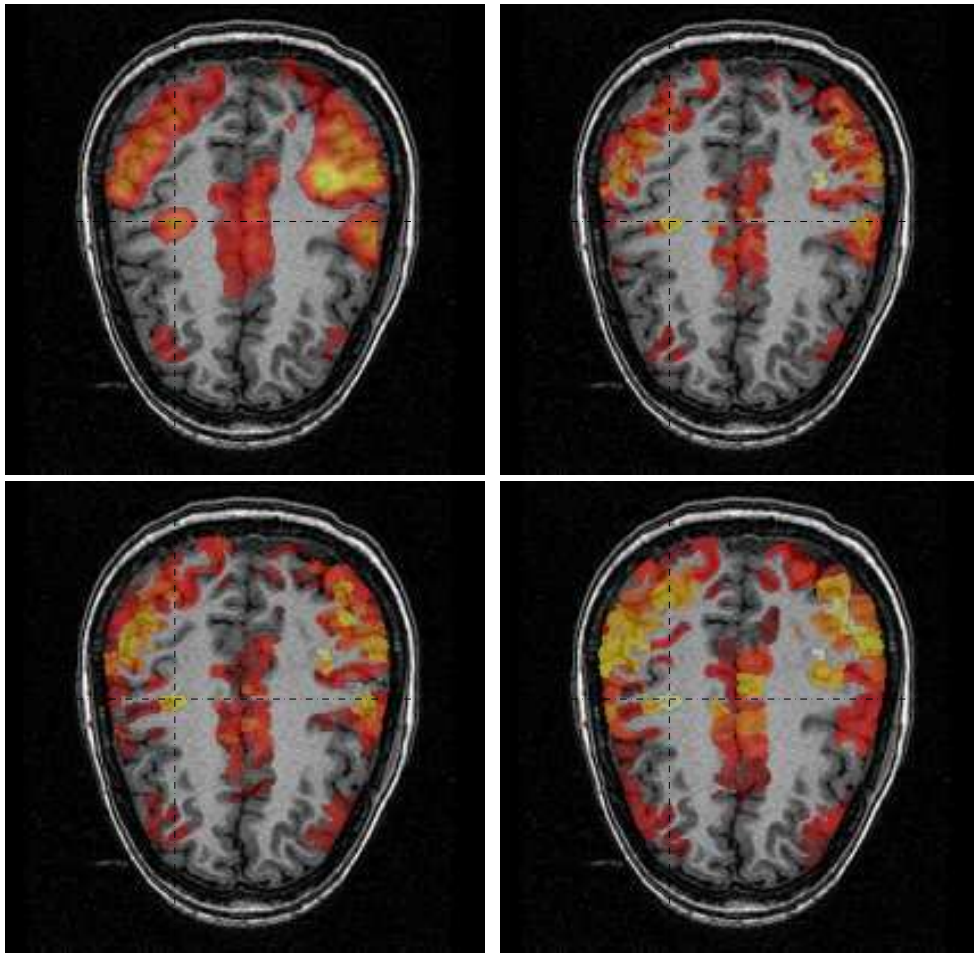


FIG. 6.6 – Coupe IRM pondérée en T_1 axiale avec les activations détectées superposées ($p_c < 0.05$). Carte en T d'après SPM avec 8mm de lissage (en haut à gauche). Carte en T par la méthode parcelles avec 4900 parcelles (en haut à droite), 1700 parcelles (en bas à gauche) et 340 parcelles (en bas à droite).

Ce choix ne se fonde pas sur des considérations théoriques mais sur les aspects de détection et de localisation observés ici.

La figure 6.7 présente des cartes SPM et des parcellisations avec trois niveaux de résolution aux mêmes seuils statistiques ($p_c < 0.05$). On observe d'une part que la sensibilité cesse de croître pour SPM quand le lissage devient trop important et que ce phénomène apparaît plus tardivement avec le principe de parcellisation.

6.2.5 Discussion

Une comparaison plus directe et qui semble, de prime abord, plus juste aurait pu être faite avec la technique *Cortical Surface Mapping* (CSM) car celle ci prend aussi en compte l'anatomie du cortex. Cependant, plusieurs arguments plaident en faveur de la comparaison que nous avons proposée. D'une part il est difficile de comparer les techniques 2D et 3D (prise en compte des structures internes, relation entre un volume et

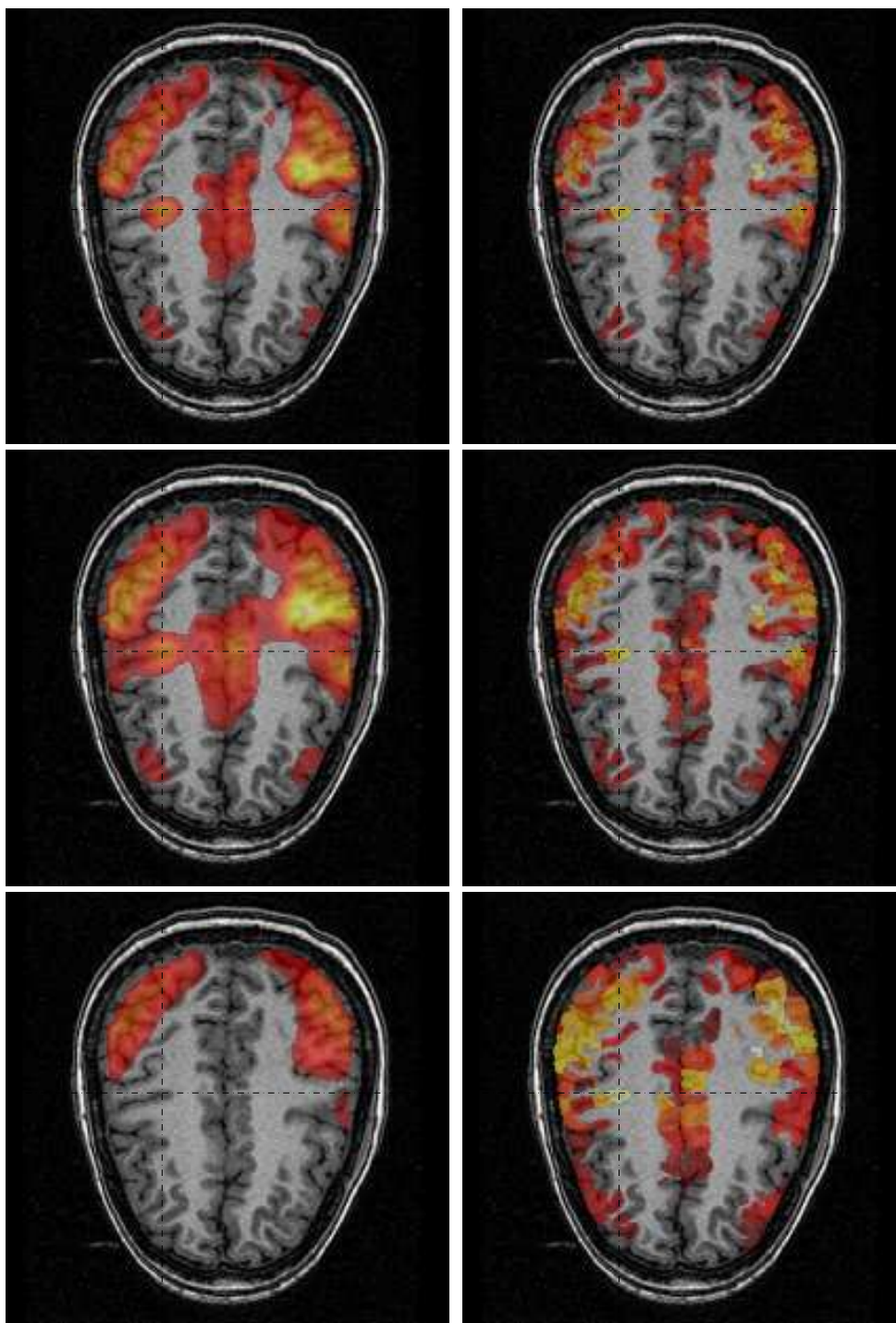


FIG. 6.7 – Coupe IRM T_1 axiale avec les activations détectées superposées ($p_c < 0.05$). Carte en T d'après SPM avec 8mm de lissage (en haut à gauche), 12 mm de lissage (milieu à gauche) et 16mm de lissage (en bas à gauche). Carte en T d'après la méthode parcelle avec 4900 parcelles (en haut à droite), 1700 parcelles (au milieu à droite) et 340 parcelles (en bas à droite).

une surface), d'autre part il est intéressant de se comparer aux techniques effectivement utilisées en standard.

Bien que proche dans l'esprit des travaux antérieurs d'Andrade et Kiebel ([Andrade *et al.*, 2001; Kiebel *et al.*, 2000]), il semble que, d'une part, la technique soit plus facile d'utilisation car incluant directement l'ensemble des structures contenant de la matière grise. Les gains de sensibilité observés sont meilleurs que ceux observés avec CSM. Les techniques reposent toutes deux sur un bon ajustement signal anatomique T_1 et signal fonctionnel T_2^* mais la parcellisation est plus directe (évite les problèmes de maillage) et est de manière générale construite sur un principe plus robuste, moins directement sujette aux artefacts de mouvement et de distorsion. La localisation, en revanche, devrait être plus précise avec CSM. De plus, comme nous allons le voir dans les chapitres suivant, la parcellisation peut se généraliser par l'inclusion d'information fonctionnelle.

6.3 Sélection de modèle par parcelle

Les travaux présentés dans ce paragraphe ont été réalisés en collaboration avec Alexis Roche (CEA/SHFJ) et ont fait l'objet des publications [Roche *et al.*, 2003; Kherif *et al.*, 2002c].

6.3.1 Motivations

Les données fonctionnelles sont classiquement analysées dans un but de détection d'activation en définissant un modèle *global* constitué de la prédiction d'une réponse BOLD. Le modèle est global dans le sens où les signaux observés en différents endroits du cerveau sont considérés comme étant diverses réalisations d'un même modèle. Ainsi, dans le cadre statistique du modèle général linéaire, la réponse BOLD est modélisée en tout voxel comme la combinaison linéaire d'un nombre prédéfini de régresseurs, constituant la matrice de plan expérimental.

Les motivations poussant à l'utilisation d'un modèle global reposent sur des considérations pratiques de temps de calcul mais également sur des considérations statistiques liées à la correction du problème de multiples comparaisons par la théorie des champs aléatoires gaussiens qui ne pourrait s'appliquer avec un modèle différent par voxel.

Le choix du modèle *global* est alors important car il va conditionner la sensibilité de l'analyse l'utilisant. Le choix de l'utilisateur se porte souvent sur un modèle flexible constitué d'un nombre élevé de régresseurs afin de pouvoir modéliser l'incertitude sur la variabilité de la réponse BOLD. En effet, non seulement cette réponse est mal connue mais il est très probable qu'elle varie spatialement en fonction des aires cérébrales. Les conséquences de l'utilisation d'un tel modèle global sont des pertes d'efficacité statistique et de sensibilité de détection.

Nous nous proposons ici de mettre en évidence la pertinence de l'utilisation de modèles régionaux, distincts selon les aires cérébrales considérées.

6.3.2 Une modélisation régionale des données

Le modèle linéaire général appliqué au décours temporel issu du voxel k s'écrit :

$$Y_k = X_k \beta_k + \epsilon_k$$

Remarquons spécifiquement l'indice k associé à la matrice de plan expérimental, mettant en évidence que le modèle utilisé n'est pas identique pour l'ensemble des voxels analysés.

La difficulté réside dans la définition du modèle local X_k . Une première hypothèse simplificatrice peut être que le signal BOLD a une certaine régularité spatiale et qu'il n'est ainsi pas nécessaire de définir un modèle par voxel mais plutôt un modèle régional. Une parcellisation anatomique permet de définir une partition de l'espace en un nombre donné d'éléments (inférieur au nombre de voxels considérés) pour lesquels on peut définir séparément un modèle.

Le choix d'un *bon* modèle est un problème délicat. [Kherif, 2003] détaille les principales propriétés devant être vérifiées : validité, parcimonie et simplicité. Les caractéristiques de la réponse BOLD associée à une régions n'étant pas connues a priori, cette incertitude préconise l'utilisation d'un modèle large et flexible afin de ne pas omettre la modélisation d'une partie du signal. Cependant, si le modèle utilisé est trop vaste, une partie du bruit contenu dans le signal va être expliqué par le modèle ce qui va biaiser les estimations. Par ailleurs, un nombre élevé de régresseurs nécessite de recourir à un test F, plus souple mais moins sensible qu'un test T et également plus difficilement interprétable. Un *bon* modèle peut se définir comme un modèle contenant des variables explicatives ajustant au mieux le signal observé et dont le nombre est réduit. Une stratégie envisageable pour construire un tel modèle consiste à effectuer une sélection de modèle.

Une procédure de sélection de modèle a pour objectif de passer d'une modélisation de la forme $Y = X_L \beta + \epsilon$ à une autre déterminée par $Y = X_N \beta' + \epsilon'$ avec $\text{rang}(X_N) < \text{rang}(X_L)$. Autrement dit, il s'agit de trouver un sous-espace à l'espace engendré par les colonnes de X_L , permettant de modéliser fidèlement les données. Nous avons utilisé la procédure décrite dans [Kherif *et al.*, 2002c] et représentée sur la figure 6.8.

Disposant de Y contenant l'ensemble des signaux temporels contenus dans une parcelle et de X_L une matrice du plan expérimental contenant un modèle large et flexible, la sélection de modèle proposée par [Kherif *et al.*, 2002c] consiste à étudier l'inter-dépendance entre les données et le modèle et à sélectionner les composantes les plus significatives d'une décomposition en valeurs singulières de la matrice $X_L' Y$ normalisée : c'est la méthode MLM (Modèle Linéaire Multivarié).

$$(X_L' \Sigma X_L)^{-\frac{1}{2}} X_L' Y \Xi^{-\frac{1}{2}} = U S V'$$

où Σ représente la matrice de covariance temporelle et $\Xi = \text{diag}(\epsilon' \epsilon)$ la métrique sur l'espace des voxels contenus dans la parcelle. Les décours temporels propres correspondant à cette décomposition se déduisent alors avec la formule suivante :

$$Z_x = X_L (X_L' X_L)^{-\frac{1}{2}} U'$$

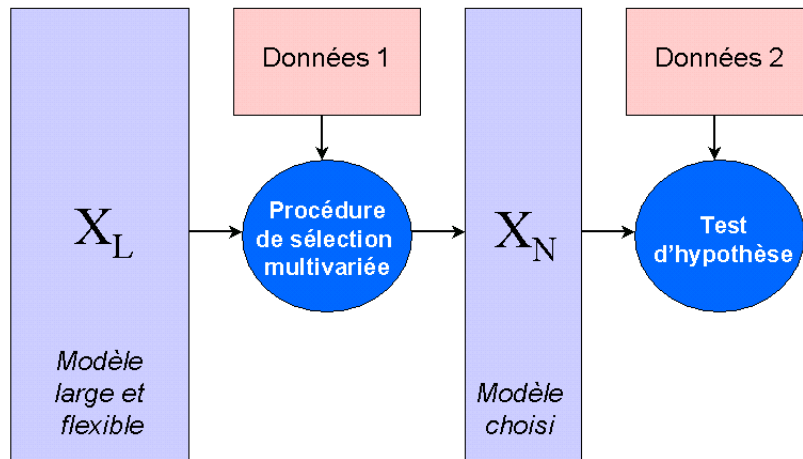


FIG. 6.8 – Diagramme de la procédure de sélection de modèle (d’après [Kherif et al., 2002c]). Un modèle large et flexible X_L est tout d’abord défini. Puis un sous-modèle X_N est sélectionné selon une approche multivariée en utilisant un jeu de données correspondant à une base d’apprentissage. Le reste des données est alors analysé en utilisant le modèle X_N .

La sélection de modèle consiste alors à construire un nouveau modèle X_N constitué des p premiers décours temporels propres, avec $p < \text{rang}(X_L)$. Si l’on effectue cette opération au niveau de chaque parcelle avec un premier jeu de données (une base d’apprentissage), on obtient un ensemble de modèles $\{X_N\}$ optimisés pour les parcelles pour lesquelles ils ont été définis, qui peuvent alors être utilisés comme modèles de signal dont on cherche la présence dans un second jeu de données.

6.3.3 Application sur un jeu de données

Cette approche a été testée sur un des jeux de données détaillé dans le paragraphe 6.2.2 et étudiant les aires fonctionnelles impliquées dans une tâche de calcul mental. Ce jeu de données est constitué de deux sessions de 93 images, ce qui permet d’utiliser une session comme base d’apprentissage et la seconde comme jeu de données effectivement utilisé pour l’établissement de cartes statistiques.

Un modèle constitué de trois régresseurs est défini ; il est constitué des trois décours temporels présentés sur la figure 6.3 : le paradigme convolué par un modèle de réponse hémodynamique canonique et ses dérivées premières et secondes. Cette modélisation assez lâche permet de prendre en compte une incertitude sur la forme exacte de la réponse hémodynamique. L’inconvénient d’un tel modèle est que, pour interroger les données, il faut alors utiliser un test F au lieu d’un test T, responsable d’une perte de sensibilité.

La procédure de sélection de modèle est alors appliquée en chaque parcelle sur la session base d’apprentissage, en ne sélectionnant qu’un seul régresseur expliquant le maximum de variance dans les données. La figure 6.9 présente deux exemples de signaux sélectionnés dans deux parcelles distinctes. Les résultats obtenus suggèrent que la forme de la

réponse BOLD peut être très différente d'une région à l'autre. En particulier, on observe une variabilité importante dans l'échelle temporelle de la réponse, ce qui suggère que la dynamique de l'effet BOLD est dépendante de la région étudiée. Plusieurs phénomènes peuvent expliquer cela : des différences intrinsèques de la dynamique de l'activité neuronale ou des hétérogénéités vasculaires qui n'affectent que le processus hémodynamique. Dans l'exemple présenté sur la figure 6.9, les deux parcelles choisies correspondent à des parcelles identifiées comme ayant des valeurs élevées et similaires pour un test F effectué avec le modèle global. Les deux modèles de réponse sélectionnés présentent des contenus fréquentiels très différents, soulignant les importantes variabilités que peut présenter le signal BOLD selon la région dans laquelle il est mesuré.

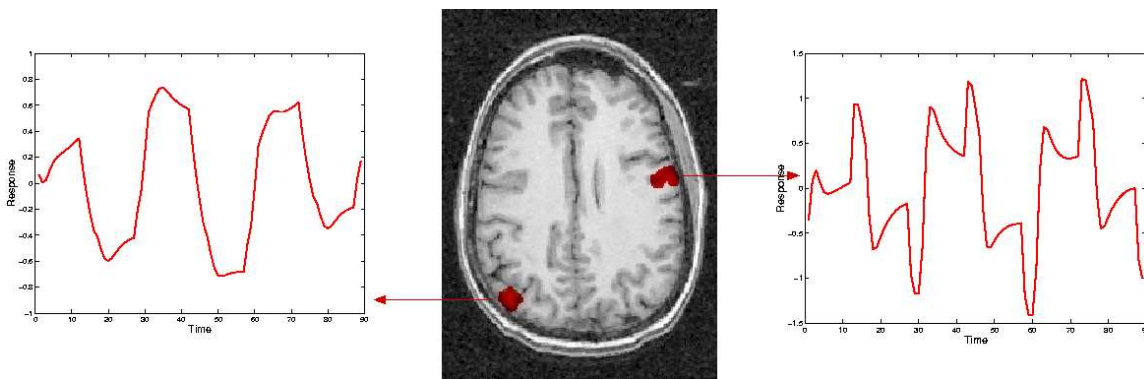


FIG. 6.9 – Exemples de régresseurs sélectionnés par la procédure MLM dans deux parcelles situées dans le lobe fronto-latéral droit (impliqué dans diverses tâches attentionnelles) et le lobe pariétal gauche (impliqué dans les tâches de calcul).

Par ailleurs, quand les modèles régionaux sélectionnés avec la base d'apprentissage sont utilisés pour analyser la deuxième session, on observe une augmentation de la sensibilité pour la détection des activations. Il apparaît donc pertinent de remettre en cause l'hypothèse selon laquelle la réponse dynamique est identique en tout point du cerveau et sa conséquence que représente l'utilisation un modèle global appliqué à l'ensemble des voxels, pour se diriger vers des approches régionales permettant de spécifier des modèles distincts selon l'aire cérébrale étudiée.

6.4 Mesures robustes de la connectivité fonctionnelle

Les travaux présentés dans ce paragraphe font référence aux publications suivantes [Lahaye *et al.*, 2003a; 2003b] dont le principal investigateur est Pierre-Jean Lahaye (CEA/SHFJ).

6.4.1 La connectivité fonctionnelle

Les études de connectivité fonctionnelle ont pour objet la compréhension des relations entre les régions du cerveau, dépendant par exemple de l'état induit par le paradigme expérimental. Les applications de ces techniques sont nombreuses et concernent par exemple l'étude de la plasticité neuronale par apprentissage ou les altérations de la connectivité entre différentes populations (sujets sains versus patients). La plupart des méthodes proposées dans la littérature reposent sur l'étude de la corrélation ou de la matrice de covariance des données [Biswal *et al.*, 1995]. L'hypothèse sous-jacente effectuée est que l'information significative concernant les interactions entre les différentes régions du cerveau est reflétée par une relation linéaire entre les valeurs des deux signaux aux mêmes instants temporels. Les travaux présentés dans [Lahaye *et al.*, 2003a] visent à étudier la pertinence de cette hypothèse en quantifiant l'importance des interactions non-linéaires ou asynchrones entre deux signaux BOLD.

6.4.2 Problème de dimensionnalité

Une des difficultés concernant les études de connectivité est la quantité d'information à traiter. En effet, si l'on souhaite étudier toutes les interactions entre tous les voxels fonctionnels dans le cerveau (au nombre de N), cela représente $\frac{N(N-1)}{2}$ interactions dans le cas d'une mesure symétrique comme le coefficient de corrélation et N^2 sinon. Il est ainsi difficile d'étudier la connectivité entre toutes les paires de décours temporels associés aux voxels. Qui plus est, la corrélation spatiale entre voxels voisins entraîne une corrélation temporelle entre les signaux correspondant, ce qui a une incidence sur les études de connectivité [Kiebel *et al.*, 1999].

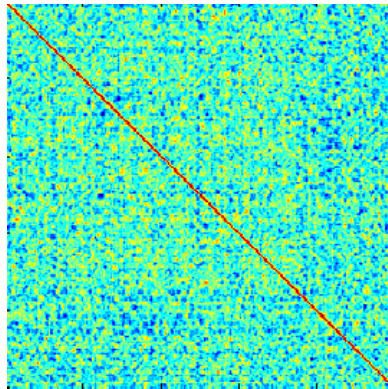


FIG. 6.10 – *Exemple de matrice de connectivité fonctionnelle basée sur la corrélation d'un nombre restreint de signaux BOLD. Il est important de noter que la matrice carrée ainsi construite est de la dimension du nombre de signaux considérés, qui peut ainsi être très importante si l'on considère l'ensemble des signaux fonctionnels mesurés dans le cerveau : environ 1.5 Go pour une matrice 20000×20000 ! Une telle matrice ne peut alors être stockée dans la mémoire d'un ordinateur standard.*

Une solution fréquemment utilisée pour réduire la dimensionnalité du problème

consiste à sélectionner manuellement un nombre restreint de voxels (ou de régions d'intérêt) et à étudier la connectivité des seuls décours temporels associés. Cette étape n'est pas entièrement satisfaisante car la sélection manuelle des voxels/régions d'intérêt peut introduire un biais dans l'analyse et s'avère de plus coûteuse en temps si l'on souhaite échantillonner tout le cortex. Par ailleurs, le signal BOLD associé à des voxels isolés peut être perturbé par du bruit, ce qui diminue la sensibilité de l'analyse qui en découle.

La parcellisation du cortex est ici une alternative intéressante permettant d'obtenir une nouvelle représentation des données fonctionnelles à une résolution intermédiaire quelque entre le voxel et la région d'intérêt. Le nombre de parcelles n'a donc pas ici de sens précis (il n'y a pas de nombre optimal) mais permet plutôt de définir la résolution effective à laquelle les données sont analysées. La construction de la parcellisation est entièrement automatique et permet de prendre en compte la totalité du cortex.

6.4.3 Etude de la nature des interactions entre parcelles

Dans l'étude présentée dans [Lahaye *et al.*, 2003a], une centaine de parcelles ont été définies par cerveau, pour un volume moyen de 6.3 cm³. C'est une résolution spatiale assez grossière mais qui permet d'être robuste face aux différents artefacts (distorsions, mouvements, couplage vasculaire). Des décours temporels sont ensuite assignés à chacune des parcelles en suivant la méthode proposée dans le paragraphe 6.2.3.

Considérons deux parcelles a et b associées aux signaux BOLD $x_a(t)$ et $x_b(t)$ et étudions les différentes formes possibles des interactions de la parcelle b sur la parcelle a , notées $b \rightarrow a$. Chaque hypothèse effectuée sur la nature des interactions peut s'exprimer par un certain nombre de régresseurs au sein du modèle général linéaire :

- *terme d'interaction linéaire instantanée* : ce terme correspond à $\{x_b(t)\}$ et modélise les hypothèses formulées avec l'utilisation de la corrélation,
- *terme d'interaction non-linéaire instantanée* : l'influence non-linéaire de $x_b(t)$ sur $x_a(t)$ peut être modélisée par un ensemble de polynômes $\{x_b(t)^k\}_{1 \leq k \leq n_{NL}}$,
- *terme d'interaction linéaire retardée* : il s'agit ici de modéliser l'influence des valeurs passées de $x_b(t)$ sur $x_a(t)$. Ceci peut s'écrire par un ensemble de régresseurs de la forme $\{x_b(t - l\tau)\}_{1 \leq l \leq n_{DT}}$,
- *terme d'interaction non-linéaire retardée* : la combinaison des effets de non-linéarité et de décalages temporels s'expriment par les termes $\{x_b(t - l\tau)^k\}_{1 \leq l \leq n_{DT}, 1 \leq k \leq n_{NL}}$.

Le signal $x_a(t)$ peut alors être décomposé en plusieurs termes selon le modèle linéaire général :

- *un terme déterministe* correspondant à la modélisation du paradigme,
- *des termes autorégressifs* modélisant la corrélation temporelle du signal $x_a(t)$, c'est-à-dire $\{x_a(t - l\tau)\}_{1 \leq l \leq n_{DT}}$,
- *des termes d'interactions avec $x_b(t)$* : ce sont les termes explicités ci-dessus,
- *un résidu* supposé être un bruit blanc gaussien.

La figure 6.11 de gauche présente le pourcentage moyen de variance expliqué par chacun de ces 4 facteurs pour l'ensemble des signaux extraits des parcelles de 6 sujets ayant participé au protocole de [Simon *et al.*, 2002].

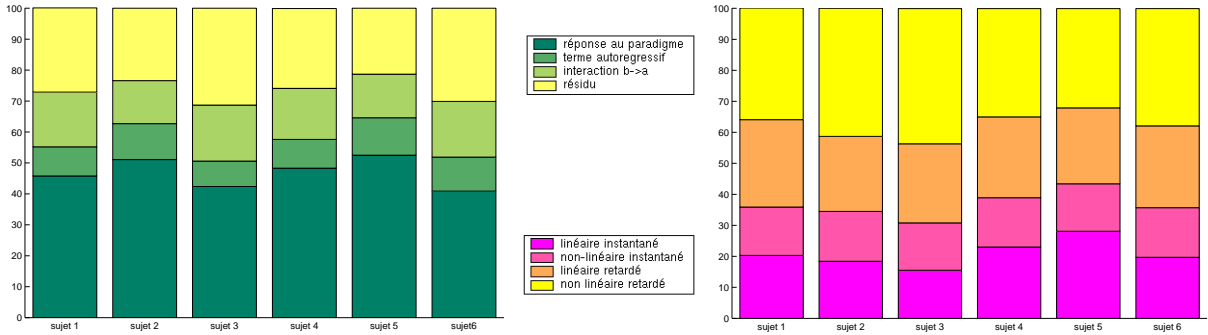


FIG. 6.11 – *Etude de l'importance des différents termes d'interactions estimés sur 6 sujets. (à gauche) Pourcentage moyen de la variance expliquée par les différents régresseurs du modèle. (à droite) Pourcentage moyen de la variance expliquée par chaque type d'interaction, normalisée par la variance totale de l'ensemble des interactions.*

On remarque que les proportions des divers termes sont assez reproductibles entre sujets, les interactions $b \rightarrow a$ représentant en moyenne 10% de la variance totale des données. La figure 6.11 de droite indique alors comment les différents termes d'interaction se décomposent selon leur origine. En particulier, on remarque que les termes non-linéaires et retardés expliquent une part non négligeable de l'influence de b sur a . Il apparaît ainsi que la corrélation, qui ne teste que pour les interactions linéaires instantanées et ne capture pas l'information provenant de non-linéarité ou de retards temporels, n'est pas le meilleur critère pour étudier la connectivité fonctionnelle par signal BOLD.

Une étude plus fine de ces différents termes montre que le terme de retard linéaire joue un rôle important dans la modélisation de la connectivité fonctionnelle.

6.4.4 Une nouvelle mesure de connectivité

[Lahaye *et al.*, 2003a] ont donc proposé une nouvelle mesure de connectivité $\mathcal{F}_{b \rightarrow a}$ prenant en compte l'histoire temporelle du signal b , à la différence de la corrélation. Cette mesure correspond à un test F, testant pour la significativité du modèle de dépendance temporelle historique de b sur a par rapport à un modèle autorégressif de a .

Pour tester cette nouvelle mesure, une parcelle de référence est choisie au maximum d'activité pour un sujet (la carte d'activation correspondante est présentée sur la figure 6.12.a.). La carte de connectivité associée à cette parcelle avec une mesure de corrélation est disponible sur la figure 6.12.b tandis que celle obtenue avec la nouvelle mesure \mathcal{F} est présentée en 6.12.c. On note une augmentation de la sensibilité avec la mesure \mathcal{F} .

6.4.5 Discussion

Les études menées sur les signaux BOLD extraits d'une parcellisation anatomique montrent que l'histoire temporelle des signaux joue un rôle important pour évaluer la connectivité fonctionnelle entre des régions du cerveau. Une nouvelle mesure de connecti-

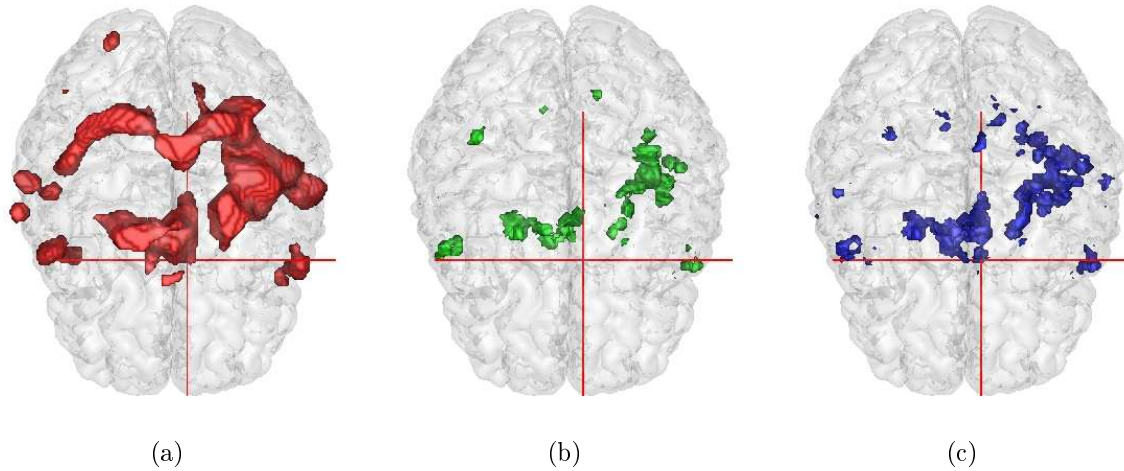


FIG. 6.12 – (a) Carte d’activation pour la tâche de grasping avec un risque $\alpha = 10^{-10}$ non corrigé. (b) et (c) Cartes de connectivité construites pour un seuil $\alpha = 10^{-3}$ avec pour mesure de connectivité la corrélation C (b) et la nouvelle mesure proposée \mathcal{F} (c).

tivité permettant de prendre en compte ces termes a été proposée et les cartes de connectivité obtenues démontrent que cette mesure a une meilleure sensibilité que la corrélation. Les valeurs passées (strictement inférieures à t) du signal d’une région peuvent ainsi partiellement prédire la réponse d’une autre aire à l’instant t , indépendamment du passé de cette aire.

6.5 Conclusion

Nous avons présenté dans ce chapitre trois applications faisant intervenir une parcellisation anatomique, telle que définie dans le chapitre précédent. Les motivations sous-jacentes à l’utilisation d’une parcellisation sont variées. Dans la première application, la parcellisation permet d’effectuer un lissage spatial des données utilisant une information anatomique évitant de moyennner des signaux intervenant dans des tissus distincts comme le fait un filtrage gaussien. Ainsi, en ce cas, une parcellisation anatomique apparaît comme une technique de restauration des données bruitées. Le bon comportement obtenu en détection d’activation (augmentation de la sensibilité) montre que malgré la diminution de la résolution effective des données, la parcellisation permet d’obtenir une représentation compacte et concise des données, en limitant la perte d’informations pertinentes. C’est cette qualité qui est utilisé dans les deux autres applications. Dans le cas de la sélection de modèle régional, la parcellisation permet de construire une partition homogène de l’espace sur laquelle peuvent être conduits des analyses indépendantes. Ainsi, nous avons pu mettre en évidence la pertinence de l’idée de remettre en cause la notion de modélisation globale des données pour favoriser la mise en œuvre de techniques permettant de définir des modèles variant selon la position spatiale. Enfin, la parcellisation anatomique permet de pouvoir tester de nouvelles techniques sur les données en restreignant la quantité de données à analyser. Ceci a été mis en évidence pour l’étude de connectivité que nous

avons présentée. Typiquement, ce type d'analyse suppose de comparer tous les signaux fonctionnels deux à deux, ce qui s'avère quasiment irréalisable en pratique. Une parcellisation propose alors un moyen simple et efficace pour réduire la quantité d'information à traiter.

Il y a cependant une hypothèse qui est formulée lors de l'assignation d'un signal fonctionnel à une parcelle en moyennant les signaux contenus dans celle-ci, qui n'est pas forcément vérifiée si les parcelles sont relativement volumineuses. En effet, cette méthode d'assignation n'est valide que si les parcelles ont des signaux homogènes au sein de chacune d'elle. Plus précisément, l'hypothèse qui est faite est que les signaux observés au sein d'une parcelle ne sont que des réalisations bruitées d'un même modèle, correspondant alors à la moyenne. La technique de parcellisation que nous avons proposée n'impose pas explicitement de vérifier cette hypothèse par construction, mais elle s'avère vérifiée en pratique si le nombre de parcelles est suffisamment élevé. Nous proposons dans le chapitre suivant un nouveau modèle générateur pour les données fonctionnelles qui prend en compte ces considérations. Par ailleurs, la parcellisation anatomique se prête difficilement à son utilisation dans des analyses multisujets ; c'est un autre point sur lequel nous allons revenir dans le chapitre 7.

Chapitre 7

Modélisation anatomo-fonctionnelle des données IRMf

Sommaire

7.1	Motivations	160
7.2	Un modèle de mélange spatio-temporel	160
7.2.1	Le modèle mathématique général	161
7.2.2	Adaptations pratiques	162
7.3	Exemples d'applications	164
7.3.1	Classification anatomique pure	164
7.3.2	Classification fonctionnelle pure	164
7.3.3	Classification spatio-temporelle	167
7.4	Extension multi-sujets	172
7.4.1	Motivations et modèle proposé	172
7.4.2	Classification fonctionnelle multi-sujets	173
7.4.3	Modèle anatomo-fonctionnel multi-sujets	174
7.5	Discussion	175

7.1 Motivations

Dans le chapitre 5, nous avons présenté un algorithme de parcellisation anatomique du cortex en un nombre ajustable d'éléments, reposant sur un critère de compacité spatiale. L'utilisation de ces parcellisations au chapitre 6 pour l'analyse de données IRMf fait intervenir l'affectation de signaux fonctionnels à chaque parcelle. La solution que nous avons retenue consiste à moyenniser les signaux contenus dans chaque parcelle, ce qui suppose une homogénéité du signal au sein de chaque parcelle. Cette contrainte n'étant pas incluse lors de la construction de la parcellisation (reposant sur des considérations uniquement géométriques), une extension naturelle consiste à introduire l'information fonctionnelle lors de la définition de la parcellisation.

Christopher Genovese évoque cette solution dans [Genovese, 1999], dont voici un extrait :

Modeling [spatial relationships among the model parameters associated with different voxels] increases the precision of inferences about the parameters because multiple voxels contribute information about features they have in common. [...] The challenge is to find the “dependence neighborhoods” over which we can borrow strength in estimating the model parameters.

One straightforward approach involves segmenting each image into grey-matter (cortex), white matter, and blood vessel voxels using the baseline (time-averaged) signal magnitude along with physical constraints. This is motivated by the different imaging properties of these tissue types. Such a categorization does not depend on the task-related signal changes, but it is still a somewhat computationally intensive procedure. A plausible dependence neighborhood for a given voxel could then be defined by taking nearby voxels in the same category.

A better approach is to let the data (particularly the temporal changes in the signal) determine what the neighborhoods should be. This requires an adaptive partitioning of the voxels that is tied to the temporal structure of the signal at each voxel. The partition consists of sets of contiguous voxels whose union is the entire image. Voxels in the same partition element would share the response amplitude and shape parameters.

La parcellisation anatomique que nous avons proposée permet une implémentation efficace de la première approche évoquée par Genovese. Dans ce chapitre, nous proposons d'apporter une réponse à la seconde approche, en étendant le modèle précédent. Pour ce faire, nous avons choisi de nous placer dans un cadre spatio-temporel permettant d'utiliser les outils théoriques présentés dans le chapitre 4. Nous présentons le modèle choisi dans le paragraphe suivant, suivi de plusieurs illustrations mettant en évidence sa pertinence. Enfin, nous montrons que le formalisme utilisé par ce modèle se généralise aisément dans le cadre d'analyses multi-sujets.

7.2 Un modèle de mélange spatio-temporel

Le concept de parcellisation anatomo-fonctionnelle s'explique aisément dans le cadre de classification de données et plus particulièrement par les approches de type *agrégation*.

tion : il s'agit de regrouper les voxels du cerveau qui ont à la fois une position spatiale proche et dont l'activation présente un décours temporel similaire. Bien sûr, il faut définir précisément le sens de la notion de *similarité* spatiale et fonctionnelle.

7.2.1 Le modèle mathématique général

Un modèle de mélange de K gaussiennes de paramètres $\theta_k = \{\mu_k, \Sigma_k, \pi_k\}$ (cf section 4.2.2) dans l'espace joint signal et position apporte un formalisme à cette idée générale. Pour un voxel i situé en la position v_i et ayant pour paramètres de régression β_i (associés à une matrice de plan expérimental X au sein du modèle linéaire général), on a :

$$p(\beta_i, v_i | \Theta) = \sum_{k=1}^K p(k) p(\beta_i, v_i | k; \theta_k) \quad \text{avec} \quad p(\beta_i, v_i | k; \theta_k) = G_{\mu_k, \Sigma_k}(\beta_i, v_i)$$

Ce modèle associe donc explicitement les coordonnées spatiales aux valeurs fonctionnelles, en considérant un échantillon comme étant la concaténation des coordonnées du voxel dans l'espace avec les coefficients de régressions β du décours temporel. Le choix de cet espace (*feature space*) dans lequel effectuer la classification repose sur deux considérations :

- d'un point de vue fonctionnel, il s'agissait de choisir quelle information utiliser. L'utilisation directe des décours temporels nécessite de travailler dans des espaces très grands, ce qui pose des problèmes d'estimation. Le choix de projeter les données sur une base est alors une solution satisfaisante permettant de réduire la dimension des vecteurs : la projection des données sur le modèle contenu dans la matrice de plan expérimental X apparaît comme étant la base de projection la plus pertinente vis-à-vis des données. Ainsi au lieu de classer directement les décours temporels BOLD, nous proposons de classer les coefficients de régression estimés (c'est-à-dire les tailles des effets estimés). Par extension, la classification peut se faire sur des contrastes obtenus à partir de combinaisons linéaires de ces coefficients. Enfin, nous avons vu dans la section 3.3.2.0 que les coefficients β estimés suivent une loi normale si l'on fait l'hypothèse que les résidus sont les réalisations d'un bruit gaussien, d'où la validité du modèle.
- d'un point de vue spatial, nous avons relaté dans le chapitre 4 deux principales techniques pour rajouter des a priori spatiaux à un modèle de mélange. La première se propose de travailler dans l'espace joint position et signal, tandis que la seconde consiste à utiliser une approche markovienne. Nous avons choisi la première approche car elle définit une extension spatiale gaussienne à chaque classe, ce qui correspond à notre notion de parcellisation (agrégat de voxels proches spatialement). L'approche markovienne ne définit pas de modèle a priori d'extension spatiale, et des classes peuvent ainsi avoir une étendue spatiale moins compacte. De plus, les temps de calcul associés peuvent s'avérer relativement élevés lorsque la dimension des éléments à classer est élevée.

7.2.2 Adaptations pratiques

Le modèle proposé présente un nombre élevé de degrés de liberté, rendant difficile l'estimation des paramètres. Nous proposons dans ce paragraphe diverses hypothèses simplificatrices, dont nous présentons également les conséquences.

Nombre de classes K :

Le nombre de classes est un hyperparamètre du modèle. Comme dans le cas de la parcellisation anatomique, spécifier sa valeur revient à définir une résolution de travail à laquelle on observe les données. Nous laissons donc le choix à l'utilisateur de définir ce nombre de classes K .

Contraintes sur les matrices de covariance :

Nous formulons plusieurs hypothèses sur la forme des matrices de covariance, permettant d'éviter les singularités mais également de réduire le nombre de paramètres.

Tout d'abord, nous considérons les matrices de covariance diagonales par blocs, ce qui supprime tous les termes de corrélation croisée entre les termes liés au signal et les coordonnées spatiales. On est, en ce cas, en présence d'un modèle conditionnellement séparable :

$$p(\beta_i, v_i | k; \theta_k) = p(\beta_i | k; \theta_k) p(v_i | k; \theta_k)$$

ce qui donne :

$$\begin{aligned} p(\beta_i, v_i | \Theta) &= \sum_k p(k) p(\beta_i | k; \theta_k) p(v_i | k; \theta_k) \\ &= \sum_k p(\beta_i | k; \theta_k) p(k | v_i; \theta_k) p(v_i; \theta_k) \quad \text{avec la relation de Bayes} \end{aligned}$$

et donc en particulier :

$$p(\beta_i | v_i; \Theta) = \sum_k p(\beta_i | k; \theta_k) p(k | v_i; \theta_k) \tag{7.1}$$

On retrouve le modèle présenté dans [Penny and Friston, 2003] : précisément, il est intéressant de noter que cette formulation conditionnelle met en évidence le fait que le modèle définit implicitement une densité sur les coefficients de régression β qui est dépendante de la position spatiale. Autrement dit, l'équation 7.1 stipule que le signal β en une position de l'espace donnée v est obtenu comme étant le mélange (la combinaison linéaire) de K signaux pondérés par les probabilités d'appartenance du voxel v à chacune des classes associées aux K signaux. Cependant, le modèle tel que nous l'avons présenté conduit à des techniques d'optimisation beaucoup plus simples que 7.1 et leur efficacité permet de rechercher plus de quelques classes.

Nous avons fait une deuxième hypothèse simplificatrice concernant les matrices de covariance réduisant de manière drastique le nombre de paramètres. Il s'agit de considérer

l'ensemble des classes comme étant homogènes et donc de leur affecter la même matrice de covariance. Ainsi, en particulier, ce choix formule l'hypothèse que toutes les classes présentent la même étendue spatiale.

Une remarque naturelle concerne les distances mixtes intervenant dans le modèle, mélangeant des coefficients de régression et des positions spatiales, d'unités distinctes et a priori incomparables. En fait, le mélange de gaussiennes fait intervenir des distances de Mahalanobis entre deux éléments de l'espace joint : ainsi, les matrices de covariance permettent de normaliser les influences de chacune des composantes selon leur variance.

$$d = \begin{bmatrix} \beta^T & v^T \end{bmatrix} \Sigma^{-1} \begin{bmatrix} \beta^T \\ v^T \end{bmatrix} = \begin{bmatrix} \beta & v \end{bmatrix} \begin{bmatrix} \Sigma_\beta^{-1} & 0 \\ 0 & \Sigma_v^{-1} \end{bmatrix} \begin{bmatrix} \beta \\ v \end{bmatrix} = \beta^T \Sigma_\beta^{-1} \beta + v^T \Sigma_v^{-1} v$$

Il est alors possible d'introduire artificiellement un paramètre λ permettant de pondérer les influences relatives des deux termes :

$$d = \lambda \beta^T \Sigma_\beta^{-1} \beta + (1 - \lambda) v^T \Sigma_v^{-1} v \quad \text{avec} \quad 0 \leq \lambda \leq 1$$

Par "défaut", ce paramètre est fixé à $\frac{1}{2}$, donnant aux deux termes – signal et spatial – une importance équivalente. L'ajout de ce paramètre dans le modèle permet à l'utilisateur de favoriser des classes plutôt bien localisées spatialement ($0 \leq \lambda \leq \frac{1}{2}$) ou plutôt très représentatives d'un même signal ($\frac{1}{2} \leq \lambda \leq 1$). Nous verrons l'influence de ce paramètre par la suite.

Estimation des paramètres :

La formulation du problème par un modèle de mélange dans l'espace joint met en évidence la possibilité d'utiliser directement un algorithme EM pour estimer les paramètres (cf paragraphe 4.2.2). Nous rappelons brièvement les deux étapes de l'algorithme EM pour les mélanges de gaussiennes avec des données $x_i = \begin{bmatrix} \beta_i & v_i \end{bmatrix}$:

- **Etape E** : Estimation des données manquantes γ_i^k :

$$\gamma_i^k = \frac{\pi_k G_{\mu_k, \Sigma_k}(x_i)}{\sum_{l=1}^K \pi_l G_{\mu_l, \Sigma_l}(x_i)}$$

- **Etape M** : Maximisation de la vraisemblance des données complétées :

$$\begin{aligned} \pi_k &= \frac{1}{N} \sum_{i=1}^N \gamma_i^k \\ \mu_k &= \frac{\sum_{i=1}^N \gamma_i^k x_i}{\sum_{i=1}^N \gamma_i^k} \\ \Sigma &= \frac{1}{K} \sum_{k=1}^K \frac{\sum_{i=1}^N \gamma_i^k (x_i - \mu_k)(x_i - \mu_k)^T}{\sum_{i=1}^N \gamma_i^k} \end{aligned}$$

En pratique, l'initialisation est réalisée en effectuant au préalable une parcellisation anatomique du masque du cerveau extrait des images fonctionnelles (cf paragraphe 5.2.2). Les moyennes des gaussiennes sont alors initialisées comme étant le barycentre de chaque parcelle pour la partie spatiale et comme étant la moyenne des coefficients de régression contenus dans chaque parcelle pour la partie fonctionnelle. Les matrices de covariance sont elles aussi initialisées en calculant la covariance des signaux observés au sein de chaque parcelle. Enfin, les classes sont initialisées équiprobables ($\pi_k = \frac{1}{K}$).

Le critère d'arrêt que nous utilisons consiste à étudier la variation relative du critère (la vraisemblance). Il serait cependant éventuellement plus pertinent d'étudier les variations des paramètres, cependant cela nécessite de définir une métrique dans l'espace des paramètres.

Après convergence, une classification "dure" peut être obtenue par maximum a posteriori, en affectant chacune des données signal/spatial à la classe la plus probable, selon les probabilités a posteriori γ_i^k .

7.3 Exemples d'applications

Dans ce paragraphe, nous proposons d'étudier les propriétés du modèle dans deux cas particuliers puis de présenter des parcellisations anatomo-fonctionnelles obtenues avec le modèle complet.

7.3.1 Classification anatomique pure

Tout d'abord, il est à noter que le modèle proposé généralise la parcellisation anatomique présentée au chapitre 5. En effet, choisir $\lambda = 0$ équivaut à ne prendre en compte aucune information fonctionnelle et à appliquer un modèle de mélange aux seules coordonnées spatiales des voxels situés dans le masque du cerveau. Utiliser les hypothèses simplificatrices détaillées ci-dessus concernant les matrices de covariance et estimer les paramètres du modèle via un algorithme CEM revient alors à appliquer un algorithme K-Means, comme nous l'avons remarqué dans le paragraphe 4.2.4. La principale différence réside dans le choix de la métrique : ici, la distance entre deux échantillons de l'espace à classifier est euclidienne, au lieu d'être géodésique.

7.3.2 Classification fonctionnelle pure

A l'opposé, choisir $\lambda = 1$ revient à ignorer toute information spatiale, ce qui nous place dans le cadre des méthodes de type *agrégation* purement fonctionnel, déjà proposées dans la littérature [Goutte *et al.*, 1999]. La principale difficulté de ce type d'algorithme pour l'analyse exploratoire des données est de déterminer quels sont les paramètres pertinents à extraire des données, qui seront utilisés au cours de l'algorithme. En effet, les décours temporels associés à chaque voxel correspondent à des vecteurs de dimension éventuellement élevée (plusieurs centaines d'instant temporels) rendant

problématique la mise en œuvre pratique de tels algorithmes. L'utilisation des coefficients de régression β permet de réduire la dimension des paramètres et donc de pouvoir traiter tous les voxels simultanément, tout en conservant l'essentiel de l'information intéressante contenue dans les signaux (liée au modèle sur lequel les données sont projetées).

En particulier, nous avons utilisé ce type de *meta-clustering* pour étudier le jeu de données correspondant à la cartographie du signal pariétal [Simon *et al.*, 2002] dont nous avons présenté une partie au paragraphe 6.2.2. Les 10 sujets prenant part à l'expérience ont effectué six tâches impliquant le sillon pariétal : tâches de préhension, de pointage, de saccades oculaires, d'attention visuo-spatiales, de calcul mental et de langage (détection de phonème).

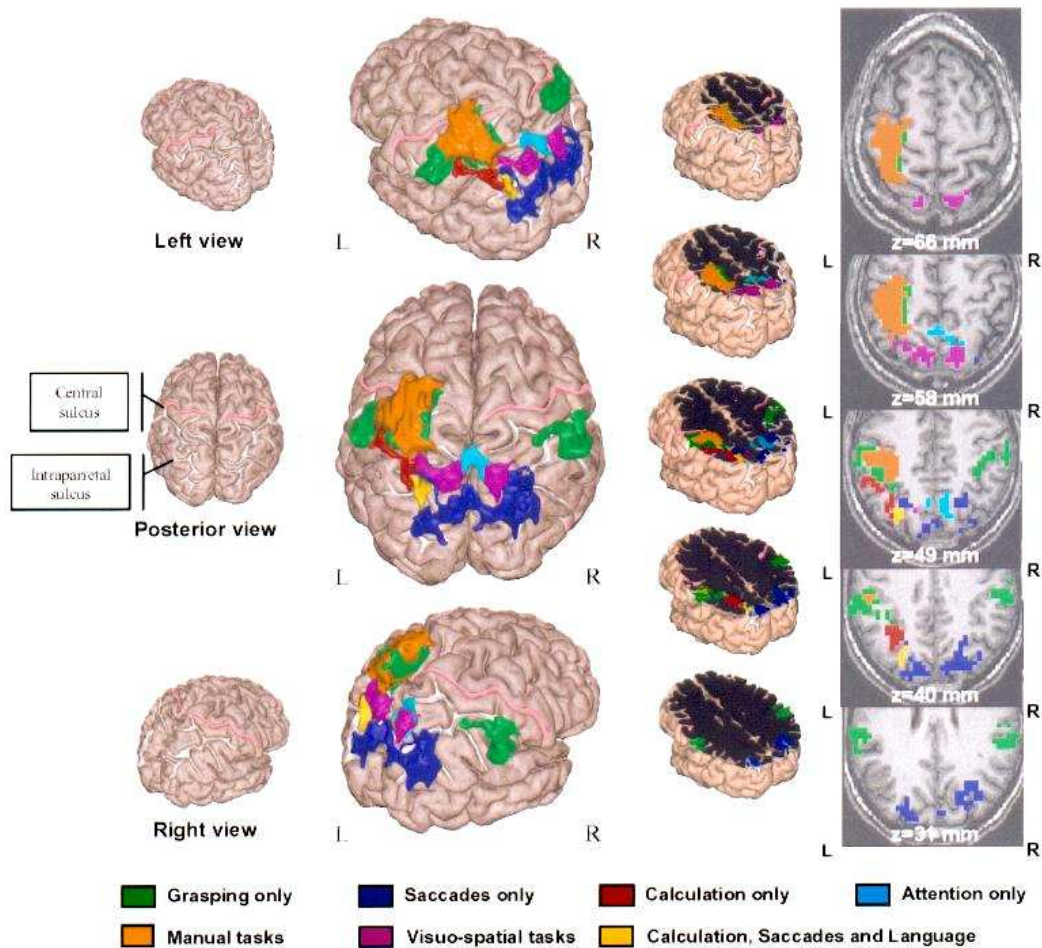


FIG. 7.1 – Cartographie du sillon pariétal [Simon *et al.*, 2002]. La figure présente les intersections des différentes tâches (préhension, pointage, saccades, attention, calcul mental, langage) autour du sillon pariétal pour l'analyse de groupe.

Les résultats de l'analyse de groupe présentés dans [Simon *et al.*, 2002] mettent en

évidence l'activation d'un large réseau d'aires comprenant les aires pariétales mais aussi le cortex frontal latéral et les régions fronto-médianes. Ainsi, les auteurs présentent pour chaque tâche le motif global d'activation observé dans la tâche primaire par rapport à son contrôle. Une autre carte d'intérêt présente les régions activées par une seule tâche (c'est-à-dire significativement plus que chacun des autres tâches) ainsi que les régions activées de façon commune par deux tâches ou plus. La figure 7.1 présente la carte correspondante, pour laquelle cette étude a été effectuée manuellement pour les régions situées à proximité du sillon pariétal.

La difficulté de ce type d'analyse réside dans le fait qu'il existe de nombreuses conditions expérimentales et qu'il devient fastidieux de comparer toutes les conditions deux à deux : pour la figure 7.1, les auteurs se sont limités à la région du sillon pariétal et n'ont étudié qu'un sous-ensemble de toutes les conditions possibles, selon des a priori sur les réseaux engendrés par chacune des tâches.

Nous proposons de rechercher de manière exploratoire et automatique les conditions expérimentales pertinentes à partir des données. Pour cela, nous avons considéré pour chaque tâche l'image représentant le contraste *tâche - contrôle* dans l'analyse de groupe. Ces images étant dans le même espace, on obtient en chaque voxel un vecteur de dimension 6 (*feature vector*) représentatif de l'activité liée à chaque tâche. L'application d'un algorithme d'*agrégation* dans cet espace permet de définir un ensemble de contrastes faisant intervenir une ou plusieurs tâches, associé à des aires cérébrales. Nous présentons sur la figure 7.2 le résultat de classification obtenu en estimant les paramètres d'un modèle de mélange de gaussiennes plongé dans cet "espace des tâches", avec 6 classes (nombre optimal selon le critère BIC).

Voici une rapide analyse des regroupements de tâches détectés²⁴, et présentés sur la figure 7.2 :

- la *classe 2* (en jaune) regroupe les activations communes aux tâches manuelles (préhension et pointage), situées dans le lobe pariétal supérieur et le sillon postcentral.
- la *classe 5* (en bleu) rassemble les activations communes aux tâches visuo-spatiales (préhension, pointage, saccades et attention), situées de manière bilatérales dans les lobules pariétaux supérieurs.
- la *classe 4* (en bleu clair) correspond lui aux tâches calcul et langage, situé dans des régions plus frontales.
- la *classe 1* (en rouge) regroupe les deux tâches manuelles avec la tâche d'attention visuo-spatiale, de manière unilatérale sur le sillon intrapariétal droit mais aussi de manière bilatérale dans les lobes frontaux.
- enfin, on note la présence de deux classes ne faisant intervenir chacune qu'une seule tâche : calcul et saccades.

Il est intéressant de noter que l'on retrouve certains des regroupements présentés sur la figure 7.1, en particulier pour les tâches manuelles et visuo-spatiales. En revanche, cette classification automatique fait également apparaître des contrastes non testés comme celui correspondant aux activations simultanées préhension, pointage et attention. Par ailleurs,

24. Une analyse plus détaillée a été entamée avec Stanislas Dehaene en vue d'une exploitation des résultats.

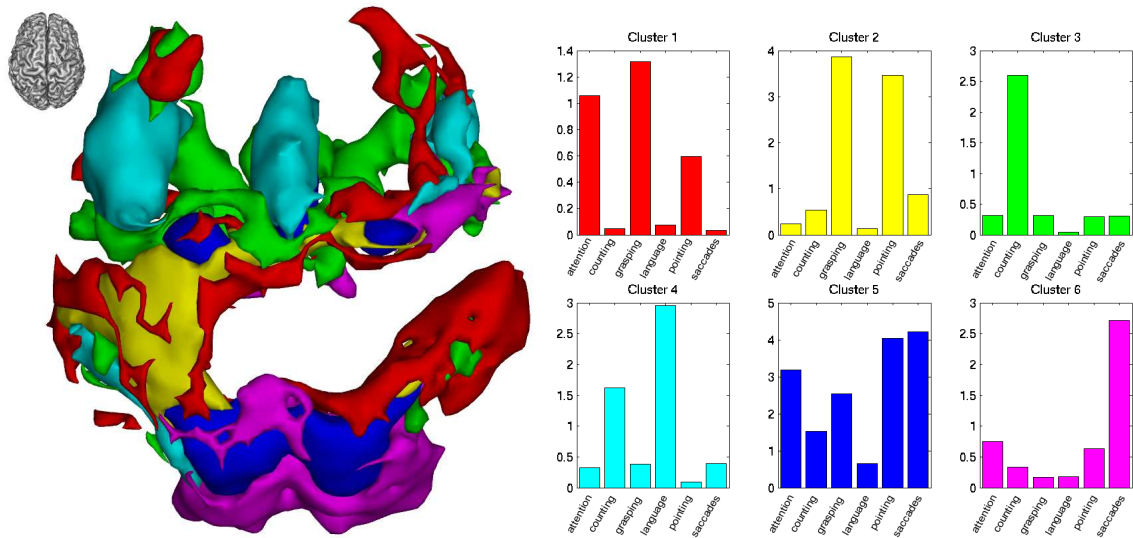


FIG. 7.2 – Classification des régions cérébrales du jeu de données étudiant le sillon pariétal selon les tâches impliquées. Six classes ont été estimées par un algorithme EM : l'image de gauche présente un rendu surfacique des 6 classes déterminés. Les courbes de droite présentent les représentants de chaque classe (associés par couleur), correspondant au vecteur moyen des gaussiennes estimées.

nous avons effectué d'autres classifications en modifiant le nombre de classes : choisir un nombre de classes plus élevé semble également indiquer d'autres clusters intéressants.

L'utilisation d'un algorithme de classification appliqué dans l'espace des tâches permet ainsi d'indiquer de manière exploratoire au neuroscientifique quels sont les contrastes d'intérêt à tester sur les jeux de données. Dans l'expérience que nous avons présentée, les auteurs avaient restreint leurs investigations sur le sillon pariétal ; la méthode que nous proposons est globale, travaillant sur la totalité du volume imagé, ce qui a permis de mettre en évidence d'autres aires dans les lobes frontaux. Cette méthode permet ainsi d'augmenter la quantité d'information que l'on peut extraire des données et participe à une meilleure compréhension des réseaux mis en œuvre lors de tâches cognitives complexes.

7.3.3 Classification spatio-temporelle

Etant donné les résultats prometteurs obtenus avec les données signal seul, on peut s'attendre à des performances intéressantes avec le modèle spatio-temporel complet. Dans ce paragraphe, nous illustrons les principaux apports de la méthode sur des données simulées, puis nous présentons des résultats préliminaires sur des données réelles 2D et 3D.

Dans un but purement illustratif, nous présentons sur la figure 7.3 une image simulée bidimensionnelle représentant un motif de "haricot" noyé dans un bruit gaussien. Cette image n'a bien sûr rien de réaliste mais nous présentons les résultats obtenus dans un intérêt didactique. La valeur scalaire affectée à chaque pixel est considérée comme étant

un résumé temporel d'un signal.

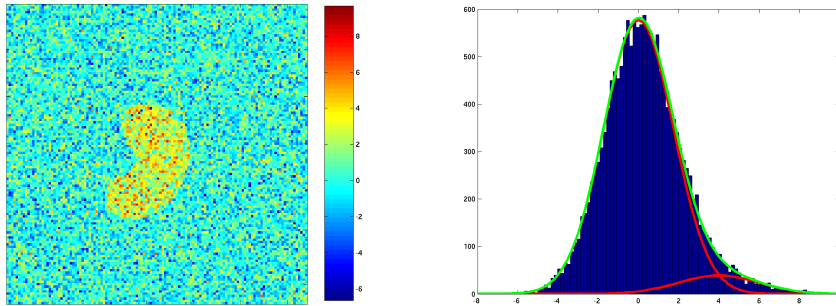


FIG. 7.3 – Image 2D simulée présentant un fond composé d'un bruit gaussien de moyenne nulle et de variance 3 et d'un objet au centre composé également d'un bruit gaussien de moyenne 4 et de variance 4. L'histogramme est présenté à droite avec en rouge les deux distributions sous-jacentes.

La segmentation de cette image en deux classes sans information spatiale par algorithme K-Means ou EM ne donne pas de résultats satisfaisants (cf figure 7.4). En effet, le K-Means fait l'hypothèse que les deux classes sont équiprobables, ce qui n'est pas vérifié ici (le fond a une étendue beaucoup plus grande que l'objet). L'algorithme EM suivi d'un maximum a posteriori fournit un meilleur résultat, en relâchant les hypothèses fortes du K-Means mais la segmentation reste incorrecte, en raison des estimations trop approximatives des paramètres des gaussiennes. Ainsi, sans information spatiale, il est difficile de segmenter cette image en deux classes.

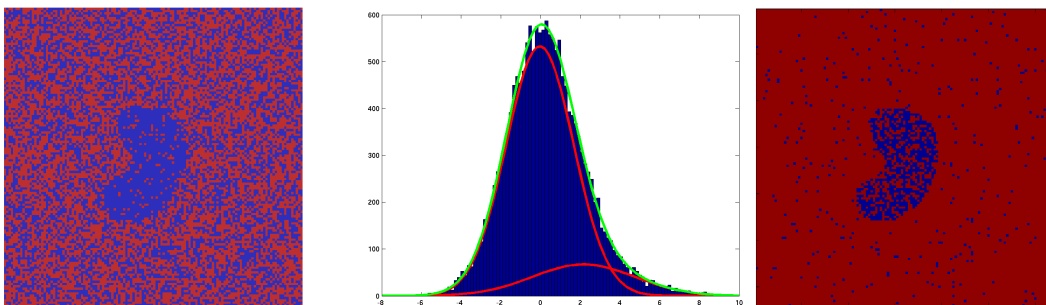


FIG. 7.4 – L'image de gauche présente le résultat de segmentation de l'image 7.3 utilisant l'algorithme K-Means tandis que l'image de droite présente le résultat obtenu par algorithme EM avec 2 classes. Les gaussiennes estimées sont représentées sur l'histogramme.

Nous avons ensuite utilisé l'algorithme EM pour estimer les paramètres du modèle de mélange 2D+t avec 150 classes : les résultats sont présentés sur la figure 7.5. L'image de gauche présente la parcellisation obtenue (image de label), tandis que sur l'image de droite le signal moyen a été affecté à chaque parcelle. On observe que les frontières spatiales des

parcelles correspondent à celles de l'objet "haricot" présent dans l'image. De plus, on vérifie sur l'histogramme de la composante signal des moyennes que le signal à l'intérieur de chaque parcelle est homogène : on obtient deux classes distinctes centrées en 0 et en 4. Un deuxième algorithme de classification travaillant uniquement sur les intensités (tel que le K-Means ou un algorithme de classification hiérarchique) peut alors aisément séparer les deux classes. La classification de l'image se fait donc en deux étapes : une première classification spatio-temporelle résume l'information pertinente en un nombre restreint d'éléments, qui peuvent alors être aisément classifiés, même par un algorithme imposant de fortes contraintes. Les intérêts de ce type d'approche en deux étapes ont été mis en évidence dans [Johnson, 1994].

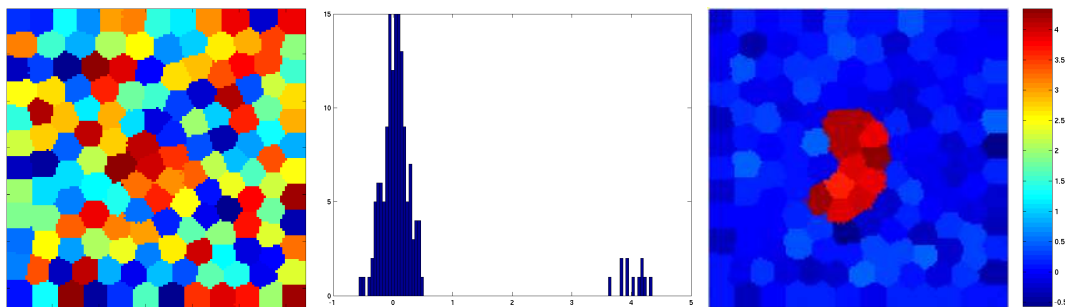


FIG. 7.5 – Parcellisation de l'image 7.3 avec 150 classes en utilisant en utilisant le modèle présenté au paragraphe 7.2.1. L'histogramme présente la distribution de la composante signal des moyennes des gaussiennes estimées. Ces moyennes sont alors reportées sur chaque parcelle dans l'image de droite.

Afin de mettre en évidence l'influence de l'aspect signal lors de la construction de la parcellisation, nous présentons sur la figure 7.6 d'autres résultats de parcellisation avec diverses pondérations de cette composante dans le calcul des distances.

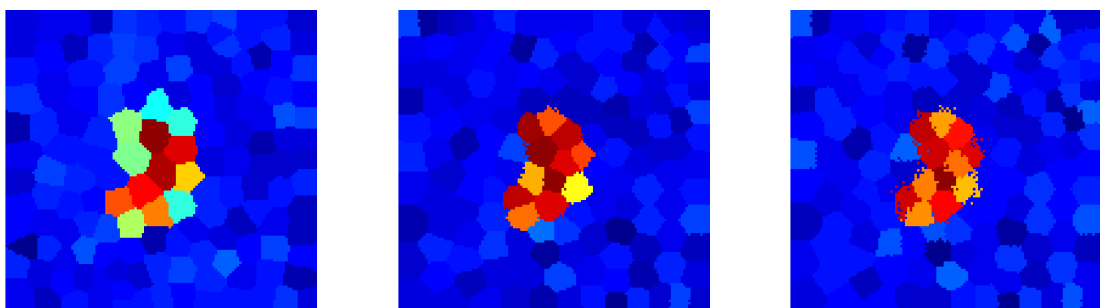


FIG. 7.6 – Exemples de classification avec différentes valeurs du coefficient λ pondérant l'influence relative des termes signaux et spatiaux dans le calcul des distances entre éléments de l'espace joint. Les images correspondent à $\lambda = 0$ (à gauche), $\lambda \leq \frac{1}{2}$ (au milieu) et $\lambda \geq \frac{1}{2}$ (à droite).

En particulier, l'image de gauche présente le résultat d'une parcellisation anatomique, sans aucune information temporelle : la construction des parcelles se fait donc uniquement

sur un critère de compacité, d'où des signaux moyens quelconques au sein de chaque parcelle. Cette parcellisation sert toutefois d'initialisation pour l'estimation des paramètres du modèle joint. Les deux images suivantes prennent en compte l'information signal avec une influence croissante (grâce au paramètre λ introduit ci-dessus) : on observe que plus ce terme est élevé, plus les frontières des parcelles correspondent à celles des objets présents dans l'image. Ce paramètre λ permet ainsi de régler le compromis entre compacité spatiale (relié à la notion de connexité) et représentation fine du signal.

Nous présentons maintenant la parcellisation d'une coupe axiale d'une image de contraste *tâche - contrôle* pour le jeu de données étudiant le sillon pariétal avec une tâche de préhension. Les résultats sont résumés sur la figure 7.7.

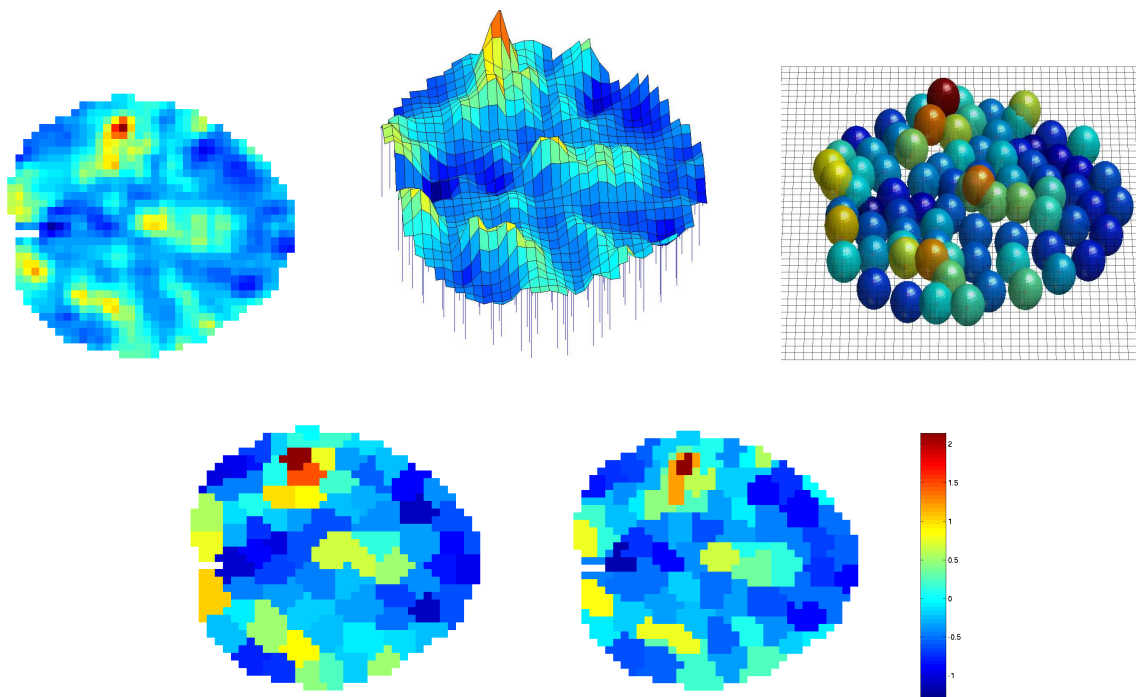


FIG. 7.7 – Exemple 2D de parcellisation anatomo-fonctionnelle pour une coupe axiale d'un contraste de la tâche de préhension du protocole pariétal [Simon et al., 2002]. La figure en haut à droite présente une visualisation du modèle de mélange $2D+t$ estimé, tandis que les deux images du bas présentent deux résultats de parcellisation : celle de gauche est la parcellisation anatomique servant d'initialisation pour la parcellisation dans l'espace joint présenté en bas à droite.

L'image de contraste 2D est représentée en haut à gauche de la figure 7.7. Utiliser le modèle joint spatio-temporel revient à la plonger dans un espace 3D $[x,y,t]$ comme l'illustre la figure du milieu. Nous avons estimé les paramètres d'un modèle de mélange de 80 gaussiennes, grâce à l'algorithme EM : le modèle obtenu est présenté en haut à droite. Chaque gaussienne est représentée par l'ellipsoïde correspondant à sa matrice de covariance (toutes identiques par construction) centrée sur la moyenne de chaque

classe. La couleur associée à chaque ellipsoïde est elle reliée à la valeur moyenne de la composante signal. Ce modèle obtenu est un modèle générateur : on peut considérer que l'image utilisée en entrée est une réalisation de ce modèle probabiliste. L'application d'un maximum a posteriori permet d'obtenir une parcellisation : celle-ci est présentée sur la figure 7.7 en bas à droite. En comparaison, nous avons également présenté à sa gauche la parcellisation anatomique pure utilisée comme initialisation du modèle. Ceci met en évidence la pertinence du modèle joint, permettant d'obtenir une parcellisation spatiale ayant également un sens du point de vue fonctionnel en favorisant l'homogénéité du signal au sein de chaque parcelle.

Enfin, nous présentons un exemple de parcellisation tridimensionnel sur le même jeu de données. La figure 7.8 permet de visualiser les modèles estimés avec 100 et 500 classes. L'estimation des paramètres par algorithme EM nécessite entre 20 et 50 itérations pour converger avec une initialisation obtenue grâce à une parcellisation anatomique par K-Means. En fait, l'ajustement du modèle joint sur les données ne sert qu'à raffiner la parcellisation anatomique déjà obtenue en adaptant localement la moyenne de chaque gaussienne afin de modéliser au mieux le signal, la parcellisation anatomique ayant déjà placé les classes de manière régulière dans l'espace.

Comme précédemment, l'application d'un maximum a posteriori sur les modèles obtenus permet d'obtenir un parcellisation. Celle obtenue avec le modèle à 500 classes est présenté sur la figure 7.8 à droite.

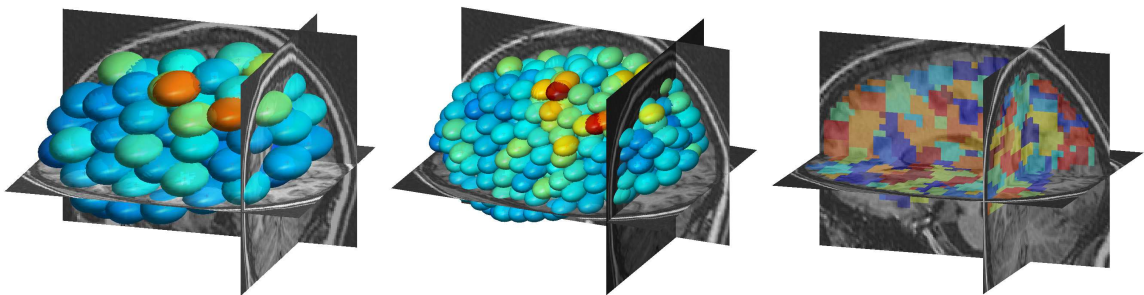


FIG. 7.8 – Exemples 3D de parcellisation anatomo-fonctionnelle du jeu de données correspondant à la tâche de préhension [Simon et al., 2002]. L'image de gauche présente le modèle de mélange de 100 gaussiennes tandis que l'image du milieu permet de visualiser celui estimé avec 500 classes. La parcellisation correspondante obtenue par maximum a posteriori est présentée à sa droite.

Les résultats présentés ici restent préliminaires mais montrent que l'estimation des paramètres du modèle joint est possible numériquement grâce à l'efficacité de l'algorithme EM. En particulier, les contraintes sur les matrices de covariance jouent un rôle de régularisation important dans le comportement de l'algorithme.

La validation de ce modèle peut être effectuée par son utilisation pour la détection d'activations sur une parcellisation, comme nous l'avons fait dans le paragraphe 6.2. À côté des gains potentiels en détection, l'hypothèse d'un signal homogène par parcelle

ouvre de nombreuses autres perspectives : ainsi, en affectant un sujet moyen par parcelle, on obtient une représentation concise des données, permettant de mettre en œuvre effectivement des algorithmes considérés comme coûteux en temps et en mémoire (méthodes multivariées par ACP à noyau, classification hiérarchique, méthodes non paramétriques, etc). Dans la suite, nous avons choisi de nous intéresser à l'adaptation de notre modèle aux analyses multi-sujets, car c'est le domaine où les retombées sont potentiellement les plus intéressantes du point de vue des neurosciences.

7.4 Extension multi-sujets

7.4.1 Motivations et modèle proposé

Dans ce paragraphe, nous proposons d'étendre le modèle présenté au début de ce chapitre au niveau multi-sujet. En effet, l'essentiel des expériences menées en neurosciences mettent en jeu plusieurs sujets, afin que les résultats des analyses puissent se généraliser à une population : il s'agit de déterminer les activations stables à travers les individus. Pour ce faire, la méthodologie standard recherche les tâches moyennes d'activation en normalisant les sujet dans un espace standard et en lissant les données afin de palier le manque de précision des algorithmes de normalisation et permettre d'augmenter le recouvrement des signaux fonctionnels entre sujets.

Nous avons vu au chapitre 3 que l'étape de normalisation spatiale est un problème majeur car les techniques proposées reposent sur l'hypothèse de corrélation anatomo-fonctionnelle. Du point de vue des neurosciences, le problème de la normalisation spatiale ne devrait pas être défini par rapport à l'anatomie sulco-gyrale mais comme une mise en correspondance des aires corticales. Il n'y a cependant pas de consensus à l'heure actuelle sur la définition des aires corticales. Ainsi, la mise en correspondance des cerveaux entre différents sujets reste à l'heure actuelle un problème essentiellement non résolu.

De plus, même en considérant le recalage entre sujets comme valide (diminution de la variabilité anatomique), il reste une variabilité de la localisation fonctionnelle non négligeable à travers les sujets, comme cela a pu être mis en évidence pour les aires cytoarchitectoniques avec des variations importantes de l'étendue des aires selon les sujets [Amunts *et al.*, 1999], ainsi qu'une variabilité fonctionnelle pure : la réponse d'un sujet varie d'un essai à un autre mais également de sujet à sujet. Le faible recouvrement des aires fonctionnelles entraîne une baisse de sensibilité lors des analyses de groupe. L'utilisation d'un lissage spatial gaussien important des images fonctionnelles permet d'augmenter ce recouvrement mais ceci se fait au détriment du pouvoir de localisation.

Ainsi, une *bonne* propriété d'un outil d'analyse de données fonctionnelles multi-sujets serait de pouvoir assouplir l'hypothèse de mise en correspondance de voxels homologues à travers les sujets. Nous proposons ici de placer l'analyse de groupe dans le cadre de la classification en cherchant à regrouper les voxels des différents sujets présentant des décours temporels similaires en des positions spatiales similaires dans un espace standard. Pour ce faire, les données de chaque sujet sont normalisées dans l'espace de Talairach en utilisant un algorithme non rigide [Ashburner and Friston, 1999] mais par la suite on ne suppose pas exacte la correspondance voxel à voxel entre les sujets.

Nous proposons ainsi d'utiliser le modèle décrit dans le paragraphe 7.2.1, en estimant ses paramètres à partir de l'ensemble des données de tous les sujets. La figure 7.9 présente un exemple de vecteurs d'entrée à classifier : cela revient à considérer simultanément les données de tous les sujets normalisés comme s'il s'agissait de elles d'un unique sujet placé dans l'espace de Talairach. On peut relier cette notion aux analyses de groupe à effet fixe (cf paragraphe 3.3.5) où l'effet sujet n'est pas modélisé : le signal est supposé identique pour tous les sujets, la variabilité résiduelle étant assimilée à du bruit. Un des problèmes principaux est que l'on observe une variabilité inter-sujet parasite importante concernant les paramètres fonctionnels (coefficients de régression) due à l'amplitude différente de la réponse de chaque sujet et à la dynamique de l'imager. Pour s'abstraire de ces variables de non-intérêt, nous avons choisi de normaliser l'intensité de chacune des composantes "signal" en les contraignant à avoir mêmes moyennes et variances.

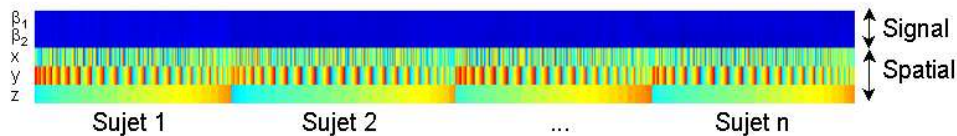


FIG. 7.9 – *Exemple de données (feature vectors) à classifier : chaque colonne de cette matrice représente un voxel d'un sujet avec ses coordonnées spatiales dans l'espace de Talairach et un certain nombre de régresseurs modélisant le signal fonctionnel.*

Le modèle estimé avec les données de tous les sujets correspond alors à un modèle moyen, et n'est spécifique à aucun des sujets. Un *agrégat de voxels* est donc intrinsèquement défini de manière multi-sujet. Il est en revanche possible, comme dans le paragraphe 7.3.3, d'utiliser les données d'un sujet et de calculer un maximum a posteriori du modèle global pour obtenir une parcellisation spécifique au sujet en question. Pour résumer, on définit un unique modèle global correspondant à autant de parcellisations qu'il n'y a de sujets. Un point important de cette modélisation est que les parcelles obtenues pour chacun des sujets sont homologues entre elles (par l'intermédiaire du modèle de mélange) sans pour autant être identiques d'un point de vue de leur localisation spatiale : la contrainte d'homologie est intrinsèquement floue.

7.4.2 Classification fonctionnelle multi-sujets

Nous présentons tout d'abord la version multi-sujet de l'étude 7.3.2 réalisée ci-avant. Pour cela, nous avons normalisé dans l'espace standard du modèle MNI 5 sujets ayant pris part au protocole pariétal avec 6 tâches. L'analyse univariée par modèle linéaire général associée à un lissage spatial de 8 mm permet de construire une image de contraste tâche - contrôle pour chaque sujet, dans le même espace spatial. Nous avons alors appliqué l'algorithme EM, ici sans utiliser les coordonnées spatiales, avec 6 classes. Les résultats sont présentés sur la figure 7.10.

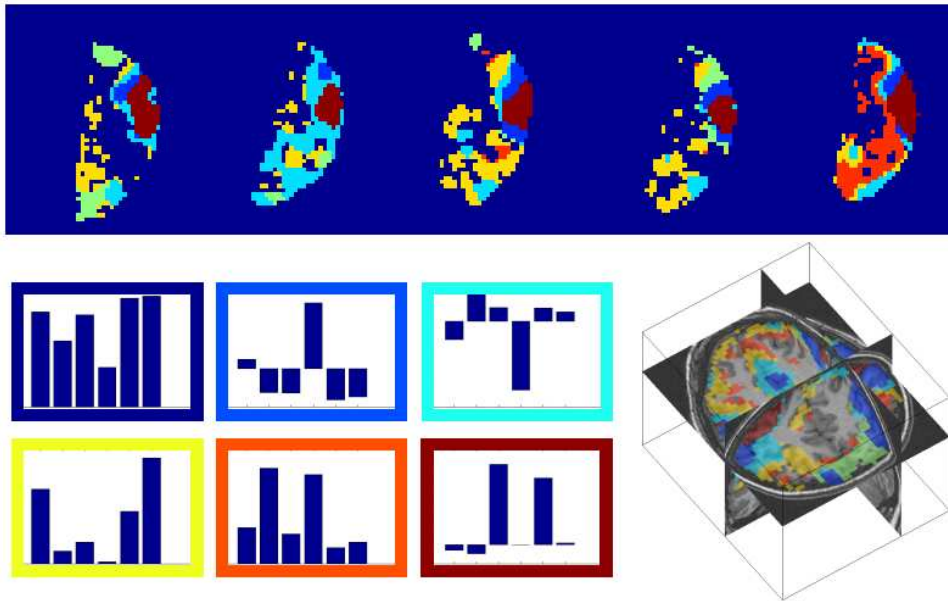


FIG. 7.10 – Classification des régions cérébrales associées à plusieurs tâches à travers 5 sujets. Après avoir été normalisés dans l'espace de Talairach, les données fonctionnelles associées à chacun des sujets ont été analysées par un algorithme de classification avec 6 classes. Les représentants de chaque classe sont affichés en bas à gauche de la figure, en utilisant le même ordre que la figure 7.2 pour l'agencement des différentes tâches. On remarque qu'on obtient certains motifs similaires à travers les sujets, en particulier pour les tâches manuelles, avec cependant une assez grande variabilité spatiale!

On retrouve en multi-sujet certains des contrastes obtenus lors de l'analyse mono-sujet 7.3.2. En particulier, on retrouve à travers les sujets un large *agrégat* regroupant les tâches manuelles (classe 6). Il est intéressant cependant de noter que la localisation ainsi que l'étendue spatiale de cet *agrégat* varie de manière significative à travers cette population de sujets. Ceci met en évidence l'intérêt discuté précédemment des méthodes de classification pour les analyses multi-sujets, car celles-ci évitent de satisfaire à l'hypothèse contraignante de mise en correspondance parfaite voxel à voxel.

7.4.3 Modèle anatomo-fonctionnel multi-sujets

Nous présentons ici de premiers résultats concernant l'estimation du modèle anatomo-fonctionnel complet pour l'analyse de groupe faisant intervenir 5 des 10 sujets ayant pris part au protocole pariétal [Simon *et al.*, 2002] (nous avons pour l'instant exclu de l'analyse les sujets pour lesquels il apparaissait nécessaire de corriger les mouvements). Cette expérience est le pendant multi-sujet des résultats présentés dans le paragraphe 7.3.3. Comme dans celui-ci, nous avons pris le résultat du contraste *tâche - contrôle* pour définir la composante signal des vecteurs à classer. Le modèle utilisant 500 gaussiennes et estimé par algorithme EM en une trentaine d'itérations est présenté sur la figure 7.11.

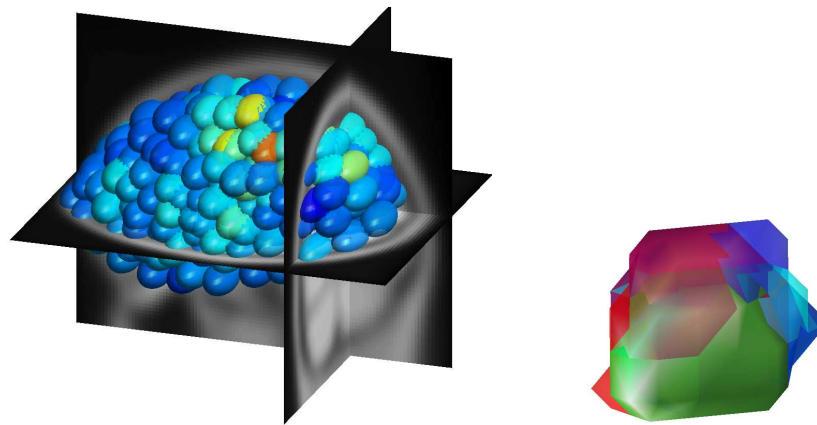


FIG. 7.11 – *Modèle anatomo-fonctionnel obtenu à partir des données de 5 sujets ayant participé au protocole étudiant la tâche de préhension. Les paramètres de 500 gaussiennes ont été estimés par algorithme EM : celles-ci sont visualisées ici dans l'espace de Talairach sur l'anatomie du modèle MNI. L'application du maximum a posteriori sur les données de chaque sujet permet de définir une parcellisation pour chacun d'entre eux. L'image de droite présente la superposition dans le repère du modèle MNI des agrégats obtenu pour chacun des sujets, correspondant au maximum d'activation. On vérifie ainsi que notre modèle autorise une certaine variabilité spatiale pour les activations entre sujets.*

Nous avons choisi de représenter le modèle sur l'anatomie de l'espace standard MNI afin de mettre en évidence que le modèle obtenu n'est pas spécifique à un des sujets mais se place plutôt dans l'espace moyen. Nous pouvons cependant ensuite appliquer un maximum a posteriori aux données de chacun des sujets afin de déterminer pour chaque voxel quelle est la classe la plus proche (dans l'espace $3D+t$). On obtient alors 5 parcellisations distinctes, une pour chaque sujet. Nous avons choisi la classe correspondant au maximum d'activation dans le modèle et représenté sur la figure 7.11 à droite l'*agrégat* correspondant pour chacun des sujets. On observe que leur localisation spatiale ainsi que leur forme diffère selon les sujets, bien que modélisant la même tâche cognitive pour l'ensemble des sujets. Ce résultat met en évidence la pertinence de notre modèle spatio-temporel en multi-sujet, car il permet de trouver des clusters homologues entre sujets sans pour autant émettre l'hypothèse que ceux-ci se superposent exactement dans l'espace standard dans lequel les images fonctionnelles sont normalisées.

7.5 Discussion

Nous avons présenté dans ce chapitre une nouvelle modélisation anatomo-fonctionnelle des données IRMf : celle-ci se place dans le contexte de la classification de données (*clustering*) en recherchant à regrouper les voxels des images fonctionnelles qui sont similaires à la fois d'un point de vue fonctionnel et d'un point de vue spatial. Le modèle proposé

repose sur un mélange de gaussiennes dans l'espace joint spatio-temporel, dont une des qualités est que ses paramètres peuvent être estimés efficacement par un algorithme EM. Nous avons montré que la parcellisation anatomique avec distances euclidiennes est un cas particulier de ce modèle : nous présentons donc un modèle plus général ici, permettant de définir plus finement la position des parcelles en fonction du signal local.

Pour la définition des distances fonctionnelles intervenant dans l'algorithme, nous avons choisi d'utiliser les régresseurs estimés avec un modèle linéaire : cela permet d'obtenir avec quelques valeurs scalaires un résumé pertinent des décours temporels. En particulier, nous avons montré, dans le cadre de classification de données utilisant uniquement cette information, que nous pouvions utiliser le modèle pour déterminer de manière exploratoire quels sont les contrastes d'intérêt pour les analyses faisant intervenir un grand nombre de conditions.

Enfin, nous avons montré que le modèle proposé s'étend naturellement aux analyses multi-sujets, en concaténant les données issues de chaque sujet. Plus précisément, ce modèle ne fait pas l'hypothèse forte d'homologie entre voxels de plusieurs sujets. L'objectif ici est de regrouper les voxels similaires d'un point de vue spatial et fonctionnel à travers les sujets. Quelques résultats préliminaires montrent que le modèle choisi permet de déterminer des *agrégats* homologues entre sujets, ne possédant pas toutefois la même localisation spatiale.

Nous n'avons toutefois montré dans ce chapitre que le potentiel de notre modèle en analyse multi-sujet et une validation rigoureuse reste à réaliser ; une possibilité consiste à comparer les résultats d'analyses de groupe sur des données réelles, effectuées avec l'approche standard et la parcellisation. Les deux principes mis en opposition sont d'une part la normalisation spatiale des sujets agrémentée d'un lissage spatial pour augmenter le recouvrement des signaux et d'autre part le regroupement des signaux proches spatialement et fonctionnellement grâce à la parcellisation. Pour cela, nous proposerions de construire une parcellisation multi-sujets à partir d'un jeu de données et d'effectuer une analyse à effet fixe au niveau des parcelles homologues en testant la présence ou non d'une activation significative. Le résultat obtenu pourrait alors être comparé avec la carte paramétrique obtenue directement à partir des données par une analyse à effet fixe avec normalisation spatiale et lissage gaussien spatial élevé. La principale difficulté de comparaison vient du fait que dans notre cas, on n'obtient pas une carte résultante moyenne des effets de groupe mais au contraire autant de cartes qu'il n'y a de sujet car le modèle sur lequel l'analyse de groupe est menée n'est pas relié à une parcellisation globale. Toutefois, une possibilité pour obtenir une seule carte reviendrait à moyennner les cartes obtenues pour chaque sujet, même si cela nécessite de supposer l'homologie voxel à voxel. On s'attend comme dans le cas de l'expérience 6.2 à une augmentation de la sensibilité, même si elle sera limitée par le dernier moyennage nécessaire à la comparaison. Cette expérience réductrice vis-à-vis de notre modèle a pour but de montrer que les résultats obtenus sont au moins équivalents à l'analyse multi-sujets classique. Avec ce modèle non contraint par l'hypothèse d'homologie, on peut s'attendre par la suite à des résultats beaucoup plus intéressants, bien que difficilement comparables à ceux des techniques actuelles. Nous évoquerons dans le chapitre suivant ces perspectives.

Chapitre 8

Conclusions et Perspectives

Synthèse

Au cours de cette thèse, nous nous sommes intéressés à l’apport de l’information anatomique et des techniques de classification pour l’analyse des données fonctionnelles (IRMf), avec pour perspective une meilleure compréhension de la relation entre la structure cérébrale et sa fonction. Notre objectif a été de proposer un cadre de travail pour l’analyse des données fonctionnelles à une résolution “amas de voxels” (*cluster*), correspondant à la résolution utilisée couramment lors de la description et de l’interprétation des résultats d’analyses, tant pour les analyses intra- que inter-sujets. La construction de ces regroupements de voxels est alors comparable à l’opération de lissage spatial classiquement effectué sur les images fonctionnelles.

Dans un premier temps, nous avons proposé de définir ces regroupements de voxel grâce à une parcellisation reposant sur des critères uniquement anatomiques. Pour cela, nous avons proposé de décomposer l’ensemble du cortex en un nombre ajustable de parcelles permettant d’obtenir une résolution intermédiaire quelconque entre le voxel et la région d’intérêt de grande taille (par exemple un gyrus). Nous avons présenté trois exemples d’utilisation de cette parcellisation – pour la détection d’activations, pour la mise en œuvre de modèle d’analyse spécifique à chaque région cérébrale et pour des études de connectivité – mettant en évidence sa pertinence, avec en particulier une augmentation de la sensibilité lors de la détection d’activations en comparant approche voxel et parcelle.

Par la suite, nous avons élargi le problème en proposant la construction de parcelles anatomo-fonctionnelles, grâce à des techniques de classification génériques. Ce modèle a en particulier l’avantage de proposer une solution originale au problème fondamental de l’analyse de plusieurs sujets présentant nécessairement une certaine variabilité anatomo-fonctionnelle qui rend peu optimale les techniques classiques de repères stéréotaxiques. Lors de la généralisation aux analyses multi-sujets, on recherche à travers les sujets les voxels proches spatialement (anatomiquement) et qui partagent une certaine similarité fonctionnelle. Cette technique fait intervenir la notion de distance entre voxels inter-sujets permettant de relâcher l’hypothèse selon laquelle les structures anatomiques des différents sujets doivent se correspondre exactement. D’autre part, la similarité fonctionnelle peut se fonder sur plusieurs conditions expérimentales, permettant une classification plus inté-

ressante des zones cérébrales. Ce modèle anatomo-fonctionnel autorise donc l'ajustement sujet par sujet des régions étiquetées de façon similaires à travers l'ensemble du groupe de sujets. Ce modèle devient un outil d'investigation des informations pertinentes pour la mise en correspondance de différents sujets. Ces techniques ont été implantées dans des bibliothèques logicielles complètes et souples qui seront mises à disposition de la communauté s'intéressant à la cartographie cérébrale.

Validation

Un point important des travaux à venir est la validation du modèle anatomo-fonctionnel (spatio-temporel) proposé. Une méthodologie d'évaluation pourrait comparer la détection des effets moyens à travers les sujets dans l'espace standard de Talairach aux parcelles détectées sujet à sujet à partir de l'ensemble du groupe. On pourrait alors mettre en évidence la capacité du modèle présenté au chapitre 7 à prendre en compte la variabilité anatomique résiduelle observée après une mise en correspondance des anatomies des sujets, par exemple sur un protocole mettant en jeu une tâche reliée à un sillon connu (sensitif versus moteur pour le sillon central, paradigme visuel pour la scissure calcarine etc). Une fois les signaux fonctionnels résumés sur chaque parcelle, il est alors possible de détecter le signal "d'activation" sur les parcelles homologues sujet à sujet et d'étudier la variabilité spatiale dans l'espace de Talairach de ces régions par rapport au repère anatomique (eg sillon central). Si ces deux informations sont corrélées, on aura montré que l'on est capable d'ajuster la position du signal par rapport à celle de l'amer anatomique (cf la variabilité de la position du sillon central visualisée sur la figure 3.7).

Notons qu'indépendamment des amers anatomiques, les aires fonctionnelles sont elles aussi variables en taille et en position et [Roland and Zilles, 1998] décrivent que la taille des aires cytoarchitectoniques V1 peuvent varier du simple au double. Notre modèle devrait là aussi être capable de prendre en compte cette variabilité, et est donc *fondamentalement mieux adapté que les modèles classiques*.

Perspectives

Les perspectives que nous envisageons dans le prolongement de ces travaux de thèse sont multiples, mais peuvent toutes s'inscrire dans le cadre que nous avons proposé, en raison de la généralité du modèle de classification, qui permet de gérer des informations aussi bien spatiales que temporelles. En particulier, nous proposons d'étudier les pistes suivantes :

- la définition de nouvelles distances géométriques par rapport à des amers,
- la transposition de nos travaux sur la surface corticale,
- la généralisation des données d'entrées avec différentes modalités fonctionnelles,
- la distinction entre variabilité intra-sujet et inter-sujet.

Une amélioration prometteuse consiste à définir des distances géodésiques par rapport à des **amers géométriques** anatomiques ou fonctionnels spécifiques de chaque sujet,

en faisant l'hypothèse sous-jacente d'invariance de ces amers. Ainsi, l'information anatomique associée à chaque échantillon fonctionnel ne serait plus définie de manière absolue par la position dans l'espace de Talaraich mais plutôt de manière relative par rapport à des amers géométriques spécifiques à chacun des sujets, comme par exemple les principaux sillons étiquetés par l'algorithme de [Rivière *et al.*, 2002]. De même, ces distances pourraient également être définies vis-à-vis d'amers fonctionnels définis par un protocole préalable (ceci rejoint la discussion du paragraphe 3.2.4). On peut alors s'attendre à une meilleure prise en compte de la variabilité anatomique. Pour ce faire, il semble pertinent de vouloir réintroduire l'utilisation de **distances géodésiques** contraintes au ruban cortical pour le calcul de distances anatomiques, comme nous l'avons fait au chapitre 5. Il y a cependant plusieurs difficultés liées à leur mise en œuvre. La première est due à la différence de résolution entre images anatomiques et fonctionnelles ; si la résolution de travail choisie est la résolution fonctionnelle, le domaine anatomique dans lequel est calculé les distances géodésiques s'avère imprécis. Par ailleurs, la définition de matrices de covariance reste à étudier dans le cas de distances géodésiques. Cependant, le choix de matrices de covariance diagonales isotropes identiques évite d'avoir à gérer ce cas difficile.

Une autre perspective, liée au point précédent, consiste en la transposition de nos travaux sur la **surface du cortex**. En effet, nous avons utilisé une approche volumique permettant de travailler directement avec les données acquises et minimiser ainsi les éventuelles conséquences de mise en correspondance imparfaites des images fonctionnelles et anatomiques. Cependant, les techniques d'imagerie récentes tendent à atténuer fortement les distorsions géométriques observées en EPI et on peut ainsi s'attendre à pouvoir facilement interpoler des données fonctionnelles sur les nœuds d'un maillage construit à partir d'une IRM anatomique. Il serait alors judicieux de mettre en œuvre nos algorithmes de parcellisation à partir des données projetées sur la surface. Ceci nécessite ici aussi de travailler avec des distances géodésiques surfaciques (telles que celles utilisées pour la figure 5.19). Nous pourrions également prendre avantages des travaux de [Cachia, 2003; Cachia *et al.*, 2003b] qui permettent de définir des parcellisations lobaires surfaciques que nous pourrions subdiviser en cellules élémentaires ayant un sens du point de vue fonctionnel et de localisation dans le gyrus.

Le modèle anatomo-fonctionnel que nous avons décrits n'est pas forcément restreint à son utilisation en IRM fonctionnelle. En particulier, il peut être avantageux de l'associer avec d'**autres modalités comme la MEG ou l'EEG**. En effet, de plus en plus d'expériences ont lieu permettant de fusionner les informations apportées par chaque modalités comme par exemple les acquisitions simultanées IRMf/EEG [Lahaye *et al.*, 2004]. Ce type d'enregistrement permet d'étudier les relations qui existent entre les données métaboliques (IRMf BOLD) et l'activité neuronale/synaptique (EEG). En particulier, l'EEG permet d'apporter au signal BOLD une résolution temporelle permettant d'étudier précisément – grâce à la résolution spatiale de l'IRM – les réseaux cérébraux et leur dynamique. Pour ce faire, une parcellisation prenant en compte les informations fonctionnelles multimodales permet d'obtenir un cadre de travail performant pour mettre en œuvre d'autres techniques de recherche de réseaux de connectivité au sein du cerveau.

Enfin, il apparaît important de **distinguer les variabilités intra- et inter-sujets** d'un point de vue fonctionnel. En effet, dans le chapitre 7, nous avons normalisé les données fonctionnelles avant de les classifier afin de réduire la variabilité inter-sujet. Or, comme nous l'avons indiqué lors de la présentation de la distinction entre analyses à effets fixes et analyses à effets aléatoires, il est essentiel de calculer ces deux types de variance pour pouvoir généraliser les résultats d'une analyse de groupe à une population parente. Une voie peut être de modéliser cette variabilité fonctionnelle inter-sujets explicitement dans le modèle.

Pour conclure, nous avons introduit dans cette thèse l'idée d'un système de classification assez souple pour résoudre en partie le très délicat problème de la mise en correspondance anatomo-fonctionnelle de sujets différents. Cette proposition reste à étudier plus en profondeur mais une voie est ouverte vers de nombreuses applications pour l'étude du lien fonction-structure dans le cerveau humain (sain ou pathologique). Le choix des coordonnées (paramètres à classifier) est crucial, et c'est l'étude de la variabilité ou reproductibilité des résultats fonctionnels inter-sujets, dépendant du système de coordonnées choisi, qui devrait guider ce choix.

Annexe A

L'inégalité de Jensen

Théorème A.1 Si f est une fonction convexe et X est une variable aléatoire discrète²⁵ alors

$$E[f(X)] \geq f(E[X])$$

Autrement dit,

$$\sum_{i=1}^M p_i f(x_i) \geq f\left(\sum_{i=1}^M p_i x_i\right) \quad \text{avec} \quad \sum_{i=1}^M p_i = 1 \quad \text{et} \quad p_i \geq 0 \quad \forall i$$

Si f est strictement convexe alors il y a égalité si et seulement si $X = E[X]$, c'est-à-dire si et seulement si X est une variable déterministe.

Preuve: La démonstration²⁶ est obtenue par récurrence pour $M \geq 2$.

Quand X peut avoir deux valeurs distinctes x_1 et x_2 , l'inégalité est:

$$p_1 f(x_1) + p_2 f(x_2) \geq f(p_1 x_1 + p_2 x_2)$$

Ceci est vrai avec $p_1 + p_2 = 1$ par définition de fonction convexe.

Maintenant supposons que le théorème est vrai pour des distributions avec $M-1$ valeurs et définissons $p'_i = \frac{p_i}{1-p_M}$ pour $i = 1, \dots, M-1$.

$$\begin{aligned} \sum_{i=1}^M p_i f(x_i) &= p_M f(x_M) + (1-p_M) \sum_{i=1}^{M-1} p'_i f(x_i) \\ &\geq p_M f(x_M) + (1-p_M) f\left(\sum_{i=1}^{M-1} p'_i x_i\right) \\ &\geq f\left(p_M x_M + (1-p_M) \sum_{i=1}^{M-1} p'_i x_i\right) \\ &= f\left(\sum_{i=1}^M p_i x_i\right) \end{aligned}$$

25. L'inégalité de Jensen est valable pour des variables aléatoires discrètes et continues mais la démonstration est ici faite avec la notation discrète.

26. D'après le tutorial de Todd K. Moon, <http://www.engineering.usu.edu/classes/ece/7680/>

Le théorème est donc vrai pour les distributions avec M valeurs et par récurrence, le théorème est vrai pour tout $M \geq 2$. \square

Si on applique cette inégalité à $f = -\log$, qui est strictement convexe, on peut démontrer plusieurs résultats concernant la divergence de Kullback. En effet, l'inégalité de Jensen donne :

$$\begin{aligned} -KL(p||q) &= -\sum_x p(x) \log \frac{p(x)}{q(x)} \\ &= \sum_x p(x) \log \frac{q(x)}{p(x)} \\ &\leq \log \left(\sum_x p(x) \frac{q(x)}{p(x)} \right) = \log \sum_x q(x) = \log 1 = 0 \end{aligned}$$

Ainsi la divergence de Kullback entre deux distributions p et q est toujours positive : $KL(p||q) \geq 0$. Et puisque $-\log$ est strictement convexe, l'égalité a lieu si et seulement si $\frac{p(x)}{q(x)} = cst$ c'est-à-dire si $p(x) = q(x)$.

Cependant la divergence de Kullback n'est pas une distance puisqu'elle n'est pas symétrique et qu'elle ne vérifie pas l'inégalité triangulaire. Toutefois elle est souvent interprétée comme une sorte de mesure de distance entre distributions de probabilité.

Anecdotiquement, l'inégalité de Jensen permet aussi de prouver que la moyenne arithmétique est toujours supérieure à la moyenne géométrique :

$$\begin{aligned} \log \left(\frac{1}{N} \sum_{i=1}^N x_i \right) &\geq \frac{\sum_{i=1}^N \log x_i}{N} \\ &\geq \frac{1}{N} \log \left(\prod_{i=1}^N x_i \right) \\ &\geq \log (x_1 \dots x_N)^{\frac{1}{N}} \end{aligned}$$

Le comportement monotone croissant de \log donne :

$$\frac{x_1 + \dots + x_N}{N} \geq \sqrt[N]{x_1 \dots x_N}$$

Annexe B

L'algorithme EM vu comme une minimisation alternée

L'algorithme EM appliqué à un modèle de mélange de gaussiennes peut être vu comme une minimisation alternée [Hathaway, 1986] du critère explicité par l'équation 4.10 :

$$\mathcal{L}(\Theta) = \underbrace{\sum_Z p(Z|X) \log p(X, Z|\Theta)}_{\text{fonction Q}} - \underbrace{\sum_Z p(Z|X) \log p(Z|X)}_{\text{entropie}} \quad (\text{B.1})$$

Ceci peut être interprété comme la log-vraisemblance complète pénalisée par un terme d'entropie qui mesure le recouvrement des composantes du mélange. Avec un modèle de mélange gaussien, le critère est :

$$\begin{aligned} \mathcal{L}(\Theta, \gamma) &= \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log p(Z_i = k, X = x_i | \theta_k) - \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log \gamma_i^k \\ &= \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log [\pi_k G_{\mu_k, \Sigma_k}(x_i)] - \sum_{i=1}^N \sum_{k=1}^K \gamma_i^k \log \gamma_i^k \end{aligned} \quad (\text{B.2})$$

Effectuons une minimisation alternée du modèle de mélange sur les paramètres π_k, μ_k, Σ_k et sur la partition floue γ_i^k .

Première étape : Le paramètre de partition est mis à jour pour maximiser le critère :

$$\gamma^{(t+1)} = \arg \max_{\gamma} \mathcal{L}(\Theta^{(t)}, \gamma)$$

Il faut ajouter la contrainte $\sum_{k=1}^K \gamma_i^k = 1$, grâce à la technique des Multiplicateurs de Lagrange :

$$\widehat{\mathcal{L}}(\Theta, \gamma) = \mathcal{L}(\Theta, \gamma) + \sum_{i=1}^N \lambda_i \left(\sum_{k=1}^K \gamma_i^k - 1 \right)$$

Rechercher les zéros de la dérivée selon γ_i^k donne :

$$\frac{\partial \widehat{\mathcal{L}}(\Theta, \gamma)}{\partial \gamma_i^k} = \log [\pi_k G_{\mu_k, \Sigma_k}(x_i)] - \log \gamma_i^k - 1 + \lambda_i = 0 \quad (\text{B.3})$$

Cette équation ainsi que la contrainte forment le système suivant :

$$\begin{cases} \gamma_i^k = \exp \{-1 + \lambda_i + \log [\pi_k G_{\mu_k, \Sigma_k}(x_i)]\} \\ \sum_{k=1}^K \gamma_i^k = 1 \end{cases} \quad (\text{B.4})$$

On obtient alors que :

$$\lambda_i = 1 - \log \left[\sum_{k=1}^K \pi_k G_{\mu_k, \Sigma_k}(x_i) \right]$$

ce qui donne :

$$\gamma_i^k = \frac{\pi_k G_{\mu_k, \Sigma_k}(x_i)}{\sum_{l=1}^K \pi_l G_{\mu_l, \Sigma_l}(x_i)} \quad (\text{B.5})$$

On retrouve exactement la même équation que dans l'étape E (équation 4.17) de l'algorithme EM.

Deuxième étape: les paramètres du mélange sont mis à jour pour maximiser :

$$\Theta^{(t+1)} = \arg \max_{\Theta} \mathcal{L}(\Theta, \gamma^{(t+1)})$$

Il est évident que :

$$\frac{\partial \mathcal{L}(\Theta, \gamma)}{\partial \Theta} = \frac{\partial Q(\Theta, \gamma)}{\partial \Theta}$$

de telle sorte que la dérivée des paramètres du mélange est exactement la même que celle effectuée dans l'étape M de l'algorithme EM.

Nous avons ainsi montré que chaque étape de l'EM peut être interprétée comme une maximisation alternée des paramètres du modèle de mélange et des paramètres de partition floue.

Annexe C

Dérivation matricielle

Nous répertorions dans cette annexe quelques résultats de dérivation matricielle utilisés dans le chapitre 4 pour la formulation des équations de mise à jour des paramètres intervenant dans l'algorithme EM. Les démonstrations de ces résultats sont disponibles dans [Pennec, 1995], annexe B.

Dans le cas de matrices A , B et C quelconques et d'un vecteur a , on a :

$$\frac{\partial}{\partial a} [a^T B a] = (B + B^T) a$$

$$\frac{\partial}{\partial A} [\text{tr}(BAC)] = B^T C^T$$

La dérivée du logarithme du déterminant est :

$$\frac{\partial \log|A|}{\partial A} = (A^{-1})^T$$

Enfin, dans le cas d'une matrice symétrique Σ , on a :

$$\frac{\partial}{\partial \Sigma} [x^T \Sigma y] = xy^T$$

$$\frac{\partial}{\partial \Sigma} [x^T \Sigma^{-1} y] = -\Sigma^{-T} xy^T \Sigma^{-T}$$

Annexe D

Développements logiciel sous MATLAB

Au cours de cette thèse, plusieurs outils logiciel ont été développés, majoritairement sous MATLAB²⁷ (The MathWorks, Natick, MA). Le choix de cet environnement de programmation a été grandement conditionné par le logiciel SPM, lui-même entièrement développé sous MATLAB. Choisir le même langage permet ainsi de pouvoir directement utiliser certaines des fonctions SPM mais également d'être, de fait, directement compatible et pouvoir ainsi être facilement distribuable à une large part de la communauté des neurosciences. Enfin, il est à noter que MATLAB a également des qualités intrinsèques, indépendamment de SPM, ce qui en fait un environnement de développement de choix : simplicité d'utilisation, multi-plateforme, prototypage rapide d'algorithmes, accès à de nombreuses boîtes à outils, etc. Sa relative lenteur qui peut lui être reprochée lors de la mise en œuvre de certains algorithmes peut être contournée grâce à l'utilisation d'une API "MEX" permettant de programmer en C les routines difficilement vectorisables et par la suite de les utiliser directement dans MATLAB.

Nous présentons ici quelques-uns des outils développés sous MATLAB, certains encore à l'état de prototypes et d'autres déjà distribués sous licence GNU/GPL.

Parcel It Up!

Parcel It Up! est une bibliothèque de fonctions MATLAB implémentant les algorithmes de classification reposant sur le modèle de mélange présenté dans le chapitre 4 : K-Means (utilisant la technique de *memory mapping* pour être à même de gérer d'importantes quantités de données), fuzzy c-means, algorithme EM. *Parcel It Up!* implémente également l'algorithme de parcellisation anatomique du cortex décrite au chapitre 5 ainsi que son extension anatomo-fonctionnelle décrite dans le chapitre 7.

Parcel It Up! fait également appel à certaines routines de SPM concernant les entrées/sorties et les analyses statistiques. A l'avenir, cette boîte à outils sera intégrée à BrainVisa²⁸, afin de pouvoir utiliser simplement les résultats de segmentation des images anatomiques.

27. <http://www.mathworks.com/>

28. <http://www.brainvisa.info/>

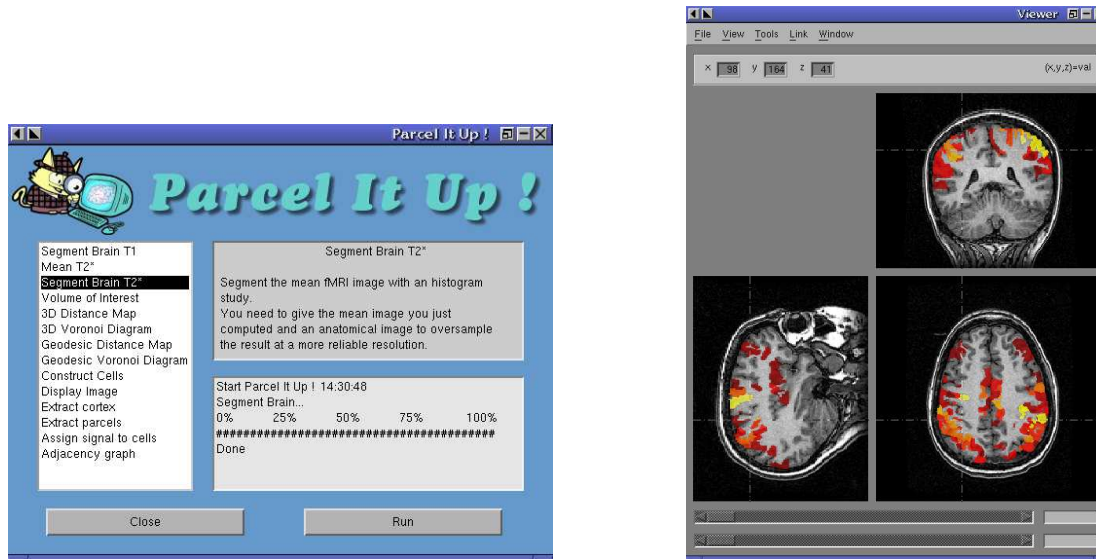


FIG. D.1 – Interface graphique de Parcel It Up! et affichage d'une carte statistique de résultat avec le logiciel View3D développé par Jonathan Stoeckel.

XMLTree

XMLTree²⁹ est une boîte à outils XML³⁰ pour MATLAB, permettant de lire, écrire, modifier un fichier XML grâce à l'implémentation d'un processeur XML mais également d'une partie du langage XPath³¹ pour effectuer des requêtes dans un fichier et ainsi utiliser XML comme une base de données.

XML est un méta-langage permettant de définir d'autres langages. Il permet ainsi de structurer, poser le vocabulaire et la syntaxe des données qu'il va contenir. Ses spécifications ont été mises au point par le World Wide Web Consortium (W3C) en 1998 et il a connu depuis un essor inégalé, car il est particulièrement bien adapté à l'échange de données et de documents. Ses principaux atouts sont :

- lisibilité : simple fichier texte,
- autodéscriptif et extensible,
- structure arborescente (arbre hiérarchique),
- universalité et portabilité,
- déployable : il peut être facilement distribué par tout protocole à même de transporter du texte, tel HTTP,
- intégrabilité : il existe de nombreux processeurs XML disponibles pour chaque langage de programmation, il est ainsi inutile pour le développeur de devoir créer son format de donnée spécifique associé à ses propres routines de lecture/écriture.

29. <http://www.artefact.tk/software/matlab/xml/>

30. *Extensible Markup Language*, <http://www.w3.org/XML/>

31. <http://www.w3.org/XPath/>

Nous avons ainsi utilisé le format de fichier XML pour stocker les paramètres d'entrée et de sortie d'applications telles que Parcel It Up! Les avantages sont multiples : utilisation d'un format ouvert et répandu, stockage sous forme de fichier texte facilement éditable par l'utilisateur, communication avec d'autres logiciels facilitée, possibilité de fonctionnement en mode *batch*, etc [Barillot and Achard, 2000].

```
<parcel>
  <input>
    <anatomy>/home/flandin/data/S4115/mri/S4115.img</anatomy>
    <segmentation>/home/flandin/data/S4115/segment/?cortex.img</segmentation>
    <sulci></sulci>
    <nb_parcel>500</nb_parcel>
  </input>
  <output>
    <dir>/home/flandin/parcel/</dir>
  </output>
  <defaults>
    <display>on</display>
    <nb_iter_max>50</nb_iter_max>
  </defaults>
</parcel>
```

FIG. D.2 – Exemple de fichier XML utilisé dans Parcel It Up!

M2HTML

M2HTML³² est un système de documentation automatique de code MATLAB en HTML. A partir d'un ensemble de fichiers .m, il génère un ensemble de pages HTML contenant un descriptif de chaque fichier (nom, description, synopsis) ainsi que ses dépendances avec d'autres fichiers. En particulier, cela permet de construire des graphes de dépendance de boîtes à outils. Un point intéressant est l'utilisation d'un système de templates HTML permettant d'avoir une séparation complète entre contenant (mise en forme HTML) et contenu (fichier MATLAB). Il existe par ailleurs de nombreuses autres options facilement accessibles, détaillées sur le site internet de M2HTML.

Calcul distribué et SPM

Ce travail est plus prospectif mais le prototype déjà réalisé a permis de montrer l'intérêt de pouvoir distribuer les calculs SPM sur plusieurs machines.

Lors d'une analyse SPM, le nombre d'images à traiter est relativement conséquent, ce qui entraîne des temps de calcul élevés. Or, nombre de traitements sont appliqués sur chaque image indépendamment et itérativement. Ceci se vérifie particulièrement pour

32. <http://www.artefact.tk/software/matlab/m2html/>

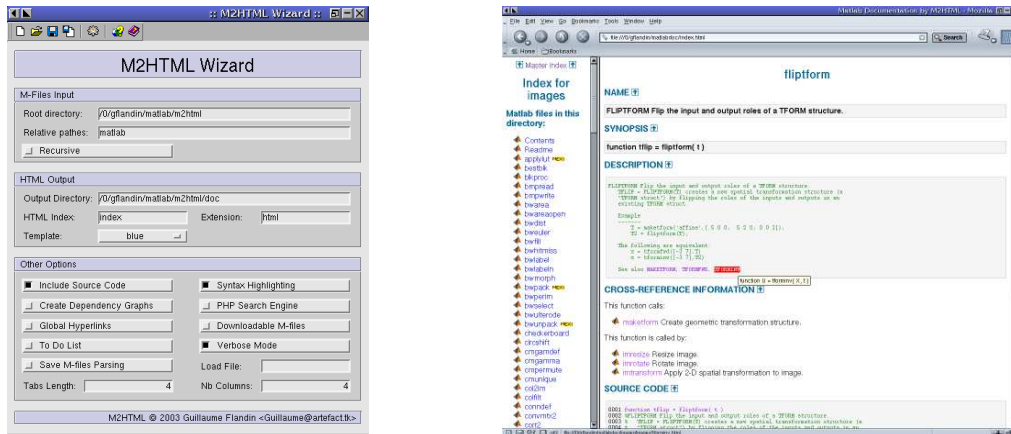


FIG. D.3 – Interface graphique de M2HTML et exemple de documentation générée.

les prétraitements (cf chapitre 3) : correction du délai temporel intercoupe, réalignement, normalisation, lissage. Il est alors tout à fait possible de distribuer ces calculs sur plusieurs machines afin d'accélérer les temps de calcul. Plusieurs laboratoires ont déjà proposé des solutions : Parallel SPM³³, SPM Parallel³⁴, PSPM³⁵.

Notre choix s'est porté sur une boîte à outils MATLAB ParMatlab³⁶, développé par Lucio Andrade [Andrade and Manolakos, 2002], permettant de distribuer des processus sur des clients MATLAB disponibles à travers un intranet ou internet. La discussion entre les différentes machines se fait en utilisant le protocole TCP-IP, permettant de faire communiquer des machines tournant sur des systèmes différents (Windows, Unix, Linux, Solaris) avec des versions de système d'exploitation différentes et des versions de MATLAB éventuellement différentes.

Le principe est représenté sur le schéma de la figure D.4 : c'est un modèle SPMD (*Single Program-Multiple Data*) possédant une architecture client/serveur. Les processus sont lancés à partir d'une machine, faisant office de serveur : elle dispose alors d'une liste de traitements associés à des images situées dans une base de données. Le serveur se met alors en situation d'attente de connections de clients MATLAB. Pour cela, il suffit d'exécuter MATLAB sur une autre machine et d'envoyer au serveur un signal indiquant la disponibilité de la machine, à travers le réseau intranet ou internet. Le serveur enregistre l'information et, s'il a un traitement à faire effectuer, envoie les informations concernant le traitement au client (paramètres d'entrée et de sortie). Le client effectue alors la tâche qui lui a été attribuée et recontacte le serveur à la fin pour déclarer que la tâche demandée a bien été effectuée et qu'il est à nouveau libre pour en effectuer une nouvelle. Un des points forts de ce modèle est qu'il est possible d'ajouter et de supprimer autant de machine client que l'on souhaite à tout moment : c'est le serveur qui se charge de distribuer les processus

33. <http://www.metacube.org/pspm/>

34. <http://spm.parallel.online.fr/>

35. <http://sourceforge.net/projects/parallelsmpm/>

36. <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectType=file&objectId=217>

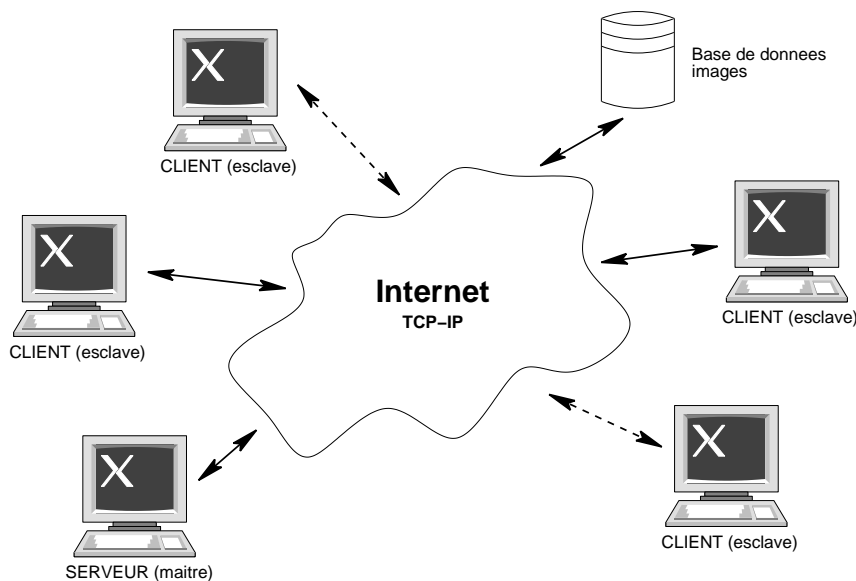


FIG. D.4 – Calcul distribué selon une architecture client/serveur au travers d’un réseau internet ou intranet. D’après [Andrade and Manolakos, 2002].

aux machines disponibles.

Nous avons utilisé ce schéma lors de l’étape de lissage spatial des données dans SPM. Pour cela, nous avons choisi de lisser les données d’un sujet du protocole utilisé au cours de cette thèse, correspondant à 186 images, distribuées sur 8 machines du projet Epidaure reliées par réseau Ethernet 100 MBits, les images étant elles-même archivées sur une 9^{ième} machine. La figure D.5 présente la courbe d’efficacité (*speed up*) en fonction du nombre de processeurs mis en œuvre :

$$CPU^*(N) = \frac{\text{temps de calcul pour 1 CPU}}{\text{temps de calcul sur N CPU}}$$

Les pertes observées par rapport à la courbe idéale s’expliquent principalement par les temps de latence de la machine sur laquelle sont stockées les images, lorsque plusieurs images sont demandées à la fois. Il faut noter que cette version parallèle n’est responsable que d’une très faible augmentation du trafic réseau car même dans le cas monoprocasseur, les images doivent transiter par le réseau. Par ailleurs, on peut s’attendre à de meilleurs résultats avec les opérations de réalignement et de normalisation car ce sont des étapes plus gourmandes en temps de calcul avec des accès en lecture/écriture sur les images moins fréquent. Le véritable inconvénient de la méthode proposée est le fait que cela nécessite autant de licence MATLAB qu’il n’y a de machines en jeu, cependant les laboratoires utilisant de manière intensive SPM disposent fréquemment de licence-site, éliminant de fait le problème.

Il serait intéressant à l’avenir, de mettre en œuvre ce système de calcul distribué pour tous les traitements concernés, en utilisant éventuellement XML lors des communications entre machines. L’objectif de ce système n’est pas d’obtenir un logiciel temps-réel (ceci semble difficilement réalisable) mais plutôt de réduire les temps de calcul en un temps

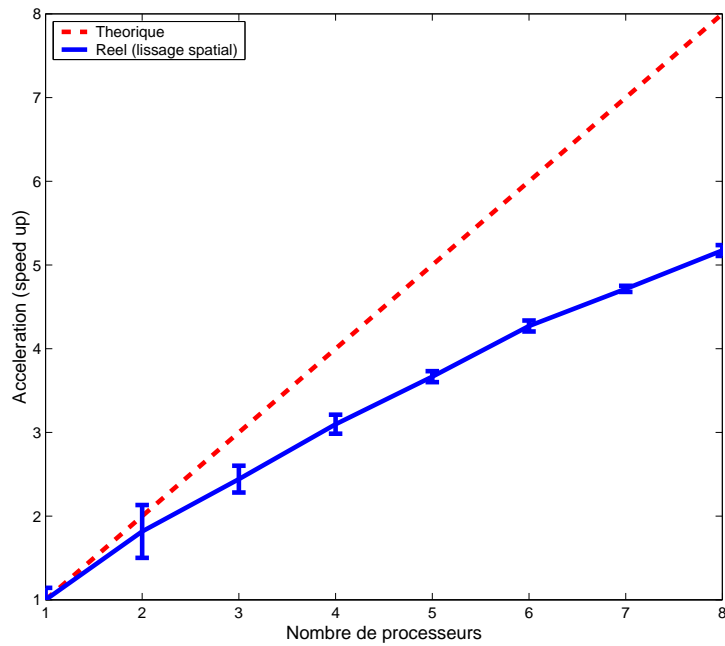


FIG. D.5 – Efficacité du calcul distribué de l’opération de lissage 3D spatial (FWHM 5 mm) appliqué sur 186 images en impliquant de 1 à 8 machines. Les mesures de temps de calcul ont été effectuées 10 fois pour chaque nombre de machine et les barres d’erreur indiquent les écarts-types.

humain, permettant à l’utilisateur de pouvoir relancer les analyses avec différents paramètres, au lieu de se satisfaire de ceux chèrement acquis après une nuit de calcul.

A terme, seule la machine serveur pourrait être physiquement dans le laboratoire de neuroimagerie, les machines client étant alors disposées en une ferme de PCs et accessibles par internet (voir tous les travaux récents concernant le *Grid Computing*).

Bibliographie

- [Akaike, 1973] Hirotugu Akaike. Information theory as an extension of the maximum likelihood principle. In B.N. Petrov and F. Csaki, editors, *Second International Symposium on Information Theory*, pages 267–281, Budapest, 1973. Akademiai Kiado.
- [Akaike, 1974] Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.
- [Almeida and Ledberg, 2001] Rita Almeida and Anders Ledberg. A spatially constrained clustering algorithm with no prior knowledge of the number of clusters. In *NeuroImage*, volume 13, 2001.
- [Ambroise *et al.*, 1997] Christophe Ambroise, Mo Van Dang, and Gérard Govaert. Clustering of spatial data by the EM algorithm. In A. Soares, J. Gomez-Hernandez, and R. Froidevaux, editors, *geoENV I - Geostatistics for Environmental Applications*, volume 9, pages 493–504. Quantitative Geology and Geostatistics, Kluwer Academic Publisher, 1997.
- [Amunts *et al.*, 1999] Katrin Amunts, Axel Schleicher, Uli Bürgel, Hartmut Mohlberg, Harrg B. Uylings, and Karl Zilles. Broca’s region revisited: Cytoarchitecture and inter-subject variability. *Journal of Comparative Neurology*, 412:319–341, 1999.
- [Andersson *et al.*, 2001] Jesper L.R. Andersson, Chloe Hutton, John Ashburner, Robert Turner, and Karl J. Friston. Modeling geometric deformations in EPI time series. *NeuroImage*, 13:903–919, 2001.
- [Andrade and Manolakos, 2002] Lucio Andrade and Elias S. Manolakos. ParMatlab: Coarse parallelization of Matlab processes in heterogeneous networks. In *XIII Research Workshop, Communications and Digital Signal Processing Center*, Northeastern University, May 2002.
- [Andrade *et al.*, 2001] Alexandre Andrade, Ferath Kherif, Jean-François Mangin, Keith J. Worsley, Anne-Lise Paradis, Olivier Simon, Stanislas Dehaene, Denis Le Bihan, and Jean-Baptiste Poline. Detection of fMRI activation using cortical surface mapping. *Human Brain Mapping*, 12:79–93, 2001.
- [Ashburner and Friston, 1997] John Ashburner and Karl J. Friston. Multimodal image coregistration and partitioning - a unified framework. *NeuroImage*, 6(3):209–217, 1997.
- [Ashburner and Friston, 1999] John Ashburner and Karl J. Friston. Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7:254–266, 1999.
- [Ashburner *et al.*, 1997] John Ashburner, Peter Neelin, Louis Collins, Alan C. Evans, and Karl J. Friston. Incorporating prior knowledge into image registration. *NeuroImage*, 6:344–352, 1997.

- [Bahrami, 2000] Stéphane Bahrami. Correction des distorsions des images echo-planaires dues aux inhomogénéités de champ. Technical report, Ecole Polytechnique - CEA/SHFJ, 2000.
- [Baillet and Garnero, 1997] Sylvain Baillet and Line Garnero. A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Transactions on Biomedical Engineering*, 44(5):374–385, 1997.
- [Baillet *et al.*, 1999] Sylvain Baillet, Line Garnero, Gildas Marin, and Jean-Paul Hugonin. Combined MEG and EEG source imaging by minimization of mutual information. *IEEE Transactions on Biomedical Engineering*, 46:522–534, 1999.
- [Bajcsy and Kovačič, 1989] R. Bajcsy and S. Kovačič. Multiresolution Elastic Matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, 1989.
- [Bandettini *et al.*, 1992] Peter A. Bandettini, Eric C. Wong, R. Scott Hinks, Ronald S. Tikofsky, and James S. Hyde. Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*, 25:390–397, 1992.
- [Bannister *et al.*, 2002] Peter Bannister, Stephen Smith, Michael Brady, and Mark Jenkinson. Spatio-temporal realignment of fMRI data. In *NeuroImage, Human Brain Mapping Conference (HBM'02)*, page 380, 2002.
- [Barillot and Achard, 2000] Emmanuel Barillot and Frédéric Achard. XML: a lingua franca for science. *Trends in Biotechnology*, 18(8):331–333, 2000.
- [Barillot *et al.*, 2002] Christian Barillot, Pierre Hellier, and Isabelle Corouge. Non rigid registration in neuroimaging: from a retrospective evaluation of global approaches to new advances in local and hybrid methods. In *IEEE International Symposium on Biomedical Imaging (ISBI'02)*, Washington, USA, 2002.
- [Baumgartner *et al.*, 2000] Richard Baumgartner, Lawrence Ryner, Mathias W. Richter, Randy Summers, Mark Jarmasz, and Ray Somorjai. Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis. *Magnetic Resonance in Medicine*, 18:89–94, 2000.
- [Baune *et al.*, 1999] Axel Baune, Friedrich T. Sommer, Michael Erb, Dirk Wildgruber, Bernd Kardatzki, Günther Palm, and Wolfgang Grodd. Dynamical cluster analysis of cortical fMRI activation. *NeuroImage*, 9:477–489, 1999.
- [Belliveau *et al.*, 1991] Jack W. Belliveau, David N. Kennedy, R.C. McKinstry, Bradley R. Buchbinder, Robert M. Weisskoff, Mark S. Cohen, Jack M. Vevea, Thomas J. Brady, and Bruce R. Rosen. Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*, 25:716–719, 1991.
- [Bertin, 1994] Etienne Bertin. *Diagrammes de Voronoï 2D et 3D: Applications en Analyse d'Images*. PhD thesis, Université Joseph Fourier, Grenoble, France, Janvier 1994.
- [Besag, 1974] Julian E. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*, pages 192–236, 1974.
- [Bezdek, 1981] James C. Bezdek. *Pattern Recognition with Fuzzy Objective Functions Algorithms*. Plenum, New York, 1981.
- [Biernacki *et al.*, 1998] Christophe Biernacki, Gilles Celeux, and Gérard Govaert. Assessing a mixture model for clustering with the integrated classification likelihood. Technical Report 3521, INRIA, 1998.

-
- [Biernacki *et al.*, 1999] Christophe Biernacki, Gilles Celeux, and Gérard Govaert. An improvement of the NEC criterion for assessing the number of components arising from a mixture. *Pattern Recognition letters*, pages 267–272, 1999.
- [Bilmes, 1998] Jeff A. Bilmes. A gentle tutorial of the EM algorithm and its application to parameter estimation for gaussian mixture and hidden Markov models. Technical report, ICSI, Berkeley, 1998.
- [Bishop, 1995] Christopher Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [Biswal *et al.*, 1995] Bharat B. Biswal, Zerrin Yetkin, Victor M. Haughton, and James S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*, 34(4):537–541, August 1995.
- [Bjaalie, 2002] Jan G. Bjaalie. Localization in the brain: new solutions emerging. *Nature Reviews Neuroscience*, 3:322–325, April 2002.
- [Borgefors, 1984] Gunilla Borgefors. Distance transformations in arbitrary dimensions. *Computer Vision, Graphics, and Image Processing*, 27:321–345, February 1984.
- [Borgefors, 1986] Gunilla Borgefors. Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing*, 34(3), 1986.
- [Boynton *et al.*, 1996] Geoffrey M. Boynton, Stephen A. Engel, Gary H. Glover, and David J. Heeger. Linear systems analysis of functional magnetic resonance imaging in human V1. *The Journal of Neuroscience*, 16(13):4207–4221, 1996.
- [Brett *et al.*, 2001] Matthew Brett, Kalina Christoff, Rhodri Cusack, and Jack Lancaster. Using the Talairach atlas with the MNI template. In *NeuroImage, Human Brain Mapping Conference (HBM’01)*, volume 13, page S85, 2001.
- [Brett *et al.*, 2002a] Matthew Brett, Jean-Luc Anton, Romain Valabregue, and Jean-Baptiste Poline. Region of interest analysis using an SPM toolbox. In *NeuroImage (HBM’02)*, Sendai, Japan, 2002.
- [Brett *et al.*, 2002b] Matthew Brett, Ingrid S. Johnsrude, and Adrian M. Owen. The problem of functional localization in the human brain. *Nature Reviews Neuroscience*, 3:243–249, March 2002.
- [Bro-Nielsen, 1996] Morten Bro-Nielsen. *Medical Image Registration and Surgery Simulation*. PhD thesis, Institut for Matematisk Modellering, Danmarks Tekniske Universitet, Lyngby, Denmark, 1996.
- [Broca, 1861] Paul Broca. Remarques sur le siège de la faculté du langage articulé, suivies d’une observation d’aphémie. *Bul Mém Soc. Anat.*, 2:330–357, 1861.
- [Brodmann, 1909] Korbinian Brodmann. *Vergleichende Lokalisationslehre der Großhirnrinde*. Verlag von Johann Ambrosius Barth, Leipzig, 1909.
- [Burnham and Anderson, 1998] Kenneth P. Burnham and David R. Anderson. *Model Selection and Inference*. Springer-Verlag, 1998.
- [Buxton *et al.*, 1998] Richard Buxton, Eric C. Wong, and Lawrence Frank. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magnetic Resonance in Medicine*, 39(6):855–864, 1998.
- [Cachia *et al.*, 2003a] Arnaud Cachia, Jean-François Mangin, Denis Rivière, Ferath Kherif, Nathalie Boddaert, Alexandre Andrade, Dimitri Papadopoulos-Orfanos, Jean-

- Baptiste Poline, Isabelle Bloch, Monica Zilbovicius, Pascale Sonigo, Francis Brunelle, and Jean Régis. A primal sketch of the cortex mean curvature: a morphogenesis based approach to study the variability of the folding patterns. *IEEE Transactions on Medical Imaging*, 22(6):754–765, 2003.
- [Cachia *et al.*, 2003b] Arnaud Cachia, Jean-François Mangin, Denis Rivière, Dimitri Papadopoulos-Orfanos, Ferath Kherif, Isabelle Bloch, and Jean Régis. A generic framework for parcellation of the cortical surface into gyri using geodesic Voronoï diagrams. *Medical Image Analysis*, 7(4):403–416, 2003.
- [Cachia, 2003] Arnaud Cachia. *Modèles statistiques morphométriques et structurels du cortex pour l'étude du développement cérébral*. PhD thesis, ENST, 2003.
- [Cachier *et al.*, 2001] Pascal Cachier, Jean-François Mangin, Xavier Pennec, Denis Rivière, Dimitri Papadopoulos-Orfanos, Jean Régis, and Nicholas Ayache. Multisubject non-rigid registration of brain MRI using intensity and geometric features. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'01)*, volume 2208 of *LNCS*, pages 734–742, Utrecht, The Netherlands, October 2001. Springer Verlag.
- [Cachier *et al.*, 2003] Pascal Cachier, Eric Bardinet, Didier Dormont, Xavier Pennec, and Nicholas Ayache. Iconic Feature Based Nonrigid Registration: The PASHA Algorithm. *CVIU — Special Issue on Nonrigid Registration*, 89(2-3):272–298, 2003.
- [Cachier, 2002] Pascal Cachier. *Recalage non rigide d'images médicales volumiques - contribution aux approches iconiques et géométriques*. PhD thesis, École Centrale des Arts et Manufactures, January 2002.
- [Celeux and Govaert, 1992] Gilles Celeux and Gérard Govaert. A classification EM algorithm for clustering and two stochastic versions. *Computational statistics and data analysis*, 14:315–332, 1992.
- [Christensen *et al.*, 1997] Gary Christensen, Sarang Joshi, and Michael I. Miller. Volumetric Transformation of Brain Anatomy. *IEEE Transactions on Medical Imaging*, 16(6):864–877, December 1997.
- [Chuang *et al.*, 1999] Kai-Hsiang Chuang, Ming-Jang Chiu, Chung-Chih Lin, and Jyh-Horng Chen. Model-free functional MRI analysis using Kohonen clustering neural network and fuzzy c-means. *IEEE Transactions on Medical Imaging*, 18(12):1117–1128, 1999.
- [Clerc *et al.*, 2002] Maureen Clerc, Renaud Keriven, Olivier Faugeras, Jan Kybic, and Théo Papadopoulos. The fast multipole method for the direct E/MEG problem. In *IEEE International Symposium on Biomedical Imaging (ISBI'02)*, Washington, USA, 2002.
- [Cointepas *et al.*, 2001] Yann Cointepas, Jean-François Mangin, Line Garnero, Jean-Baptiste Poline, and Habib Benali. BrainVISA: Software platform for visualization and analysis of multi-modality brain data. In *Proc. 7th HBM*, page S98, Brighton, United Kingdom, 2001.
- [Collignon *et al.*, 1995] André Collignon, Frederik Maes, Pierre Delaere, Dirk Vandermeulen, Paul Suetens, and Guy Marchal. Automated multi-modality image registration based on information theory. In *International Conference on Information Processing in Medical Images (IPMI'95)*, pages 263–274, 1995.

-
- [Collins *et al.*, 1994] Louis Collins, Peter Neelin, Terence M. Peters, and Alan C. Evans. Automatic 3D inter-subject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, 18(2):192–205, 1994.
- [Collins *et al.*, 1998] Louis Collins, Georges Le Goualher, and Alan C. Evans. Non-linear cerebral registration with sulcal constraints. In W.M. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'98)*, volume 1496, pages 974–984, Berlin, Germany, October 1998. Lecture Notes in Computer Science, Springer.
- [Corouge *et al.*, 2001] Isabelle Corouge, Christian Barillot, Pierre Hellier, Pierre Toulouse, and Bernard Gibaud. Non-linear local registration of functional data. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'01)*, volume 2208 of *LNCS*, pages 948–956, Utrecht, The Netherlands, October 2001.
- [Coulon *et al.*, 2000] Olivier Coulon, Jean-Francois Mangin, Jean-Baptiste Poline, Monica Zilbovicius, Dimitri Roumenov, Yves Samson, Vincent Frouin, and Isabelle Bloch. Structural group analysis of functional activation maps. *NeuroImage*, 11:767–782, 2000.
- [Cressie, 1993] Noel A.C. Cressie. *Statistics for Spatial Data*. Wiley, 1993.
- [Crum *et al.*, 2003] W.R. Crum, L.D. Griffin, D.L.G. Hill, and D.J. Hawkes. Zen and the art of medical image registration: correspondence, homology, and quality. *NeuroImage*, 20:1425–1437, 2003.
- [Cuisenaire, 1997] Olivier Cuisenaire. Region growing euclidean distance transforms. In *9th International Conference on Image Analysis and Processing (ICIAP'97)*, volume 1, pages 263–270, Florence, Italy, September 1997.
- [Cuisenaire, 1999] Olivier Cuisenaire. *Distance Transformations: Fast Algorithms and Applications to Medical Image Processing*. PhD thesis, Université Catholique de Louvain, Belgique, October 1999.
- [Dale and Sereno, 1999] Anders M. Dale and Martin I. Sereno. Cortical surface-based analysis. I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179–194, 1999.
- [Damasio *et al.*, 1994] Hanna Damasio, Thomas Grabowski, Randell Frank, Albert M. Galaburda, and Antonio R. Damasio. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, 264:1102–1105, 1994.
- [Dang and Govaert, 1998] Mo Van Dang and Gérard Govaert. Spatial fuzzy clustering using EM and Markov random fields. *System Research and Information Systems*, 8:183–202, 1998.
- [Danielsson, 1980] Per-Erik Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14:227–248, 1980.
- [Dellaert, 2002] Frank Dellaert. The expectation maximization algorithm. Available at <http://www.cc.gatech.edu/~dellaert/html/publications.html>, February 2002.
- [Dempster *et al.*, 1977] Arthur P. Dempster, Nan M. Laird, and Donald B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39 (Series B):1:38, 1977.
- [Descombes *et al.*, 1998a] Xavier Descombes, Frithjof Kruggel, and D. Yves von Cramon. fMRI signal restoration using a spatio-temporal Markov random field preserving transitions. *NeuroImage*, 8:340–349, 1998.

- [Descombes *et al.*, 1998b] Xavier Descombes, Frithjof Kruggel, and D. Yves von Cramon. Spatio-temporal fMRI analysis using Markov random fields. *IEEE Transactions on Medical Imaging*, 17(6):1028–1039, December 1998.
- [Du *et al.*, 1999] Qiang Du, Vance Faber, and Max Gunzburger. Centroidal Voronoï tessellations: Applications and algorithms. *Society for Industrial and Applied Mathematics Review*, 41(4):637–676, 1999.
- [Duda and Hart., 1973] Richard O. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis*. Wiley, New York, 1973.
- [Evans *et al.*, 1992] Alan C. Evans, Louis Collins, and Brenda Milner. An MRI-based stereotactic atlas from 250 young normal subjects. In *Soc. Neurosci. Abstr.*, volume 18, page 408, 1992.
- [Evans *et al.*, 1993] Alan C. Evans, Louis Collins, Steve R. Mills, E.D. Brown, R.L. Kelly, and Terence M. Peters. 3D statistical neuroanatomical models from 305 MRI volumes. In L. Klaisner, editor, *IEEE Nuclear Science Symposium and Medical Imaging Conference*, volume 18, pages 1813–1817, Piscataway, New Jersey, 1993.
- [Everitt and Bullmore, 1999] Brian S. Everitt and Edward T. Bullmore. Mixture model mapping of brain activation in functional magnetic resonance images. *Human Brain Mapping*, 7:1–14, 1999.
- [Fadili *et al.*, 2000] Mohamed Jalal Fadili, Su Ruan, Daniel Bloyet, and Bernard Mazoyer. A multi-step unsupervised fuzzy clustering analysis of fMRI time series. *Human Brain Mapping*, 10(4):160–178, 2000.
- [Fischl *et al.*, 1999a] Bruce Fischl, Martin I. Sereno, and Anders M. Dale. Cortical surface-based analysis. II. inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195–207, 1999.
- [Fischl *et al.*, 1999b] Bruce Fischl, Martin I. Sereno, Roger B.H. Tootell, and Anders M. Dale. High-resolution inter-subject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4):272–284, 1999.
- [Flandin *et al.*, 2002a] Guillaume Flandin, Ferath Kherif, Xavier Pennec, Grégoire Malandain, Nicholas Ayache, and Jean-Baptiste Poline. Improved detection sensitivity in functional MRI data using a brain parcelling technique. In Takeyoshi Dohi and Ron Kikinis, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, volume 2488 of *LNCS*, pages 467–474, Tokyo, September 2002. Springer.
- [Flandin *et al.*, 2002b] Guillaume Flandin, Ferath Kherif, Xavier Pennec, Denis Rivière, Nicholas Ayache, and Jean-Baptiste Poline. A new representation of fMRI data using anatomico-functional constraints. In *NeuroImage (HBM'02)*, Sendai, Japan, 2002.
- [Flandin *et al.*, 2002c] Guillaume Flandin, Ferath Kherif, Xavier Pennec, Denis Rivière, Nicholas Ayache, and Jean-Baptiste Poline. Parcellation of brain images with anatomical and functional constraints for fMRI data analysis. In *IEEE International Symposium on Biomedical Imaging*, pages 907–910, Washington, USA, 2002.
- [Flandin *et al.*, 2003] Guillaume Flandin, William D. Penny, Xavier Pennec, Nicholas Ayache, and Jean-Baptiste Poline. A multisubject anatomico-functional parcellation of the brain. In Tomas Paus, Ed Bullmore, and Jonathan D. Cohen, editors, *NeuroImage (HBM'03)*, New York, USA, 2003. Academic Press.

-
- [Fox *et al.*, 1986] Peter T. Fox, Mark A. Mintun, Marcus E. Raichle, Fran M. Miezin, John M. Allman, and David C. Van Essen. Mapping human visual cortex with positron emission tomography. *Nature*, 323:806–809, 1986.
- [Fox *et al.*, 1988] Peter T. Fox, Marcus E. Raichle, Mark A. Mintun, and C. Dence. Non-oxidative glucose consumption during focal physiological neural activity. *Science*, 242:462–464, 1988.
- [Fox, 1995] Peter T. Fox. Spatial normalization origins: Objectives, applications, and alternatives. *Human Brain Mapping*, pages 161–164, 1995.
- [Frackowiak *et al.*, 1997] Richard S.J. Frackowiak, Karl J. Friston, Chris D. Frith, Raymond J. Dolan, and John C. Mazziotta. *Human Brain Function*. Academic Press, 1997.
- [Frackowiak *et al.*, 2004] Richard S.J. Frackowiak, Karl J. Friston, Chris D. Frith, Raymond J. Dolan, John Ashburner, Cathy Price, William D. Penny, and Semir Zeki. *Human Brain Function II*. Academic Press, 2004.
- [Freeman *et al.*, 2003] Walter J. Freeman, Mark D. Holmes, Brian C. Burke, and Sampsa Vanhatalo. Spatial spectra of scalp EEG and EMG from awake humans. *Clinical Neurophysiology*, 114(6):1053–1068, 2003.
- [Freire and Mangin, 2001] Luis Freire and Jean-François Mangin. Motion correction algorithms may create spurious brain activations in the absence of subject motion. *Neuroimage*, 14(3):709–722, 2001.
- [Freire *et al.*, 2002] Luis Freire, Alexis Roche, and Jean-François Mangin. What is the best similarity measure for motion correction in fMRI time series? *IEEE Transactions on Medical Imaging*, 21(5):470–484, May 2002.
- [Friman *et al.*, 2002] Ola Friman, Magnus Borga, Peter Lundberg, and Hans Knutsson. Detection of neural activity in fMRI using maximum correlation modeling. *NeuroImage*, 15(2):386–395, 2002.
- [Friston *et al.*, 1994] Karl J. Friston, Peter Jezzard, and Robert Turner. The analysis of functional MRI time-series. *Human Brain Mapping*, 1:153–171, 1994.
- [Friston *et al.*, 1995a] Karl J. Friston, John Ashburner, Chris D. Frith, Jean-Baptiste Poline, Jon D. Heather, and Richard S.J. Frackowiak. Spatial registration and normalization of images. *Human Brain Mapping*, 2:165–189, 1995.
- [Friston *et al.*, 1995b] Karl J. Friston, Andrew P. Holmes, Jean-Baptiste Poline, Christopher D. Frith, and Richard S.J. Frackowiak. Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2:189–210, 1995.
- [Friston *et al.*, 1995c] Karl J. Friston, Andrew P. Holmes, Jean-Baptiste Poline, Paul J. Grasby, Steve C.R. Williams, Richard S.J. Frackowiak, and Robert Turner. Analysis of fMRI time-series revisited. *NeuroImage*, 2:45–53, 1995.
- [Friston *et al.*, 1996] Karl J. Friston, Steve C.R. Williams, Robert Howard, Richard S.J. Frackowiak, and Robert Turner. Movement-related effects in fMRI time series. *Magnetic Resonance in Medicine*, 35:346–355, 1996.
- [Friston *et al.*, 1998] Karl J. Friston, Oliver Josephs, Geraint Rees, and Robert Turner. Non-linear event-related response in fMRI. *Magnetic Resonance in Medicine*, 39(1):41–52, 1998.

- [Friston *et al.*, 2000] Karl J. Friston, Oliver Josephs, Eric Zarahn, Andrew P. Holmes, Stéphanie Rouquette, and Jean-Baptiste Poline. To smooth or not to smooth? bias and efficiency in fMRI time-series analysis. *NeuroImage*, 12:196–208, 2000.
- [Genovese, 1999] Christopher R. Genovese. Functional magnetic resonance imaging and spatio-temporal inference. *Bayesian Statistics*, 6:255–274, 1999.
- [Glover, 1999] Gary H. Glover. Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage*, 9(4):416–429, 1999.
- [Golay *et al.*, 2000] Xavier Golay, Klaas P. Pruessmann, Markus Weiger, Gérard R. Crelier, Paul J.M. Folkers, Spyros S. Kollias, and Peter Boesiger. PRESTO-SENSE: an ultrafast whole-brain fMRI technique. *Magnetic Resonance in Medicine*, 43:779–786, 2000.
- [Goldman *et al.*, 2000] Robin I. Goldman, John M. Stern, Jerome Engel Jr., and Mark S. Cohen. Acquiring simultaneous EEG and functional MRI. *Clin. Neuroph.*, 111:1974–1980, 2000.
- [Gossl *et al.*, 2001] Christoff Gossl, Dorothee P. Auer, and Ludwig Fahrmeir. Bayesian spatio-temporal inference in functional magnetic resonance imaging. *Biometrics*, pages 554–562, 2001.
- [Goutte *et al.*, 1999] Cyril Goutte, Peter Toft, Egill Rostrup, Finn Årup Nielsen, and Lars Kai Hansen. On clustering fMRI time series. *NeuroImage*, 9(3), 1999.
- [Goutte *et al.*, 2001] Cyril Goutte, Lars Kai Hansen, Matthew G. Liptrot, and Egill Rosstrup. Feature-space clustering for fMRI meta-analysis. *Human Brain Mapping*, pages 165–183, 2001.
- [Green, 1990] Peter J. Green. On use of the EM algorithm for penalized likelihood estimation. *Journal of the Royal Statistical Society*, 52(3):443–452, 1990.
- [Greenblatt, 1995] Samuel H. Greenblatt. Phrenology in the science and culture of the 19th century. *Neurosurgery*, 37:790–805, 1995.
- [Guimond *et al.*, 1998] Alexandre Guimond, Jean Meunier, and Jean-Philippe Thirion. Automatic computation of average brain models. In W.M. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'98)*, volume 1496, pages 623–630, Berlin, Germany, October 1998. Lecture Notes in Computer Science, Springer.
- [Guimond *et al.*, 2000] Alexandre Guimond, Jean Meunier, and Jean-Philippe Thirion. Average brain models: a convergence study. *Computer Vision and Image Understanding*, 77(2):192–210, February 2000.
- [Hales, 2000] Thomas C. Hales. Cannonballs and honeycombs. *Notices of the American Mathematical Society*, 47(4):440–449, 2000.
- [Hales, 2001] Thomas C. Hales. The honeycomb conjecture. *Discrete and Computational Geometry*, 25:1–22, 2001.
- [Hartigan and Wong, 1979] John A. Hartigan and M. Anthony Wong. A K-means clustering algorithm. *Applied Statistics*, 28:100–108, March 1979.
- [Hartvig and Jensen, 2000] Niels V. Hartvig and Jens L. Jensen. Spatial mixture modeling of fMRI data. *Human Brain Mapping*, 11:233–248, 2000.

-
- [Hathaway, 1986] Richard J. Hathaway. Another interpretation of the EM algorithm for mixture distributions. *Journal of Statistics & Probability Letters*, 4:53–56, 1986.
- [Hausner, 2001] Alejo Hausner. Simulating decorative mosaics. In Eugene Fiume, editor, *SIGGRAPH*, 2001.
- [Heeger and Rees, 2002] David J. Heeger and David Rees. What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience*, 3:142–151, 2002.
- [Hellier and Barillot, 2000] Pierre Hellier and Christian Barillot. Multimodal non-rigid warping for correction of distortions in functional MRI. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'00)*, volume 1935 of *LNCS*, pages 512–520, Pittsburgh, Pennsylvania, October 2000.
- [Hellier and Barillot, 2001] Pierre Hellier and Christian Barillot. Cooperation between local and global approaches to register brain images. In M.F. Insana and R.M. Leahy, editors, *International Conference on Information Processing in Medical Images (IPMI'01)*, volume 2082 of *LNCS*, pages 315–328. Springer, 2001.
- [Hellier et al., 2001a] Pierre Hellier, Christian Barillot, Isabelle Corouge, Bernard Gibaud, Georges Le Goualher, Louis Collins, Alan C. Evans, Grégoire Malandain, and Nicholas Ayache. Retrospective evaluation of inter-subject brain registration. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'01)*, volume 2208 of *LNCS*, pages 258–265, Utrecht, The Netherlands, October 2001.
- [Hellier et al., 2001b] Pierre Hellier, Christian Barillot, Etienne Mémin, and Patrick Pérez. Hierarchical estimation of a dense transformation field for 3D robust registration. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 2001.
- [Hellier et al., 2002] Pierre Hellier, John Ashburner, Isabelle Corouge, Christian Barillot, and Karl J. Friston. Inter subject registration of functional and anatomical data using SPM. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, volume 2489 of *LNCS*, pages 590–587, Tokyo, Japan, September 2002.
- [Hellier et al., 2003] Pierre Hellier, Christian Barillot, Isabelle Corouge, Bernard Gibaud, Georges Le Goualher, Louis Collins, Alan C. Evans, Grégoire Malandain, Nicholas Ayache, Gary Christensen, and Hans Johnson. Retrospective evaluation of inter-subject brain registration. *IEEE Transactions on Medical Imaging*, 22(9):1120–1130, 2003.
- [Henson et al., 1999] Rick N.A. Henson, Christian Büchel, Oliver Josephs, and Karl J. Friston. The slice-timing problem in event-related fMRI. In *NeuroImage, Human Brain Mapping Conference (HBM'99)*, 1999.
- [Hermosillo, 2003] Gerardo Hermosillo. *Variational Methods for Multimodal Image Matching*. PhD thesis, INRIA Sophia Antipolis, 2003.
- [Heskes, 2001] Tom Heskes. Self-organizing maps, vector quantization and mixture modeling. *IEEE Transactions on Neural Networks*, 12(6):1299–1305, 2001.
- [Hill et al., 2001] Derek L.G. Hill, Philipp G. Batchelor, Mark Holden, and David J. Hawkes. Medical image registration. *Phys. Med. Biol.*, 46:R1–R45, 2001.
- [Houdé et al., 2002] Olivier Houdé, Bernard Mazoyer, and Nathalie Tzourio-Mazoyer. *Cerveau et Psychologie : Introduction à l'Imagerie Cérébrale Anatomique et Fonctionnelle*. Presses Universitaires de France, 2002.

- [Ichihashi *et al.*, 2001] Hidetomo Ichihashi, Kiyotaka Miyagishi, and Katsuhiko Honda. Fuzzy C-means clustering with regularization by Kullback-Leibler information. In *FUZZ-IEEE*, pages 924–927, 2001.
- [Jebara and Pentland, 1998] Tony Jebara and Alex Pentland. Maximum conditional likelihood via bound maximization and the CEM algorithm. In *Neural Information Processing Systems 11*, 1998.
- [Jenkinson *et al.*, 2002] Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–841, 2002.
- [Jezzard and Balaban, 1995] Peter Jezzard and Robert S. Balaban. Correction for geometric distortion in echo planar images from B_0 field variations. *Magnetic Resonance in Medicine*, 34:65–73, 1995.
- [Jezzard and Clare, 1999] Peter Jezzard and Stuart Clare. Sources of distortion in functional MRI data. *Human Brain Mapping*, 8:80–85, 1999.
- [Johnson, 1994] Valent E. Johnson. A model of segmentation and analysis of noisy images. *American Statistical Association*, 89(425):230–241, 1994.
- [Jordan and Bishop, 2002] Michael I. Jordan and Christopher Bishop. *An Introduction to Graphical Models*. MIT Press, 2002. In press.
- [Kapur *et al.*, 1998] Tina Kapur, Eric Grimson, William M. Wells III, and Ron Kikinis. Enhanced spatial priors for segmentation of magnetic resonance imagery. In *MICCAI*, 1998.
- [Kapur, 1999] Tina Kapur. *Model Based Three Dimensional Medical Image Segmentation*. PhD thesis, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, May 1999.
- [Kherif *et al.*, 2002a] Ferath Kherif, Guillaume Flandin, Philippe Ciuciu, Habib Benali, Olivier Simon, and Jean-Baptiste Poline. Model based spatial and temporal similarity measures between series of functional magnetic resonance images. In Takeyoshi Dohi and Ron Kikinis, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, volume 2489 of *LNCS*, pages 509–516, Tokyo, September 2002. Springer.
- [Kherif *et al.*, 2002b] Ferath Kherif, Guillaume Flandin, Olivier Simon, Habib Benali, and Jean-Baptiste Poline. How much subjects are resembling each other? spatial and temporal similarity measures for fMRI data. In *NeuroImage (HBM'02)*, Sendai, Japan, 2002.
- [Kherif *et al.*, 2002c] Ferath Kherif, Jean-Baptiste Poline, Guillaume Flandin, Habib Benali, Olivier Simon, Stanislas Dehaene, and Keith J. Worsley. Multivariate model specification for fMRI data. *NeuroImage*, 16(4):1068–1083, August 2002.
- [Kherif *et al.*, 2003] Ferath Kherif, Jean-Baptiste Poline, Sébastien Meriaux, Habib Benali, Guillaume Flandin, and Matthew Brett. Group analysis in functional neuroimaging: selecting subjects using similarity measures. *NeuroImage*, 20(4):2197–2208, December 2003.
- [Kherif, 2003] Ferath Kherif. *Applications de modèles statistiques multivariés à la résolution de problèmes posés par les données d'imagerie fonctionnelle cérébrale*. PhD thesis, Université Paris 7 - CEA/SHFJ, 2003.

-
- [Kiebel and Friston, 2002] Stefan J. Kiebel and Karl J. Friston. Anatomically informed basis functions in multisubject studies. *Human Brain Mapping*, 16:36–46, 2002.
- [Kiebel *et al.*, 1999] Stefan J. Kiebel, Jean-Baptiste Poline, Karl J. Friston, Andrew P. Holmes, and Keith J. Worsley. Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *NeuroImage*, 10(6):756–766, 1999.
- [Kiebel *et al.*, 2000] Stefan J. Kiebel, Rainer Goebel, and Karl J. Friston. Anatomically informed basis functions. *NeuroImage*, 11(6):656–667, 2000.
- [Kim and Ugurbil, 1997] Seong-Gi Kim and Kamil Ugurbil. Comparison of blood oxygenation and cerebral blood flow effects in fMRI : Estimation of relative oxygen consumption change. *Magnetic Resonance in Medicine*, 38:59–65, 1997.
- [Klein *et al.*, 2003] Isabelle Klein, Ferath Kherif, Jean-François Mangin, Guillaume Flandin, Jean-Baptiste Poline, Michel Denis, Stephen M. Kosslyn, and Denis Le Bihan. Imagined stimulus orientation alters activations in area V1. Technical report, CEA/SHFJ, 2003.
- [Kohonen, 2001] Teuvo Kohonen. *Self-Organizing Maps*, volume 30. Springer, 2001.
- [Kruggel *et al.*, 2003] Frithjof Kruggel, Martina K. Brückner, Thomas Arendt, Christopher J. Wiggins, and D. Yves von Cramon. Analyzing the neocortical fine-structure. *Medical Image Analysis*, 7:251–264, 2003.
- [Kwong *et al.*, 1992] Kenneth Kwong, Jack W. Belliveau, D.A. Chesler, I.E. Goldberg, Robert M. Weisskoff, B.P. Poncelet, David N. Kennedy, B.E. Hoppel, Mark S. Cohen, Robert Turner, H.-M. Cheng, Thomas J. Brady, and Bruce R. Rosen. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad. Sci. USA*, 85:5675–5679, 1992.
- [Kybic *et al.*, 2000] Jan Kybic, Philippe Thévenaz, Arto Nirkko, and Michael Unser. Unwarping of unidirectionally distorted EPI images. *IEEE Transactions on Medical Imaging*, 19(2):80–93, 2000.
- [Lahaye *et al.*, 2003a] Pierre-Jean Lahaye, Jean-Baptiste Poline, Guillaume Flandin, Silke Dodel, and Line Garnero. Functional connectivity: studying nonlinear, delayed interactions between BOLD signals. *NeuroImage*, 20(2):962–974, October 2003.
- [Lahaye *et al.*, 2003b] Pierre-Jean Lahaye, Jean-Baptiste Poline, Guillaume Flandin, and Line Garnero. Functional connectivity: studying non-linear, asynchronous interactions between BOLD signals. In Tomas Paus, Ed Bullmore, and Jonathan D. Cohen, editors, *NeuroImage (HBM’03)*, New York, USA, 2003. Academic Press.
- [Lahaye *et al.*, 2004] Pierre-Jean Lahaye, Sylvain Baillet, Jean-Baptiste Poline, and Line Garnero. Fusion of simultaneous fMRI/EEG data based on the electro-metabolic coupling. In *Proc. 2th Proc. IEEE ISBI*, Arlington, Virginia, Apr. 2004.
- [Lancaster *et al.*, 2000] Jack Lancaster, Marty G. Woldorff, Lawrence M. Parsons, Mario Liotti, Catarina S. Freitas, L. Rainey, Peter V. Kochunov, D. Nickerson, Shawn A. Mikiten, and Peter T. Fox. Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, 10:120–131, March 2000.
- [Lauterbur, 1973] Paul Lauterbur. Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature*, 242:190–191, 1973.

- [Le Bihan *et al.*, 2001] Denis Le Bihan, Jean-François Mangin, Cyril Poupon, C.A. Clark, Sabine Pappata, Nicolas Molko, and Hugues Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:543–546, 2001.
- [Le Goualher *et al.*, 1997] Georges Le Goualher, Christian Barillot, and Yves Bizais. Modeling cortical sulci with active ribbon. *Journal of Pattern Recognition and Artificial Intelligence*, 11:1295–1315, 1997.
- [Leemput *et al.*, 2001] Koen Van Leemput, Frederik Maes, Dirk Vandermeulen, Alan Colchester, and Paul Suetens. Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Transactions on Medical Imaging*, 20(8):677–688, August 2001.
- [Leemput, 2001] Koen Van Leemput. *Quantitative Analysis of Signal Abnormalities in MR Imaging*. PhD thesis, Katholieke Universiteit Leuven, May 2001.
- [Lester and Arridge, 1999] Hava Lester and Simon R. Arridge. A survey of hierarchical non-linear medical image registration. *Pattern Recognition*, 32:129–149, 1999.
- [Logothetis *et al.*, 2001] Nikos K. Logothetis, Jon Pauls, Mark Augath, Torsten Trinath, and Axel Oeltermann. Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412:150–157, 2001.
- [Maintz and Viergever, 1998] J.B. Antoine Maintz and Max A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.
- [Maldjian *et al.*, 2003] Joseph A. Maldjian, Paul J. Laurienti, Robert A. Kraft, and Jonathan H. Burdette. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19:1233–1239, 2003.
- [Mangin *et al.*, 1995] Jean-François Mangin, Vincent Frouin, Isabelle Bloch, Jean Régis, and Jaime Lopez-Krahe. From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. *Journal of Mathematical Imaging and Vision*, 5:297–318, 1995.
- [Mangin *et al.*, 1998] Jean-François Mangin, Jean Régis, Jean-Baptiste Poline, Denis Rivière, Cyril Poupon, Fabrice Poupon, Dimitri Papadopoulos-Orfanos, Fabien Delaye, Olivier Pizzato, Olivier Coulon, S. Crozier, Pascal Belin, Monica Zilbovicius, C. A. Clark, Denis Le Bihan, Dominique Hasboun, Stéphane Lehericy, Didier Dormont, O. de Dreuille, H. Foehrenbach, J.-F. Gaillard, Isabelle Bloch, Yves Samson, and Vincent Frouin. Place de l’anatomie dans la cartographie fonctionnelle du cerveau. *Annales de l’Institut Pasteur*, 9:243–258, 1998.
- [Mangin *et al.*, 2003] Jean-François Mangin, Denis Rivière, Olivier Coulon, Cyril Poupon, Arnaud Cachia, Yann Cointepas, Jean-Baptiste Poline, Denis Le Bihan, Jean Régis, and Dimitri Papadopoulos-Orfanos. Coordinate-based versus structural approaches to brain image analysis. *Artificial Intelligence in Medicine*, 2003. In press.
- [Mangin, 1995] Jean-François Mangin. *Mise en Corrépondance d’Images Médicales 3D Multi-modalités Multi-individus pour la Corrélation Anatomico-fonctionnelle Cérébrale*. PhD thesis, ENST, Mars 1995.
- [Mangin, 2000] Jean-François Mangin. Entropy minimization for automatic correction of intensity nonuniformity. In *Mathematical Method in Biomedical Image Analysis (MMBIA’00)*, pages 162–169, Hilton Head Island, South Carolina, USA, 2000.

-
- [Mansfield, 1977] Peter Mansfield. Multi-planar image formation using NMR spin echoes. *J. Phys. C.*, 10:L55–L58, 1977.
- [Marquardt, 1963] Donald W. Marquardt. An algorithm for least squares estimation of nonlinear parameters. *SIAM Journal Applied Math.*, 11:431–441, 1963.
- [Mazoyer, 2001] Bernard Mazoyer. *L’Imagerie Cérébrale Fonctionnelle*. Presses Universitaires de France, 2001.
- [McLachlan and Peel, 2000] Geoffrey McLachlan and David Peel. *Finite Mixture Models*. Wiley Series in Probability and Statistics, 2000.
- [Minka, 1998] Thomas P. Minka. Expectation-maximization as a lower bound maximization. Available at <http://www.stat.cmu.edu/~minka/papers/em.html>, November 1998.
- [Modersitzki, 2004] Jan Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, New York, 2004.
- [Neal and Hinton, 1998] Radford M. Neal and Geoffrey E. Hinton. A view of the EM algorithm that justifies incremental, sparse, and other variants. In *Learning in Graphical Models*, pages 355–368. Kluwer Academic Publishers, 1998.
- [Nielsen and Hansen, 2002] Finn Årup Nielsen and Lars Kai Hansen. Modeling of activation data in the BrainMap: Detection of outliers. *Human Brain Mapping*, 15(3):146–156, March 2002.
- [Nieto-Castanon *et al.*, 2003] Alfonso Nieto-Castanon, Satrajit S. Ghosh, Jason A. Tourville, and Frank H. Guenther. Region of interest based analysis of functional imaging data. *NeuroImage*, 19:1303–1316, 2003.
- [O’Craven and Kanwisher, 2000] Kathy M. O’Craven and Nancy Kanwisher. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, 12(6):1013–1023, 2000.
- [Ogawa *et al.*, 1992] S. Ogawa, D.W. Tank, R. Menon, J. Ellerman, S.-G. Kim, H. Merkle, and Kamil Ugurbil. Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proc. Natl. Acad. Sci. USA*, 89:5951–5955, 1992.
- [Ono *et al.*, 1990] Michio Ono, Stefan Kubik, and Chad D. Abernathy. *Atlas of the Cerebral Sulci*. Thieme Medical, New York, 1990.
- [Ormoneit and Tresp, 1997] Dirk Ormoneit and Volker Tresp. Improved gaussian mixture density estimates using bayesian penalty terms and network averaging. Technical report, Technische Universität München, 1997.
- [Ormoneit and Tresp, 1998] Dirk Ormoneit and Volker Tresp. Averaging, maximum penalized likelihood and bayesian estimation for improving gaussian mixture probability density estimates. *IEEE Transactions on Neural Networks*, 9(4):639–650, 1998.
- [Paradis, 2001] Anne-Lise Paradis. *Exploration par IRM fonctionnelle de la perception visuelle des formes tridimensionnelles ; dessin de paradigmes d’acquisition et outils d’analyse appliqués au domaine de la vision*. PhD thesis, Université Paris 6, 2001.
- [Passingham *et al.*, 2002] Richard E. Passingham, Klaas E. Stephan, and Rolf Kötter. The anatomical basis of functional localization in the cortex. *Nature Reviews Neuroscience*, 3:606–616, August 2002.

- [Pelat, 1998] Didier Pelat. Bruits et signaux. Introduction aux méthodes de traitements des données. Technical report, Observatoire de Paris, section de Meudon, 1998.
- [Penfield and Rasmussen, 1950] Wilder Penfield and Théodore Rasmussen. *The cerebral cortex of man*. Macmillan, 1950.
- [Pennec, 1995] Xavier Pennec. *L'incertitude dans les problèmes de reconnaissance et de recalage. Application en imagerie médicale et biologie moléculaire*. PhD thesis, Ecole Polytechnique, 1995.
- [Penny and Friston, 2003] William D. Penny and Karl J. Friston. Mixture of general linear models for functional neuroimaging. *IEEE Transactions on Medical Imaging*, 22(4):504–514, 2003.
- [Pettersson *et al.*, 1999a] Karl Magnus Pettersson, Thomas E. Nichols, Jean-Baptiste Poline, and Andrew P. Holmes. Statistical limitations in functional neuroimaging. I. non-inferential methods and statistical models. *Philosophical Transactions of the Royal Society, London: Biological Sciences*, 354:1239–1260, 1999.
- [Pettersson *et al.*, 1999b] Karl Magnus Pettersson, Thomas E. Nichols, Jean-Baptiste Poline, and Andrew P. Holmes. Statistical limitations in functional neuroimaging. II. signal detection and statistical inference. *Philosophical Transactions of the Royal Society, London: Biological Sciences*, 354:1261–1281, 1999.
- [Piper and Granum, 1987] Jim Piper and Erik Granum. Computing distances transformations in convex and non-convex domains. *Pattern Recognition*, 20(6):599–615, 1987.
- [Poline and Mazoyer, 1993] Jean-Baptiste Poline and Bernard Mazoyer. Analysis of individual PET activation maps by detection of high signal-to-noise-ratio pixel clusters. *Journal of Cerebral Blood Flow and Metabolism*, 13:425–437, 1993.
- [Poline and Mazoyer, 1994a] Jean-Baptiste Poline and Bernard Mazoyer. Analysis of individual brain activation maps using hierarchical description and multiscale detection. *IEEE Transactions on Medical Imaging*, 13:702–710, 1994.
- [Poline and Mazoyer, 1994b] Jean-Baptiste Poline and Bernard Mazoyer. Enhanced detection in brain activation maps using a multifiltering approach. *Journal of Cerebral Blood Flow and Metabolism*, 14:639–642, 1994.
- [Prima *et al.*, 2001] Sylvain Prima, Nicholas Ayache, Tom Barrick, and Neil Roberts. Maximum Likelihood Estimation of the Bias Field in MR Brain Images: Investigating Different Modelings of the Imaging Process. In W.J. Niessen and M.A. Viergever, editors, *4th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'01)*, volume 2208 of *LNCS*, pages 811–819, Utrecht, The Netherlands, October 2001.
- [Pruessmann *et al.*, 1999] Klaas P. Pruessmann, Markus Weiger, Markus B. Scheidegger, and Peter Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42:952–962, 1999.
- [Raichle, 1994] Marcus E. Raichle. Visualising the mind. *Scientific American*, 270-41:36–42, 1994.
- [Ridolfi and Idier, 2000] Andrea Ridolfi and Jérôme Idier. Penalized maximum likelihood estimation for univariate normal mixture distributions. In *Bayesian Inference and Maximum Entropy Methods. MaxEnt Workshops*, 2000.

-
- [Rissanen, 1978] Jorma Rissanen. Modelling by shortest data description. *Automatica*, 14:465–471, 1978.
- [Rivière *et al.*, 2002] Denis Rivière, Jean-François Mangin, Dimitri Papadopoulos-Orfanos, Jean-Marc Martinez, Vincent Frouin, and Jean Régis. Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. *Medical Image Analysis*, 6(2):77–92, 2002.
- [Rivière *et al.*, 2003] Denis Rivière, Jean Régis, Yann Cointepas, Dimitri Papadopoulos-Orfanos, Arnaud Cachia, and Jean-François Mangin. A freely available Anatomist/BrainVISA package for structural morphometry of the cortical sulci. In *Proc. 9th HBM*, Neuroimage 19(2), page 934, 2003.
- [Rivière, 2000] Denis Rivière. *Apprentissage de la variabilité inter-individuelle de l'anatomie corticale cérébrale pour la reconnaissance automatique des sillons*. PhD thesis, Université d'Evry Val d'Essonne, 2000.
- [Roche *et al.*, 1998] Alexis Roche, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In W.M. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'98)*, volume 1496, pages 1115–1124, Berlin, Germany, October 1998. Lecture Notes in Computer Science, Springer.
- [Roche *et al.*, 2003] Alexis Roche, Ferath Kherif, Guillaume Flandin, and Jean-Baptiste Poline. Shoud fMRI data be analyzed using a single BOLD response model across regions? In Tomas Paus, Ed Bullmore, and Jonathan D. Cohen, editors, *NeuroImage (HBM'03)*, New York, USA, 2003. Academic Press.
- [Roche, 2001] Alexis Roche. *Recalage d'images médicales par inférence statistique*. PhD thesis, Université de Nice-Sophia Antipolis, Projet Epidaure, INRIA, 2001.
- [Rohde *et al.*, 2003] Gustavo K. Rohde, Akram Aldroubi, and Benoit M. Dawant. The adaptive bases algorithm for intensity-based nonrigid image registration. *IEEE Transactions on Medical Imaging*, 22(11):1470–1479, 2003.
- [Roland and Zilles, 1994] Per E. Roland and Karl Zilles. Brain atlases – a new research tool. *Trends in Neuroscience*, 17:458–467, 1994.
- [Roland and Zilles, 1998] Per E. Roland and Karl Zilles. Structural divisions and functional fields in the human cerebral cortex. *Brain Research Reviews*, 26:87–105, 1998.
- [Roland *et al.*, 1997] Per E. Roland, S. Geyer, Katrin Amunts, Thorsten Schormann, Axel Schleicher, Aleksandar Malikovic, and Karl Zilles. Cytoarchitectural maps of the human brain in standard anatomical space. *Human Brain Mapping*, 5:222–227, 1997.
- [Roy and Sherrington, 1890] Claude S. Roy and Charles S. Sherrington. On the regulation of the blood supply of the brain. *Journal of Physiology (London)*, 11:85–108, 1890.
- [Rueckert *et al.*, 1999] Daniel Rueckert, Luke I. Sonoda, Carmel Hayes, Derek L.G. Hill, Martin O. Leach, and David J. Hawkes. Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, pages 712–721, 1999.
- [Russell, 1998] Stuart Russell. The EM algorithm, 1998. Available at <http://citeseer.nj.nec.com/russell98em.html>.

- [Régis *et al.*, 1995] Jean Régis, Jean-François Mangin, Vincent Frouin, Francesca Sastre, Jean-Claude Peragut, and Yves Samson. Generic model for the localization of the cerebral cortex and preoperative multimodal integration in epilepsy surgery. *Stereotactic and Functional Neurosurgery*, 65:72–80, 1995.
- [Salli *et al.*, 2001] Eero Salli, Hannu J. Aronen, Sauli Savolainen, Antti Korvenoja, and Ari Visa. Contextual clustering for analysis of functional MRI data. *IEEE Transactions on Medical Imaging*, 20(5):403–414, 2001.
- [Salli, 2002] Eero Salli. *Contextual detection of fMRI activations and multimodal aspects of brain imaging*. PhD thesis, Helsinki University of Technology, Laboratory of Biomedical Engineering, Department of Engineering Physics and Mathematics, 2002.
- [Schwarz, 1978] G. Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6:461–464, 1978.
- [Sijbers *et al.*, 1998] Jan Sijbers, Arnold J. den Dekker, Paul Schneunders, and Dirk Van Dyck. Maximum-likelihood estimation of rician distribution parameters. *IEEE Transactions on Medical Imaging*, 17(3), June 1998.
- [Simon *et al.*, 2002] Olivier Simon, Jean-François Mangin, Laurent Cohen, Denis Le Bihan, and Stanislas Dehaene. Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron*, 31(33(3)):475–87, Jan 2002.
- [Sodickson and Manning, 1997] Daniel K. Sodickson and Warren J. Manning. Simultaneous acquisition of spatial harmonics (SMASH): ultra-fast imaging with radiofrequency coil arrays. *Magnetic Resonance in Medicine*, 38:591–603, 1997.
- [Stark and Bradley, 1992] David D. Stark and William G. Bradley. *Magnetic Resonance Imaging*. Mosby-Year Book, 1992.
- [Studholme *et al.*, 1996] Colin Studholme, Derek L.G. Hill, and David J. Hawkes. Automated 3D registration of MR and CT images of the head. *Medical Image Analysis*, 1:163–175, 1996.
- [Studholme *et al.*, 2000] Colin Studholme, R. Todd Constable, and James S. Duncan. Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model. *IEEE Transactions on Medical Imaging*, 19(11):1115–1127, 2000.
- [Surazhsky *et al.*, 2003] Vitaly Surazhsky, Pierre Alliez, and Craig Gotsman. Isotropic remeshing of surfaces: a local parameterization approach. In *12th International Meshing Roundtable*, September 2003.
- [Swallow *et al.*, 2003] Khena M. Swallow, Todd S. Braver, Abraham Z. Snyder, Nicole K. Speer, and Jeffrey M. Zacks. Reliability of functional localization using fMRI. *NeuroImage*, 20:1561–1577, 2003.
- [Talairach and Tournoux, 1988] Jean Talairach and Pierre Tournoux. *Co-planar Stereotaxic Atlas of the Human Brain*. Georg Thieme Verlag, Stuttgart, 1988.
- [Talairach and Tournoux, 1993] Jean Talairach and Pierre Tournoux. *Referentially Oriented Cerebral MRI Anatomy*. Georg Thieme Verlag, Stuttgart, New York, 1993.
- [Talairach *et al.*, 1967] Jean Talairach, G. Szikla, Pierre Tournoux, A. Prosalenti, M. Bornas-Ferrier, L. Covello, M. Jacob, and A. Mempel. *Atlas d’Anatomie Stéréotaxique du Télencéphale*. Masson, Paris, 1967.

-
- [Thiel, 1994] Edouard Thiel. *Les Distances de Chanfrein en Analyse d'Images: Fondements et Applications*. PhD thesis, Université Joseph Fourier, Grenoble, France, 1994.
- [Thirion and Faugeras, 2003a] Bertrand Thirion and Olivier Faugeras. Dynamical components analysis of fMRI data through kernel PCA. *NeuroImage*, 20(1):34–49, 2003.
- [Thirion and Faugeras, 2003b] Bertrand Thirion and Olivier Faugeras. Feature detection in fMRI data: The information bottleneck approach. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'03)*, volume 2879 of *LNCS*, pages 83–91, Montreal, Canada, November 2003.
- [Thirion, 1998] Jean-Philippe Thirion. Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.
- [Thirion, 2003] Bertrand Thirion. *Analyse de données d'IRM fonctionnelle: statistiques, information et dynamique*. PhD thesis, ENST, 2003.
- [Thompson and Toga, 1996] Paul M. Thompson and Arthur W. Toga. A surface-based technique for warping 3-dimensional images of the brain. *IEEE Transactions on Medical Imaging*, 15:1–16, 1996.
- [Thompson *et al.*, 1996] Paul M. Thompson, Craig Schwartz, and Arthur W. Toga. High-resolution random mesh algorithms for creating a probabilistic 3D surface atlas of the human brain. *NeuroImage*, 3(1):19–34, 1996.
- [Thompson *et al.*, 1997] Paul M. Thompson, David MacDonald, Michael S. Mega, Colin J. Holmes, Alan C. Evans, and Arthur W. Toga. Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces. *Journal of Computer Assisted Tomography*, 21(4):567–581, 1997.
- [Toro and Burnod, 2003] Roberto Toro and Yves Burnod. Geometric atlas: modeling the cortex as an organized surface. *NeuroImage*, 20:1468–1484, 2003.
- [Turner, 2002] Robert Turner. How much cortex can a vein drain? downstream dilution of activation-related cerebral blood oxygenation changes. *NeuroImage*, 16:1062–1067, 2002.
- [Tzourio *et al.*, 1998] Nathalie Tzourio, Bob Nkanga-Ngila, and Bernard Mazoyer. Left planum temporale surface correlates with functional dominance during story listening. *Neuroreport*, 9:829–833, 1998.
- [Tzourio-Mazoyer *et al.*, 2002] Nathalie Tzourio-Mazoyer, B. Landeau, Dimitri Papathanassiou, Fabrice Crivello, Olivier Etard, N. Delcroix, Bernard Mazoyer, and Marc Joliot. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15:273–289, 2002.
- [Van de Moortele, 1999] Pierre-François Van de Moortele. *IRM Fonctionnelle à 3 Tesla: développements méthodologiques*. PhD thesis, Université Paris XI, 1999.
- [Van Essen and Drury, 1997] David C. Van Essen and Heather A. Drury. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *Journal of Neuroscience*, 17(18):7079–7102, 1997.
- [Van Essen *et al.*, 1998] David C. Van Essen, Heather A. Drury, Sarang Joshi, and Michael I. Miller. Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proc. Natl. Acad. Sci. USA*, 95:788–795, 1998.

- [Verwer *et al.*, 1989] Ben J.H. Verwer, Piet W. Verbeek, and Simon T. Dekker. An efficient uniform cost algorithm applied to distance transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(4):425–429, April 1989.
- [Viola and Wells III, 1997] Paul Viola and Williams M. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997.
- [Vogt, 1991] Oskar Vogt. Die myeloarchitektonik des isocortex parietalis. *Journal of Psychology and Neurology*, 18:107–118, 1991.
- [Welker, 1990] Wally Welker. Why does the cerebral cortex fissure and fold? a review of determinants of gyri and sulci. *Cerebral Cortex*, 8B:3–136, 1990.
- [Wernicke, 1874] Carl Wernicke. *Der Aphasische Symptomenkomplex*. Coh and Weigart, 1874.
- [Wiegell *et al.*, 2003] Mette R. Wiegell, David S. Tuch, Henrik W.B. Larson, and Van J. Wedeen. Automatic segmentation of thalamic nuclei from diffusion tensor magnetic resonance imaging. *NeuroImage*, 19:391–402, 2003.
- [Woods *et al.*, 1993] Roger P. Woods, John C. Mazziotta, and Simon R. Cherry. MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography*, 17(4):536–546, 1993.
- [Woods *et al.*, 1998] Roger P. Woods, S.T. Grafton, J.D. Watson, N.L. Sicotte, and John C. Mazziotta. Automated image registration. II. intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography*, pages 153–165, 1998.
- [Woolrich *et al.*, 2001] Mark W. Woolrich, Brian D. Ripley, Michael Brady, and Stephen Smith. Temporal autocorrelation in univariate linear modelling of fMRI data. *NeuroImage*, 14(6):1370–1386, 2001.
- [Worsley and Friston, 1995] Keith J. Worsley and Karl J. Friston. Analysis of fMRI time-series revisited—again. *NeuroImage*, 2:173–181, 1995.
- [Worsley *et al.*, 1992] Keith J. Worsley, Alan C. Evans, Sean Marrett, and Peter Neelin. A three-dimensional statistic analysis of CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, 12:900–918, 1992.
- [Worsley *et al.*, 1998] Keith J. Worsley, Jean-Baptiste Poline, Karl J. Friston, and Alan C. Evans. Characterizing the response of PET and fMRI data using multivariate linear models. *NeuroImage*, 6:305–319, 1998.
- [Worsley *et al.*, 2002] Keith J. Worsley, Chuanhong Liao, John Aston, Valentina Petre, Gary H. Duncan, Frank Morales, and Alan C. Evans. A general statistical analysis for fMRI data. *NeuroImage*, 15:1–15, 2002.
- [Zarahn *et al.*, 1997] Eric Zarahn, Geoffrey Aguirre, and Mark D’Esposito. Empirical analyses of BOLD fMRI statistics. I. spatially unsmoothed data collected under null-hypothesis conditions. *NeuroImage*, 5:179–197, 1997.
- [Zhang *et al.*, 2000] Yongyue Zhang, Stephen Smith, and Michael Brady. Hidden Markov random field model and segmentation of brain MR images. Technical report, FMRIB, 2000.

[Zhang *et al.*, 2001] Yongyue Zhang, Michael Brady, and Stephen Smith. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57, January 2001.

Résumé

L'imagerie par résonance magnétique fonctionnelle (IRMf) est une modalité récente permettant de mesurer *in vivo* l'activité neuronale chez le sujet sain ou le patient et d'étudier le lien entre la structure et la fonction cérébrale.

Nous nous sommes intéressés à l'utilisation de l'information de l'anatomie cérébrale pour l'analyse des données fonctionnelles. Ainsi, nous avons reconsidéré l'analyse classique typiquement réalisée voxel par voxel avec lissage spatial pour proposer un modèle de représentation des données reposant sur une parcellisation anatomo-fonctionnelle du cortex. Cette représentation permet de réduire la dimension des données en un nombre restreint d'éléments pertinents du point de vue des neurosciences.

Nous présentons des exemples d'utilisation de cette approche de parcellisation fondée sur l'anatomie seulement dans un premier temps. Une étude de détection d'activations par modèle linéaire met en évidence une sensibilité accrue comparée à une approche voxel par voxel. Nous présentons également deux autres applications utilisant des parcellisations, concernant la sélection de modèle régional et les études de connectivité fonctionnelle.

Cette description permet en outre de proposer une solution au problème de l'analyse d'un groupe de sujets qui peuvent présenter une forte variabilité anatomique et fonctionnelle. Pour s'affranchir du délicat problème de mise en correspondance des différents sujets, nous avons présenté une parcellisation regroupant entre les sujets les régions homogènes du point de vue à la fois anatomique et fonctionnel. L'application de cette méthode sur un protocole fonctionnel montre sa robustesse pour les analyses multi-sujets.

Mots-clés: IRM fonctionnelle, neuro-imagerie, imagerie médicale, parcellisation, apprentissage non supervisé, classification.

Abstract

Functional magnetic resonance imaging (fMRI) is a recent modality allowing to measure *in vivo* the neuronal activity of healthy subjects or patients and thus to investigate the link between cerebral structure and function.

We are interested in introducing cerebral anatomy to analyse functional data. We thus reconsidered the classical analysis usually performed voxel-by-voxel following a spatial smoothing to propose a representation of data relying on an anatomo-functional parcellation of the cortex. This representation allows to reduce the dimensionality of the data into a small number of elements, more relevant from a neuroscience point of view.

We present several examples of application of this parcellation approach, at first based on anatomy only. An activity detection study based on a linear model highlights an increased sensitivity compared to the voxel-by-voxel approach. We also present two other applications using parcellations, dealing with the selection of regional models and functional connectivity studies.

This description also allows to propose a solution to the problem of group analyses, where subjects may exhibit an important anatomo-functional variability. To bypass the difficult problem of registration between different subjects, we propose a parcellation approach grouping the homogeneous areas from an anatomical and functional point of view, between subjects. The application of this method on a functional protocol reveals that it actually can take into account the anatomo-functional variability and thus offers some robustness to multi-subject analyses.

Keywords: Functional MRI, neuro-imagery, medical imaging, parcellation, unsupervised learning, clustering.

