

Université de Montréal

École Polytechnique

Classification d'images SPECT 3D du cerveau

par

Pablo Chamas

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.) en Génie Biomédical

21 novembre 2002

©Pablo Chamas, 2002

Université de Montréal
École Polytechnique
Faculté des études supérieures
Ce mémoire intitulé

Classification d'images SPECT 3D du cerveau

présenté par :

Pablo Chamas

a été évalué par un jury composé des personnes suivantes:

Pierre Savard

(président-rapporteur)

Jean Meunier

(directeur de recherche)

Max Mignotte

(codirecteur de recherche)

Farida Cheriet

(membre du jury)

Mémoire accepté le : 5 novembre 2002

Abstract

The SPECT imagery is one of the most used techniques in nuclear medicine because it is relatively costless and the collected data contains useful information for the detection of diffuse diseases. But for the brain, the reconstructed SPECT images have poor resolution. These diseases are not visible (or only lightly visible) to the naked eye at an early stage and physicians have difficulties to detect precociously the disease.

In order to improve the early diagnostic, two methods have been developed and tested.

The first one was based on a comparison of the left and right hemisphere by some techniques which can be as simple as the difference of the grey level to more complex attributes like the analysis of cooccurrence matrix. This approach does not produce a good classification, all the attributes have a detection rate inferior to 65%.

The second approach was based on the segmentation of all the brains according to macroscopic structures such as the cerebro-spinal fluid and the white and grey matter. This allowed the extraction of more specific attributes and better classification between healthy and diseased brains. The second method allows a good discrimination between the two classes. The rate of good classification is near 85% and this, with only one attribute.

Key Words : Image processing, Attribute, Learning, Classification, SPECT images

Sommaire

L'imagerie SPECT est l'une des techniques les plus utilisées en médecine nucléaire dans la mesure où elle est relativement peu coûteuse et les données collectées contiennent beaucoup d'informations pouvant être utiles à la détection de maladies diffuses ou focales. Mais pour le cerveau, les images SPECT reconstruites sont de faible résolution. Ces maladies ne sont pas ou très peu visibles à l'oeil nu à un stade précoce et les médecins ont ainsi des difficultés à diagnostiquer précocement la maladie.

Afin de les aider dans leur diagnostic, deux méthodes ont été développées et testées pour permettre une détection précoce de ces maladies.

La première était basée sur une comparaison des hémisphères droit et gauche du cerveau, à partir d'une série de mesures allant d'une simple différence des niveaux de gris à des attributs plus complexes telle que l'analyse des matrices de cooccurrence. Mais la classification obtenue à partir des attributs extraits de chacun des cerveaux présentait un taux d'erreur très important de 35 %.

Ainsi une autre méthode semblant plus appropriée à cette recherche a été développée dans un second temps. Elle s'appuyait sur une segmentation de chaque cerveau selon les structures macroscopiques du cerveau, liquide cephallorachidien, matière blanche et grise, permettant une extraction d'attributs plus spécifiques et l'obtention au final de meilleurs résultats pour la classification des cerveaux.

Les résultats obtenus avec la seconde méthode permettent une bonne différenciation entre les cerveaux atteints d'une maladie diffuse et les cerveaux sains avec un taux d'erreur de classification inférieur à 15% et ce avec un seul attribut.

Mots clés : Traitement d'images, Attribut, Apprentissage, Classification, Image SPECT

Table des matières

Sommaire	i
Abstract	ii
Table des matières	iii
Liste des figures	vii
Liste des tableaux	ix
Abréviation	x
Liste des abréviations	xi
Remerciements	xii
Introduction	1
1 Principe de l'imagerie par émission de photon unique contrôlée par ordinateur (SPECT)	3
1.1 Historique de l'imagerie SPECT	3
1.2 Principe de l'imagerie SPECT	5
1.2.1 Introduction	5

1.2.2	Particules et rayonnements	6
1.2.3	Les diverses formes d'interactions des photons avec la matière	7
1.2.3.1	Effet Photoélectrique	10
1.2.3.2	Diffusion (Effet) Compton	11
1.2.4	Caractéristiques des radio-isotopes utilisés en imagerie SPECT	12
1.2.5	Photons gamma γ	13
1.2.6	Les radio-pharmaceutiques	13
1.3	détection des rayonnements	16
1.3.1	Principe général	16
1.3.2	Les collimateurs	17
1.3.3	Les cristaux	18
1.3.4	Les tubes photomultiplicateurs	18
2	Aquisition et traitement des données	20
2.1	L'électronique de détection	20
2.2	Aquisition des données	22
2.3	Traitement des données	23
2.3.1	La méthode de rétro-projection	23
2.3.2	La méthode de rétro-projection filtrée	25
3	Méthodes basées sur la comparaison des activités des hémisphères gauche et droit du cerveau	29
3.1	Introduction et préparation des données	29
3.2	Somme des différences pondérée D	31
3.3	Le coefficient de corrélation ρ	33
3.4	Matrice de cooccurrence	34

3.5	Analyse des résultats	37
3.5.1	Tests statistiques	37
3.5.1.1	test F de Fisher, $H_0 : \sigma_X^2 = \sigma_Y^2$	38
3.5.1.2	test t de Student, $H_0 : \mu_X = \mu_Y$	38
3.5.2	Algorithme de classification	39
3.5.2.1	Algorithme des K-moyennes	39
3.5.2.2	Algorithme EM	41
3.5.3	Résultats	41
3.5.4	Diffus-Normaux	44
3.5.5	Autres résultats	45
4	Segmentation markovienne d'images SPECT du cerveau	49
4.1	Introduction	49
4.2	Estimation des paramètres	50
4.2.1	Définition théorique du problème	51
4.2.2	Algorithme ECI	52
4.3	Segmentation	54
4.3.1	L'estimateur du MAP	55
4.3.2	Segmentation à l'aide de l'algorithme ICM	57
4.4	Applications	58
4.5	Résultats	58
4.5.1	Attributs testés	60
4.5.2	Diffus-Normaux	61
4.5.3	Autres Résultats	61
4.5.3.1	Diffus-Négatifs	61

<i>TABLE DES MATIÈRES</i>	vi
4.5.3.2 Focaux-Négatifs et Focaux-Normaux	63
4.5.3.3 Focaux-Diffus	63
Conclusion	66
Bibliographie	68

Liste des figures

1.1	(A)Scanner de Cassen, (B)Caméra d'Anger[17]	4
1.2	Classification des rayonnements	6
1.3	Les différents événements d'interaction des photons avec la matière	8
1.4	Effet photoélectrique	10
1.5	Effet Compton	11
1.6	Importance relative des différents effets des rayonnements sur la matière	14
1.7	Principe de détection	16
1.8	Les différents types de collimateur	17
1.9	Influence du diamètre de collimateur	18
1.10	Principe de fonctionnement d'un tube photo-multiplicateur	19
2.1	Détermination de la position de la scintillation	21
2.2	Détermination de la position et de l'intensité de la scintillation	21
2.3	Les différents mode d'acquisition des données	22
2.4	Principe de la rétro-projection	24
2.5	Démonstration de l'effet d'étoile [22]	25
2.6	Principe de la rétro-projection filtrée [22]	26
2.7	Principe de la rétro-projection filtrée [22]	28
3.1	différentes étapes du traitement préliminaire	30

3.2	Différentes étapes pour la réalisation de la Somme des différences pondérée D	32
3.3	Exemples de matrice de cooccurrence	35
3.4	matrice de cooccurrence obtenue avec l'image de cerveau sain (à gauche) et à l'aide d'un cerveau atteint de tumeur focale (à droite)	37
4.1	Histogramme des niveaux de gris d'une image SPECT	50
4.2	Résultat de la segmentation, en haut, l'image originale, au milieu l'image segmentée et en bas, l'histogramme des niveaux de gris avec les deux lois gaussiennes et la loi exponentielle	59

Liste des tableaux

2.1	Résumé de la rétro-projection filtrée	27
3.1	Résumé du test F de Fisher	39
3.2	Résumé du test t de Student	40
3.3	Algorithme des K-moyennes.	42
3.4	Algorithme EM pour l'estimation des paramètres d'un mélange de deux lois Gaussiennes.	43
3.5	Algorithme EM. Formule de ré-estimation dans le cas d'un mélange de deux lois Gaussiennes.	44
3.6	Résultats de classification entre la classe Normaux et la classe Diffus, avec 43 degrés de liberté ($m + n - 2 = 43$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.48, 2.03]$ pour être significatif et un résultat au test t de Student > 2.0167 ou < -2.0167 pour être révélateur . . .	45
3.7	Résultats de classification entre la classe Négatif et la classe Diffus, avec 34 degrés de liberté ($m + n - 2 = 34$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.383, 2.21]$ pour être significatif et un résultat au test t de Student > 2.0323 ou < -2.0323 pour être révélateur . . .	46
3.8	Résultats de classification entre la classe Normaux et la classe Focaux, avec 29 degrés de liberté ($m + n - 2 = 29$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.290, 2.51]$ pour être significatif et un résultat au test t de Student > 2.0518 ou < -2.0518 pour être révélateur . . .	46

3.9	Résultats de classification entre la classe Négatif et la classe Focaux, avec 20 degrés de liberté ($m + n - 2 = 20$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.278, 3.01]$ pour être significatif et un résultat au test t de Student > 2.1009 ou < -2.1009 pour être révélateur . . .	47
3.10	Résultats de classification entre la classe Diffus et la classe Focaux, avec 32 degrés de liberté ($m + n - 2 = 32$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.293, 2.42]$ pour être significatif et un résultat au test t de Student > 2.0423 ou < -2.0423 pour être révélateur . . .	48
4.1	Résultats de classification entre la classe Normaux et la classe Diffus, avec 43 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.48, 2.03]$ pour être significatif et un résultat au test t de Student > 2.0167 ou < -2.0167 pour être révélateur	62
4.2	Résultats de classification entre la classe Négatif et la classe Diffus, avec 34 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.383, 2.21]$ pour être significatif et un résultat au test t de Student > 2.0323 ou < -2.0323 pour être révélateur	63
4.3	Résultats de classification entre la classe Normaux et la classe Focaux, avec 29 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.290, 2.51]$ pour être significatif et un résultat au test t de Student > 2.0518 ou < -2.0518 pour être révélateur	64
4.4	Résultats de classification entre la classe Négatifs et la classe Focaux, avec 20 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.278, 3.01]$ pour être significatif et un résultat au test t de Student > 2.1009 ou < -2.1009 pour être révélateur	64
4.5	Résultats de classification entre la classe Diffus et la classe Focaux, avec 32 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.293, 2.42]$ pour être significatif et un résultat au test t de Student > 2.0423 ou < -2.0423 pour être révélateur	65

Liste des abréviations

ECI	Estimation Conditionnelle Itérative
EM	Expectation Maximization
SEM	Stochastique Expectation Maximization
MAP	Maximum A Posteriori
MPM	Mode des Marginales à Posteriori
MV	Maximum de Vraisemblance
ICM	Iterative Conditionnal Mode
RS	Recuit Simulé
SPECT	Single Photon Emission Computerized Tomography
PET	Positon Emission Tomography
CSF	Cerecro-Spinal Fluid

Remerciements

Je tiens à remercier quelques personnes qui m'ont apporté leur soutien, leur aide et leur amitié et m'ont permis d'aller au terme de cette maîtrise.

En premier lieu, je tiens tout d'abord à remercier mon Directeur Monsieur Jean Meunier, pour m'avoir tout d'abord accepté lors d'un stage de six mois pour un projet de fin d'étude dans le cadre de ma maîtrise française puis dans le cadre de cette maîtrise et pour toute l'aide qu'il m'a apporté, sa disponibilité dont il a fait preuve ainsi que sa grande patience.

Un grand merci également à mon co-directeur de recherche, Monsieur Max Mignotte pour son aide en programmation, sa précieuse analyse en estimation de paramètres et en L^AT_EX.

De très sincères remerciements à mes amis, Messieurs Pierre-Marc Jodoin et Jean-François Laliberté, pour leur amitié durable et leurs conseils en programmation et système Linux.

Un grand merci également aux Messieurs Jean-Paul Soucy et Christian Janicki du département de médecine nucléaire de l'hôpital Notre-Dame pour le temps qu'ils m'ont consacré et les données qu'ils ont fournies.

Je suis également très reconnaissant à Monsieur Sébastien Roy et Monsieur Pierre Poulin pour leur entrain et leur soutien.

Un petit mot spécial à toute la “gang” de Matane, de Montréal et de l'Ontario : François-Paul, David, Mike, Daniel, Caro, Jojo, Juliette, Richard, Kyle, Chantal, Mat, Vincenzo, Julio, Gwyneth, Bertrand et Nancy

Un mot à la femme qui a su me soutenir, m'aider et faire battre mon cœur lors de cette dernière année, Aurelie.

Enfin , merci à ma famille, Marie, Toto et Clairette pour leurs visites nombreuses qui ont su me faire aller de l'avant.

Introduction

Depuis toujours, la médecine s'est employée à explorer l'anatomie humaine. Les technologies ont évolué au cours des siècles, et ainsi à la fin du XIX^{ème} siècle, une découverte a révolutionné le monde de la médecine : la découverte des rayons X. Les médecins vont enfin pouvoir voir des images anatomiques d'êtres vivants.

Depuis, de multiples techniques de visualisation ont été développées, telle que l'imagerie par émission de photon unique contrôlée par ordinateur (SPECT). L'imagerie SPECT est à l'heure actuelle la méthode la plus simple et la moins coûteuse (si on la compare à la tomographie d'émission de positon ou à la résonance magnétique nucléaire) pour obtenir des images fonctionnelles d'organes.

Le chapitre 1 de ce rapport sera consacré à l'imagerie SPECT, à ses principes physiques, ses points forts, mais également ses points faibles.

Les techniques d'acquisition et de reconstruction en imagerie SPECT seront décrites au chapitre 2, avec une description de l'appareillage permettant la détection des photons γ (photons utilisés en imagerie SPECT), et les méthodes de reconstruction les plus utilisées.

L'objectif de cette étude est d'utiliser cette technique pour permettre une détection précoce de maladies diffuses ou focales dans le cerveau à partir de l'analyse de la texture des images du cerveau. Dans un premier temps, au chapitre 3, l'analyse de ces images s'appuiera sur une comparaison de l'activité des deux hémisphères du cerveau. Puis au chapitre 4, une autre méthode d'analyse de ces images sera développée, celle-ci permettant d'extraire plus spécifiquement les données relatives à certaines parties du cerveau (matière blanche, matière grise).

La conclusion s'attardera sur les problèmes rencontrés et sur les différentes voies ouvertes par cette recherche et dont la poursuite pourra mener à une amélioration de la classification obtenue.

Chapitre 1

Principe de l'imagerie par émission de photon unique contrôlée par ordinateur (SPECT)

1.1 Historique de l'imagerie SPECT

La médecine nucléaire exploite les différences dans la distribution d'un traceur radioactif administré dans le corps du patient.

Le premier système d'imagerie en médecine nucléaire fut inventé par Cassen durant les années 40 (Fig 1.1.A). Le système en question était un scanner mécanique constitué d'un détecteur à scintillation capable de se déplacer en mouvements rectilignes au dessus du patient. Les intensités des rayons gamma détectés sont représentées comme une fonction de la position du scanner, afin de générer une image représentative de la distribution du traceur. L'invention de la caméra à scintillation en 1957 par Anger (Fig 1.1.B) allait changer le cours de la médecine nucléaire. À la différence du scanner de Cassen, la caméra d'Anger est constituée d'un détecteur à scintillations stationnaire, formée d'un cristal à large diamètre comprenant une matrice de tubes photo-multiplicateur, ce qui permet d'obtenir une vue globale de l'organe [17].

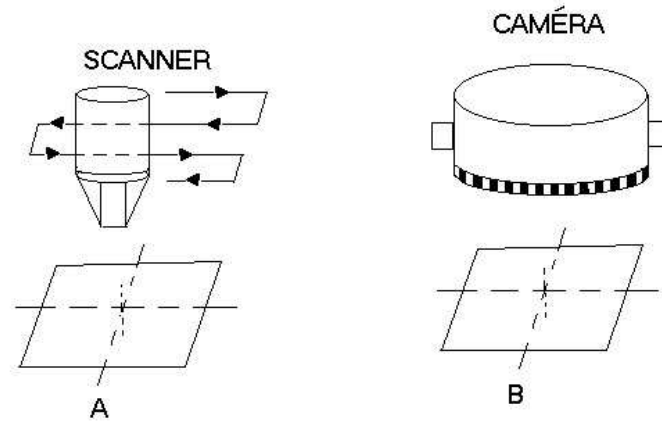


FIG. 1.1 – (A)Scanner de Cassen, (B)Caméra d'Anger[17]

Vers 1958, Khul remarqua qu'il pouvait tirer parti du scanner à collimateur focalisé s'il faisait "tourner" ce dernier autour de l'organe. Il pouvait ainsi apercevoir les gros organes (le foie, par exemple) sous tous les angles de vues différents. En 1963, Khul et Edwards proposèrent d'utiliser le scanner focalisé en tomographie. Ceci fut possible en plaçant l'instrument sur un support mécanique à déplacement complexe, permettant de faire des rotations et d'ajouter à cet appareil un oscilloscope. L'année suivante, ils purent ainsi élaborer le balayage stéréoscopique, longitudinal et transversal. Grâce à eux, à partir de ce moment l'imagerie par émission devint l'un des outils les plus importants en diagnostic médical.

En 1968, Khul et Edwards suggérèrent l'utilisation de l'ordinateur numérique à des fins d'acquisition et d'enregistrement des données obtenues par leurs instruments. L'ordinateur devient alors l'outil par excellence en tomographie par émission, mais seulement pour le stockage et le rehaussement du contraste. À la même époque, Cormack proposa l'utilisation des mathématiques afin de synthétiser une image transaxiale à partir d'images projetées.

En 1973, Hounsfield utilisa pour la première fois la tomographie par rayons X assistée par ordinateur pour prendre des images de la tête. La même année, Budin-

ger et Gullberg firent une excellente revue des différents développements théoriques touchant la reconstruction et discutèrent des applications possibles de chaque théorie à la médecine nucléaire. Tandis que la transmission tomographique par rayons X assistée par ordinateur devenait un succès mondial, le chemin s'ouvrait à de nouvelles inventions en médecine nucléaire. Ceci commença par l'installation d'une unité tomographique en 1971 à l'hôpital Alkinson Morley en Angleterre. Au cours des années suivantes, le développement de la caméra d'Anger rotative et des équipements informatiques a permis à la tomographie par émission de photon unique contrôlée par ordinateur (SPECT¹) de devenir une réalité.

Aujourd'hui l'utilisation de systèmes multicristaux à collimateur focalisé SPECT², le descendant de l'instrument original de Khul, de même que du plus récent *single photon ring tomograph* (SPRINT) sont choses courantes dans les hôpitaux. L'utilisation croissante du SPECT est due à sa capacité de localiser et d'étudier les structures d'organes en fonction de leurs propriétés physiques et biochimiques [26].

1.2 Principe de l'imagerie SPECT

1.2.1 Introduction

Une des techniques les plus répandues utilise un radio-isotope émetteur de photons gamma qui, après injection dans le corps du patient, est détecté par émission à l'aide d'une gamma caméra ou caméra à scintillation. Ce radio-isotope passe d'un état instable à un état stable en émettant des particules nucléaires (α et β) et des rayons γ . Nous allons détailler les différents classements des rayonnements puis nous développerons les différentes formes d'interactions des photons γ avec les électrons des différents atomes présents dans les tissus traversés. Ces photons γ seront ensuite détectés à l'aide d'une gamma caméra.

¹Single Photon Emission Computed tomography

²Multicrystal Focused Collimator SPECT system

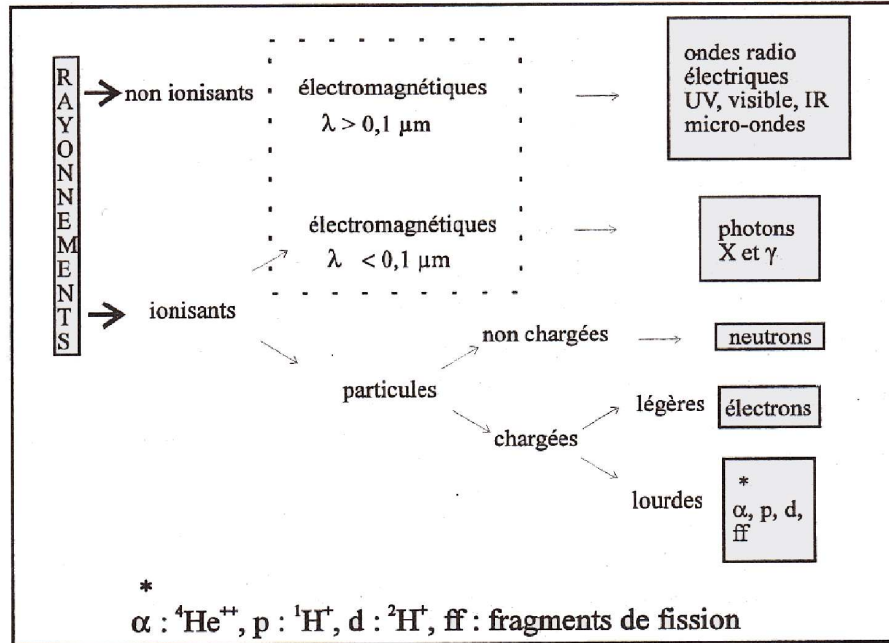


FIG. 1.2 – Classification des rayonnements

1.2.2 Particules et rayonnements

On peut définir un rayonnement [8] comme un mode de propagation de l'énergie dans l'espace, sous forme d'ondes électromagnétiques ou de particules. Les rayonnements ne peuvent être détectés et caractérisés que grâce à leurs interactions avec la matière dans laquelle ils se propagent. Ils cèdent au milieu qu'ils traversent, en totalité ou en partie, leur énergie au cours de ces interactions et la matière, en contrepartie, subit des modifications.

On peut classer, selon leur nature, les rayonnements en deux catégories :

- Les rayonnements particuliers formés d'un flux de particules matérielles douées de masse au repos.
- Les rayonnements électromagnétiques qui, constitués par un flux de photons, peuvent être considérés comme de l'énergie à l'état pur.

La matière rencontrée par les rayonnements est composée d'atomes, donc d'une part de noyaux, chargés d'électricité positive, lourds, d'autre part d'électrons, beaucoup plus nombreux, légers et chargés négativement.

Les rayonnements électromagnétiques produisent des effets différents sur la matière selon qu'ils interagissent avec les noyaux ou les électrons.

Une autre classification (Fig. 1.2), fondée sur les effets des interactions des rayonnements sur la matière conduit à distinguer :

- Les rayonnements ionisants dont l'énergie est suffisante pour arracher un électron aux atomes du milieu rencontré et les transformer ainsi en ions positifs.
- Les rayonnements non ionisants dont l'énergie est insuffisante pour déclencher cet arrachement.

On désigne par énergie d'ionisation, l'énergie minimale nécessaire pour détacher complètement un électron orbital des atomes ou des molécules du milieu irradié, ces derniers se trouvant primitivement à l'état fondamental. L'examen des valeurs des énergies d'ionisation des principaux atomes constituant la matière vivante (à savoir l'hydrogène, le carbone, l'azote et l'oxygène) permet de fixer le seuil énergétique qui fonde cette distinction : sont ionisantes les radiations dont l'énergie est supérieure à 12.4 eV. Tous les rayonnements particuliers sont donc ionisants. Cette limite oppose les ondes radio-électriques (radio, télévision), l'ultraviolet, la lumière visible, l'infrarouge et les micro-ondes, non ionisantes, aux rayons X et γ , rayonnements électromagnétiques photoniques qui, eux, entraînent l'ionisation des atomes des milieux biologiques soumis à leur action.

1.2.3 Les diverses formes d'interactions des photons avec la matière

Quatre destinées possibles (Fig. 1.3) s'offrent à un photon d'énergie $h\nu$ traversant une certaine épaisseur de matière [8] :

- 1- N'être affecté par aucune interaction : le photon est alors transmis.
- 2- Subir une déviation, sans communiquer d'énergie au milieu rencontré : le photon est alors diffusé sans transfert d'énergie.
- 3- Subir une déviation avec transfert partiel d'énergie à la substance traversée :

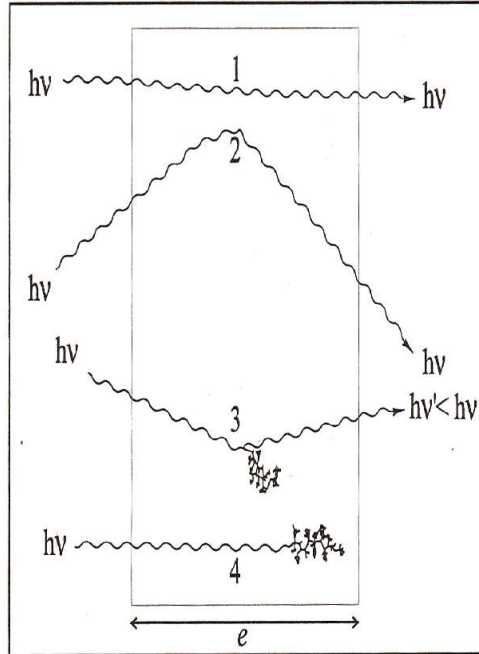


FIG. 1.3 – Les différents événements d'interaction des photons avec la matière

son énergie est alors $h\nu' < h\nu$ ($\lambda' > \lambda$).

- 4- Disparaître dans l'épaisseur de la cible en y cédant toute son énergie : le photon est alors absorbé.

L'énergie globale du faisceau incident se partage entre trois composantes, l'énergie transmise, l'énergie diffusée et l'énergie transférée.

Des quatre cas de figure précédents, seuls les deux derniers nous intéressent dans ce paragraphe car ils correspondent à une véritable interaction des rayonnements sur la matière. Dans le milieu traversé, les photons peuvent interagir avec les électrons ou avec les noyaux.

Les interactions avec les électrons sont de trois types :

- l'effet photoélectrique.
- la diffusion Compton.
- la diffusion Rayleigh (dont nous ne parlerons pas).³

Alors que les interactions avec les noyaux peuvent donner lieu :

³quasiment inexistante en imagerie SPECT.

- soit à la production de paires d'électrons (effet de matérialisation).
- soit à une réaction photo nucléaire (dont nous ne parlerons pas).⁴

On peut donc considérer, en pratique, que le photon interagit avec la matière traversée selon trois effets principaux (effet photoélectrique, diffusion Compton, effet de matérialisation). Au cours de chacun d'entre eux, il transfère, en totalité ou en partie, son énergie au milieu rencontré. En imagerie SPECT, l'effet de matérialisation est absent car les photons γ utilisés ont une énergie trop basse pour interagir avec les noyaux des atomes. Nous détaillerons simplement l'effet photoélectrique et la diffusion Compton, qui sont les deux facteurs importants à prendre en compte en imagerie SPECT.

⁴quasiment inexistante en imagerie SPECT.

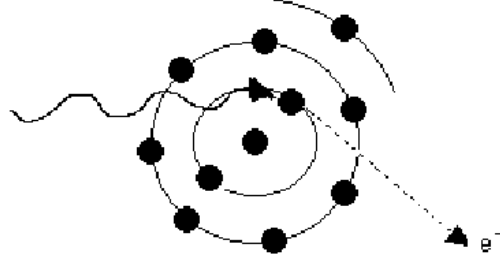


FIG. 1.4 – Effet photoélectrique

1.2.3.1 Effet Photoélectrique

L'effet photoélectrique (Fig. 1.4) est la collision d'un photon incident avec un électron du cortège électronique d'un atome[1]. Ce phénomène se traduit par l'absorption par l'atome de la totalité de l'énergie E du photon incident. Cette énergie excédentaire est transférée à un électron qui est éjecté de sa couche électronique, d'énergie de liaison W_i , grâce à une partie de l'énergie E et, lors de son éjection, il emporte le reste de cette énergie sous forme d'énergie cinétique :

$$E_a = E - W_i$$

Il manque donc pendant un bref instant un électron sur une couche électronique. Ce manque est comblé très rapidement par un électron qui vient d'une autre couche plus périphérique et très rarement d'un électron libre du même atome ou d'un atome différent. Cette ionisation sur la couche i à laquelle appartenait l'électron est suivie de l'émission d'un rayonnement (le retour à l'état fondamental de l'atome excité). L'énergie E_a du photoélectron est absorbée par le milieu.

L'effet photoélectrique ne peut avoir lieu avec un électron d'une couche i , que si E est supérieure ou égale à W_i . Il se produit en fait presque uniquement sur les électrons de la couche pour laquelle W_i est la plus proche de E par valeur inférieure ; plus l'atome est riche en électrons, plus la probabilité que l'effet photoélectrique se produise est grande.

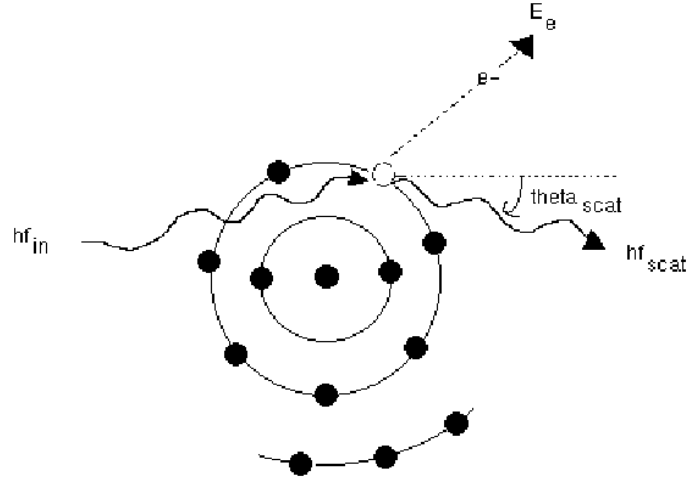


FIG. 1.5 – Effet Compton

1.2.3.2 Diffusion (Effet) Compton

Le photon incident (Fig. 1.5) d'énergie $h\nu$ entrant en collision avec un électron au repos lui transfère une partie de son énergie, E , et le reste de son énergie $h\nu'$ est emportée par le photon diffusé d'énergie $h\nu'$ avec [1] :

$$h\nu = h\nu' + E$$

Un photon incident ne peut pas transférer toute son énergie à une particule isolée, il existe toujours un photon diffusé d'énergie $h\nu'$, si infime soit cette énergie .

Pour un choc tangentiel nous avons :

$$E = E_{min} \quad \text{et} \quad h\nu' = E_{max}$$

et pour un choc frontal nous avons :

$$E = E_{max} \quad \text{et} \quad h\nu' = E_{min}$$

La diffusion Compton a pour effet de dévier les photons émis de la vraie source vers les tissus environnants. Ces derniers agissent comme une deuxième source de

rayonnement ce qui diminue le contraste de l'image obtenue. Il existe des techniques particulières (par exemple, les collimateurs) pour masquer ce rayonnement qui ne seront pas détaillées dans ce mémoire.

1.2.4 Caractéristiques des radio-isotopes utilisés en imagerie SPECT

En imagerie SPECT, les rayons utilisés sont les photons γ [2] car ce sont les seuls à :

- pouvoir être détectés de l'extérieur du patient (les rayons α et β , ne traversent respectivement que quelques micromètres et quelques millimètres).
- avoir une énergie comprise entre 70 et 500 keV, pour permettre une détection efficace.
- que l'émission soit monochromatique, ce qui permet de ne pas comptabiliser les photons ayant subits des interactions avec le milieu.
- Que la demi-vie⁵ de l'élément soit comparable au temps nécessaire pour la préparation de l'élément additionné au temps qu'il faut pour réaliser l'examen, ce qui correspond a une demie-vie d'environ 6 heures.

Les plus utilisés des radio-isotopes sont, un isotope métastable⁶ du technecium, le ^{99m}Tc et un isotope du thallium, le ^{201}Tl , qui produisent respectivement des photons de 140 keV et 70 keV.

⁵temps au bout duquel une grandeur (physique, biologique) atteint la moitié de sa valeur initial.(LAROUSSE 1992)

⁶Se dit d'un système qui n'est pas stable en théorie, mais qui paraît stable en raison d'une vitesse de transformation très faible.(LAROUSSE 1994)

1.2.5 Photons gamma γ

Qu'est ce qu'un photon gamma γ ?

Les rayons γ sont des quanta électromagnétiques ou photons [11]. Ils sont d'origine nucléaire, donc très énergétiques. Les photons γ ne produisent guère d'ionisation de façon directe, au contraire, ils perdent de l'énergie en faveur d'électrons qui provoquent à leur tour des ionisations. Les photons γ ont par conséquent de très longs parcours dans la matière. Un photon de 1 MeV, par exemple, a un parcours moyen de 10cm dans l'eau.

Les photons transfèrent de l'énergie aux électrons par trois processus (Fig. 1.6). Pour des énergies inférieures à 0.1 MeV, l'effet Photoélectrique est le plus important. Aux environs de 1 MeV, c'est la diffusion Compton qui prédomine. A des énergies plus élevées, il devient possible pour un photon de produire une paire électron-positron. Ceci peut se produire au voisinage d'un noyau lorsque l'énergie du photon est supérieure à l'énergie de masse totale des deux particules, c'est-à-dire à :

$$2mc^2 = 1.02MeV \quad (1.1)$$

La probabilité d'absorption du photon diminue généralement quand son énergie augmente. Par conséquent, quand l'énergie croît, l'absorption diminue graduellement et le rayonnement devient plus pénétrant ou plus dur.

1.2.6 Les radio-pharmaceutiques

Un radio-isotope est un élément qui va émettre le rayon γ mais il faut également un vecteur qui va permettre d'atteindre spécifiquement la zone à étudier (foie, poumon ou cerveau etc.) et qui va également dénoter une différence relative entre un tissu sain et un tissu pathologique. Il faut que ce composé soit non-toxique pour les patients mais également facilement liable chimiquement au radio-isotope sur le site de l'hôpital car sinon le temps de demi-vie requis serait trop important.

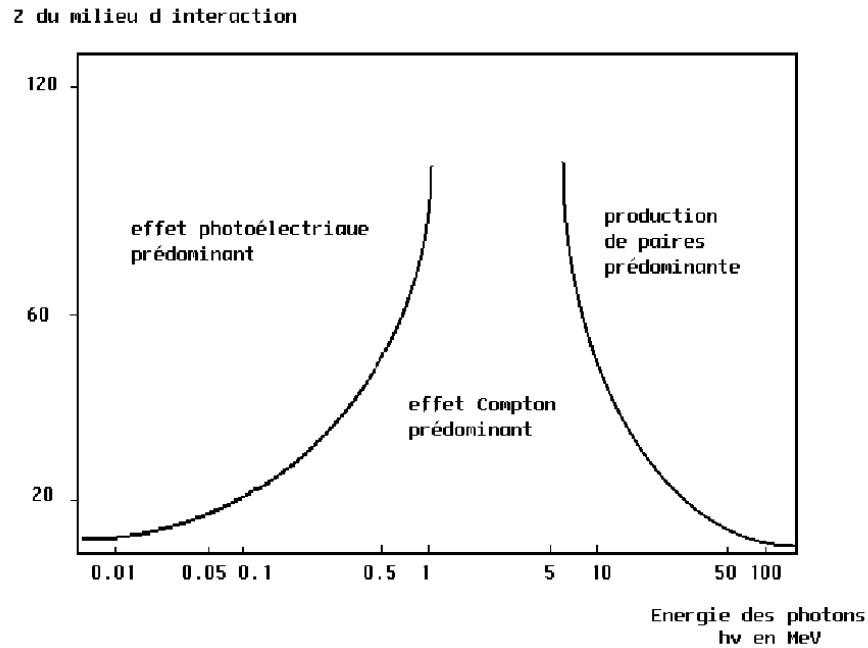
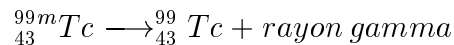


FIG. 1.6 – Importance relative des différents effets des rayonnements sur la matière

Avec toutes ces nouvelles contraintes, le radio-isotope le plus utilisé est le ^{99m}Tc , car il a une demi-vie de 6 heures, il se combine avec une grande variété de composés et il est facilement produit par un générateur qui est situé en général au service de médecine nucléaire de l'hôpital.

Le ^{99m}Tc est un isotope excité métastable du technecium ($Z=43$) qui contient 56 neutrons, il se désexcite en émettant de l'énergie sous la forme d'un rayon gamma d'une énergie de 140 keV.



Le $^{99}_{43}\text{Tc}$ résultant a le même nombre de protons et de neutrons, ce qui est appelé une transformation isométrique.

Cet isotope est couplé à une molécule qui va lui permettre de traverser la barrière entre le sang et le cerveau, car sinon il resterait dans le sang. Ces molécules permettent ce passage par, soit un transport actif, soit un passage favorisé par un transporteur où

soit par une diffusion passive.

Le ^{99m}Tc permet de mesurer le flot sanguin [5] au niveau du cerveau.

Ces radio-pharmaceutiques sont injectés dans le patient avant l'examen (le temps d'attente avant l'acquisition peut varier en fonction du temps de diffusion des composés pour atteindre l'organe cible), puis il y a une étape d'acquisition où l'on va détecter les rayons γ émis depuis l'organe cible. Le prochain chapitre décrit l'appareillage nécessaire à une telle acquisition ainsi que ses différents modes possibles.

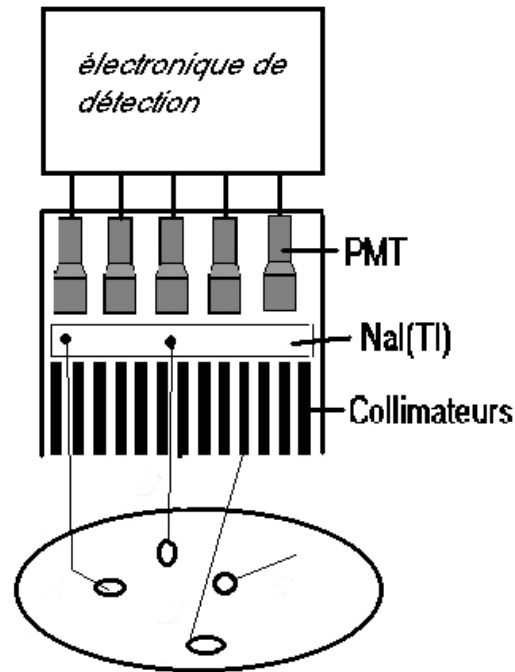


FIG. 1.7 – Principe de détection

1.3 détection des rayonnements

1.3.1 Principe général

La plupart des détecteurs utilisés en imagerie SPECT utilisent un ou plusieurs détecteurs à scintillation. Le développement le plus significatif en médecine nucléaire a été la caméra à scintillation⁷ qui consiste en un large cristal d'iodure de sodium dopé au thallium (NaI(Tl))(typiquement 40 centimètres) derrière lequel est placé une série de tubes photomultiplicateurs (Fig 1.7) [1].

Cette caméra (Fig. 1.7) possède :

- Un collimateur qui va permettre de trier en direction tous les photons venant du plan de coupe étudié.

⁷aussi appelé caméra d'Anger.

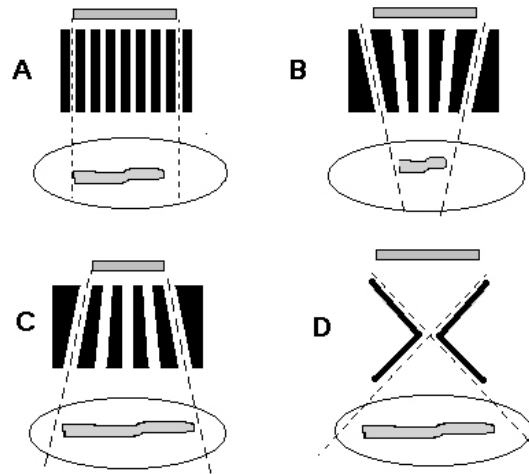


FIG. 1.8 – Les différents types de collimateur

- Un cristal scintillant effectue la conversion des photons γ détectés en photons lumineux.
- Des tubes photo-multiplicateurs dont le rôle est d'amplifier le signal des photons lumineux.
- Un système d'analyse électronique permet de caractériser spatialement l'origine de l'émission de chaque photon et de rejeter le signal émis par des photons avec des énergies erronées.

1.3.2 Les collimateurs

Etant donné que les rayons γ ne peuvent pas être focalisés, le rôle de la lentille est joué par des collimateurs ce qui permet de les trier en direction. Il en existe différents types [1] :

- Le collimateur à trous parallèles, qui est le plus utilisé dans la pratique courante (Fig. 1.8 A).
- Le collimateur à trous convergents qui produit des images agrandies (Fig. 1.8 B).

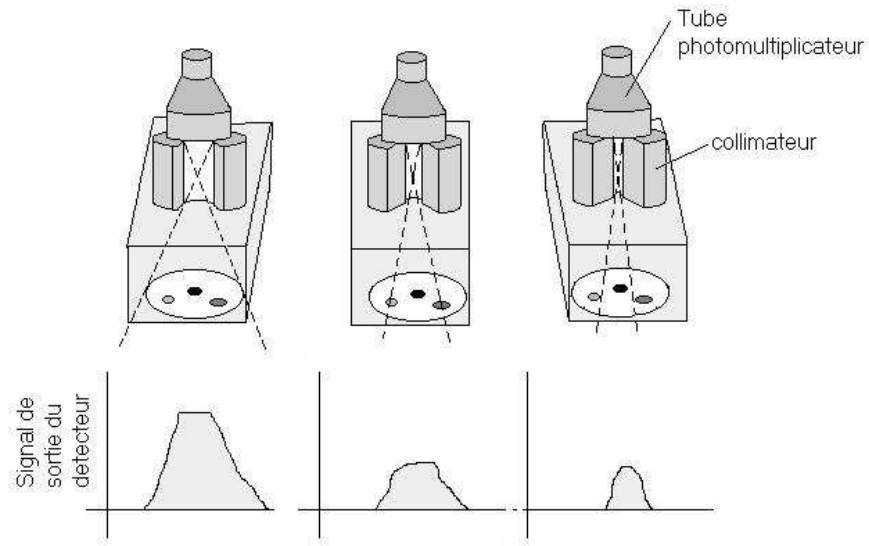


FIG. 1.9 – Influence du diamètre de collimateur

- Le collimateur à trous divergents qui réduit et permet l'imagerie de gros organes avec une petite caméra de détection comme par exemple une gamma caméra mobile (Fig. 1.8 C).
- Le collimateur à stenopé qui pourra agrandir ou rétrécir des images selon la distance entre la caméra et l'organe étudié (Fig. 1.8 D).

Il existe des collimateurs réversibles qui peuvent être utilisés en mode convergent ou divergent. Plus le diamètre du collimateur augmente, plus la sensibilité est grande mais moins la résolution est bonne (Fig. 1.9).

1.3.3 Les cristaux

Le NaI(Tl) est le cristal le plus utilisé pour l'imagerie SPECT car c'est lui qui scintille le mieux aux alentours de 140 keV qui est l'énergie du photon γ délivré par le ^{99m}Tc ⁸. Ce cristal transforme les photons γ en photons lumineux [1].

1.3.4 Les tubes photomultiplicateurs

⁸Rappelons que le ^{99m}Tc est le radio-isotope le plus utilisé en imagerie SPECT.

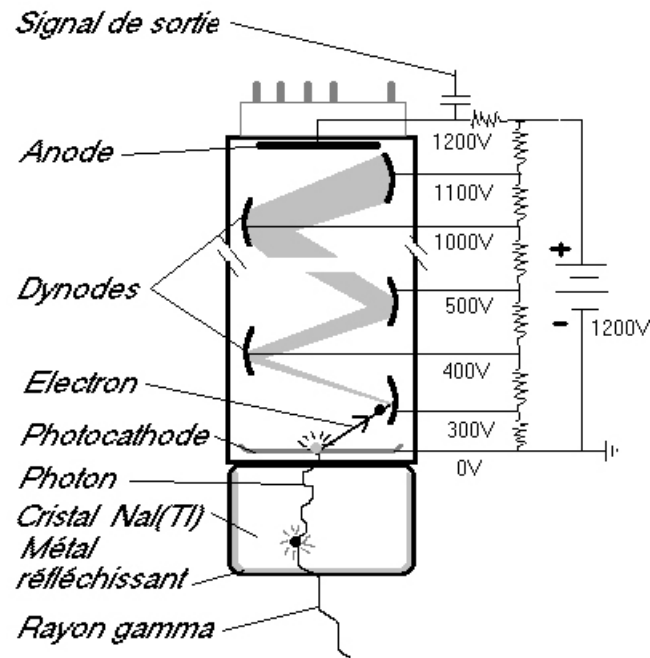


FIG. 1.10 – Principe de fonctionnement d'un tube photo-multiplicateur

Prenons le cas d'un seul photon lumineux pour cette explication (Fig. 1.10) [2].

La photo-cathode du tube photo-multiplicateur absorbe ce photon lumineux, elle va relâcher un photo-électron négatif par photons lumineux. Le photo-électron négatif va aller frapper la première dynode qui lors de la collision avec l'électron va en libérer plusieurs qui vont aller sur la deuxième dynode et ainsi de suite jusqu'à l'anode. Les dynodes sont à des différences de potentiel différentes. Cela permet d'amplifier le signal et d'avoir, à partir d'un photon γ , un signal mesurable.

Chapitre 2

Aquisition et traitement des données

2.1 L'électronique de détection

Cette électronique permet de déterminer à partir des signaux détectés par les PMT¹ l'origine spatiale de l'événement sur le cristal de scintillation (Fig. 2.1). La figure 2.1 ne considère qu'un seul axe mais il en va de même pour les deux axes X et Y [2].

Quand un rayon γ excite le cristal scintillant, il y a émission de lumière qui va être captée par les différents PMT, le voltage que chacun des PMT produira sera proportionnel à la distance qui le sépare de la source lumineuse. De là, le circuit de position produira deux potentiels V_x et V_y .

Ceci nous permet d'avoir les positions X et Y de cette scintillation. Il reste à déterminer l'intensité de cette scintillation (Fig. 2.2), pour cela on utilise un analyseur de pics qui va analyser l'intensité de chaque scintillation et la transformer en potentiel électrique V_z proportionnel à l'intensité de la scintillation.

¹Tubes photo-multiplicateurs

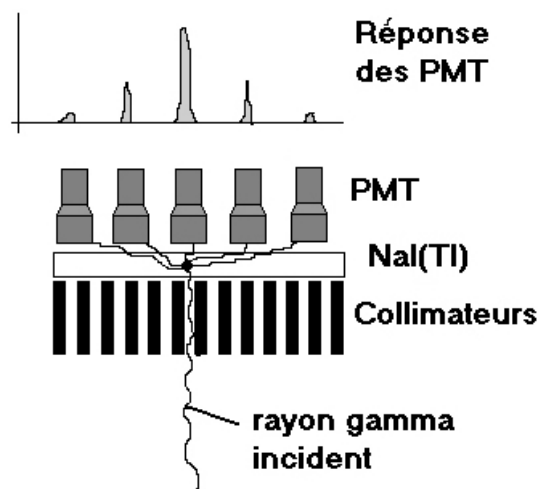


FIG. 2.1 – Détermination de la position de la scintillation

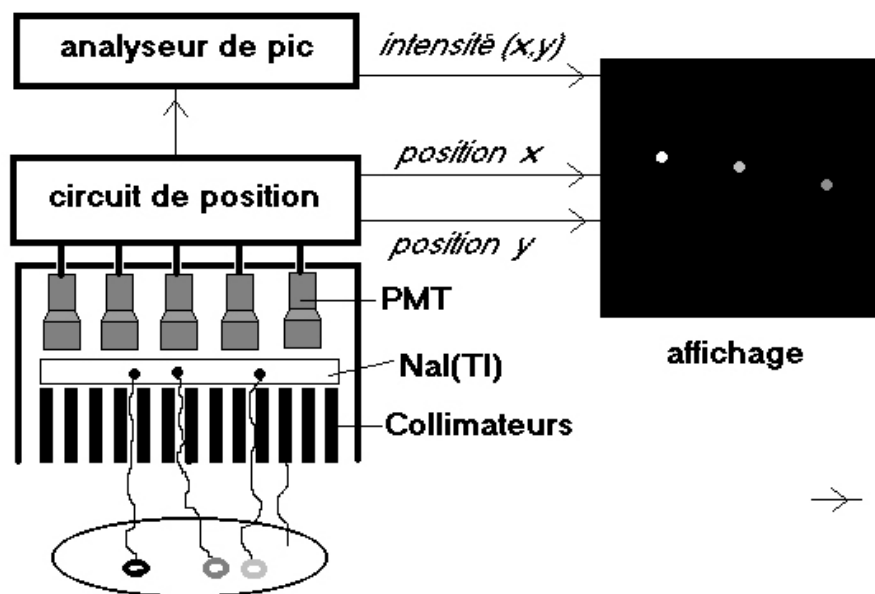


FIG. 2.2 – Détermination de la position et de l'intensité de la scintillation

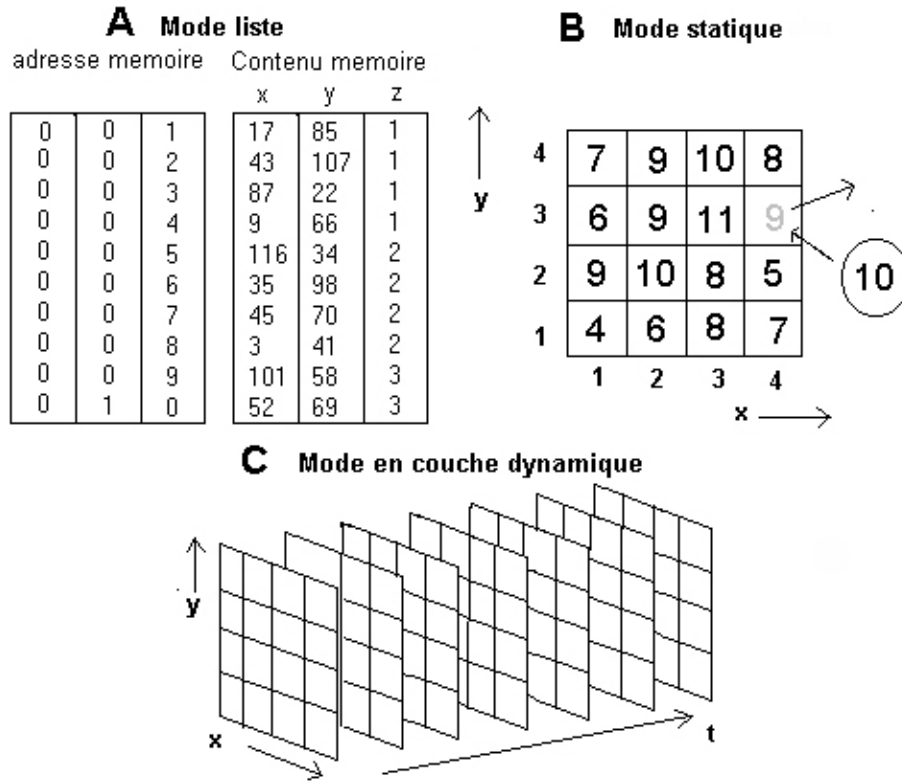


FIG. 2.3 – Les différents mode d'acquisition des données

2.2 Aquisition des données

L'acquisition des données se fait en plaçant la (ou les) gamma caméra à différents angles par rapport au patient [4] [2] [22].

La gamma caméra est couplée à un ordinateur qui va permettre le stockage de l'information et son traitement ultérieur. Pour le stockage, vu le débit d'information il faut qu'il y ait un bon convertisseur analogique-numérique qui va permettre la digitalisation des voltages V_x et V_y pour chaque scintillation.

Il existe trois modes d'acquisition des données[2] :(Fig. 2.3)

- Le mode liste (Fig. 2.3 A), qui va simplement lister dans un tableau toutes les différentes scintillations selon leurs positions sur les axes x et y mais aussi selon l'angle d'obtention de l'évènement. Cela permet de garder toutes les scintillations qui se sont produites lors de l'acquisition, ce qui est très intéressant pour

certains types de recherche. L'inconvénient de cette méthode est qu'elle nécessite une très grande mémoire car le nombre de scintillation est très important.

- Le mode statique (Fig. 2.3 B), pour chaque angle de prise de vue il y a une matrice de 64x64 à 512x512 correspondant à une grille selon les axes x et y où chaque case représente un couple de valeur (x,y) donné. Chaque scintillation va incrémenter de un le nombre de coups sur le pixel où ce sera produit la scintillation. Cette méthode ne requiert pas beaucoup de mémoire et en plus, est directement utilisable pour générer une image.
- Le mode dynamique (Fig. 2.3 C), si le temporel a une importance pour la recherche en cours, on peut utiliser ce mode. Il est basé sur le même principe que le mode statique mais au bout d'un temps déterminé à l'avance l'ordinateur va changer de matrice, et en prendre une nouvelle. Cette procédure sera répétée jusqu'à la fin de l'acquisition. Ce mode peut être intéressant pour suivre le radio-isotope injecté dans le corps du patient.

2.3 Traitement des données

Il existe beaucoup de techniques de reconstruction des images ou des volumes 3D à partir de données brutes [22] [4] .

Nous pouvons citer trois grand types de méthodes :

- Les méthodes de rétro-projection.
- Les méthodes itératives.
- Les méthodes analytiques comme l'analyse de Fourier et la rétro-projection filtrée.

Nous décrirons tout d'abord la méthode de rétro-projection puis la rétro-projection filtrée qui découle directement de celle-ci.

2.3.1 La méthode de rétro-projection

Cette méthode est souvent appelée "méthode des sommations", c'est une des plus

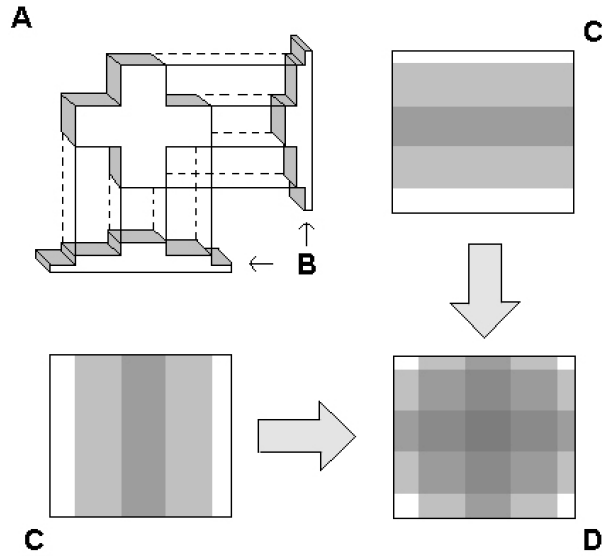


FIG. 2.4 – Principe de la rétro-projection

anciennes méthodes de reconstruction d'images [22].

C'est une méthode qui n'est pas employée en pratique mais elle va servir à poser les bases de la rétro-projection filtrée. Le principe est montré dans la figure 2.4 qui démontre la reconstruction bidimensionnelle d'une croix émettrice de photons γ (Fig. 2.4 A).

On réalise deux mesures sur les côtés de la croix avec des collimateurs(il ne sont pas sur la figure). Le résultat de chaque mesure ressemble à un escalier car la partie du récepteur en face de la partie centrale de la croix reçoit plus de rayonnement que celle qui est juste en face d'un bras (Fig. 2.4 B).

La hauteur de la marche est proportionnelle à un niveau de gris. Ces niveaux de gris sont rangés en colonne (Fig. 2.4 C) dont la largeur est la même que celle de la marche (mais dans la pratique l'échelle peut être changée). Quand les deux projections sont superposées il apparaît une représentation grossière de la croix (Fig. 2.4 D) qui peut être améliorée en faisant plus de projections.

Tous les points de l'image rétroprojetées reçoivent une contribution de leur voisinage à leur niveau de gris. Ce qui a pour effet (Fig. 2.5), par exemple lors de la rétro-

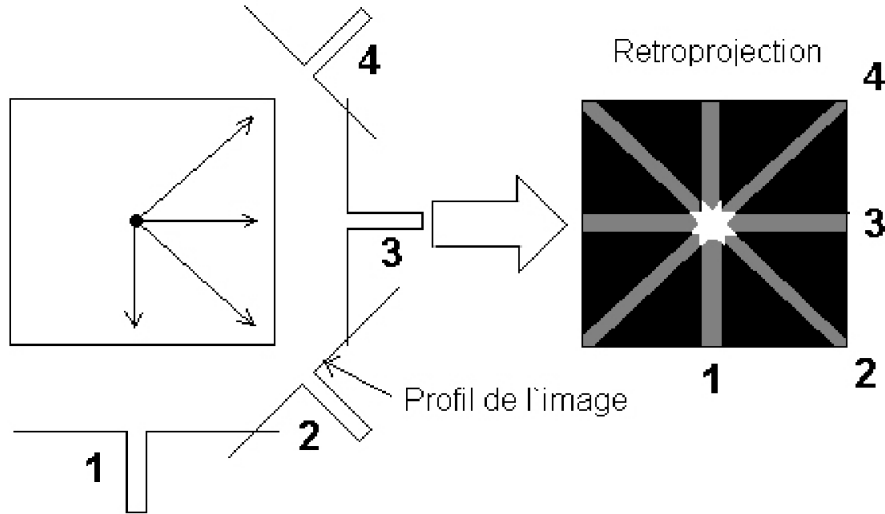


FIG. 2.5 – Démonstration de l'effet d'étoile [22]

projection de 4 images d'un point source selon des angles différents qu'il apparaisse en étoile au lieu d'en un point, ceci peut être atténué par un grand nombre de projections, mais il subsistera toujours une tache (réponse à l'impulsion du système ou PSF : Point Spread Function) plutôt qu'un point.

Cette méthode est assez simple, elle permet aisément de reconstruire des images. Cependant la qualité de l'image reconstruite peut laisser à désirer, c'est pour cela qu'une autre méthode a été introduite, cette méthode s'appelle la rétro-projection filtrée.

2.3.2 La méthode de rétro-projection filtrée

La rétro-projection filtrée est strictement basée sur la rétro-projection sauf que l'image est filtrée, ce qui permet d'éliminer l'effet d'étoile.[22]

C'est une des méthodes de reconstruction les plus employées en imagerie SPECT mais également en imagerie par tomodensitométrie (scanner à rayons X). Le filtrage a pour but d'éliminer les fréquences responsables de l'effet d'étoile pour augmenter les fréquences utiles (Fig. 2.6).

Les limites intérieures des zones où le niveau de gris est très fort sont augmentées

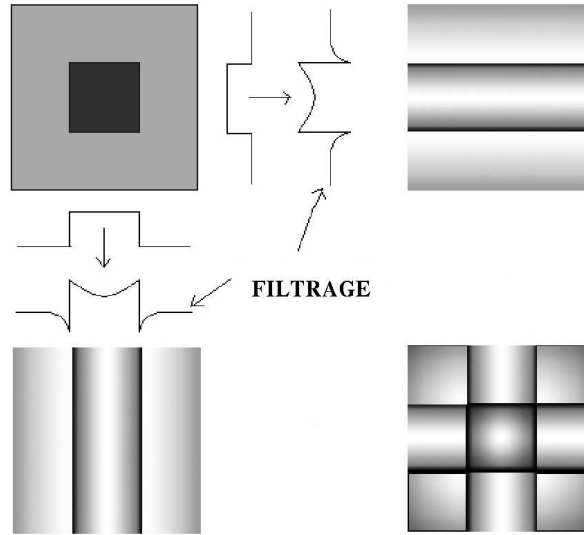


FIG. 2.6 – Principe de la rétro-projection filtrée [22]

tandis que leurs centres et les zones adjacentes voient leurs niveaux de gris réduits. L'image ainsi reconstruite est plus nette et donne une meilleure ressemblance de l'image originale. Dans le cas de cette méthode la qualité de l'image augmentera avec le nombre de projections. Plus formellement on peut décrire cette méthodologie comme au tableau 2.1 [27][4].

Les images ainsi reconstruites vont servir à la classification après avoir subi d'autres traitements, qui dépendent des paramètres que l'on veut prendre en compte pour réaliser la classification. Dans le cas présent, le but est de mettre en évidence une différence d'activité entre les hémisphères gauche et droit du cerveau, survenant avec l'évolution de la maladie. Ce qui, va être fait à l'aide de différentes mesures de l'activité dans les images SPECT et pourrait permettre une classification, ceci est l'objet de notre prochain chapitre.

□ **Rétro-projection filtrée**

- Acquisition de toutes les projections $P_\theta(t)$ à différents angles sachant que :

$$P_\theta(t) = \int_l f(x, y) dl$$

$$x \cos \theta + y \sin \theta = t$$

- FFT de chaque projection.

$$P_\theta(\omega) = \int P_\theta(t) e^{(-j\omega t)} dt$$

Ceci est fait pour chaque θ , ce qui nous donne $F(\omega, \theta)$.

- On filtre par un passe-bande ce qui nous donne $F^*(\omega, \theta)$.
- On réalise la FFT inverse par :

$$f(x, y) = \int \int F^*(\omega, \theta) e^{2\pi j(x\omega \sin \theta + y\omega \cos \theta)} |\omega| d\omega d\theta$$

- On rétro-projete toutes les projections pour avoir la reconstruction.

$$\begin{aligned} f(x, y) &= \int \int \left(P_\theta(\omega) |\omega| \right) e^{j\omega(x \sin \theta + y \cos \theta)} d\omega d\theta \\ &= \int \int \int \left(P_\theta(\omega) |\omega| \right) e^{(j\omega t)} d\omega \delta(x \sin \theta + y \cos \theta - t) dt d\theta \\ &= \int \int P_\theta(t)_{\text{filtree}} \delta(x \sin \theta + y \cos \theta - t) dt d\theta \end{aligned}$$

Tableau 2.1 – Résumé de la rétro-projection filtrée

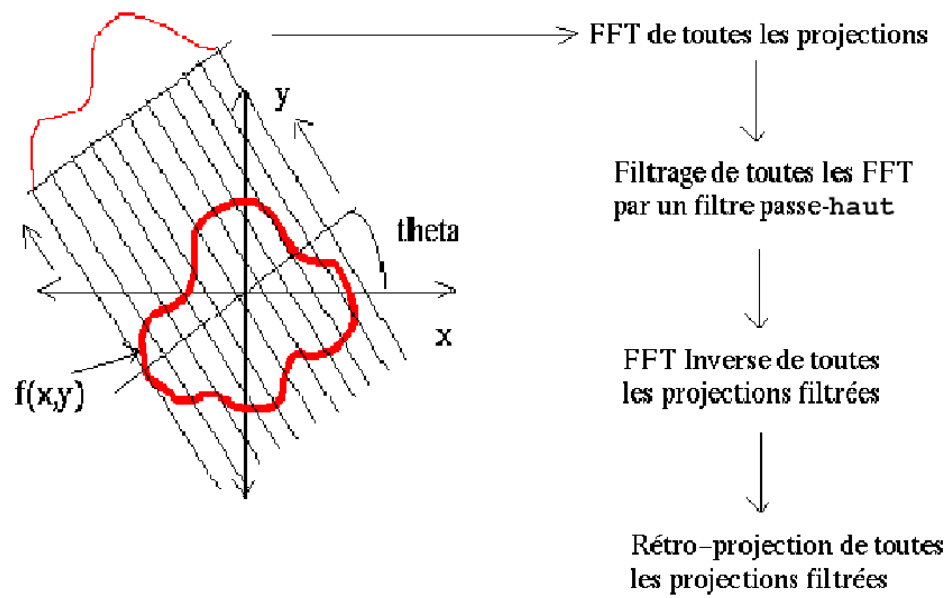


FIG. 2.7 – Principe de la rétro-projection filtrée [22]

Chapitre 3

Méthodes basées sur la comparaison des activités des hémisphères gauche et droit du cerveau

3.1 Introduction et préparation des données

Le but de cette maîtrise était de déterminer la présence de régions anormales diffuses dans les cerveaux par une analyse comparative de la texture. Dans cette optique, nous nous sommes appuyés sur des données d'imageur SPECT qui pouvaient se subdiviser en différentes classes :

- Classe Diffus : 24 cas de patients atteints de maladies diffuses. Ces maladies diffuses sont très difficilement détectables à l'oeil nu (dans les premiers stades de la maladie) car il faudrait pouvoir percevoir de très faibles variations de niveau de gris.
- Classe Focaux : 8 cas de patients atteints de tumeurs focales. Dans les images reconstruites du cerveau de ces patients, on observe une zone dans une partie quelconque du cerveau où l'activité est anormale.

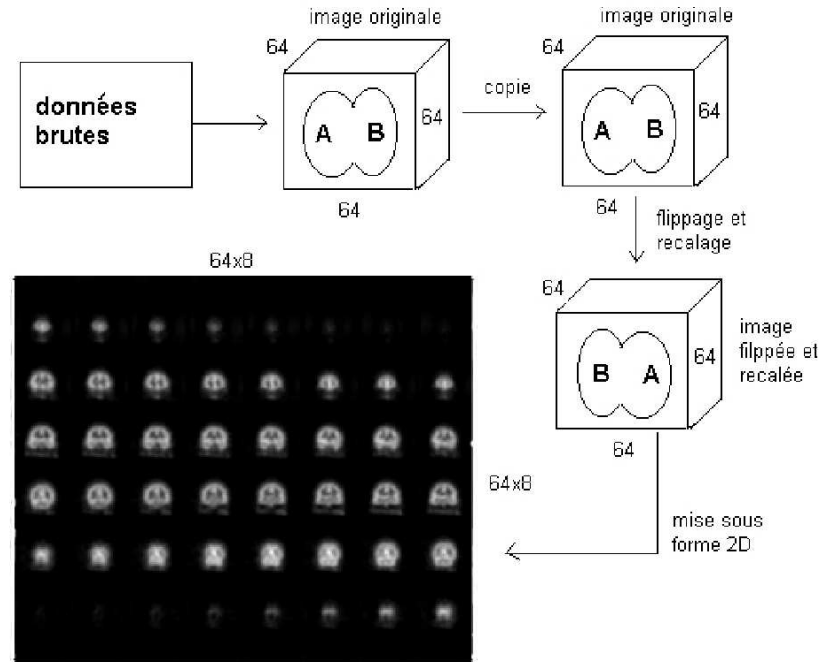


FIG. 3.1 – différentes étapes du traitement préliminaire

- Classe Négatifs : 12 cas de patients, qui après examen se sont avérés négatifs, cependant si ces patients ont passé l'examen, c'est qu'il avait des symptômes (migraines, étourdissement etc..), c'est pour cette raison que même si le diagnostic du médecin est négatif, on ne peut pas les rassembler avec la classe ci-dessous.
- Classe Normaux : 21 cas de données prises sur des médecins et chercheurs volontaires en bonne santé et sans antécédent.

La première étape du processus consiste à reconstruire les images à l'aide d'un programme permettant de réaliser une rétro-projection filtrée des données brutes de l'imageur (voir le paragraphe 2.3.2). La reconstruction aboutit à des images de 64x64 avec 64 tranches pour chaque cerveau.

Un premier objectif est de comparer les parties droites et gauches de chaque tranche d'un cerveau pour y détecter des différences d'activité, en faisant l'hypothèse que les différences d'activité sont faibles chez les sujets normaux. Pour ce faire, il

faut tout d'abord retourner ("*flipper*") une copie de l'image originale et la recalcr sur cet original, par rapport au plan sagittal (cf. Fig. 3.1). Nous avons utilisé le logiciel SPM [20] pour réaliser cette partie. L'algorithme de recalage utilisé est basé sur la minimisation de la variance de la différence des images à recalcr et à été proposée par Woods *et al.* [15].

A ce stade, nous avons donc deux groupes de données pour chaque cas :

- les données originales.
- les données retournées et recalées.

Pour plus de simplicité, au lieu de travailler en trois dimensions, il est utile pour fin de visualisation de réduire les données en mode 2D c'est à dire, mettre sur un même plan, les 64 tranches de chaque image.

Il faut maintenant comparer les parties gauches et droites du cerveau pour cela, il existe différentes méthodes qui ont pour but de chercher un attribut caractéristique à chaque classe. Nous commencerons par une description des différents attributs que nous avons recueillis à partir des images, puis nous nous intéresserons aux résultats obtenus.

3.2 Somme des différences pondérée D

Le but est de faire une différence en valeur absolue, pixel à pixel de l'image originale et de l'image flippée et recalée (cf. Fig. 3.2) [10]. En premier lieu, il faut mettre les deux images dans deux matrices 512x512 Mat1 et Mat2, respectivement image originale et image flippée et recalée.

Puis, on réalise un masque de l'image originale en prenant tout d'abord son maximum (que nous appellerons Max) et en comparant la valeur de ces pixels avec celle du Max (en prenant un pourcentage du Max), on peut attribuer à chaque pixel du masque une valeur (cf. Fig. 3.2 A) :

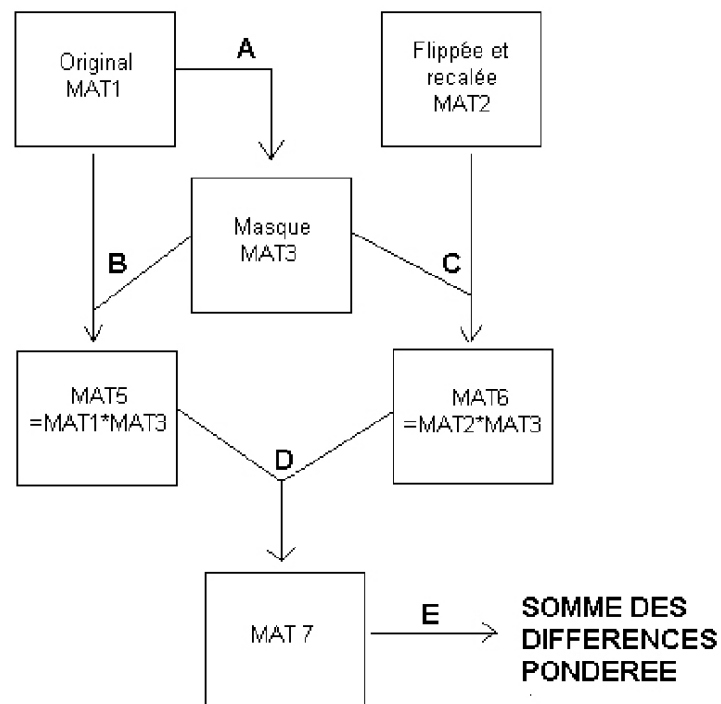


FIG. 3.2 – Différentes étapes pour la réalisation de la Somme des différences pondérée

D

$$\text{Mat3}[i][j] = \begin{cases} 1 & \text{si } \text{Mat1}[i][j] \geq 30\% \text{ du Max} \\ 0 & \text{si } \text{Mat1}[i][j] < 30\% \text{ du Max} \end{cases}$$

Le choix de 30 % est dictée par l'expérience mais au besoin le pourcentage du Max peut être changé pour affiner le masque.

Cette matrice de masque (Mat3) sera ensuite multipliée avec Mat1 et Mat2, (ce qui nous donne $\text{Mat5} = \text{Mat1} * \text{Mat3}$ et $\text{Mat6} = \text{Mat2} * \text{Mat3}$) (cf. Fig. 3.2 B et C) ce qui permet d'éliminer l'arrière-plan avec le bruit et les artefacts de reconstruction. La comparaison peut alors se faire en additionnant dans une variable $\wp$, la valeur absolue de la différence de chaque pixel(i,j) de Mat5 et Mat6 (cf Fig 3.2 D), ce qui nous donne (cf Eq 3.1) :

$$\wp = \frac{1}{N} \sum_{i=0, j=0}^{i=511, j=511} | \text{Mat5}[i][j] - \text{Mat6}[i][j] | \quad (3.1)$$

où N est le nombre de pixels du masque utilisé pour fin de normalisation.

3.3 Le coefficient de corrélation ρ

On peut le définir ainsi (cf Eq 3.2)[19] [16] :

$$\rho = \frac{\sum_{i=0, j=0}^{i=511, j=511} (x_{ij} - \bar{x})(y_{ij} - \bar{y})}{\sqrt{\sum_{i=0, j=0}^{i=511, j=511} (x_{ij} - \bar{x})^2} \sqrt{\sum_{i=0, j=0}^{i=511, j=511} (y_{ij} - \bar{y})^2}} \quad (3.2)$$

Où x et y représentent respectivement l'image originale et l'image flippée recalée. $\bar{x}$ et $\bar{y}$ représentent la moyenne des niveaux de gris de tous les pixels dans les matrices après multiplication par le masque, respectivement, image originale et image flippée recalée.

Ce coefficient peut prendre théoriquement une valeur comprise entre -1 et 1, sachant que si l'on a une image qui est le négatif de l'autre, le coefficient de corrélation aura une valeur de -1 et que plus deux images sont identiques plus leur coefficient sera près de 1.

La réalisation d'un masque pour le calcul de cet attribut est également nécessaire pour les raisons citées ci-dessus.

Le coefficient de corrélation se calcule aisément à partir des données des matrices, il suffit juste de décomposer les calculs de sorte que la programmation soit plus aisée.

3.4 Matrice de cooccurrence

La matrice de cooccurrence est couramment utilisée en traitement d'images pour mesurer certaines propriétés comme la texture [19] [16] [25] [6] [14] [10]. Pour nos travaux, elle servira à obtenir de nouvelles mesures de la variabilité entre les hémisphères gauche et droit. Pour nos images de 256x256 niveaux de gris, la matrice de cooccurrence est une matrice de 256x256 (unités) que l'on remplit en comptabilisant le nombre d'occurrences des niveaux de gris i et j dans l'image originale et dans l'image flippée recalée respectivement. Par exemple $\text{Mat8}[10][20] = 100$ indique qu'il existe 100 pixels de niveaux de gris égale à 10 dans la première image et de niveaux de gris égale à 20 dans la seconde image pour un pixel de même coordonné.

Tous les pixels sont ainsi classés selon leur valeur dans l'image originale et dans l'image flippée recalée. Si l'on réalise la matrice de cooccurrence avec deux images strictement identiques, on obtiendra une ligne dans la diagonale de Mat8 où $i=j$ (cf. Fig. 3.3 en haut). Plus les deux images sont différentes, plus il y aura un nuage de point s'éloignant de la diagonale (cf. Fig. 3.3 milieu et bas).

On voit sur les images des figures 3.3 et 3.4 que le nuage de point en diagonale ne commence pas tout à fait à gauche, cela est dû au fait que l'on ne prend que des valeurs de niveau de gris supérieur à un certain seuil (masque).

Après normalisation, cette matrice représente les probabilités de coocurrences $\mathcal{P}(i,j)$ des niveaux de gris (i,j) , dans les images comparées.

A partir de cette matrice, on peut calculer différents attributs comme par exemple :



FIG. 3.3 – Exemples de matrice de cooccurrence

$$\text{L'énergie} = \sum_{i=0, j=0}^{i=511, j=511} \mathcal{P}(i, j)^2 \quad (3.3)$$

Cette mesure définit *l'hétérogénéité* de la texture [9]. Elle est relativement petite lorsque toutes les valeurs des $\mathcal{P}(i, j)$ sont égales. Par contre, elle est grande, lorsque les valeurs des $\mathcal{P}(i, j)$ sont différentes [10].

$$\text{L'entropie} = \sum_{i=0, j=0}^{i=512, j=512} \mathcal{P}(i, j) \text{Log} \mathcal{P}(i, j) \quad (3.4)$$

Cet attribut calcule la quantité d'information contenue dans la texture. La valeur de l'attribut est grande lorsque les $\mathcal{P}(i, j)$ sont égaux.

$$\text{L'inertie} = \sum_{i=0, j=0}^{i=512, j=512} (i - j)^2 \mathcal{P}(i, j) \quad (3.5)$$

Cet attribut est grand lorsqu'il y a une dispersion du nuage de points par rapport à la diagonale, ce qui peut être intéressant lorsque par exemple l'on s'intéresse aux maladies focales qui ne sont situées que d'un côté du cerveau.

Les deux matrices de la figure 3.4 illustrent l'effet sur une matrice de cooccurrence de l'augmentation d'activité d'un côté seulement du cerveau ; qui a été volontairement retouchée pour qu'apparaisse une tumeur focale du côté droit (on peut le voir par une augmentation d'activité), on peut noter la dispersion plus importante au coin inférieur droit. L'image non modifiée donne un attribut de 96.206509, tandis que la modification donne 1418.098642. Cet attribut devrait théoriquement permettre une meilleure séparation des tumeurs focales.

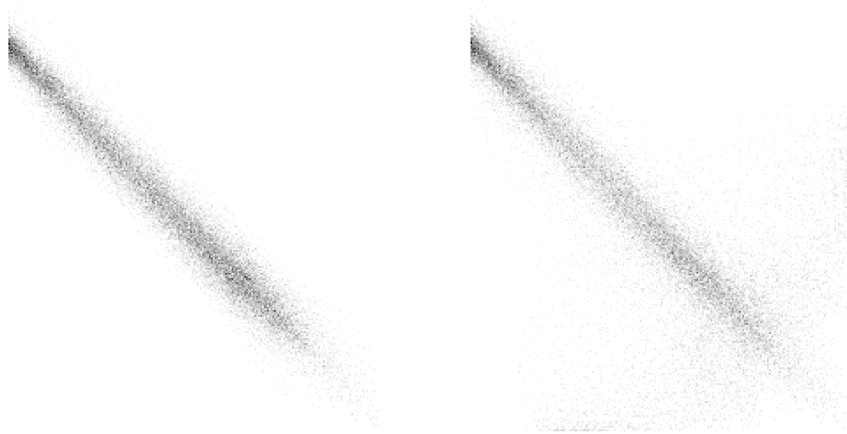


FIG. 3.4 – matrice de cooccurrence obtenue avec l'image de cerveau sain (à gauche) et à l'aide d'un cerveau atteint de tumeur focale (à droite)

3.5 Analyse des résultats

Afin d'analyser nos résultats, nous allons effectuer différents tests statistiques. Cependant il faut prendre en compte que le nombre de données que nous possédons dans chaque catégorie, c'est à dire le nombre d'image de cerveau 3D SPECT pour chaque cas considéré est relativement faible (cf. section 3.1).

Nous avons utilisé différents tests pour valider la séparabilité des données et la classification ainsi obtenue :

- Le test t de Student et le test F de Fisher [18] sont appropriés pour de petits ensembles de données : ils permettent de tester respectivement, la moyenne et la variance de 2 ensembles de données.
- Des algorithmes de classification, tel que l'algorithme des K -moyennes et l'EM¹.

3.5.1 Tests statistiques

Le test t de Student et le test F de Fisher [18] sont appropriés pour de petits ensembles de données, ils permettent de tester respectivement, la moyenne et la variance de 2 ensembles de données X et Y . Les 2 tests fonctionnent sur le même principe :

¹ *Estimation, Maximisation*

- Hypothèse égalité $\mu_X = \mu_Y$ ou $\sigma_X^2 = \sigma_Y^2$.
- L'hypothèse peut être rejetée ou acceptée, ce qui dénotera si les 2 groupes de données sont significativement différents, donc pour notre intérêt si les deux groupes peuvent être classifiés avec un taux d'erreur acceptable.

Il faut rappeler que la procédure du test t de Student pour tester l'égalité de 2 moyennes est conditionnelle au fait que les variances des 2 populations soient égales. Cela implique que l'hypothèse $H_0 : \sigma_X^2 = \sigma_Y^2$ doit précéder tout test de $H_0 : \mu_X = \mu_Y$, car si $H_0 : \sigma_X^2 = \sigma_Y^2$ est rejeté alors le test de Student est inapproprié. Dans les deux cas, les données sont 2 échantillons aléatoires de taille n et m . Le premier $\{X_1, X_2, \dots, X_n\}$, est sensé avoir une distribution gaussienne, de moyenne μ_X et de variance σ_X^2 et le second $\{Y_1, Y_2, \dots, Y_m\}$ est supposé avoir une distribution gaussienne de moyenne μ_Y et de variance σ_Y^2 .

3.5.1.1 test F de Fisher, $H_0 : \sigma_X^2 = \sigma_Y^2$

Ce test est un rapport de vraisemblance pour vérifier l'hypothèse de $H_0 : \sigma_X^2 = \sigma_Y^2$ par rapport à $H_1 : \sigma_X^2 \neq \sigma_Y^2$.

On estime le rapport de vraisemblance généralisé par rapport aux valeurs de la fonction de Fisher $F_{\frac{\alpha}{2}, m-1, n-1}$ et $F_{1-\frac{\alpha}{2}, m-1, n-1}$, donné par des tables en fonction de la probabilité d'erreur α voulue et du degré de liberté $n+m-2$ pour rejeter ou accepter H_0 . Une description est donnée dans le tableau 3.1.

3.5.1.2 test t de Student, $H_0 : \mu_X = \mu_Y$

Cette fois ci, on compare deux hypothèses sur les moyennes $H_0 : \mu_X = \mu_Y$ et $H_1 : \mu_X \neq \mu_Y$. La comparaison se fait à partir de $-t_{\frac{\alpha}{2}, m+n-2}$ et $t_{\frac{\alpha}{2}, m+n-2}$ données à partir de tables en fonction de l'erreur désirée et du degré de liberté. Le tableau 3.2 en fournit une description.

□ test F de Fisher

Avec $\{X_1, X_2, \dots, X_n\} \sim N(\mu_X, \sigma_X^2)$

et $\{Y_1, Y_2, \dots, Y_m\} \sim N(\mu_Y, \sigma_Y^2)$, les X et Y sont indépendants.

$$H_0 : \sigma_X^2 = \sigma_Y^2$$

par rapport à

$$H_1 : \sigma_X^2 \neq \sigma_Y^2$$

selon α , la probabilité d'erreur voulue et des degrés de liberté $n - 1$ et $m - 1$, H_0 sera rejeté si :

$$\frac{\sigma_Y^2}{\sigma_X^2} \text{ est } \begin{cases} \leq F_{\frac{\alpha}{2}, m-1, n-1} \\ \text{ou} \\ \geq F_{1-\frac{\alpha}{2}, m-1, n-1} \end{cases} \quad (3.6)$$

Tableau 3.1 – Résumé du test F de Fisher

3.5.2 Algorithme de classification

Nous avons utilisé deux algorithmes pour la classification des données à partir des attributs recueillis : K-moyenne et EM. Ce choix est essentiellement justifié par la simplicité d'implémentation de ces algorithmes ainsi que leur bonne propriété de (rapide) convergence. Nous décrirons tout d'abord brièvement ces algorithmes avant de présenter les résultats.

3.5.2.1 Algorithme des K-moyennes

L'algorithme K-Moyennes est une des méthodes les plus utilisées, il permet de déterminer une partition des données en supposant que ces données se distribuent selon des hypersphères selon un nombre de classes K , il procède selon les étapes décrites dans le tableau 3.3. Le cas présenté donne la procédure pour une dimension mais on peut également utiliser cet algorithme pour faire une partition des données

□ **Test t de Student**

Avec $\{X_1, X_2, \dots, X_n\} \sim N(\mu_X, \sigma_X^2)$

et $\{Y_1, Y_2, \dots, Y_m\} \sim N(\mu_Y, \sigma_Y^2)$, les X et Y sont indépendants.

$$H_0 : \mu_X = \mu_Y$$

par rapport à

$$H_1 : \mu_X \neq \mu_Y$$

selon α , la probabilité d'erreur voulue et des degrés de liberté $n + m - 2$, H_0 sera rejeté si :

$$t = \frac{(\bar{x} - \bar{y})}{\sqrt{(n\sigma_X^2) + (m\sigma_Y^2)}} \sqrt{\frac{(n + m - 2)nm}{n + m}} \text{ est } \begin{cases} \leq t_{\frac{\alpha}{2}, m+n-2} \\ \text{ou} \\ \geq t_{1-\frac{\alpha}{2}, m+n-2} \end{cases} \quad (3.7)$$

Tableau 3.2 – Résumé du test t de Student

selon plusieurs dimensions, donc dans notre cas, un vecteur d'attributs.

3.5.2.2 Algorithme EM

L'algorithme EM va nous permettre d'estimer ce mélange de 2 lois que nous supposons dans notre application comme étant gaussiennes. Il s'agit d'un algorithme itératif qui peut être utilisé pour calculer ces paramètres au sens du maximum de vraisemblance quand les observations sont vues comme étant des données incomplètes. A chaque itération l'algorithme suit deux étapes, la première que l'on appelle l'étape d'estimation et une deuxième étape que l'on appelle l'étape de maximisation.

Soit l'échantillon $Y = \{y_1, y_2, \dots, y_N\}$ avec $N = 64^3$ réalisations particulières du vecteur de variables aléatoires Y (les données que l'on cherche à partitionner) et $x = \{x_1, x_2, \dots, x_N\}$, réalisation particulière du champs des étiquettes associés à chaque donnée $x_i \in \{e_0, e_1, \dots, e_K\}$ (les K étiquettes ou classes possibles). La loi de Y s'écrit :

$$P_Y(y) = \sum_{k=0}^K \pi_k P_{Y_i|X_i}(y_i|e_k), \quad (3.11)$$

en notant π_k , la probabilité qu'un point de l'échantillon y suive la loi $P_{Y_i|X_i}(y_i|e_k)$. Connaissant y , le problème consiste à estimer les paramètres Φ d'un mélange de densités de probabilité [3]. Dans notre application, puisque $K = 2$ le problème consiste plus précisément à estimer les paramètres des lois $P_{Y_i|X_i}(y_i|e_0)$ et $P_{Y_i|X_i}(y_i|e_1)$ ainsi que les proportions π_0 et π_1 de ce mélange.

3.5.3 Résultats

L'analyse des résultats doit se faire en essayant de comparer les attributs de deux classes. Cela peut être fait pour différents cas :

- Diffus-Normaux
- Diffus-Négatifs
- Focaux-Normaux

□ **Algorithme K-Moy.**

1. On associe aléatoirement chaque échantillon y_i à un des K groupements, et on calcule les centres c_k des K groupements :

$$c_k^{[1]} = \frac{1}{N_k} \sum_{y \in C_k^{[0]}} y \quad \forall k \quad (3.8)$$

avec N_k le nombre d'échantillons de $C_k^{[p]}$.

2. A l'itération p , on associe l'échantillon y_l , ($1 \leq l \leq N$) au groupement C_k (noté $C_k^{[p]}$) si :

$$\| \mathbf{y}_l - c_i^{[p]} \| < \| \mathbf{y}_l - c_j^{[p]} \| \quad \forall j \neq i \quad (3.9)$$

En fait, on associe l'échantillon $\mathbf{y}_l$ au groupement dont le centre lui est le plus proche.

3. On détermine le nouveau centre de chaque groupement par :

$$c_i^{[p+1]} = \frac{1}{N_i} \sum_{y \in C_i^{[p]}} y \quad \forall i \quad (3.10)$$

avec N_k le nombre d'échantillons de $C_k^{[p]}$.

4. Si les centres des groupements à l'itération p sont les mêmes que ceux obtenus à l'itération précédente, alors les groupements constitués sont identiques et l'algorithme a convergé, sinon on retourne à l'étape 2.

Tableau 3.3 – Algorithme des K-moyennes.

□ **Algorithme** EM (POUR L'ESTIMATION D'UN MÉLANGE DE 2 LOIS GAUSSIENNES).

$\hat{\Phi}^{[0]} = (\hat{\Phi}_1^{[0]}, \hat{\Phi}_2^{[0]})$	vecteur de paramètres initial
$\hat{\Phi}_1$	vecteur de paramètres de la première loi Gaussienne
$\hat{\Phi}_2$	vecteur de paramètres de la deuxième loi Gaussienne

- Calcul de $\hat{\Phi}^{[p+1]}$ à partir de $\hat{\Phi}^{[p]}$

- Étape d'*Estimation*

Pour chaque y_i , on calcule, en se basant sur les paramètres calculés à l'itération précédente, les probabilités $\hat{\mathbf{x}}_{i[k=\text{classe } 1]}^{[p]}$ et $\hat{\mathbf{x}}_{i[k=\text{classe } 2]}^{[p]}$ (cf. tableau 3.5) avec :

$$P_{Y_i|X_i, \Phi_1}(y_i | e_0 = \text{classe } 1, \hat{\Phi}_1^{[p]}) \sim \mathcal{N}(y_i; \Phi_1^{[p]})$$

$$P_{Y_i|X_i, \Phi_2}(y_i | e_1 = \text{classe } 2, \hat{\Phi}_2^{[p]}) \sim \mathcal{N}(y_i; \Phi_2^{[p]})$$

- Étape de *Maximisation*

▷ Pour $\hat{\Phi}_1^{[p+1]} = (\hat{\pi}_1^{[p+1]}, \hat{\mu}_1^{[p+1]}, \hat{\sigma}_1^{[p+1]})$ et $\hat{\Phi}_2^{[p+1]} = (\hat{\pi}_2^{[p+1]}, \hat{\mu}_2^{[p+1]}, \hat{\sigma}_2^{[p+1]})$

on ré-estime la moyenne, la variance et la proportion en utilisant les relations données dans le tableau 3.5.

Retour à l'étape d'estimation jusqu'à la réalisation d'un critère d'arrêt.

Généralement :

si $\hat{\Phi}^{[p+1]} \not\approx \hat{\Phi}^{[p]}$ $p \leftarrow (p + 1)$ et retour à l'étape d'estimation.

Tableau 3.4 – Algorithme EM pour l'estimation des paramètres d'un mélange de deux lois Gaussiennes

$$\hat{\mu}_k^{[p+1]} = \frac{\sum_{i=1}^N \hat{\mathbf{x}}_{ik}^{[p]} y_i}{\sum_{i=1}^N \hat{\mathbf{x}}_{ik}^{[p]}} \quad (3.12)$$

$$(\hat{\sigma}_k^{[p+1]})^2 = \frac{\sum_{i=1}^N \hat{\mathbf{x}}_{ik}^{[p]} (y_i - \hat{\mu}_k^{[p+1]})^2}{\sum_{i=1}^N \hat{\mathbf{x}}_{ik}^{[p]}} \quad (3.13)$$

avec

$$\hat{\mathbf{x}}_{ik}^{[p]} = \frac{\hat{\pi}_k^{[p]} P_{Y_i|X_i, \Phi_k}(y_i|e_k, \hat{\Phi}_k^{[p]})}{\sum_{j=1}^k \hat{\pi}_j^{[p]} P_{Y_i|X_i, \Phi_k}(y_i|e_k, \hat{\Phi}_k^{[p]})} \quad (3.14)$$

$$\hat{\pi}_k^{[p+1]} = \frac{\sum_{i=1}^N \hat{\mathbf{x}}_{ik}^{[p]}}{N} \quad (3.15)$$

Tableau 3.5 – Algorithme EM. Formule de ré-estimation dans le cas d'un mélange de deux lois Gaussiennes.

- Focaux-Négatifs
- Focaux-Diffus

Les différents attributs testés sont : Somme des différences pondérée D (att1), le coefficient de corrélation avec deux seuils différents à 15% et 30% (att2 et att3) et les attributs extraits des matrices de cooccurrence, l'énergie (att4), l'entropie (att5), l'inertie (att6).

3.5.4 Diffus-Normaux

Le cas le plus intéressant pour notre étude est la comparaison des attributs des classes Diffus et Normaux, et ce pour deux raisons. Tout d'abord, le but de ce projet était une différenciation entre ces deux classes précisément et également à cause du fait que le nombre de ces données pour ces deux classes est le plus important, 45 cas en tout, ce qui est déjà relativement faible pour les algorithmes de classification considérés. Le tableau 3.6 montre les résultats pour les meilleurs attributs. Trois attributs ont un résultat au test de Fisher (avec un $\alpha=0.05$, il faut que le résultat au F-test soit comprise entre 0.48 et 2.03) qui permet de continuer l'étude, c'est à dire que les variances des 2 classes sont comparables. La comparaison de leurs moyennes par le

test de t-Student nous montre que celle-ci ne sont pas significativement différentes. La classification à l'aide des algorithmes a confirmé le résultat en ne donnant que des pourcentages de classification inférieurs à 60%, ce qui n'est pas un taux de classification acceptable.

Attributs	F-test	t-test	K-moyenne	EM
att1	0.252	1.103	0.518	0.518
att2	0.127	1.403	0.524	0.524
att3	0.392	0.529	0.518	0.539
att4	0.987	0.559	0.577	0.503
att5	0.821	0.639	0.509	0.515
att6	0.092	0.855	0.524	0.524

Tableau 3.6 – Résultats de classification entre la classe Normaux et la classe Diffus, avec 43 degrés de liberté ($m + n - 2 = 43$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.48, 2.03]$ pour être significatif et un résultat au test t de Student > 2.0167 ou < -2.0167 pour être révélateur

3.5.5 Autres résultats

Les différents tableaux ci-après (3.7, 3.8, 3.9, 3.10) indiquent les résultats obtenus en essayant de séparer les autres classes entre elles. Les résultats montrent que même si les attributs ont des variances comparables, leur moyenne l'est aussi ce qui ne permet pas de faire une classification avec un taux d'erreur acceptable.

Attributs	F-test	t-test	K-moyenne	EM
att1	11.498	1.410	0.625	0.542
att2	12.94	1.887	0.667	0.542
att3	8.425	1.152	0.604	0.563
att4	2.724	0.726	0.625	0.625
att5	3.084	0.279	0.521	0.583
att6	9.885	0.944	0.563	0.563

Tableau 3.7 – Résultats de classification entre la classe Négatif et la classe Diffus, avec 34 degrés de liberté ($m + n - 2 = 34$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.383, 2.21]$ pour être significatif et un résultat au test t de Student > 2.0323 ou < -2.0323 pour être révélateur

Attributs	F-test	t-test	K-moyenne	EM
att1	0.605	0.293	0.539	0.539
att2	0.171	0.341	0.539	0.539
att3	0.564	0.490	0.539	0.539
att4	0.523	0.692	0.574	0.531
att5	0.539	0.491	0.607	0.539
att6	0.261	0.314	0.539	0.539

Tableau 3.8 – Résultats de classification entre la classe Normaux et la classe Focaux, avec 29 degrés de liberté ($m + n - 2 = 29$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.290, 2.51]$ pour être significatif et un résultat au test t de Student > 2.0518 ou < -2.0518 pour être révélateur

Attributs	F-test	t-test	K-moyenne	EM
att1	27.591	1.535	0.563	0.563
att2	17.5	1.760	0.563	0.563
att3	12.128	1.471	0.563	0.563
att4	1.444	0.524	0.521	0.542
att5	2.024	1.094	0.646	0.563
att6	27.958	1.559	0.563	0.563

Tableau 3.9 – Résultats de classification entre la classe Négatif et la classe Focaux, avec 20 degrés de liberté ($m + n - 2 = 20$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.278, 3.01]$ pour être significatif et un résultat au test t de Student > 2.1009 ou < -2.1009 pour être révélateur

Compte-tenu de ces résultats négatifs, nous proposerons dans le chapitre suivant une autre approche pour la classification. Cette nouvelle approche inclus une segmentation des images de cerveau selon les structures macroscopiques (liquide céphalo-rachidien, matière blanche et matière grise) ce qui permettra de faire une analyse régionale.

Attributs	F-test	t-test	K-moyenne	EM
att1	2.4	1.246	0.521	0.583
att2	1.352	0.941	0.521	0.5211
att3	1.439	1.062	0.5	0.5
att4	0.530	1.137	0.604	0.604
att5	0.656	1.012	0.583	0.521
att6	2.828	1.556	0.542	0.563

Tableau 3.10 – Résultats de classification entre la classe Diffus et la classe Focaux, avec 32 degrés de liberté ($m + n - 2 = 32$), avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.293, 2.42]$ pour être significatif et un résultat au test t de Student > 2.0423 ou < -2.0423 pour être révélateur

Chapitre 4

Segmentation markovienne d'images SPECT du cerveau

4.1 Introduction

Notre but est de segmenter ces images SPECT 3D selon les 3 structures macroscopiques du cerveau qui sont :

- Le liquide céphalo-rachidien (CSF), qui entoure le cerveau. Cette région n'est pas vascularisée et n'émet donc pas de radiation. Son activité sera donc quasi-nulle.
- La matière blanche qui, quant à elle, est vascularisée et aura une activité normale.
- La matière grise qui est très fortement vascularisée et qui aura une activité très importante en conséquence.

Cette segmentation permettra d'extraire ultérieurement des attributs selon ces trois régions.

La répartition statistique des niveaux de gris dans toutes ces classes peut être modélisée par différentes lois [13]. Le CSF peut être modélisé suivant une loi exponentielle de type [22] : $\frac{1}{\alpha}e^{-\frac{x}{\alpha}}$, α étant le paramètre qui permet de contrôler la décroissance

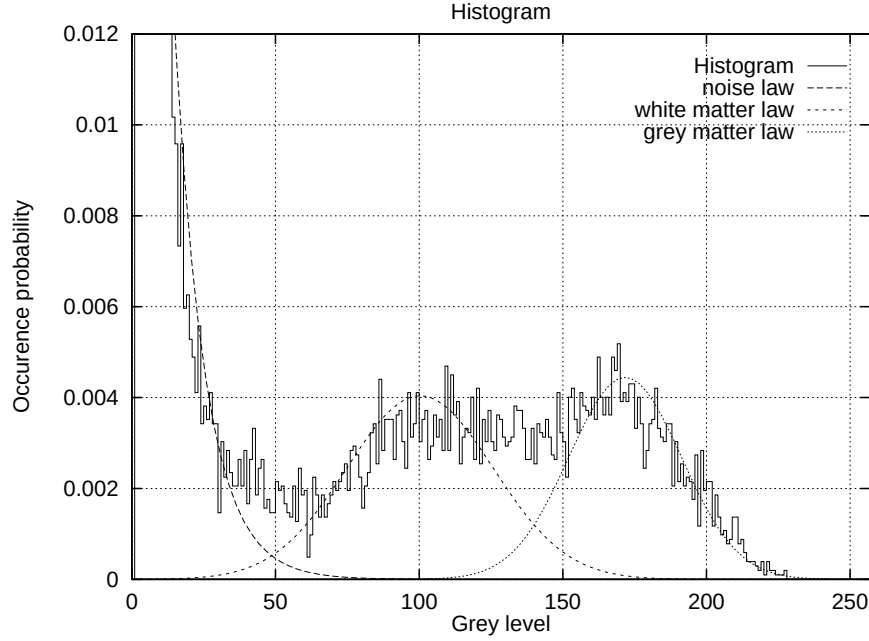


FIG. 4.1 – Histogramme des niveaux de gris d'une image SPECT

de l'exponentielle. La matière blanche et la matière grise seront modélisées par deux gaussiennes de paramètres (μ_1, σ_1) et (μ_2, σ_2) (cf Fig. 4.1).

La première partie consiste à faire l'estimation de ces paramètres puis dans un second temps, on utilisera ces paramètres estimés pour réaliser une segmentation markovienne de ces images SPECT.

4.2 Estimation des paramètres

Le problème de l'estimation des paramètres réside dans deux facteurs majeurs [12] :

- Le fait qu'il y ait deux types de lois différentes dont les paramètres sont à estimer : une de type exponentielle et deux de type gaussienne. Ceci limite les algorithmes que l'on peut utiliser.
- De plus, il faut prendre en compte le caractère *a priori* markovien de la carte de segmentation, ceci limite une fois de plus le choix de l'algorithme pour l'estimation des paramètres.

4.2.1 Définition théorique du problème

Il faut définir correctement le problème. Nous avons une image de 64x64x64 voxels, soit $y = \{y_1, y_2, \dots, y_N\}$, une réalisation du vecteur de variables aléatoires Y associé à chacun de ces voxels ($N=64 \times 64 \times 64$) indexé sur une grille cubique S , avec pour chacun d'eux un niveau de gris qui peut être compris entre 0 et 255. Soit $x = \{x_1, x_2, \dots, x_N\}$, une réalisation du champs d'étiquette ou du vecteur de variables aléatoires X traduisant la classe ou l'étiquette de chaque y_i . $x_i \in e_0, e_1, e_2$, sont les trois classes selon lesquelles on désire segmenter.

On va chercher à estimer $P_{X,Y}(x, y)$, cette probabilité peut être décomposée comme suit :

$$\begin{aligned} P_{X,Y}(x, y) &\propto P_X(x) P_{Y|X}(x|y) \\ &= P_X(x) \prod_{s \in S} P_{Y_s|X_s}(y_s|x_s) \end{aligned}$$

La distribution de la probabilité conditionnelle pour chaque classe est modélisée par ces lois :

- $P_{Y_s|X_s}(y_s|e_0)$: qui correspond au CSF et qui est modélisé par une exponentielle de paramètre α , $\Phi_0 = (\alpha)$.
- $P_{Y_s|X_s}(y_s|e_1)$: qui correspond à la matière blanche, et qui est modélisée par une gaussienne de paramètres μ_{e_1} et $\sigma_{e_1}^2$, $\Phi_1 = (\mu_{e_1}, \sigma_{e_1}^2)$
- $P_{Y_s|X_s}(y_s|e_2)$: qui correspond à la matière grise, et qui est modélisée par une gaussienne de paramètres μ_{e_2} et $\sigma_{e_2}^2$, $\Phi_2 = (\mu_{e_2}, \sigma_{e_2}^2)$

L'ensemble des paramètres à estimer constitue donc un vecteur Φ de 5 paramètres, $\alpha, \mu_{e_1}, \sigma_{e_1}^2, \mu_{e_2}, \sigma_{e_2}^2$.

Nous ne connaissons pas les classes d'appartenance de chaque voxel, ce qui nous place dans le cas d'estimateurs de données incomplètes.

Ce problème peut être vu une nouvelle fois comme l'estimation d'un mélange de trois distributions (2 lois gaussiennes et une loi exponentielle) dont la somme pondérée par la proportion de chaque classe est aussi l'histogramme de l'image. La répartition

des niveaux de gris dans l'image 3D SPECT se traduit par la loi :

$$P_Y(y) = \sum_{k=1}^K \pi_k P_{Y|X_i}(y|e_k, \hat{\Phi}_k)$$

C'est ici qu'il faut choisir l'algorithme pour calculer $\hat{\Phi}$ (le vecteur de paramètres). Les algorithmes capables d'estimer les paramètres de ce mélange de distribution peuvent être divisés en 2 classes :

- Les algorithmes locaux de type EM ou SEM, ces algorithmes supposent que les variables aléatoires X_i sont indépendantes entre elles, donc de ce fait, ne tiennent pas compte du caractère markovien de la carte de segmentation.
- Les algorithmes globaux du type ECI¹ qui quant à eux tiennent compte du caractère markovien de la carte de segmentation. Ces derniers utilisent les informations *à priori* que l'on a sur X .

Nous détaillerons le plus efficace d'entre eux dans le cadre de cette application, à savoir l'algorithme ECI, la description de l'algorithme EM a déjà été faite précédemment et se trouve au tableau 3.5.

4.2.2 Algorithme ECI

Cet algorithme va chercher (au sens des moindres carrés), le meilleur vecteur de paramètres Φ à partir d'une estimation initiale [13], $\Phi^{[0]}$, le vecteur de paramètres à la première itération (qui doit être une solution assez proche de la solution optimale). Pour ce faire l'algorithme ECI va chercher une relation du type $\hat{\Phi}^{[p+1]} = f(\hat{\Phi}^{[p]})$ et va définir le paramètre Φ optimal comme la solution du point fixe de cette relation [24].

En notant E_p l'espérance conditionnelle de $\hat{\Phi}^{[p]}$, l'ECI définit $\hat{\Phi}^{[p+1]}$ par cette

¹Estimation conditionnelle itérative

relation :

$$\begin{aligned}\hat{\Phi}^{[p+1]} &= E_p[\hat{\Phi}(X, Y)|Y = y] \\ &= \int \hat{\Phi}(X, Y)P(X|Y)dx\end{aligned}$$

Ce calcul est impossible en pratique car on ne peut pas calculer $P(X|Y)$. On peut cependant approcher $\hat{\Phi}^{[p+1]}$ par la loi des grands nombres, ce qui nous donne :

$$\hat{\Phi}^{[p+1]} = \frac{1}{n}[\hat{\Phi}(x_{(1)}, y) + \dots + \hat{\Phi}(x_{(n)}, y)]$$

.

Les $x_{(i)}$, $i = 1, \dots, n$ sont des tirages de X selon la distribution $P_{X|Y}(x|y, \hat{\Phi}^{[p]})$. Expérimentalement il est suffisant de prendre $n = 1$ pour avoir une estimation satisfaisante. L'étape réalisant les estimations de X selon $P_{X|Y}$ est réalisée à partir d'un échantillonneur de Gibbs.

Nous allons maintenant détailler d'un point de vu algorithmique les différentes étapes que l'on va réaliser :

- 1. Cette première estimation ($\Phi^{[0]}$) peut être donnée par une estimation au sens du MV de ces paramètres sur une segmentation donnée par l'algorithme des K-moyennes.
- 2. On assigne une première fois les classes aux voxels à l'aide des paramètres de $\Phi^{[0]}$, par une segmentation au maximum de vraisemblance.
- 3a. On balaye de droite à gauche et de haut en bas, tous les voxels s de l'image 3D. On calcule pour chaque voxel s $P_{X_s}(x_s)$, c'est à dire son appartenance (*à priori*) à chaque classe, soit $P_{X_s}(e_o)$, $P_{X_s}(e_1)$, $P_{X_s}(e_2)$ en regardant les classes associés aux voisins du voxel s obtenu à l'itération précédente (dessus, dessous, droite, gauche, devant et derrière), on en déduit la probabilité :

$$P_{X_s|X\eta_s}(x_s|\eta_s) = \frac{1}{Z}e^{-\sum_{<s,t>}(1-\delta(x_s,x_t))}$$

où x_t est une étiquette du voisinage du site, X_{η_s} désigne l'ensemble des variables aléatoires au voisinage du premier ordre du voxel s , η_s une configuration particulière de ce voisinage et Z est une constante de normalisation pour que $P_{X_s|X_{\eta_s}}$, soit une distribution.

- 3b. On calcule $P_{Y_s|X_s}(y_s|x_s)$ pour les trois classes et selon les paramètres de $\Phi^{[0]}$.
- 4. Vu que l'on a $P_{X_s}(x_s)$ et $P_{Y_s|X_s}(y_s|x_s)$, on peut calculer $P_{X_s|Y_s}(x_s|y_s) \propto P_{X_s,Y_s}(x_s, y_s)$, on aura donc pour chaque voxel y_s trois probabilités, d'appartenir à e_0 , e_1 et e_2 , dont la somme est normalisée à 1.
- 5. Pour chaque site s , la variable aléatoire X_s est une variable trinomiale (possédant trois probabilités différentes d'appartenance à chaque classe). Le but de l'échantillonneur de Gibbs est de tirer sur la base de ces trois probabilités d'appartenance, une des trois classes.
- 6. Il faut maintenant réestimer Φ , ce qui est fait à l'aide des estimateurs MV des données complètes. Donc dans notre cas il s'agit de faire des calculs de moyennes empiriques et de variances empiriques pour trouver les paramètres des deux gaussiennes et de faire également la moyenne des niveaux de gris pour trouver le paramètre α de la loi exponentielle.
- 7. Nous avons donc maintenant de nouveaux paramètres pour $\hat{\Phi}$ ce qui nous donne $\Phi^{[1]}$. Il faut recommencer la procédure à l'étape 3a jusqu'au critère d'arrêt :

$$\hat{\Phi}^{[p+1]} \approx \hat{\Phi}^{[p]}$$

4.3 Segmentation

Une fois que l'on a calculé les paramètres optimaux des trois lois avec l'algorithme ECI, nous pouvons réaliser la segmentation proprement dite.

Il nous faut estimer les classes (x_s) associées à chaque voxel s . Une segmentation bayésienne considère comme critère de minimisation de l'espérance d'une fonction de coût conditionnellement aux observations :

$$\hat{x} = \arg \min_x E[C(X, x)|Y = y]$$

où X est un vecteur aléatoire représentant la solution exacte et $C(X, x)$ représente le coût à payer si l'on prend la solution x au lieu de la solution exacte X .

Cette estimation peut se faire par plusieurs types d'estimateurs :

- L'estimateur du MAP², il a l'inconvénient de pénaliser de la même façon une erreur sur un seul site et une erreur sur plusieurs sites
- L'estimateur du MPM³, qui quant à lui, pénalise la configuration x au prorata du nombre de sites associés à de mauvaises classes.

Nous décrirons l'estimateur du MAP, car c'est celui que nous avons utilisé dans le cadre de cette application.

4.3.1 L'estimateur du MAP

La fonction de coût associée au MAP est :

$$C(X, x) = 1 - \delta_X(x)$$

Où $\delta_X(x)$ est la masse de Dirac en X . Donc l'équation $\hat{x} = \arg \min_x E[C(X, x)|Y = y]$ devient :

$$\begin{aligned} \hat{x}_{MAP} &= \arg \min_x \left(\sum_{x' \in E} C(X, x') P_{X|Y}(x'|y) \right) \\ &= \arg \min_x \left(1 - P_{X|Y}(x|y) \right) \\ &= \arg \max_x \left(P_{X|Y}(x|y) \right) \end{aligned}$$

Il faut donc choisir un estimateur qui maximise $P_{X|Y}(x|y)$. Nous avons déjà vu que l'on ne peut pas réaliser cette maximisation directement car l'espace des configurations E possibles serait trop grand. Dans notre cas cela reviendrait à calculer 3^{64^3} probabilités pour une image de 64x64x64 voxels avec trois classes possibles.

On peut appliquer la règle de Bayes ce qui donne :

²Maximum A Posteriori.

³Mode des Marginales a Posteriori.

$$\begin{aligned}\hat{x}_{MAP} &\propto \arg \max_x \left(P_{X,Y}(x, y) \right) \\ &= \arg \max_x \left(P_X(x) P_{Y|X}(y|x) \right)\end{aligned}$$

En notant $V(x, y) \triangleq -\ln(P_{Y|X}(y|x))$ et en vertu du theoreme d'Hammersley et Clifford, X est markovien et suit une distribution de Gibbs.

$$P_X(x) = \frac{1}{Z} \exp^{(-U(x))}$$

Ce qui donne :

$$\begin{aligned}\hat{x}_{MAP} &= \arg \max_x \left(\frac{1}{Z} \exp^{-(U(x)+V(x,y))} \right) \\ &= \arg \min_x \left(U(x) + V(x, y) \right)\end{aligned}$$

La recherche d'une solution approchée de ce minimum peut se faire soit à l'aide d'algorithmes déterministes (comme l'algorithme ICM⁴), soit à l'aide d'algorithmes stochastiques comme le Recuit simulé (RS). Nous avons implanté l'algorithme ICM car nous avons observé expérimentalement qu'il donne une bonne solution dans le cas de la segmentation 3D des images SPECT de cerveaux [13] et un gain de temps appréciable comparativement au RS.

⁴Iterative Conditionnal Mode

4.3.2 Segmentation à l'aide de l'algorithme ICM

Cet algorithme est adapté si l'on a une segmentation initiale (comme celle donnée par une segmentation du Maximum de Vraisemblance (MV)) proche de la configuration optimale [7].

On peut le décrire de la façon suivante :

- 1. Il faut partir de $\hat{x}^{[0]}$ donné par une segmentation au sens du MV basée sur les paramètres $\hat{\Phi}$ estimés de l'algorithme ECI, i.e.,

$$\hat{x}_s = \arg \max_{x_s \in \{e_0, e_1, e_2\}} \{P_{Y_s|X_s}(y_s|x_s)\}$$

- 2. On calcule ensuite $\hat{x}^{[k+1]}$ à partir de $\hat{x}^{[k]}$:
 - 2a. Pour chaque site s , on va calculer $P_{X_s|X_{\eta_s}, Y_s}(x_s|\eta_s, y_s)$ pour nos trois classes possibles.

$$P_{X_s|X_{\eta_s}, Y_s}(x_s|\eta_s, y_s) = \frac{1}{Z_s} \exp \left(- \left(\sum_{c \in C} V_c(x_s) + V(x_s, y_s) \right) \right)$$

Où $\sum_{c \in C} V_c(x_s)$ désigne la somme des potentiels locaux sur les cliques $c \in C$. Le premier terme calcule la probabilité d'appartenance du voxel à chaque classes par rapport au voisinage. Le second terme calcule la probabilité d'appartenance du voxel à chaque classe au sens du MV.

- 2b. On va sélectionner la classe x_s qui donne la probabilité maximale et la réassigner au voxel :

$$\hat{x}_s = \arg \max_{x_s \in \{e_0, e_1, e_2\}} \{P_{X_s|X_{\eta_s}, Y_s}(x_s|\eta_s, y_s)\}$$

- Nous avons donc maintenant un nouveau champ d'étiquettes pour $\hat{x}$ ce qui nous donne $\hat{x}^{[1]}$. Il faut recommencer la procédure à l'étape 2 jusqu'au critère d'arrêt :

$$\hat{x}_s^{[p+1]} \approx \hat{x}_s^{[p]}$$

4.4 Applications

Les applications de cette segmentation sont principalement liées au prétraitement des données, on peut citer quelques exemples :

- Pour la restauration d'images SPECT, elle constitue une étape nécessaire à la différenciation de zones utiles dans le cerveau, cela permet en effet de différencier la matière blanche et grise du CSF et du bruit du fond de l'image. Et ceci est intéressant par exemple pour la restauration de l'image par déconvolution [13].
- Pour la classification d'images SPECT, elle permet d'extraire des attributs plus caractéristiques des zones de matières blanche et grise, ce qui ultérieurement permet une meilleure classification des données.

Plus généralement la segmentation markovienne peut servir pour :

- La simplification d'images.
- L'extraction de formes.

4.5 Résultats

La figure 4.2 montre le résultat final de la segmentation, l'image originale ainsi que l'histogramme associé à cette image. Sur l'image segmentée on peut voir trois zones :

- en noir, l'arrière plan et le CSF.
- en gris, la matière blanche.
- en blanc, la matière grise.

Il est à noter que compte tenu de la faible résolution des images SPECT, certaines régions internes de CSF n'apparaissent pas. Toutefois, nous pensons que cela aura un impact limité puisque les attributs significatifs se retrouveront dans la matière grise (en blanc dans l'image segmentée).

Les images étant segmentées, il faut maintenant extraire des attributs qui pourraient permettre une bonne classification des données.

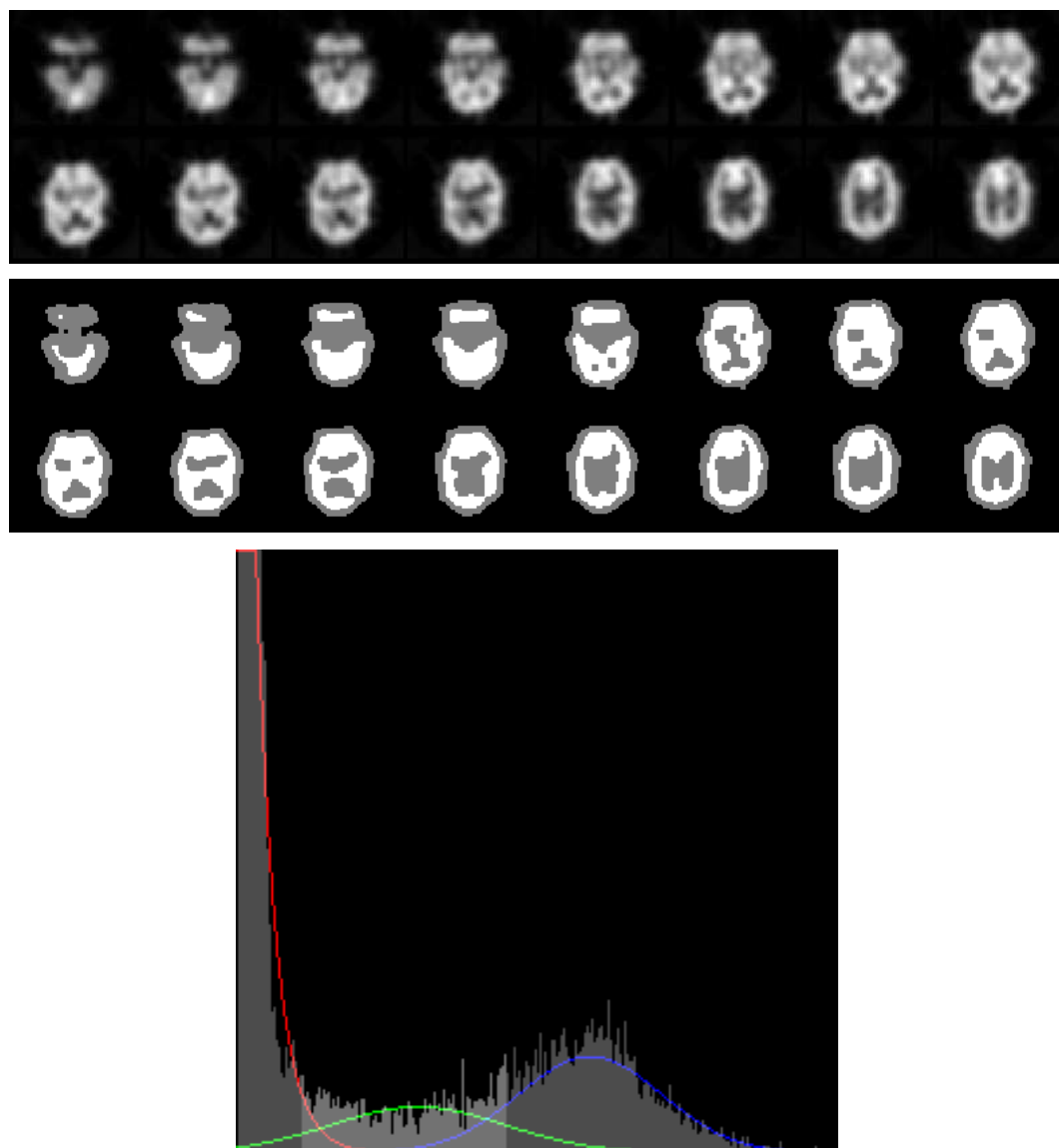


FIG. 4.2 – Résultat de la segmentation, en haut, l'image originale, au milieu l'image segmentée et en bas, l'histogramme des niveaux de gris avec les deux lois gaussiennes et la loi exponentielle

4.5.1 Attributs testés

Les attributs obtenus dans chacune des trois zones (le liquide cephallo-rachidien qui est la zone 1, la matiere blanche zone 2 et la matiere grise qui est la zone 3) sont :

- μ_1, μ_2 et μ_3 , les moyennes de l'intensite des niveaux de gris dans chaque zone. (att1, att2, att3)
- P_1, P_2 et P_3 , les proportions du nombre de pixels dans chaque zone. (att4, att5, att6)
- G_{X_1}, G_{X_2} et G_{X_3} , les gradients en X , dans chaque zone. (att7, att10, att13)
- G_{Y_1}, G_{Y_2} et G_{Y_3} , les gradients en Y , dans chaque zone. (att8, att11, att14)
- G_{Z_1}, G_{Z_2} et G_{Z_3} , les gradients en Z , dans chaque zone. (att9, att12, att15)
- G_1, G_2 et G_3 , les gradients dans chaque zone. (att16, att17, att18)
- $\frac{G_{X_2}}{G_{X_1}}, \frac{G_{X_2}}{G_{X_1}}$. (att41, att42)
- $\frac{G_{Y_2}}{G_{Y_1}}, \frac{G_{Y_3}}{G_{Y_1}}$. (att43, att44)
- $\frac{G_{Z_2}}{G_{Z_1}}, \frac{G_{Z_3}}{G_{Z_1}}$. (att45, att46)
- $\frac{G_2}{G_1}, \frac{G_2}{G_1}$. (att39, att40)
- σ_1, σ_2 et σ_3 , les écart-types dans chaque zone. (att36, att37, att38)
- σ_1^2, σ_2^2 et σ_3^2 , les variances dans chaque zone. (att19, att20, att21)
- $\frac{\mu_2}{\mu_1}, \frac{\mu_3}{\mu_1}, \frac{\mu_2}{\mu_3}$. (att22, att23, att49)
- $\frac{\sigma_1}{\mu_1}, \frac{\sigma_2}{\mu_2}, \frac{\sigma_3}{\mu_3}$. (att24, att25, att26)
- $\frac{\sigma_2^2}{\sigma_1^2}, \frac{\sigma_3^2}{\sigma_1^2}$. (att47, att48)
- $\sigma_{G_{X_1}}^2, \sigma_{G_{X_2}}^2$ et $\sigma_{G_{X_3}}^2$, les variances des gradients en X , dans chaque zone. (att27, att30, att33)
- $\sigma_{G_{Y_1}}^2, \sigma_{G_{Y_2}}^2$ et $\sigma_{G_{Y_3}}^2$, les variances des gradients en Y , dans chaque zone. (att28, att31, att34)
- $\sigma_{G_{Z_1}}^2, \sigma_{G_{Z_2}}^2$ et $\sigma_{G_{Z_3}}^2$, les variances des gradients en Z , dans chaque zone. (att29, att32, att35)

Tous ces attributs vont être testés avec les mêmes tests que ceux présentés au paragraphe 3.5, nous commencerons tout d'abord par la classification des cerveaux des classes Diffus et Normaux.

4.5.2 Diffus-Normaux

La matière grise ou cortex est le siège des fonctions cognitives, et sa dégradation comme par exemple, la maladie d'Alzheimer qui est une démence neurodégénérative à prédominance corticale, il n'est donc pas étonnant que cette zone donne les meilleurs attributs.

Avec l'algorithme des K-moyennes, nous obtenons un taux de classification correcte de 84.8 %, ce qui est un très bon taux pour une classification ne se basant que sur un seul attribut. Cet attribut (att26) est $\frac{\sigma_3}{\mu_3}$, c'est à dire l'écart-type sur la moyenne dans la zone de matière grise. Ces résultats sont présentés au tableau 4.1

La séparabilité des données est confirmée par les tests statistiques qui nous informent que les deux classes ont pour cet attribut des variances similaires (F-test) et des moyennes significativement différentes.

D'autres attributs provenant de cette région donnent des taux supérieurs à 80 %, comme la variance (att21), l'écart-type (att38) ou la variance du gradient en Z (att35) (d'avant en arrière).

4.5.3 Autres Résultats

4.5.3.1 Diffus-Négatifs

Nous avons essayé également de classer d'autres populations entre elles. Les meilleurs résultats ont été obtenus lors de la classification entre la classe Diffus et la classe Négatifs. (cf tableau 4.2). Le taux de classification correcte avec certains attributs (att21, att26, att32, att35, att38) sont proches de 80%.

Les attributs 21 26 35 38 donnent encore une fois de meilleurs résultats pour la différenciation de cerveaux atteints de maladie diffuse.

attributs	F-test	t-test	K-moyenne	EM
att2	3.349	2.289	0.649	0.679
att5	1.145	3.286	0.738	0.738
att6	2.214	1.216	0.631	0.658
att11	3.465	2.252	0.649	0.726
att12	3.270	2.351	0.673	0.679
att17	1.448	3.129	0.676	0.595
att21	1.173	5.357	0.807	0.807
att23	2.573	3.580	0.637	0.616
att25	2.959	2.953	0.673	0.682
att26	1.958	4.442	0.848	0.705
att30	1.053	2.885	0.637	0.631
att31	1.273	1.458	0.622	0.640
att32	1.146	5.056	0.807	0.807
att33	1.053	2.885	0.637	0.631
att35	1.146	5.056	0.807	0.807
att37	1.208	1.626	0.622	0.682
att38	2.021	4.845	0.801	0.771
att40	2.837	3.045	0.679	0.557
att49	3.411	3.297	0.717	0.679

Tableau 4.1 – Résultats de classification entre la classe Normaux et la classe Diffus, avec 43 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.48, 2.03]$ pour être significatif et un résultat au test t de Student > 2.0167 ou < -2.0167 pour être révélateur

attributs	F-test	t-test	K-moyenne	EM
att10	6.029	1.334	0.645	0.625
att13	2.775	1.855	0.645	0.729
att21	1.985	3.562	0.791	0.812
att22	1.037	1.712	0.666	0.666
att26	2.253	2.981	0.75	0.791
att32	2.085	3.562	0.812	0.791
att35	2.385	3.562	0.812	0.812
att38	2.260	3.332	0.791	0.666 1
att40	0.805	1.321	0.625	0.625

Tableau 4.2 – Résultats de classification entre la classe Négatif et la classe Diffus, avec 34 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.383, 2.21]$ pour être significatif et un résultat au test t de Student > 2.0323 ou < -2.0323 pour être révélateur

4.5.3.2 Focaux-Négatifs et Focaux-Normaux

Pour ce qui est de la classe Focaux, elle a été relativement mal séparée (avec un taux inférieur à 80 %) des classes Normaux et Négatifs. (cf tableau 4.4 et tableau 4.3) Le taux maximum de classification correcte a été de 77.9%, avec l'attribut 17, qui est le gradient dans la matière blanche. Les meilleurs attributs pour cette classification ont été majoritairement extraits de la matière blanche.

4.5.3.3 Focaux-Diffus

Le tableau 4.5 présente les résultats de tentatives de classification entre les cerveaux de la classe Diffus et Négatifs, le taux de 66.6% a été difficilement atteint. La différenciation de ces 2 classes ne révèle pas de grand intérêt pour ce mémoire.

attributs	F-test	t-test	K-moyenne	EM
att5	0.649	3.307	0.747	0.604
att10	0.994	1.570	0.669	0.616
att17	0.410	3.002	0.732	0.779
att22	0.963	2.965	0.717	0.693
att23	1.269	3.024	0.678	0.654
att25	0.769	2.124	0.660	0.639
att26	0.313	2.206	0.660	0.598
att30	0.228	2.698	0.699	0.651
att33	0.228	2.698	0.699	0.651
att38	0.615	2.171	0.708	0.592
att40	1.278	2.727	0.678	0.678
att49	0.737	2.223	0.654	0.592

Tableau 4.3 – Résultats de classification entre la classe Normaux et la classe Focaux, avec 29 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.290, 2.51]$ pour être significatif et un résultat au test t de Student > 2.0518 ou < -2.0518 pour être révélateur

attributs	F-test	t-test	K-moyenne	EM
att2	0.101	0.456	0.604	0.562
att3	0.216	0.331	0.645	0.562
att6	0.627	0.448	0.604	0.5
att14	0.227	0.436	0.645	0.645
att15	0.233	0.332	0.645	0.562
att17	0.455	0.337	0.666	0.625

Tableau 4.4 – Résultats de classification entre la classe Négatifs et la classe Focaux, avec 20 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.278, 3.01]$ pour être significatif et un résultat au test t de Student > 2.1009 ou < -2.1009 pour être révélateur

attributs	F-test	t-test	K-moyenne	EM
att6	0.288	0.054	0.625	0.583
att15	0.095	0.331	0.666	0.604
att17	0.283	0.610	0.625	0.625
att18	0.378	0.270	0.625	0.583
att21	0.184	1.847	0.625	0.645
att26	0.093	1.543	0.645	0.520
att29	0.184	1.847	0.625	0.645
att32	0.184	1.847	0.625	0.645
att35	0.184	1.847	0.625	0.645
att38	0.234	1.488	0.645	0.645

Tableau 4.5 – Résultats de classification entre la classe Diffus et la classe Focaux, avec 32 degrés de liberté, avec un $\alpha=0.05$, il faut que le résultat au F-test $\in [0.293, 2.42]$ pour être significatif et un résultat au test t de Student > 2.0423 ou < -2.0423 pour être révélateur

Conclusion

Notre objectif était d'aider le médecin dans sa décision pour les cas où, la simple observation des images reconstruites ne suffisait pas pour conclure à un diagnostic. Les taux d'erreurs obtenus avec la première méthode (comparaison gauche droite) sont supérieurs à 30%, ce qui ne suffit pas pour une application clinique et cela dans tous les cas considérés.

L'utilisation de la segmentation Markovienne a permis l'obtention d'un meilleur masque, qui a permis une extraction des attributs plus spécifiques à la matière blanche et grise.

On peut s'étonner de ne pas voir apparaître les ventricules remplis de CSF dans les tranches centrales des coupes du cerveau. Ceci est simplement dû au fait que la résolution des images SPECT sur lesquelles nous avons travaillé n'est que de 4.2x4.2x4.2 millimètres, ce qui est tout juste suffisant pour les entrevoir sur les images originales. Ils sont donc éclipsés par le voisinage de ceux-ci.

Il faut faire preuve de prudence dans l'interprétation des résultats, et ce pour deux raisons majeures. En premier lieu, le médecin lui-même fait environ 20% d'erreur en classifiant à l'oeil les données, donc les classes associées aux données que nous avons peuvent être faussées. En second point, nous avons extrait un vaste éventail d'attributs et ce, sur un nombre peu important de données. Il faut être conscient qu'à la limite, on pourra toujours trouver un attribut qui classera les données avec un taux d'erreur satisfaisant, même s'il s'agit par exemple de la couleur des cheveux du patient !

Nous avons pris un seul attribut pour faire la classification des données, ce qui est un facteur limitant.

Cependant les taux de classification obtenus avec la deuxième méthode sont supérieurs à 80% dans le cas qui nous intéressait le plus, c'est à dire la différenciation des cerveaux atteints de maladie diffuse et des cerveaux sains.

La classification des autres couples de classe (Diffus-Focaux, Négatifs-Diffus etc.) a été faite à titre expérimental, car les autres classes ont un nombre de données encore plus restreint.

Les meilleurs attributs sont extraits de la matière grise du cerveau, siège des fonctions cognitives, cela nous rassure et dénote la plausibilité de ces attributs.

Cependant nous ne voyons pas d'explication au fait que le gradient en Z soit un bon attribut, car alors pourquoi les gradients en X et Y ne donnent pas de bons résultats ?

Les mauvais résultats de l'algorithme EM peuvent provenir du faible nombre de données qui induit une mauvaise reconnaissance des paramètres des lois Gaussiennes des ensembles de données à classifier.

L'emploi d'un vecteur d'attributs composé de plusieurs paramètres, (qui peut être réduit, par analyse factorielle discriminante ou réduction de Fisher pour éliminer l'information redondante) devrait permettre une classification plus aisée et nous envisageons de l'utiliser aussi dans des travaux futurs.

L'augmentation du nombre et la qualité des données (avec une acquisition qui n'est pas faite de routine en clinique mais faite selon une procédure standardisée) est essentielle si l'on veut que les analyses statistiques et la classification avec les différents algorithmes sur les résultats soit fiable.

Une évolution possible de cette technique serait de segmenter les images SPECT du cerveau selon un plus grand nombre de régions anatomiques à l'aide d'un atlas de cerveaux de référence comme celui de Talairach ou de Greitz [21], ce qui a été fait par une équipe suédoise pour détecter des patients atteints de la maladie d'Alzheimer ou de la démence du lobe frontal [23]

Bibliographie

- [1] BRINTON WOLBARST A. Physics of radiology. chapter 12 :pp 103–106.
- [2] BRINTON WOLBARST A. Physics of radiology. chapter 37-38-39 :pp 320–345.
- [3] DEMPSTER A.P., LAIRD N.M., and RUBIN D.B. Maximum likelihood from incomplete data via the em algorithm. *Royal Statistical Society*, pages 1–38, 1976.
- [4] BRONZINO J. D. *The Biomedical Engineering Handbook*. IEEE Press, 1995.
- [5] COSTA D.C. and ELL P.J. Brain blood flow in neurology and psychiatr. *Chur-chill Livingstone*, 1991.
- [6] SMITH G.M. Image texture analysis using zero crossing information. *Thèse de doctorat, University of Queensland*.
- [7] BESAG J. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, B48 :259–302, 1986.
- [8] FOOS J. Manuel de radioactivité tome ii : La désintégration radioactive - les interactions rayonnements matière -les applications de la radioactivité.
- [9] POSTAIRE J.G. De l'image à la décision. *Bordas*, 1987.
- [10] WESZKA J.S., DYER C.R., and ROSENFELD A. A comparative study of texture measures for terrain classification. *IEEE Transactions on System, Man and Cybernetics*, 6(4) :pp 269–285, 1976.
- [11] STERHEIM K. Physique. *INTEREDITIONS*.
- [12] MIGNOTTE M. Unsupervised hierarchical markovian segmentation of sonar images and classification of shadow shapes using statistical model. *Thèse de doctorat, Academie Navale de Brest*, 1998.

- [13] MIGNOTTE M. and MEUNIER J. Three-dimensional blind deconvolution of spect. *IEEE Transactions on Biomedical Engineering*, 47(2) :pp 274–281, February 2000.
- [14] REUZE P. Analyse d’images de résonance magnétique : Application à la segmentation des vaisseaux et à la caractérisation des textures. *Thèse de doctorat, Université de Rennes 1 ; FRANCE*, 1995.
- [15] WOODS R. P., CHERRY S. R., and MAZZIOTTA J.C. Rapid automated algorithm for aligning and reslicing pet images. *Journal of Computer Assisted Tomography*, 16 :620–633, 1992.
- [16] GONZALEZ R.C. and WOODS R.E. Digital image processing. 1992.
- [17] JASZCZAK R.J., COLEMAN R.E., and LIM C.B. Spect :single photon emission computed tomography. *IEEE Transactions on Nuclear Science*, 27 :pp.1137–1153, 1980.
- [18] LARSEN R.J. and MARX M.L. An introduction to mathematical statistics and its applications. *Prentice-Hall*, 1986.
- [19] HARALICK R.M., SHANMUGAM K., and DINSTEN I. Textural features for image classification. *IEEE Transactions on System, Man and Cybernetics*, 3(6) :pp 610–621, 1973.
- [20] site web. <http://www.fil.ion.ucl.ac.uk/spm/>.
- [21] GREITZ T., BOHM C., HOLTE S., and ERIKSSON L. A computerized brain atlas : contruction, anatomical, content and some applications. *J Comput Assist tomogr*, 15 :26–38, 1991.
- [22] CURRY T.S.III, DOWDEY J.E., and MURRY R.C.Jr. Christensen’s physics of radiology. chapter 19 :pp 289–322.
- [23] KOVALEV V., THURFEJELL L., LUNDQVIST R., and PAGANI M. Classififcation of spect scans of alzheimer’s disease and frontal lobe dementia based on intensity and gradient information.
- [24] PIECZYNSKI W. Champs de markov cachés et estimation conditionnelle itérative. *Traitement du Signal*, 11(2) :141–153, 1994.

- [25] PRATT W.K. Digital image processing. *Wiley-Interscience*, 1978.
- [26] CROFT B. Y. Spect :single photon emission computed tomography. *Yearbook Medical Publishers*, 1986.
- [27] GOUSSARD Y. cours gbm-6103.