

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mouloud MAMMERI, Tizi-Ouzou



Faculté de Génie Electrique et d'Informatique
Département d'Automatique

Mémoire de Fin d'Etudes

En vue de l'obtention du diplôme

D'Ingénieur d'Etat en Automatique

Thème

*Segmentation d'images
médicales basée sur la
classification floue*

Proposé et dirigé par : **M^r K. HAMMOUCHE**

Présenté par :

M^{elle} DJOUDREZ Kamélia

M^{elle} LARFI Hassiba

Soutenu le : **04 / 10 /2010**

Promotion 2010

Résumé

Le travail présenté dans cette thèse concerne le domaine du traitement d'images et plus précisément celui de la segmentation.

Parmi les nombreuses techniques de segmentation, nous nous sommes intéressés particulièrement à la méthode de classification floue qui est basée sur l'algorithme Fuzzy –C-Means. Cette méthode est formalisée comme un problème de minimisation d'un critère de classification sous certaines contraintes. Son inconvénient principal réside dans sa sensibilité au bruit. Nous avons alors étudié certaines variantes proposées dans la littérature pour remédier à ce problème. Celles-ci tentent soit à changer la fonctionnelle à minimiser, soit à définir une autre métrique ou soit encore à modifier l'influence du facteur flou m . Parmi toutes ces méthodes, nous avons implémenté 5 d'entre elles à savoir les algorithmes PCM, de Ahmed, de Wu, de Tang et celui de Li.

Mots clés

Images médicale, segmentation d'images, classification, la logique floue, Fuzzy-C-Means, variantes de Fuzzy-C-Means.

Remerciement

Tout d'abord, nous présentons notre remerciement les plus sincères à notre promoteur M^r HAMMOUCHE , pour sa disponibilité et son encadrement, ainsi pour son soutien tout au long de l'année. C'est grâce à lui que nous avons pu mener à bien ce travail.

Nous tenons ensuite à remercier chaque'un des membres de jury pour nous avoir fait l'insigne honneur d'accepter d'examiner notre travail.

Nos remerciements et notre gratitude vont aux professeurs et enseignants de l'UMMTO ainsi que son personnel côtoyés tout au long de notre cursus universitaire.

Nous remercions également toute personne ayant contribué à notre éducation et notre formation.

Sommaire

Introduction générale	01
------------------------------------	-----------

Chapitre 1 : Imagerie médicale

1	Introduction.....	03
2	Rayon X.....	03
2.1	Production de rayons X	03
2.2	Les rayons X et la matière.....	04
2.3	Différent types d'examens utilisant les rayons X.....	05
2.3.1	Radiographie	05
2.3.1.1	Fonctionnement de la radiographie.....	05
2.3.2	Le scanner.....	07
2.3.2.1	Principe du scanner.....	08
2.3.2.2	Fonctionnement de scanner	08
3	Les ultrasons.....	10
3.1	L'échographie.....	11
3.2	Le Doppler.....	13
4	Les rayons gamma.....	14
4.1	La scintigraphie.....	14
4.1.1	Principe de fonctionnement.....	14
4.2	Tomographie par émission de positions (TEP).....	16
5	Imagerie par résonance magnétique.....	17
5.1	Principe.....	18
5.2	Application de l'IRM.....	20
6	Endoscopie.....	20
7	Conclusion.....	22

Chapitre 2 : Segmentation d'images

1	Introduction.....	23
2	Définition de la segmentation.....	23

3	Définition de la texture.....	25
4	Paramètres extrait de la matrice de cooccurrence.....	26
5	Méthodes de la segmentation d'images.....	27
5.1	Segmentation par contour.....	27
5.2	Segmentation par région.....	28
5.2.1	Segmentation par croissance de région.....	28
5.2.2	Approche par division-fusion.....	29
5.2.3	Méthodes de classification.....	30
5.2.3.1	Méthodes supervisées.....	31
5.2.3.2	Méthodes non supervisées.....	31
5.2.3.2.1	Algorithme EM (Estimation Maximisation).....	31
5.2.3.2.2	Algorithme Mean Shift.....	32
5.2.3.2.3	K-Means.....	32
6	Intérêt de la segmentation dans le domaine médical.....	34
7	Conclusion.....	35

Chapitre 3 : Notions sur la logique floue

1	Introduction.....	36
2	La théorie des sous ensembles flous.....	36
2.1	Définition d'un sous ensemble flou.....	36
2.2	Caractéristique d'un sous ensemble flou.....	37
2.2.1	Noyau.....	37
2.2.2	Support.....	37
2.2.3	Hauteur.....	38
2.2.4	Coupe de niveau ou α coupe.....	38
2.2.5	Cardinalité.....	39
2.2.6	Sous ensemble flou convexe.....	39
2.3	Opérations sur les sous ensembles flous.....	39
2.3.1	Egalité.....	39
2.3.2	Inclusion.....	39
2.3.3	Intersection.....	39
2.3.4	Union.....	40
2.3.5	Complément d'un sous ensemble.....	41

2.4 Exemples de fonction d'appartenance	41
2.4.1 Fonction d'appartenance trapézoïde.....	41
2.4.2 Fonction d'appartenance exponentielle.....	42
2.5 Indices de comparaison.....	43
2.5.1 Indice de chevauchement.....	43
2.5.2 Indice d'inclusion.....	43
2.5.3 Indice d'égalité.....	43
2.5.4 Indice flou.....	43
2.6 Le principe d'extension.....	43
2.7 Produit cartésien.....	44
2.8 Projection.....	44
2.9 Schéma d'un traitement flou.....	44
2.9.1 Fuzzification.....	45
2.9.2 Traitement flou.....	45
2.9.3 Défuzzification.....	46
3 Conclusion.....	46

Chapitre 4 : Algorithme FCM et ses variantes

1 Introduction	47
2 FCM (Fuzzy C Means)	47
3 Variantes de la méthode FCM.....	50
3.1 Variantes de la distance.....	50
3.1.1 Algorithme de Gustafson et Kessel.....	50
3.1.2 Algorithme de Gath et Geva.....	52
3.1.3 Algorithme Shen et al.....	53
3.1.4 Algorithme Wang et al.....	53
3.2 Variantes de fuzzification.....	54
3.2.1 Rousseeuw et al.....	54
3.2.2 Klawonn et Hoppner.....	55
3.3 Variantes de la fonction objective.....	56
3.3.1 Algorithme et al Frigui et Krishnapuran.....	56
3.3.2 Algorithme de Pham.....	57
3.3.3 Possibilitic-C-Means.....	58

3.3.4	Algorithme de Wu et al.....	59
3.3.5	Algorithme d’Ahmed et al.....	60
3.3.6	Algorithme de Kang et al.....	61
3.3.7	Algorithme de Li and Mukaidono.....	62
3.3.8	Algorithme de Yang et al.....	63
3.3.9	Algorithme de Tang et al.....	64
4	Conclusion.....	65

Chapitre 5 : Tests et résultats

1	Introduction.....	66
2	Méthodes de segmentation.....	66
2.1	Influence des paramètres de la méthode.....	67
3	Segmentation d’images médicale.....	73
4	Temps de calcul.....	76
5	Conclusion.....	77
	Conclusion générale.....	78

Introduction générale

Le traitement des images constitue actuellement l'une des grandes orientations du traitement de l'information acquise à partir des capteurs. Il s'agit donc d'un domaine très vaste qui trouve de plus en plus d'application, notamment dans la reconnaissance d'objets, l'image satellitaire et l'imagerie médicale.

L'identification et la reconnaissance des éléments d'une scène ou d'une image donnée requièrent une extraction préalable des différents objets composant cette dernière. De ce besoin est née la segmentation d'image qui constitue une étape cruciale de l'analyse d'image. Son objectif consiste à décomposer des scènes plus complexes en des éléments individuellement identifiables ou plus aisément traitables que l'image entière.

L'information apportée par l'imagerie médicale est d'un apport considérable en matière de diagnostic. Il s'agit d'une image à très haute résolution issue d'un signal non invasif qui reproduit l'anatomie avec une grande fidélité, et qu'on peut facilement interpréter. La segmentation d'image permet d'apprécier les formes et de les interpréter de la manière la plus juste possible pour aboutir à une plus grande information de l'image sans altérer sa fidélité.

Les informations fournies par les images sont incertaines (bruit, rayons cosmiques, blocs manquants) et imprécises (transitions entre régions). Nous gérons ces ambiguïtés en utilisant une approche floue, et plus particulièrement en modélisant l'ensemble des informations dans le cadre de la logique possibiliste, introduite par Zadeh à partir de ses travaux princeps sur la logique floue.

De nombreux algorithmes de segmentation d'images ont été proposés et ces techniques de segmentation ne sont efficaces que pour un type d'image donné, pour un type d'application donné, et dans un contexte informatique donné. En raison de ces contraintes, les diverses stratégies de segmentation qui ont été proposées ont affirmé leurs insuffisances et leur limitations, il est donc tout à fait normal d'exploiter de nombreux horizons et trouver de nouvelles méthodes plus souples et plus efficaces.

Dans ce mémoire, nous nous sommes intéressés à la segmentation d'images par une méthode de classification basée sur la notion de la logique floue qui est Fuzzy-c means. Pour cela nous avons structuré notre travail en cinq chapitres.

Le premier chapitre est consacré à l'étude de l'image médicale qui regroupe les moyens d'acquisitions et de restitution d'images à partir de différents phénomènes physiques ou chimiques (rayons x, gamma, ultrasons...etc.)

Le second chapitre aborde des définitions et quelques méthodes de segmentation.

Le troisième chapitre décrit la notion de la logique floue.

Le quatrième chapitre est consacré à l'étude de l'algorithme FCM et ses différentes variantes.

Le dernier chapitre est dédié aux tests et aux résultats expérimentaux obtenus par les méthodes implémentées.

En fin, nous terminerons par une conclusion générale.

Chapitre 1 : Imagerie médicale

1.1 Introduction

L'imagerie médicale regroupe un ensemble de méthodes permettant de visualiser des processus biologiques au sein même des organismes vivants sans procéder à une opération. Elle est essentielle à la compréhension de leur physiologie et de leurs pathologies afin de mieux les diagnostiquer, les pronostiquer, et les soigner. L'imagerie médicale regroupe les moyens d'acquisition et de restitution d'images à partir de différents phénomènes physiques ou chimiques (rayons x, gamma, ultrasons...etc.) [Web1], [Web2], [Web3].

1.2 Rayon x

Les Rayon x ont été découverts en 1895 par le physicien allemand **Wilhem Conrad Rontgen**. Ils ont été nommés ainsi car ils étaient d'une nature inconnue. Les rayons X permettent au médecin de faire un diagnostic non seulement plus rapide, mais aussi plus sûr. Les principales propriétés des rayons x sont : l'invisibilité, la propagation rectiligne, la pénétration différente selon la matière et l'absorption variable selon l'épaisseur du corps traversé. Les rayons X sont des radiations électromagnétiques dont la longueur d'onde est comprise entre $\lambda = 0.03\text{nm}$ et $\lambda = 10\text{nm}$ (1nm vaut 10^{-9}m).

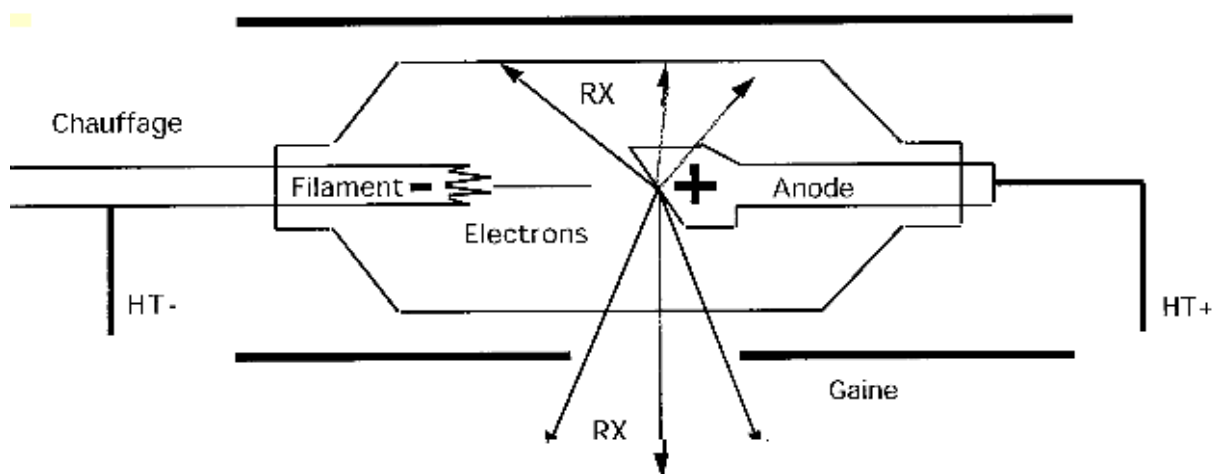
L'ensemble des radiations électromagnétiques a fondamentalement une nature commune : ce qui les distingue dans leur interaction avec la matière est le paramètre qui caractérise chacune : la longueur d'onde λ ou l'énergie du photon E . Un photon est une particule de masse nulle, non chargée, se déplaçant à la vitesse de la lumière et transportant une quantité d'énergie E déterminée. A un rayon X de fréquence f (donc de longueur d'onde dans le vide $\lambda = c/f$ donc $f = c/\lambda$), on associe un photon d'énergie :

$$E = h \cdot f = h \cdot c/\lambda$$

h correspond à la constante universelle de Planck ($h = 6,626 \cdot 10^{-34} \text{ J.s}$), λ est en mètre, E en joule, f en Hz et c en m/s. Plus l'épaisseur et la densité du corps traversé sont importantes, plus le rayon est atténué.

1.2.1 Production de rayons x

La production des rayons X nécessite l'utilisation de tensions très élevées, de l'ordre de 40000 volts. Cela impose l'utilisation d'appareils électriques (Fig 1.1), appelés générateurs. Un générateur est donc relié à une ampoule de verre vide et des électrons sont émis par un filament dans lequel circule un courant électrique. Ceux-ci sont accélérés en direction d'une anode (ou anticathode) par un champ électrique créé par une différence de potentiel élevée (généralement de 10 à 150 kV) entre le filament qui sert de cathode et l'anode. Ces électrons entrent en collision avec la cible que constitue le métal de l'anode.



HT : Haute tension

RX : Rayons X émis par l'anode

Fig 1.1:Schéma d'un tube à rayon x

1.2.2 Les rayons X et la matière

Lorsqu'un faisceau de rayons X pénètre dans un milieu matériel, on constate une diminution progressive de son intensité. Cette diminution du nombre de photons, l'atténuation du faisceau, est due essentiellement à l'interaction des photons avec les électrons. Dans un tel processus, l'énergie perdue se retrouve sous deux formes :

- Une partie est absorbée par le milieu.
- Une partie est diffusée et sort de la matière dans une direction différente de la direction du faisceau initial.

Ce sont ces phénomènes d'atténuation et d'absorption qui sont à l'origine des applications des rayons X dans le domaine médical.

Plus la longueur d'onde λ du rayonnement est faible (donc que la fréquence est grande) plus le rayonnement est pénétrant : rayonnement dur contrairement au rayonnement mou, de grande longueur d'onde λ (donc de fréquence faible), qui est très fortement absorbé.

Pour un faisceau de rayons X de longueur d'onde λ des atomes lourds seront plus absorbants que des atomes légers (qui ont un petit numéro atomique Z donc un petit nombre d'électrons.) Par exemple, l'absorption par l'air ou par les chairs du corps humain est faible tandis que l'absorption par les os est plus importante.

1.2.3 Différent types d'examens utilisant les rayons x

1.2.3.1 Radiographie

La radiologie, découverte voici 100 ans par **Wilhem Conrad RONTGEN** garde actuellement un intérêt diagnostique de premier plan dans beaucoup de domaines de la médecine. Initialement, une radiographie était simplement un film radiographique sur lequel on pouvait observer des contrastes qui correspondaient aux différentes structures internes de l'organisme radiographié. Différentes technologies comme l'informatique ou les produits de contrastes ont ensuite donné naissance à plusieurs types de radiographies. Pour des questions pratiques, les appareils utilisés diffèrent légèrement selon leur zone d'application (dentition, poumons, squelette...). Ceci explique que la radiologie est utilisée pour le système osseux car il s'agit du système le plus visible sur une radiographie du corps et elle est utilisée aussi bien en chirurgie dentaire qu'en médecine ou en kinésithérapie.

1.2.3.1.1 Fonctionnement de la radiographie

Le principe de la radiologie est une impression photographique par les rayons X sur un film argentique (plus précisément une cassette contenant un film sensible à la lumière visible et recouverte généralement de un ou deux écrans renforçateurs qui s'illuminent lorsqu'ils sont exposés aux rayons X) (Fig 1.2). La quantité de rayons X reçue dans les différentes zones du film dépendant de l'absorption des rayons X par les différents tissus traversés soit leur densité.

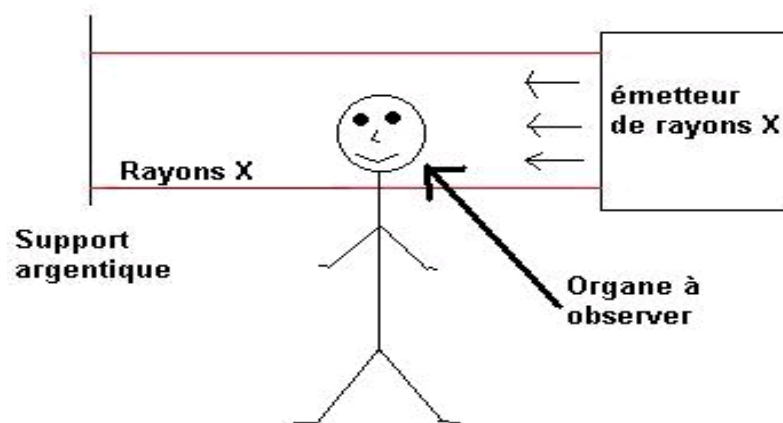


Fig 1.2 : Principe de fonctionnement de la radiographie

Donc les os très denses apparaîtront en clair, tandis que les organes tels que les poumons seront visibles sur la radiographie d'une couleur beaucoup plus sombre.

La figure (1.3) montre un exemple d'image radiographique.



Fig 1.3 : Une radiographie classique d'un crâne humain

Notons que la radiologie dite de contraste se différencie de la radiologie standard par un aspect technique : avant la radiographie, on effectue une injection (intraveineuse, rectale, orale ou intra articulaire) d'un produit de contraste radio opaque, c'est à dire une substance qui absorbe fortement les rayons X. Les milieux qui contiennent cette substance apparaîtront donc sur la radiographie en teinte très claire, ce qui augmentera fortement les contrastes entre les différents organes et les différents milieux qui contiendront ou non les éléments radio opaques injectés. La

radiographie obtenue sera alors plus contrastée, elle paraîtra plus nette, plus fine, et sera donc plus facilement interprétable. Le produit de contraste le plus souvent utilisé est l'iode.

1.2.3.2 Le scanner

Le scanner x, appelé également scanner à rayons X, scanographe ou tomodensitomètre est apparu à la fin des années soixante. Il utilise les rayons X pour fonctionner tout comme la radiographie dont il en est une évolution. Le scanner se base également sur la densité des organes. L'idée initiale vient de deux médecins : le **Dr Oldendorf** et le **Dr Ambrose**. La mise au point du procédé est due à un Américain, biophysicien nucléaire, **Allan Maclead Cormack** et à un britannique, ingénieur spécialiste des radars, **Godfrey Newbold Hounsfield**, qui réalisèrent le premier prototype industriel en 1968, mis au point en 1972, et commercialisé en 1975. Tous deux ont reçu le Prix Nobel de Médecine, en 1979, pour ces travaux.

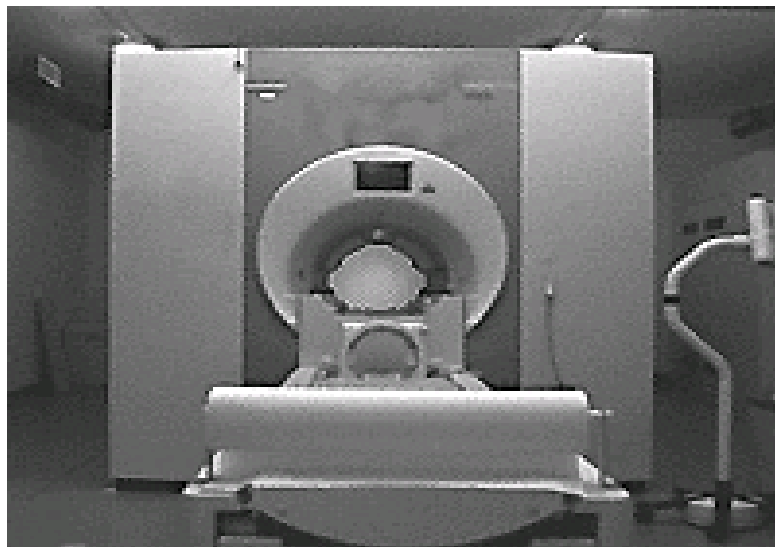


Fig 1.4 : Le scanner

1.2.3.2.1 Principe du scanner

Comme la radiographie classique, le scanner s'appuie sur l'absorption plus ou moins importante des rayons X selon le milieu traversé (les os, par exemple, étant beaucoup plus absorbants que les tissus mous). La technique du scanner permet l'exploitation précise de nombreux organes. L'avantage de cette technique par rapport à la radiologie classique est l'obtention, grâce à

sa sensibilité, de résultats, coupe par coupe, des éléments jusqu'alors confondus sur les clichés radiographiques standards. En effet, en radiologie classique, le faisceau de rayons X projette sur une plaque radiographique les ombres des organes traversés en les confondant. Les zones entourées par des tissus plus denses (comme les os) ne sont donc pas visibles. Le scanner X pallie cet inconvénient : il permet de visualiser tous les éléments profonds de l'organisme, alors qu'une radiographie ordinaire n'offre qu'une vue en projection du volume irradié. Son principe est en effet de choisir un plan de coupe et d'effectuer de multiples projections sous différents angles afin de connaître le coefficient d'atténuation en chaque point du plan.

1.2.3.2.2 Fonctionnement de scanner

Tout d'abord le patient est placé au centre d'un court anneau d'environ 70 cm de diamètre. Il est allongé sur une table qui se déplace dans le sens longitudinal à l'intérieur de l'anneau (Fig.1.4). La figure (1.5) illustre son principe de fonctionnement : un tube émetteur de rayons X tourne autour du patient et génère un faisceau d'une épaisseur de 1 à 10 millimètres. En face du tube sont disposés des milliers de détecteurs qui vont mesurer l'intensité résiduelle du faisceau qui a traversé le corps. Un détecteur électronique (ici le capteur jaune C_0) mesure l'intensité de départ du faisceau des rayons X émis par le tube à rayons X. Ensuite il balaye point par point la tranche du corps à examiner en effectuant une rotation complète degré par degré. Une partie des rayons incidents (ceux qui entrent en contact avec le corps) sont absorbés par les tissus traversés. Sur la figure (2.6) le rayonnement émergent (celui qui ressort du corps) est capté par un détecteur électronique (ici le capteur violet C_1 quand la source est en S_1 et le capteur violet C_2 quand la source est en S_2). Le détecteur tourne en même temps que le tube, de façon synchrone. Au cours de la rotation, les rayons X incidents et les rayons X émergents captés sont comparés et convertis en signaux électriques. Un peu plus de 2 millions de données sont enregistrés en quelques secondes par l'ordinateur. Le programme de celui-ci permet de calculer l'absorption du rayonnement en chaque point de la coupe. Le scanner utilise l'absorption des rayons X en relation directe avec la densité des tissus que les rayons ont rencontrés. Un traitement informatique permet de faire apparaître sur l'écran l'image (Fig 1.7) reconstituée d'une coupe axiale de 1 à 10 millimètres d'épaisseur. Cette image traduit les variations d'absorption des tissus traversés auxquelles sont associées des variations de nuances (noir, gris, blanc) ou des couleurs conventionnelles.

FONCTIONNEMENT D'UN SCANOGRAPHE

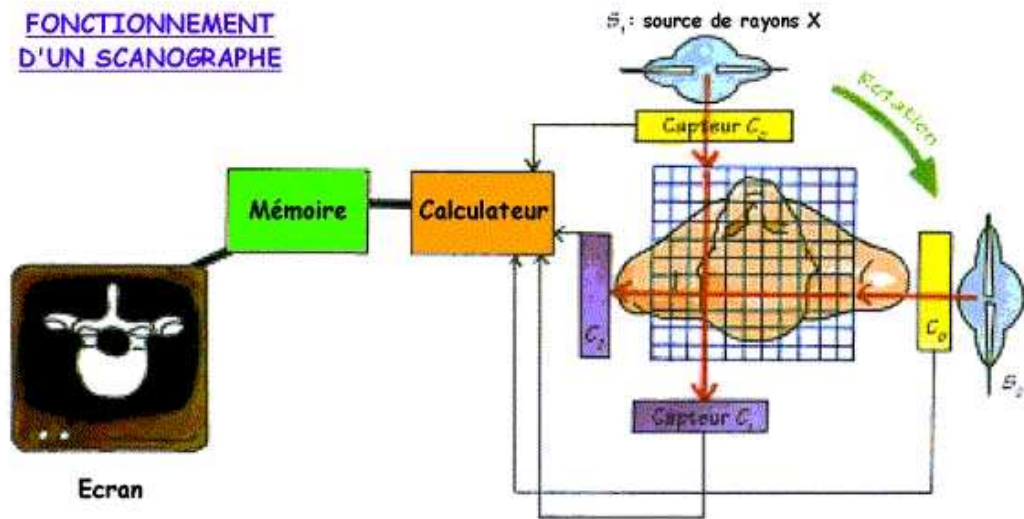


Fig 1.5 : Schéma de principe d'un scanner

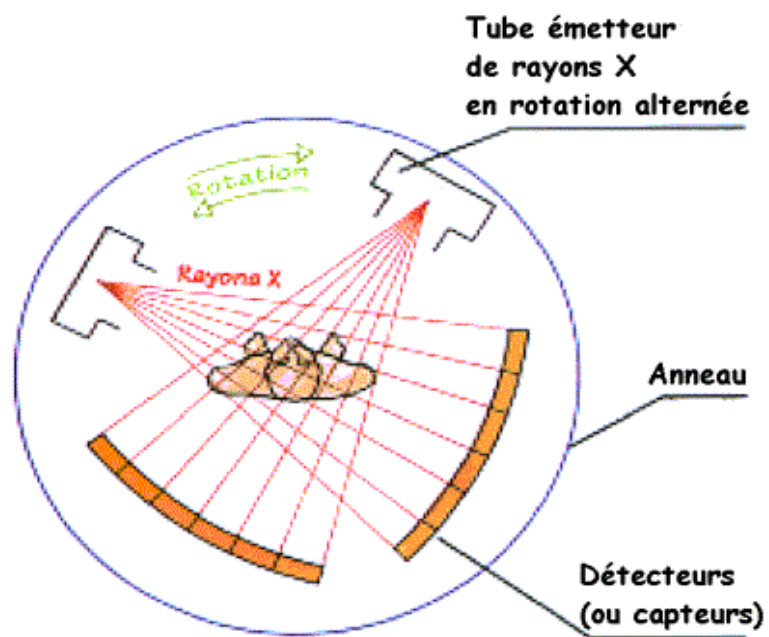


Fig 1.6 : Schéma du principe de fonctionnement d'un scanner



Fig 1.7 : Image d'une coupe axiale

Avantages et inconvénients

Le scanner est un appareil simple et pratique d'emploi. Son faible prix de celui-ci et l'apparition récente de la 5ème génération permettant en une rotation la prise de 4 à 8 coupes fines, laissent un bel avenir au scanner X. Les images numériques obtenues sont des images « de coupe », voire même des images en 3D. Il sert surtout pour diagnostiquer des tumeurs, lésions et fractures. Sa résolution temporelle est de l'ordre de quelques secondes ou millisecondes suivant la partie du corps à étudier et sa résolution spatiale est excellente puisqu'elle permet des coupes du corps humain de l'ordre du millimètre. Toutefois cette technique est purement anatomique et ne donne pas de renseignement par exemple sur les fonctions cérébrales, Mais l'inconvénient majeur du scanner X est la grande dose de rayons X reçu par le patient. Pour l'étude de certains organes, il peut être nécessaire d'injecter (par voie intra-veineuse) ou d'ingérer un produit de contraste à base d'iode, opaque aux rayons X.

1.3 Les ultrasons

Les ultrasons sont des ondes sonores imperceptibles à l'oreille humaine. Comme toutes les ondes sonores, les ultrasons sont absorbés ou réfléchis par les substances qu'ils rencontrent. Les ultrasons sont utilisés dans l'imagerie médicale et principalement dans le domaine de l'échographie et du doppler. L'imagerie par ultrasons permet d'explorer toutes sortes d'organes qu'ils soient superficiels (thyroïde, sein, muscle, articulation) ou profond (foie, vésicule biliaire, pancréas, rein, vessie, gros vaisseaux utérus).

1.3.1 L'échographie

L'échographie est une technique d'imagerie médicale récente puisqu'elle est apparue il y a vingt ans. L'appareil utilisé pour réaliser les échographies s'appelle un échographe (Fig 1.8).



Fig 1.8 : Un échographe

On place une sonde sur la région à examiner après avoir appliqué sur la peau un gel qui permet de supprimer la présence d'air entre la sonde et la peau (l'air empêche la transmission d'ultrasons). Cette sonde est constituée de plusieurs émetteurs disposés côte à côte. Quand l'un d'entre eux émet un ultrason, celui-ci pénètre donc les tissus jusqu'à ce qu'il soit arrêté par une structure et réfléchi (l'ultrason qui est émis dans cette structure est en partie absorbé par les éléments qui la compose et en partie réfléchi). Il est alors renvoyé en direction de la sonde comme un faisceau lumineux par un miroir. Plus la structure est éloignée plus il mettra du temps à revenir. L'écho qui en résulte est enregistré et donne des informations sur la position et la densité des tissus rencontrés. La vitesse de propagation de l'onde sonore étant connue, il suffit de mesurer le temps qui s'écoule entre l'émission du message et la réception de l'onde réfléchie pour connaître la position d'un organe par exemple.

Les images d'échographie sont le plus souvent monochromes. Notons que les liquides simples, dans lesquels il n'y a pas de particules en suspension, se contentent de laisser traverser les sons. Ils ne se signalent donc pas par des échos. Ils seront noirs sur l'écran. Les liquides avec particules, le sang, le mucus, contiennent de petits échos. Ils apparaîtront donc dans les tons de gris. Plus ou moins homogènes. Les structures solides, l'os par exemple, captent et renvoient beaucoup les échos. Ils n'en laissent passer que très peu. On verra donc une forme blanche avec une ombre derrière. Une exception cependant, la voûte crânienne, très fine et perpendiculaire aux échos, en

laisse passer. Les tissus mous sont plus ou moins échogènes; le placenta est plus blanc que l'utérus qui est plus blanc que les ovaires.

Le gaz et l'air apparaissent comme l'os, très blanc. La figure (1.9) montre quelques images d'échographie.



Fig 1.9 : Exemple d'images obtenues d'échographie

Remarque

Si le progrès des calculateurs a rendu possible l'obtention d'images 3D spectaculaires, principalement en obstétrique, les principales améliorations de l'échographie sont d'ordre technologique, portant sur les réseaux de transducteurs piézoélectriques ou sur l'intégration de l'électronique numérique. De nouveaux procédés de focalisation électronique des ondes ultrasonores permettent d'obtenir des images ultrasonores en temps réel, avec une résolution de quelques dizaines de microns seulement: cette biomicroscopie acoustique permet désormais l'examen de l'œil et de la peau.

Des techniques assez récentes, comme l'ostéodensitométrie ultrasonore ou l'élastographie, prennent un essor important et pourraient s'imposer comme de nouvelles modalités diagnostiques s'appuyant sur la nature élastique des ondes sonores pour évaluer la résistance des tissus solides (os) ou l'élasticité des tissus mous (sein, prostate : détection des tumeurs). L'élastographie transitoire (toujours dérivée de l'ultrasonographie) a pour but de cartographier dans une partie du corps le module de cisaillement, une des grandeurs caractérisant l'élasticité et qui peut varier d'un facteur 30 entre un tissu sain et un tissu cancéreux. La découverte récente des miroirs à retournement temporel des ondes acoustiques laisse entrevoir de nouvelles applications médicales. Le développement d'un filtre inverse spatio-temporel, ouvre notamment la voie à la mise au point d'un imageur à ultrasons du cerveau

1.3.2 Doppler

L'échographie doppler, qui utilise une sonde plus fine, permet d'explorer le système vasculaire et donne des indications majeures pour les maladies des veines et des artères. En 1842, le physicien autrichien **Christian Doppler** a observé qu'il y avait une variation apparente de la fréquence d'une onde émise par une source en mouvement par rapport à un observateur immobile. Par exemple, pour un piéton arrêté au bord d'un trottoir, le bruit d'une voiture devient plus aigu quand elle se rapproche et plus grave quand elle s'éloigne. Pour le conducteur de la voiture le bruit demeure identique ; seul le bruit perçu par le piéton varie.

Grâce à l'application de cet effet, il est possible, désormais, d'étudier un écoulement de la circulation du sang par exemple. En effet, lorsqu'une onde se réfléchit sur un objet en mouvement, comme un globule rouge, elle change de fréquence. L'analyse de ce changement permet de reconstituer les mouvements.

C'est grâce à ce système que certaines images d'échographie apparaissent en couleur : Les flux qui se rapprochent sont en rouge (ou jaune), ceux qui s'éloignent en bleu et les turbulences en vert (Fig.1.10).



Fig 1.10 : Images d'échographiques doppler

1.4 Les rayons gamma

La radioactivité, phénomène qui fut découvert en 1896 par **Henri Becquerel** sur l'uranium et très vite confirmé par **Marie Curie** pour le thorium, est un phénomène physique naturel au cours duquel des noyaux atomique instables, dits radioisotopes, se transforment spontanément en dégageant de l'énergie sous forme de rayonnement divers, pour se transformer en des noyaux atomiques plus stables ayant perdu une partie de leur masse. Les rayonnements ainsi émis sont appelés, selon le cas, des rayons α , des rayons β ou des rayons γ . Les rayons gamma sont sur le

spectre lumineux des rayonnements ayant la plus grande fréquence et la plus grande longueur d'onde, ce qui implique que les rayons gamma soient si énergétique car il faut savoir que plus un photon a une fréquence élevé et plus il est énergétique.

Deux techniques d'imagerie médicale font appel au phénomène radioactivité : la scintigraphie et la tomographie par émissions de positions.

1.4.1 La scintigraphie

Mise au point en 1958 par l'américain **Hal Hanger**, la scintigraphie utilise une petite quantité de radioactivité pour réaliser des images (d'où le nom de médecine nucléaire). Elle permet de visualiser le fonctionnement d'un organe (cœur, poumons, reins, vessie, organes digestifs) ou de tout un système (système ostéo-articulaire, système endocrinien, système immunitaire...). Elle permet en procédant au « marquage » de certaines molécules de l'organisme, d'en étudier la circulation dans le corps (scintigraphie des globules blancs par exemple). Ces cellules « marquées » peuvent être visualisées sur la caméra car elles irradient de la lumière.

1.4.1.1 Principe de fonctionnement

Cet examen consiste à injecter au malade un produit radioactif (émetteur de rayons gamma) qui varie selon l'organe examiné. Le produit va se fixer de façon passagère sur certains tissus ou organes. Une fois fixé, on mesure la radioactivité sur l'organe ou les tissus grâce à un appareil appelé gamma-caméra (Fig.1.11) : composé d'une table équipée d'une ou plusieurs caméras spéciales qui captent le rayonnement gamma.



Fig 1.11 : La gamma caméra

La figure (1.12) montre quelques images de scintigraphie.

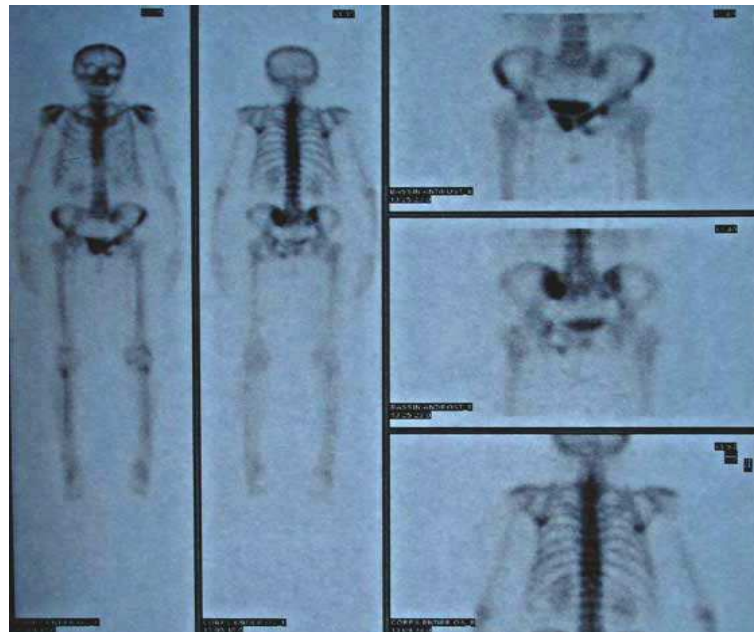


Fig 1.12 : Exemple d'images scintigraphie

Le tableau ci-dessous regroupe les différents produits radioactifs utilisés en scintigraphie

Isotope	Demi-vie (en heures)
Technétium 99	6
Thallium 201	73
Iode 123	13
Iode 131	192

Le thallium 201 est le plus souvent utilisé pour l'étude des tissus myocardiques. Il est utilisé en cardiologie pour visualiser l'irrigation du muscle cardiaque. On fait une épreuve d'effort à un patient et on lui injecte le thallium radioactif. Le produit va se fixer selon l'irrigation sanguine au niveau des fibres cardiaques. Les zones non ou mal irriguées par le sang ne recevront pas ou peu de thallium. On obtiendra ainsi une cartographie de l'irrigation du cœur. Le technétium 99 est utilisé

pour l'étude des cavités cardiaques, les poumons, les os, les reins et la thyroïde. Au niveau du poumon, la scintigraphie est surtout utilisée pour diagnostiquer l'embolie pulmonaire. Pour les os, elle sert à diagnostiquer les tumeurs osseuses, les inflammations, les infections, voire une arthrose. L'iode 131 et 123 sont utilisés pour contrôler la répartition de la fabrication de l'hormone thyroïdienne sur la glande.

1.4.2 Tomographie par émission de positons (TEP)

La tomographie par émission de positons fonctionne sur le même principe que la scintigraphie. Contrairement aux isotopes utilisés en scintigraphie, les radioéléments émetteurs de positons sont des éléments chimiques légers et abondants dans la matière vivante. Développée par deux américains, **William Sweet** et **Gordon Brownell**, la tomographie par émission de positons (**TEP** ou **PET** = **P**ositron **E**mission **T**omography en anglais) ne donne pas d'image anatomique mais renseigne sur la biochimie des organes. Elle fournit des informations sur le fonctionnement des tissus normaux et pathologiques. Suivant l'isotope radioactif utilisé, elle donne au médecin des informations spécifiques sur le métabolisme des tissus. Elle permet d'étudier in vivo le fonctionnement biochimique d'un organe, tel que le cerveau, le cœur ou les poumons, en donnant des informations spécifiques sur le métabolisme de ses tissus : flux et volume sanguin, consommation cellulaire du glucose... Les cellules cancéreuses par exemple consomment plus de glucose que les cellules bénignes, ce qui permet de repérer par cette technique d'éventuels points chauds témoins de récidives cancéreuses. Si elle est particulièrement intéressante en cancérologie, la TEP a des applications dans d'autres disciplines : en neurologie, elle permet par exemple de visualiser les pertes neuronales du système dopaminergique qui signalent la maladie de Parkinson, ou bien encore les zones hyperactives du cerveau dans l'épilepsie. Elle offre en outre la possibilité de tester in vivo l'efficacité des médicaments, ce qui intéresse notamment la cardiologie dans la mise au point de nouveaux traitements.

La figure (1.13) montre une image TEP

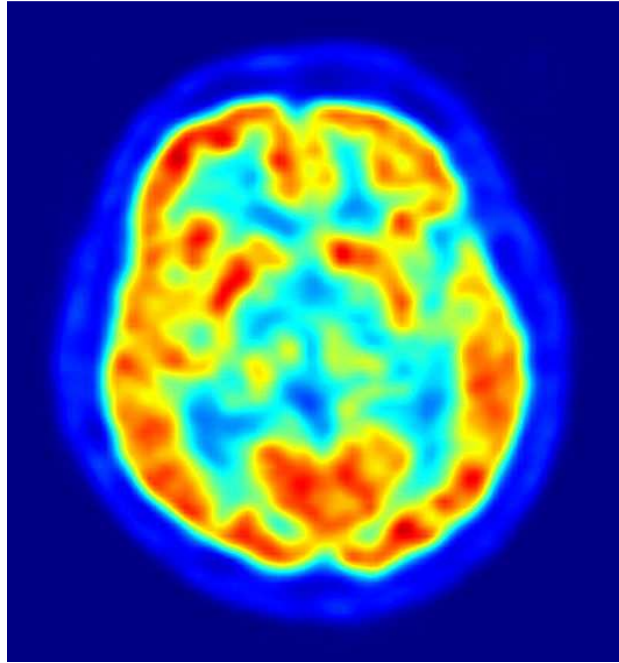


Fig 1.13 : Vue axial du cerveau par TEP

1.5 Imagerie par résonance magnétique (IRM)

L'imagerie par résonance magnétique (I.R.M.) est une technique d'imagerie médicale mise au point par **Lauterbur** et **Damadian** en 1973. Elle est utilisée pour faire un diagnostic qui se fonde sur les principes de la résonance magnétique nucléaire (RMN) découverte en 1946 par **Félix Bloch** et **Edward Purcell**. L'IRM est la méthode de diagnostic la plus puissante et la plus sensible disponible actuellement. Cet outil permet d'obtenir des images de tissus à l'intérieur du corps humain plus précises que celles obtenues par un scanner ou par ultrasons.

La figure (1.14) montre l'appareil IRM et la figure (1.15) montre une image IRM.



Fig 1.14: Appareil IRM

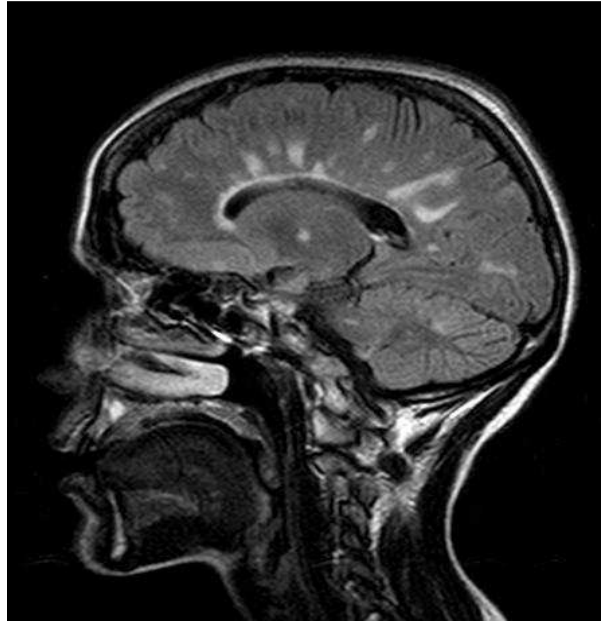


Fig 1.15 : Image IRM

1.5.1 Principe

Son principe consiste à réaliser des images du corps humain grâce aux nombreux atomes d'hydrogène qu'il contient car le corps est essentiellement formé d'eau (H_2O). Le patient est placé dans un puissant champ magnétique, produit par un gros aimant. Tous les atomes d'hydrogène s'orientent dans la même direction, ils sont alors excités par des ondes radio fournis par une antenne émettrice durant une très courte période (ils sont mis en résonance). A l'arrêt de cette stimulation les atomes reviennent à leur position initiale et restituent l'énergie accumulée en produisant un signal capté par une antenne réceptrice.

Les signaux ainsi captés sont analysés pour produire grâce à un ordinateur une image tridimensionnelle des organes, présentée en coupe successive. L'intensité du signal recueillie pour un élément de volume dépend de la concentration de l'eau à l'endroit considéré. L'image tridimensionnelle traduit de la répartition de l'eau dans le corps d'un patient. Il est donc possible d'observer des altérations des tissus (telles que des tumeurs) grâce aux différences de densité et de relaxation de l'eau.

En fonction des paramètres choisis, l'IRM permet d'obtenir des images très contrastées de certains tissus en fonction de leurs propriétés histologique, c'est donc un outil particulièrement utilisé en imagerie cérébrale.

L'examen en IRM permet de faire des coupes dans les trois (3) plans : horizontal, vertical et transversal. Les radiologues leur donnent les noms de : coupes coronale, axiale et sagittale.

Remarque

En IRM, l'intensité du signal est souvent insuffisante pour permettre d'observer une différence convenable entre les parties saines ou affectées de l'organisme. Une façon très simple d'influencer le signal en IRM est d'augmenter le contraste, soit en allongeant le temps de l'examen pour permettre de prendre plus d'acquisitions, soit en utilisant un agent de contraste spécifique ou non spécifique.

La figure 1.16 contient quelques clichés IRM de la tête d'un patient. Elles montrent la différence entre une IRM sans (clichés à gauche), et une IRM avec agent de contraste (clichés à droite). Nous voyons clairement la présence d'une tumeur, tache claire à la base de l'hémisphère gauche du cerveau après l'injection d'un agent de contraste.

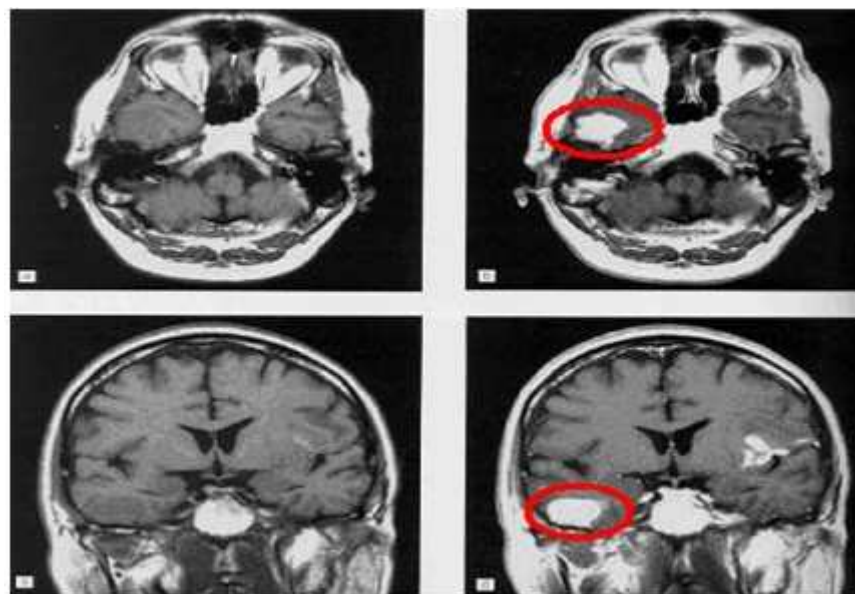


Fig 1.16: Exemple de clichés IRM

1.5.2 Application de l'IRM

Les applications de l'IRM en médecine sont très nombreuses.

Lors d'un traumatisme du cerveau, l'IRM détecte très bien les saignements ainsi que l'œdème qui en résulte. D'autres anomalies peuvent aussi être détectées, comme par exemple un anévrisme, un

infarctus, une tumeur ou bien un processus inflammatoire ou tumoral de la moelle. Les neurochirurgiens utilisent l'IRM afin d'étudier les rapports au niveau anatomique d'une lésion avec les structures saines avoisinantes mais aussi pour l'évaluation de l'intégrité de la moelle après un traumatisme. L'IRM est également utilisée pour étudier les problèmes osseux rachidiens ou discaux voire infectieux qui peuvent toucher la moelle. Elle permet aussi la détection de processus pathologiques cardiaques (notamment ischémiques) ou des gros vaisseaux sanguins et elle donne des informations sur les organes de l'abdomen. Enfin l'IRM permet d'avoir des informations sur les articulations et les parties non osseuses. Les techniques chirurgicales bénéficient alors des informations anatomiques et des informations de contraste apportées par l'IRM.

1.6 Endoscopie

L'endoscopie est une technique d'imagerie médicale très précise puisqu'elle offre des indications de texture du tissu, de volume ou de coloration. Elle permet un examen "en direct" de certains organes. L'endoscopie consiste à introduire après une incision cutanée un tube, l'endoscope, muni d'une optique (mini camera couleur), d'un système d'éclairage, d'une pince à biopsie pour la réalisation d'un prélèvement et d'une pince à corps étranger, dans l'organisme. Cette technique est souvent utilisée pour rechercher la cause d'un trouble. Les images sont conservées sur des photographies ou sur un document vidéo (vidéo-endoscopie).

L'endoscope peut être utilisé pour explorer le thorax (thoracoscopie), l'abdomen (coelioscopie) ou dans une articulation (arthroscopie), ou encore par les voies naturelles, par exemple par la bouche pour un examen du tube digestif ou des bronches (bronchoscopie), par le méat urinaire pour les uretères, la vessie et l'urètre, par l'anus pour le rectum et le côlon (coloscopie). La figure (1.18) montre une image fournie par un endoscope.



Fig. 1.17: Appareil endoscopique



Fig 1.18: Image endoscopique

1.7 Conclusion

Dans ce chapitre nous avons présenté d'une manière générale quelques notions sur l'imagerie médicale. Nous avons détaillé quelques méthodes comme celles qui sont basées sur les rayons x (radiographie, scanner, ...), sur l'ultrasons (échographie, doppler, ...), les rayons gamma (PET, scintigraphie...) ainsi que l'IRM qui est basée sur la mesure de la réponse des atomes d'hydrogènes dans l'organisme à un champ magnétique extérieur et à une onde électromagnétique. Grâce aux images médicales, les médecins peuvent voir ce qui reste invisible à l'œil nu.

Chapitre 2 : Segmentation d'images

2.1 Introduction

La segmentation des images constitue une étape importante dans le processus d'analyse des images. Elle a pour objectif de fournir une description des objets contenus dans l'image par l'extraction de différents indices visuels tels que les contours des objets et les régions homogènes.

La phase de segmentation d'images n'est pas considérée comme un but en soi, elle dépend fortement aussi bien du type de traitement fixé par l'utilisateur sur les objets présents dans l'image, que de la nature de l'image (présence de bruit, présence de zones texturées, contours flous....), ainsi que des primitives que l'on cherche à extraire de l'image et qui dépendent des opérations situées en aval de la segmentation (localisation, calcul 3D, reconnaissance de formes, interprétation).

Ces aspects, ainsi que les contraintes d'exploitation (complexité algorithmique, aspect temps réel), justifient la multiplicité des techniques développées dans le domaine de la segmentation d'image. Jusqu'à ce jour, il n'existe pas de méthode universelle de segmentation d'images. Toute technique n'est efficace que pour un type d'image donné, pour un type d'application donné, et dans un contexte informatique donné. En raison de ces contraintes, les diverses stratégies de segmentation qui ont été proposées ont affirmé leurs insuffisances et leurs limitations.

Nous allons présenter dans ce chapitre quelques principes et méthodes connues de segmentation.

2.2 Définition de la segmentation

Segmenter une image signifie décomposer cet image en régions homogènes qui peuvent être délimitée par leur contours. Ces régions sont supposées être pertinentes, c'est à dire que les régions doivent correspondre aux parties significatives des objets formant la scène, et les contours à leurs frontières apparentes.

Une définition formelle d'un algorithme de segmentation a été donnée par **Harowitz [1]** et **Pavlidis [2]** en 1975.

Soit X le domaine de l'image. Si nous définissons un prédicat P sur l'ensemble des parties de X , la segmentation de X est définie comme une partition de X en n sous-ensembles

$\{R_1, \dots, R_n\}$ tels que :

1. $X = \bigcup_{i=1}^n R_i$
2. $\forall i \in \{1, \dots, n\} R_i$ est connexe
3. $\forall i \in \{1, \dots, n\} p(R_i) = \text{vrai}$
4. $\forall i, j \in \{1, \dots, n\}^2 R_i$ est adjacent à R_j et $i \neq j \Rightarrow p(R_i \cup R_j) = \text{faux}$

Où $\cup$ représente une union d'ensemble disjoints.

Les conditions d'**Harowitz** sont résumées par **Zucker [3]** comme suit :

La première condition implique que tout pixel de l'image appartient à une région et une seule.

La seconde condition implique que toute région doit être connexe. La connexité des régions est induite par le voisinage défini sur l'image.

La troisième condition implique que chaque région doit être homogène.

La dernière condition est une condition de maximalité indiquant que la fusion de deux régions homogènes ne doit pas être homogène.

Le prédicat P est utilisé pour tester l'homogénéité des régions R_i .

Cette définition ne conduit pas à une méthode de segmentation unique et il peut exister plusieurs segmentations possibles d'une image avec l'utilisation d'un même prédicat.

La figure 2.1 montre un exemple d'image médicale segmentée :

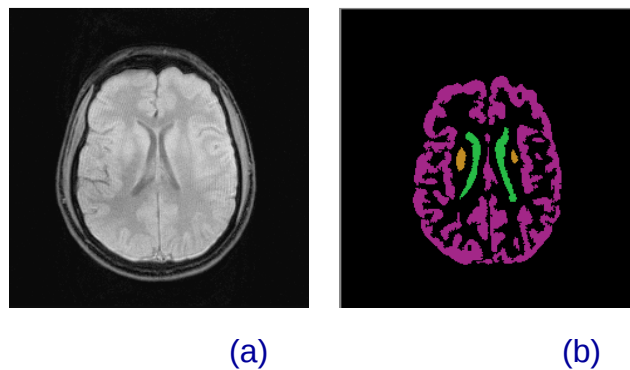


Fig 2.1 : Segmentation d'images cérébrales : (a) image d'origine
(b) image segmentée

Notons qu'une région est dite homogène si les caractéristiques des pixels la constituant sont sensiblement identiques. Or pour une image en niveaux de gris un pixel peut être caractérisé soit par son niveau de gris, soit par un ensemble de paramètres de texture. Ces derniers offrent une information plus riche que le niveau de gris seul.

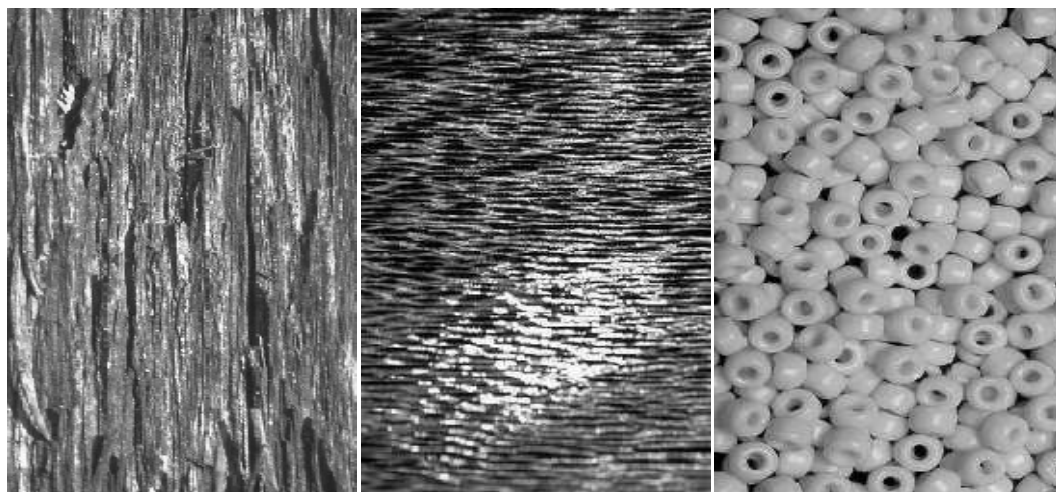
2.3 Définition de la texture

La texture est une des propriétés de base de l'image. Elle représente l'ordonnancement visuel du changement des niveaux de gris dans une surface de l'espace ou une structure d'un objet. **Maitre [4]** a proposé la définition suivante : une texture est un champ de l'image qui apparaît comme un domaine cohérent et homogène. La figure (2.2) montre quelques images de la texture.

La caractérisation d'une texture repose sur des attributs tels que le contraste, la directionnalité, la structure des lignes, la régularité et la rugosité, etc. Son but principal est l'identification de texture.

L'analyse de texture peut être employée pour effectuer la classification des tissus et la segmentation d'image IRM par exemple, mais aussi pour estimer la gravité d'une tumeur.

Plusieurs méthodes d'analyse de la texture sont proposées. La plus connue d'entre elles est celle qui est basée sur les matrices de cooccurrence



Écorce d'arbre

Poil court

Perles

Fig 2.2 : Exemples de textures naturelles

2.4 Paramètre extraits à partir d'une matrice de cooccurrence

Les matrices de cooccurrence ont été introduites par **Haralick [5]** pour l'analyse des textures. En général une matrice de cooccurrence notée $M_{(d,\theta)}$ est une estimée de la densité de probabilité conjointe de rencontrer deux niveaux de gris i et j correspondant à deux pixels séparés par une distance d et avec une orientation θ . Le nombre de ces matrices n'est pas limité puisqu'elles dépendent du choix de d et θ , d généralement pris égal à 1 et θ à $0^\circ, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$.

A partir des matrices de cooccurrences, nous pouvons extraire plusieurs paramètres caractérisant la texture. Les plus utilisés sont le moment angulaire d'ordre deux (énergie), le contraste, l'entropie et la corrélation.

- **Energie**

Ce paramètre mesure l'homogénéité des niveaux de gris de l'image, il a une valeur numérique faible quand les $M(i,j)$ de la matrice de cooccurrence ont des valeurs très proches et une valeur forte quand les niveaux de gris ne sont plus équiprobables.

$$ASM = \sum_i \sum_j M_{i,j}^2$$

- **Contraste**

Il mesure les variations locales des niveaux de gris, il permet aussi de caractériser la dispersion des valeurs de la matrice par rapport à sa diagonale principale.

$$CON = \sum_i \sum_j (i - j)^2 M_{i,j}$$

- **Entropie**

Elle caractérise le désordre dans l'image et rend compte de la complexité de la texture de celle-ci, il en résulte alors que l'entropie est maximum si la distribution spatiale des pixels est équiprobable et faible si les couples de pixels sont inégalement répartis dans l'image. De plus elle permet aussi de caractériser le degré de granulation de l'image : plus l'entropie est élevée et plus la granulation est grossière.

$$ENT = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} M(i,j) \log[M(i,j)]$$

- **Corrélation**

La corrélation a une forte valeur quand les valeurs sont uniformément distribuées dans la matrice de cooccurrence. Elle mesure la dépendance linéaire des niveaux de gris dans l'image.

$$COR = \frac{1}{\sigma_x \sigma_y} \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - \mu_x)(j - \mu_y)M(i, j)$$

$\mu_x, \mu_y, \sigma_x, \sigma_y$ sont respectivement les moyennes et les écarts-type.

$$\mu_x = \sum_{i=0}^{N-1} iP[i], \quad \mu_y = \sum_{j=0}^{N-1} jP[j],$$

$$\sigma_x^2 = \sum_{i=0}^{N-1} (i - \mu_x)P_x[i], \quad \sigma_y^2 = \sum_{j=0}^{N-1} (j - \mu_y)P_y[j]$$

2.5 Méthodes de segmentation d'image

Il existe une multitude de méthodes de segmentation. Celles-ci peuvent être de type contour ou de type région.

2.5.1 Segmentation par contour

Elle est bien adaptée aux images qui présentent des régions uniformes au sens des niveaux de gris. Mais, elle reste très sensible aux bruits et aux contours mal définis. Elle est donc utile pour détecter les frontières entre les objets mais ne s'applique pas aux frontières séparant des régions de textures différentes.

2.5.2 Segmentation par région

Contrairement à la segmentation par contour qui recherche les dissimilarités, l'approche région recherche plutôt la similarité. Cette approche fournit une carte de régions fermées. Cependant les pixels adjacents similaires, selon un certain critère d'homogénéité, sont regroupés en région distinctes. Parmi les méthodes appartenant à cette approche, on distingue les méthodes de classification, la croissance de régions et la division fusion.

2.5.2.1 Segmentation par croissances des régions [6]

Les méthodes par croissance des régions, de par leur relative simplicité de mise en œuvre, ont été l'une des premières familles d'outils utilisés pour la segmentation d'image. A partir d'un ensemble

de points initiaux, cette méthode agrège les pixels suivant un double critère : l'homogénéité et l'adjacence. Cette agrégation de pixels est contrôlée par un prédicat. Le processus s'applique sur une ou plusieurs régions. Pour chaque région, il comprend une phase d'initialisation et une phase itérative.

La phase itérative modifie la région R représentant l'objet segmenté.

Suite à la phase d'initialisation, un processus itératif de déformation ajoute progressivement des points situés à la périphérie de la région en train de croître, s'ils respectent le prédicat.

Ces nouveaux points sont regroupés dans l'ensemble que l'on note C[n].

Considérons une croissance de régions ne faisant croître qu'une seule région.

Étant donné une région initiale R[0], la région suivante R[n+1] est obtenue à partir de la région courante R[n] et de la couronne C[n] par la réunion suivante :

$$R^{[n+1]} = R^{[n]} \cup C^{[n]}$$

Comme l'itération porte sur une région croissante et bornée, la convergence est assurée au bout d'un nombre fini d'itérations.

Le processus de croissance peut s'arrêter selon deux scénarios :

- Toutes les régions satisfont le prédicat et la segmentation comprend n régions
- Il existe (n – 1) régions qui vérifient le prédicat, la n^{ième} région comprend les points ne vérifiant pas le prédicat.

La figure 2.3 montre un exemple de segmentation par croissance de région :

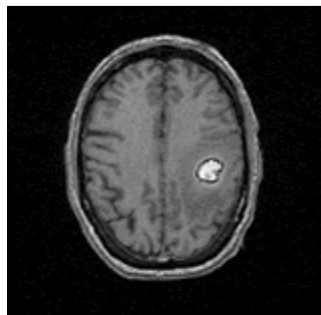


Fig 2.3 : Vue d'une tumeur au niveau du cerveau en IRM

2.5.2.2 Approche par division-fusion [7]

Cette approche combine deux étapes; division et fusion. La phase de division (split) consiste à partitionner une image en zones localement homogènes. La phase fusion (merge) consiste à regrouper des régions similaires au sens d'un prédicat de regroupement. Ces deux opérations sont répétées jusqu'à ce qu'elles ne soient plus possibles.

L'algorithme de division se résume en les étapes suivantes :

1. Diviser l'image initiale en 4 régions (principe de structure quadtree) si le:
 $H(\text{image}) = \text{Faux}$. Un critère d'homogénéité noté H .
2. Si pour une région R_i , le prédicat d'homogénéité H n'est pas vérifié ($H(R_i) = \text{faux}$), alors diviser la région en 4.
3. Aller à 2. Si toutes les régions ne satisfont en même temps le prédicat H .

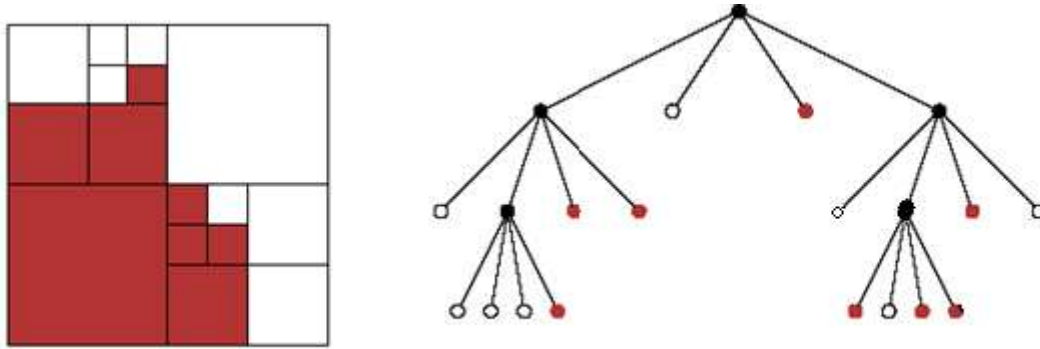


Fig 2.4 : Structure quadtree

L'algorithme de fusion commence par une image pré-segmentée, pour ensuite :

1. Fusionner tout couple de régions adjacentes qui vérifie un critère d'homogénéité.
2. Définir un prédicat Fusionne (R_i, R_j) où R_i et R_j sont deux régions adjacentes.

La labellisation est caractérisée pour sa connexité à savoir 4 ou 8 :



Fig 2.5 : Connexité

2.5.2.3 Méthodes de classification [8]

Parmi les techniques de segmentation d'images, la classification est l'une des procédures les plus utilisées. Les pixels de l'image sont regroupés en un ensemble de k classes disjointes. A chaque pixels de l'image, on attribue une étiquette parmi les k étiquettes qui correspondent chacune à une classe. On note par X l'ensemble des N pixels

$X = \{X_1, X_2 \dots X_N\}$ où chaque pixel peut être caractérisé par D paramètres $X_i = (x_{i1}, x_{i2} \dots x_{iD})$. Si on considère C l'ensemble qui représente l'image, $\{c_1, c_2, \dots, c_k\}$ l'ensemble des k classes de l'image alors $\{c_1, c_2, \dots, c_k\}$ constitue une partition de C s'il vérifie les conditions suivantes:

- $c_i \neq \emptyset \quad \forall i \in \{1, 2, \dots, k\}$
- $c_i \cap c_j = \emptyset \quad \forall i \in \{1, 2, \dots, k\} \quad \forall j \in \{1, 2, \dots, k\} \quad i \neq j$
- $\cup c_i = C$

On peut distinguer deux approches de classification:

- L'approche supervisée.
- L'approche non supervisée.

2.5.2.3.1 Méthodes supervisées

Ce sont des méthodes dans lesquelles les classes sont connues a priori. Elles demandent une première phase d'apprentissage sur l'échantillon représentatif (ensemble de pixels dont l'appartenance de chaque observation à une classe est connue a priori) dans le but d'apprendre les caractéristiques de chaque classe et une deuxième phase pour décider de l'appartenance d'un pixel à telle ou telle autre classe.

Parmi ces méthodes on peut citer : la classification Bayésienne, réseaux de neurones, etc. ce type de méthodes sont rarement utilisées pour la segmentation d'images.

2.5.2.3.2 Méthodes non supervisées

L'intérêt des méthodes non supervisées est qu'elles ne nécessitent aucune connaissance donc aucune tâche préalable d'étiquetage manuel n'est requise. La seule intervention de l'expert se situe à la fin du processus pour identifier par exemple les tissus en comparant les classes calculées avec les classes biologiques.

Les algorithmes non supervisés les plus répandus tendent généralement à minimiser une fonction coût, dépendant de la distance de chaque pixel aux prototypes (ou noyaux) des classes. Le prototype

d'une classe étant un point connu ou un ensemble de points qui sont mis à jour par l'algorithme. En suite chaque pixel est assigné à la classe qui lui est la plus proche. On distingue plusieurs méthodes de classification.

2.5.2.3.2.1 Algorithme EM (Estimation Maximisation) [9]

On suppose que la répartition de l'ensemble des pixels X peut être décrite par une fonction de densité de probabilité qui peut être assimilée à un mélange de K composantes (généralement gaussienne). Le problème de la classification revient alors à estimer les paramètres de chaque composante (moyenne, variance).

Le problème de l'estimation des paramètres des composantes dans les mélanges se base sur le principe de maximisation de la vraisemblance (maximum likelihood) : c'est-à-dire que les paramètres des composantes sont choisis de manière à maximiser la probabilité des pixels d'appartenir au mélange. C'est le but de l'algorithme *EM* (Expectation Maximization) proposé par **Dempster**.

2.5.2.3.2.2 Algorithme Mean Shift

L'algorithme du Mean Shift, introduit par **Fukunaga [10]** puis remis au goût du jour par **Comaniciu [11]**, recherche le "mode" ou point de plus haute densité d'une distribution des pixels qui corresponde au centre de la classe. Cette méthode consiste à estimer la fonction de densité de probabilité sous jacente à l'ensemble de pixels puis localiser ses maxima locaux en recherchant les zéros de son gradient.

Cette procédure est résumée comme suit :

1. Considérer les images en termes de caractéristiques (via couleur, gradient, mesures de texture, etc.)
2. Choisir une répartition uniforme des fenêtres de recherche initiales.
3. Calculer le centroïde des données pour chaque fenêtre.
4. Centrer la fenêtre de recherche sur le centroïde de l'étape 3.
5. Répéter les étapes 3 et 4 jusqu'à convergence.
6. Fusionner les fenêtres se trouvant au même point final.
7. Grouper les pixels traversés par les fenêtres fusionnées.

2.5.2.3.3 K-means

La méthode de K-means a été introduite par **J-Mc Queen** en 1971 mise en œuvre sous sa forme actuelle par **E- Forgy**. C'est la technique de classification non supervisée la plus utilisée. K-means partitionne les données (pixels) en K classes. Il est basé sur la minimisation d'une erreur quadratique entre tous les pixels et les centres des classes.

$$J = \sum_{i=1}^k \sum_{j=1}^N \|x_j - \bar{x}_i\|^2$$

Avec : $X = \{x_j / j = 1, 2, \dots, N\}$ est l'ensemble des N pixels à classer

$\bar{X} = \{\bar{x}_i / i = 1, 2, \dots, k\}$: est l'ensemble des vecteurs centres des classes.

Initialement K prototypes, ou centroides sont positionnés dans les régions de l'espace les plus peuplées. Chaque observation est alors affectée au centre le plus proche (règle dite : de la distance minimale). Chaque classe contient donc les observations qui sont plus proche d'un certain prototype que de tous autres centres. Les centres sont positionnés par une procédure itérative qui les amène progressivement dans leurs positions finales stables.

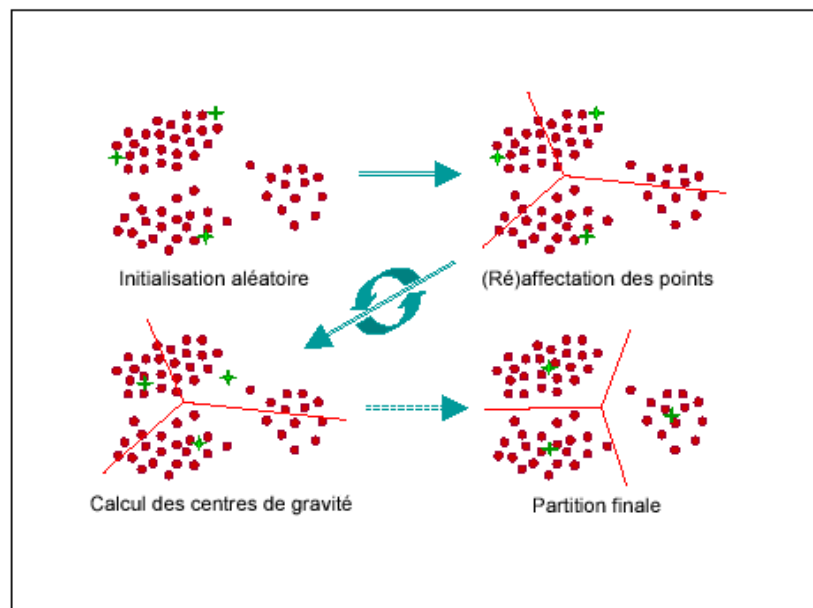


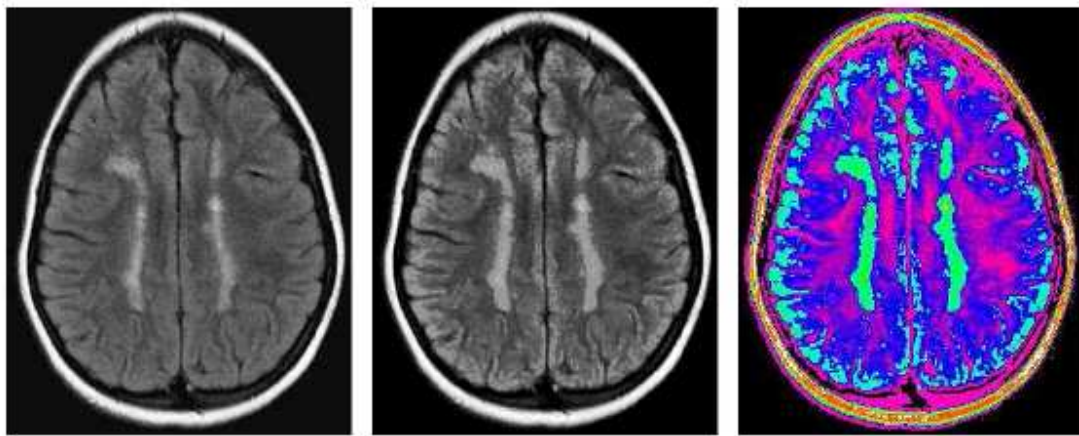
Fig 2.6 : Principe de l'algorithme de K-means.

Les principales étapes de l'algorithme *k-means* sont :

1. Choix des centres initiaux des K clusters.
2. (Ré-) Affecter les pixels à la classe la plus proche en utilisant (la distance euclidienne).
3. Recalculer les K centres des classes.
4. Répéter les étapes 2 et 3 jusqu'à ce que plus aucune réaffectation ne soit faite.

Cet algorithme est très populaire car extrêmement rapide en pratique. En effet, le nombre d'itérations est typiquement inférieur aux nombres de points. En terme de performances, cet algorithme ne garantit pas un optimum global, la qualité de la solution dépend grandement des ensembles initiaux. Le nombre de classes est également exigé mais cela n'étant pas réellement handicapant car dans certaines applications comme par exemple la segmentation cérébrale le nombre de classes est le plus souvent connu. Aussi, les pixels sont classés de façon certaine comme appartenant à une classe et une seule, quelle que soit la modalité d'imagerie, cette assertion ne reflète pas la réalité physique.

La figure (2.7) montre un ensemble de segmentation par l'algorithme *k-means*.



(a)

(b)

(c)

(a) image originale

(b) image segmentée en niveau de gris

(c) image segmentée en couleur

Fig 2.7 : Résultat de classification obtenue par *k-means* sur 16 classes

2.6 Intérêt de la segmentation dans le domaine médical

Plusieurs applications pratiques de la segmentation sont possibles aujourd'hui et touchent tout les domaines d'activité. Dans le domaine médical, la segmentation est une étape indispensable dans la chaîne de traitement dans le domaine neurologique de la recherche clinique. Nous citerons quelques exemples :

- ◆ la mesure du volume des hippocampes dans le domaine de l'épilepsie temporale ou la maladie d'Alzheimer.
- ◆ La mesure du volume des ventricules afin d'évaluer l'atrophie cérébrale dans le domaine de la maladie d'Alzheimer ;
- ◆ La mesure du volume des noyaux gris dans la maladie de Parkinson ;
- ◆ La mesure de l'atrophie du corps calleux et la détection automatique des lésions dans le cas de la Sclérose en Plaques.

Plutôt que de faire lire tous les clichés par deux médecins, il est envisageable qu'un même radiologue fasse une première lecture sans assistance et qu'ensuite le système de l'aide à la détection lui montre les zones classées suspectes pour que le médecin vérifie qu'il n'a rien raté dans ces zones.

L'avantage de cette approche est la combinaison des méthodes automatiques et humaines. Le système automatique peut faire une recherche plus exhaustive sur l'ensemble du cliché pendant que le radiologue dirige son attention sur les zones lui semblant les plus remarquables.

Krupinski et **Nishikawa** ont comparé les performances d'un système de détection automatique avec celles de six radiologues pour la détection des microcalcifications. Ils ont enregistré les positions des yeux des radiologues pendant la lecture des films et ont analysé le temps pendant lequel les radiologues se focalisent sur des lésions détectées ou ratées. Ils constatent qu'un radiologue expérimenté ne balaye pas l'image entière mais se concentre sur les zones identifiées lors de l'évaluation initiale globale comme contenant probablement des lésions. Leur évaluation sur 80 films a montré que 10% des amas de microcalcification ont été détectés par le système automatique mais par aucun des six radiologues et que 11% des amas ont été ratés par le système automatique mais détectés par au moins un des radiologues. Seulement 5% des amas ont été ratés par les radiologues et le système automatique.

2.7 Conclusion

Nous avons tenté dans ce chapitre de décrire brièvement les techniques de segmentation d'images en trois approches: approche par régions, approche par contours et une approche par classification. Les informations fournies par les images sont incertaines et imprécises. Nous gérons ces ambiguïtés en utilisant la logique floue que nous allons détailler dans le chapitre suivant.

Chapitre 3 : Notions sur la logique floue

3.1 Introduction

La logique floue est une théorie très utilisée en intelligence artificielle, elle a été formalisée par **Lotfi ZADEH** en 1965. Elle s'appuie sur la théorie mathématique des ensembles flous qui est une extension de la théorie des ensembles classiques pour la prise en compte d'ensemble défini de façon imprécise.

Les sous-ensembles flous ont été introduits afin de modéliser la représentation humaine des connaissances, et ainsi améliorer les performances des systèmes de décisions qui utilisent cette modélisation, les sous-ensembles flous utilisés soit pour modéliser l'incertitude et l'imprécision, soit pour représenter des informations précises. En autre termes la logique floue permet à une condition d'être dans un autre état que les états vrai ou faux, elle introduit des degrés dans la vérification d'une condition.

Nous présenterons dans ce chapitre quelques notions et définitions de base de la logique floue.

3.2 Théorie des sous ensembles flous

Cette théorie est basée principalement sur la graduation dans l'appartenance d'un élément à une classe ; c'est-à-dire qu'un élément peut appartenir plus ou moins fortement à cette classe.

3.2.1 Définition d'un sous ensemble flous [12]

Un sous-ensemble flou A d'un ensemble X appelé univers du discours est défini par une fonction d'appartenance μ_A qui associe à tout élément x de X une valeur réelle $\mu_A(x)$ dans l'intervalle $[0, 1]$. Ainsi, un sous-ensemble flou est toujours (et seulement) une fonction de X dans $[0,1]$.

Exemple

Soit la classification des personnes en trois «3» ensembles A “jeune”, B “entre deux âges ” et C “âgés”. Une telle classification peut être décrite en utilisant la fonction d'appartenance illustrée dans la Fig 3.1 suivante :

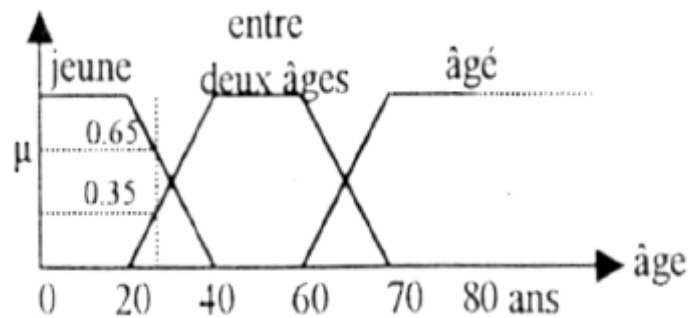


Fig 3.1 : Fonction d'appartenance

On peut alors déduire qu'une personne de 25ans appartient à l'ensemble A " jeune" avec :

$\mu_A(x) = 0.65$ et l'ensemble B " entre deux âges" avec : $\mu_B(x) = 0.35$

$$0 < \mu < 1$$

Dans le cas où $\mu_A(x)$ ne prend que des valeurs à 0 ou 1, le sous ensemble n'est rien d'autre qu'un sous ensemble classique.

3.2.2 Caractéristique d'un sous ensemble flou

Un sous ensemble flou est complètement défini par sa fonction d'appartenance. À partir d'une telle fonction, un certain nombre de caractéristiques du sous ensemble flou peuvent être déduites

3.2.2.1 Noyau

Le noyau d'un sous ensemble flou A de X, noté $\text{Noy}(A)$ est l'ensemble des éléments de X pour lesquelles la fonction d'appartenance normalisée de A vaut 1 (voir Fig 3.2)

$$\text{Noy}(A) = \{x \in X / \mu_A(x) = 1\}$$

$\text{Noy}(A) = \emptyset$ on dit que le sous ensemble flou est normalisé.

3.2.2.2 Support

Le support d'un sous ensemble flou A, noté $\text{Supp}(A)$ est la partie de X sur laquelle la fonction d'appartenance $\mu_A(x)$ n'est pas nulle (voir Fig 3.2)

$$\text{Supp}(A) = \{x \in X / \mu_A(x) \neq 0\}$$

3.2.2.3 Hauteur

La hauteur notée $h(A)$, d'un sous ensemble flou A de X est la plus grande valeur prise par sa fonction d'appartenance (voir Fig 3.2)

$$h(A) = \text{Sup } \mu_A(x) = \text{Max } (\mu_A(x)) \quad / \quad x \in X$$

$h(A)=1$, on dit que le sous ensemble est normalisée et identique à son support et à son noyau.

3.2.2.4 Coupe de niveau ou α -coupe

Pour un seuil donné α de l'intervalle $[0 \ 1]$, la α coupe du sous ensemble flou A de X est définie comme les sous ensembles : voir Fig 3.2

$$A_\alpha = \{x \in X / \mu_A(x) \geq \alpha\}$$

A_α est un sous ensemble ordinaire de fonction caractéristique :

$$\chi_{A_\alpha} = 1 \text{ si seulement si } \mu_A(x) \geq \alpha$$

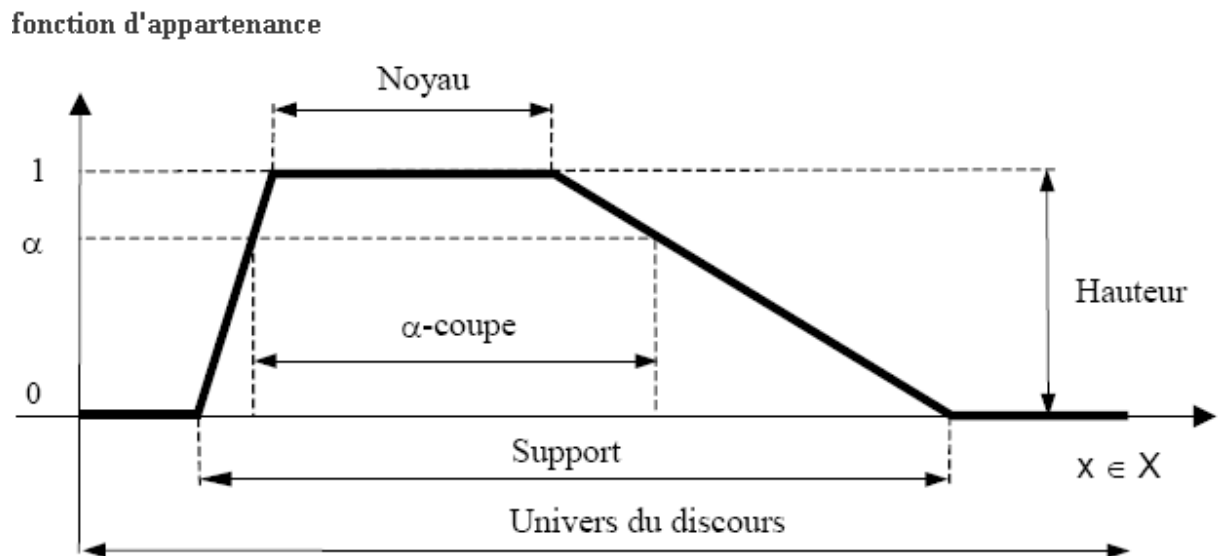


Fig 3.2 : support, hauteur, noyau et α coupe

d'un sous ensemble flou.

3.2.2.5 Cardinalité

La cardinalité d'un sous ensemble flou A de X est le degré global avec lequel des éléments de X appartiennent à A, elle est définie par :

$$|A| = \sum_{x \in X} \mu_A(x)$$

3.2.2.6 Sous ensemble flou convexe

Un sous ensemble flou A de l'ensemble X est convexe si pour tout couple d'éléments a et b de X et pour tout nombre λ de $[0, 1]$, la fonction d'appartenance de A vérifie :

$$\mu_A(\lambda a + (1-\lambda)b) \geq \min(\mu_A(a), \mu_A(b))$$

3.2.3 Opérations sur les sous ensembles flous [13]

Les opérations existant sur les sous ensembles ordinaires d'un même univers de référence peuvent être utiles aussi pour les sous ensembles flous.

3.2.3.1 Egalité

Deux sous ensembles flous A et B de X sont égaux si leurs fonctions d'appartenance prennent la même valeur pour tout élément de X :

$$\forall x \in X, \mu_A(x) = \mu_B(x)$$

3.2.3.2 Inclusion

Soit deux sous ensembles flous, A et B de X, le sous ensemble A est inclus dans le sous ensemble B, ($A \subseteq B$) si leurs fonctions d'appartenance sont telles que :

$$\forall x \in X, \mu_A(x) \leq \mu_B(x)$$

3.2.3.3 Intersection

L'intersection de deux sous ensembles flous A et B de X correspond à la propriété $\text{Prop}(A \cap B)$ selon la quelle les propriétés $\text{Prop}(A)$ et $\text{Prop}(B)$ sont satisfaites simultanément. Cette propriété, $\text{Prop}(A \cap B)$, est satisfaites avec un degré égale au plus petit degré avec lequel les deux propriétés $\text{Prop}(A)$ et $\text{Prop}(B)$ sont satisfaites.

L'intersection de deux sous ensembles flous A et B de X est un sous ensemble C noté $A \cap B$ telle que (voir Fig 3.3)

$$\forall x \in X \quad , \quad \mu_C(x) = \min (\mu_A(x) , \mu_B(x))$$

fonction d'appartenance

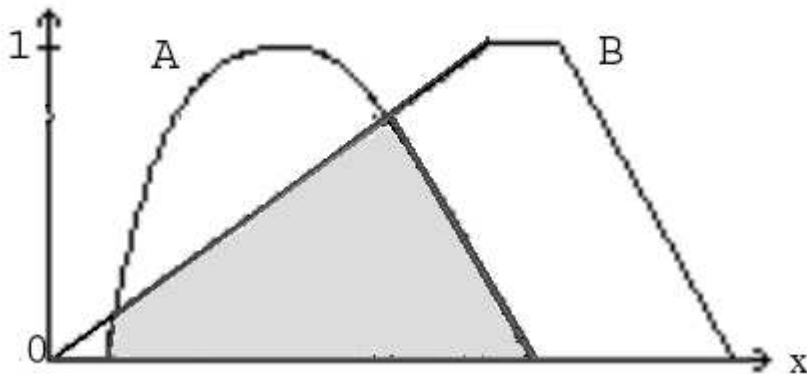


Fig 3.3: Intersection de deux sous ensembles flous avec l'opérateur min

3.2.3.4 Union

L'union de deux sous ensembles flous A et B de X correspond à la propriété $\text{Prop}(A \cup B)$ selon la quelle l'une au moins des propriétés $\text{Prop}(A)$ et $\text{Prop}(B)$ est satisfaite.

Cette propriété, $\text{Prop}(A \cup B)$, est satisfaites avec un degré égale au plus grand degré avec lequel les deux propriétés $\text{Prop}(A)$ et $\text{Prop}(B)$ sont satisfaites.

L'union de deux sous ensembles flous A et B de X est un sous ensemble D noté $A \cup B$ telle que (voir Fig 3.4)

$$\forall x \in X \quad , \quad \mu_D(x) = \max (\mu_A(x) , \mu_B(x))$$

fonction d'appartenance

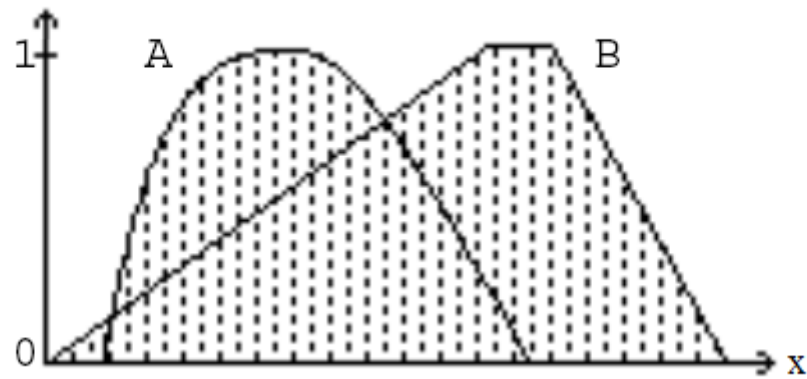


Fig 3.4 : Union de deux sous ensembles flous avec l'opérateur max

3.2.3.5 Complément d'un sous ensemble flou

Le complément A^c d'un sous ensemble flou A de X est défini comme le sous ensemble flou de X de fonction d'appartenance.

$$\forall x \in X, \mu_{A^c}(x) = 1 - \mu_A(x)$$

3.2.4 Exemple de fonction d'appartenance [13]

Plusieurs types de fonctions d'appartenance sont utilisés dans la logique floue, les plus répandus sont les fonctions d'appartenance trapézoïde et exponentielle.

3.2.4.1 Fonction d'appartenance trapézoïde

La fonction d'appartenance trapézoïde (Fig 3.5) est donnée par l'équation suivante :

$$\mu_A(x) = \begin{cases} 0 & \text{si } x \leq a-\alpha \text{ ou } x \geq b+\beta \\ x/\alpha + (1-a/\alpha) & \text{si } a-\alpha < x < a \\ 1 & \text{si } a \leq x \leq b \\ -x/\beta + (1+b/\beta) & \text{si } b < x < b+\beta \end{cases}$$

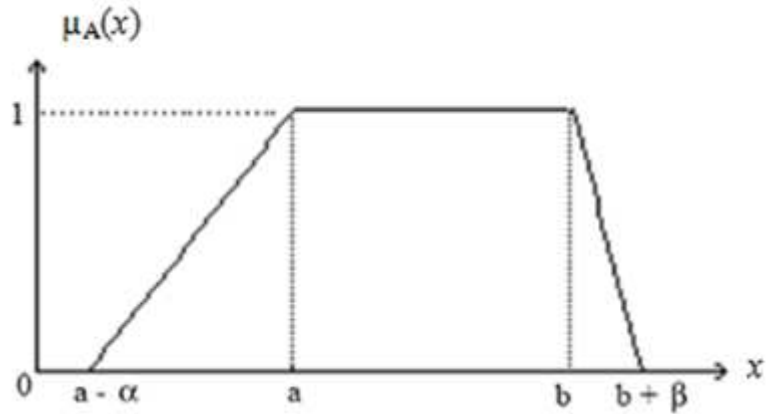


Fig 3.5 : Fonction d'appartenance.

Dans le cas ou $a=b$ la fonction est dite triangulaire

3.2.4.2 Fonction d'appartenance exponentielle

La fonction d'appartenance exponentielle (Fig 3.6) est définie par la fonction d'appartenance suivante :

$$\mu_A(x) = \begin{cases} \exp(x-a) & \text{si } x < a \\ 1 & \text{si } a \leq x \leq b \\ \exp(-x+b) & \text{si } x > b \end{cases}$$

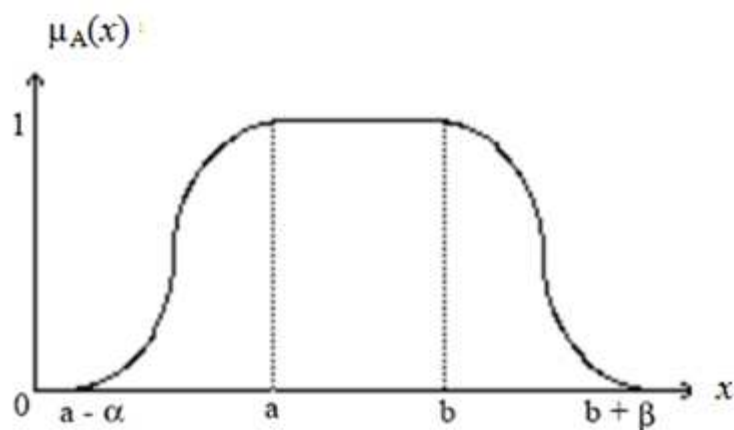


Fig 3.6 : Fonction d'appartenance.

3.2.5 Indices de comparaison [13]

3.2.5.1 Indice de chevauchement

Ils mesurent le degré de chevauchement de deux sous ensembles flous, ils sont définis par :

$$a) \quad I_1 (A \cap B) = \frac{|A \cap B|}{|A|} = |(A \cap B)|$$

$$b) \quad I_2 (A \cap B) = h (A \cap B)$$

h désigne la hauteur

3.2.5.2 Indice d'inclusion

Plusieurs définitions ont été proposées pour le degré d'inclusion de A dans B :

$$a) \quad Inc_1 (AB) = \frac{|A \cap B|}{|A|}$$

$$b) \quad Inc_2 (AB) = 1 - h |A \cap B^c|$$

3.2.5.3 Indice d'égalité

Ils permettent d'évaluer l'égalité de deux sous ensembles flous, ils sont donnés par :

$$EG (A , B) = \min (A \subseteq B , A \supseteq B)$$

3.2.5.4 Indice flou

On calcule l'indice flou pour évaluer à quel point un sous ensemble flou A et son complémentaire se chevauchent. Il est noté par $\mu(A)$ tel que :

$$\mu(A) = 1 - \frac{\sum_{x \in X} |\mu_A(x) - \mu_{A^c}(x)|}{|A|}$$

3.2.6 Le principe d'extension

Ce principe a été introduit par **Zadeh** et constitue l'un des concepts les plus importants de la théorie des ensembles flous. Il permet l'extension d'une application f d'un premier ensemble de référence X vers un second ensemble de référence Y, éventuellement identique au premier, il est alors possible de définir un ensemble flou B de Y à partir de la donnée d'un ensemble flou A de X :

$$\forall y \in Y, \mu_B(y) = \begin{cases} \sup \{ \mu_A(x) \mid y=f(x) \} & \text{si } \{ x \in X, y=f(x) \} \neq \emptyset \\ 0 & \text{sinon} \end{cases}$$

3.2.7 Produit cartésien

Soient des sous ensembles flous $A_1, A_2, A_3, \dots, A_r$ respectivement définis sur $X_1, X_2, X_3, \dots, X_r$. Leur produit cartésien $A = A_1 \times A_2 \times A_3 \times \dots \times A_r$ est défini comme un sous ensemble flou de X tel que:

$$\forall x = (x_1, x_2, \dots, x_r) \in X, \mu_A(x) = \min(\mu_{A_1}(x_1), \mu_{A_2}(x_2), \dots, \mu_{A_r}(x_r))$$

3.2.8 Projection

La connaissance d'une caractérisation floue globale définie sur un univers complexe (n dimension), doit permettre :

- De donner des informations sur les différentes composantes de cet univers.
- De définir des caractérisations floues sur chacune de ces composantes.

Soit un ensemble flou A défini sur un univers $X = X_1 \times X_2$, la projection de A sur X_1 est définie par :

$$\forall x_1 \in X_1, \mu_{proj(A, X_1)}(x_1) = \sup_{x_2 \in X_2} [\mu_A(x_1, x_2)]$$

3.2.9 Schéma d'un traitement flou

La mise en œuvre d'un traitement flou fait apparaître trois grands modules (Fig 3.7)

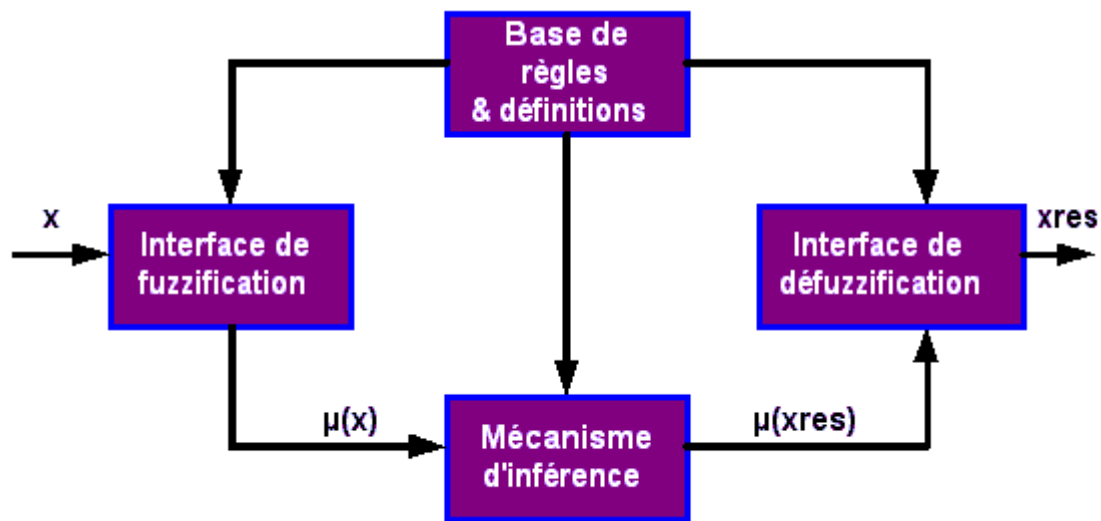


Fig 3.7 : Schéma synoptique d'un traitement flou

X : Vecteur des entrées.

X_{res} : Vecteurs de sortie.

$\mu(x)$, $\mu(x_{\text{res}})$: Les fonctions d'appartenances correspondantes.

3.2.9.1 Fuzzification

La fuzzification, consiste à attribuer à la valeur réelle de chaque entrée, au temps t , sa fonction d'appartenance à chacune des classes préalablement définies donc à transformer l'entrée réelle en un sous ensemble flou.

3.2.9.2 Traitement flou

Base de règles

C'est une base de connaissance qui contient les règles floues décrivant le comportement du système.

Mécanisme d'inférence

Elles constituent l'ensemble de règles floues permettant de décrire sous forme règles linguistiques une fonction de transfert entre les variables d'entrée et celles de sorties d'un système donné.

3.2.9.3 Défuzzification

La défuzzification est la transformation inverse de la fuzzification, elle permet de passer d'un degré d'appartenance d'une variable floue de sortie à la détermination de la valeur réelle à donner à cette sortie.

3.3 Conclusion

Dans ce chapitre, nous avons présenté quelques notions sur la logique floue qui est basée sur la théorie des sous ensemble flous.

Elle permet de considérer les classes d'objets dont les frontières ne sont pas clairement définies. Nous introduisons dans le chapitre suivant cette notion logique floue dans le processus de segmentation d'images sous forme d'un algorithme très populaire connu sous le nom de FCM (Fuzzy C-Means) car l'information fournis par les images sont souvent incertaines (bruits, blocs manquants...) et imprécises (transition entre régions).

Chapitre 4 :

L'algorithme FCM et ses variantes

4.1 Introduction

Parmi les algorithmes de classification non supervisée, l'algorithme FCM est très populaire [14]. Une étude comparative entre l'algorithme FCM et plusieurs autres méthodes de classification a été menée par **Mingoti** et **Lima** (2006) [15] en utilisant plusieurs exemples de données formées de classes chevauchantes ou non avec des variables corrélées ou non et en absence ou en présence du bruit. L'algorithme FCM s'est montré supérieur dans la plupart des cas. Il affiche les meilleures performances même en présence des classes chevauchantes et du bruit. De plus, l'algorithme FCM est très utilisé dans plusieurs applications : médicales, spatiales, diagnostic, etc.

Son inconvénient réside principalement dans l'initialisation des centres, et sa sensibilité au bruit. De plus, le nombre de classes doit être connu à priori. Pour ces raisons, d'autres variantes que nous décrirons dans ce chapitre ont été proposées afin d'améliorer ce dernier [16–28].

4.2 FCM (Fuzzy C- Means) [29]

Fuzzy C-Means est un algorithme de classification non supervisée floue, issu de l'algorithme K-Means, il est basé sur l'appartenance des données à chaque classe avec un certain degré et toutes les classes sont caractérisées par leurs centres de gravité.

Comme les autres algorithmes de classification non supervisée, il utilise un critère de minimisation des distances intra-classes, mais en donnant un certain degré d'appartenance à chaque classe pour chaque pixel.

Cet algorithme nécessite la connaissance préalable du nombre de classes et génère les classes par un processus itératif en minimisant une fonction objective. Il permet ainsi d'obtenir une partition floue des données en donnant à chaque donnée (pixel) un degré d'appartenance (compris entre 0 et 1) à une classe donnée. La classe à laquelle est associé un pixel est celle dont le degré d'appartenance sera le plus élevé.

Soit $X = \{x_j / j=1 \dots N\}$ un ensemble de N observations (pixels) à classer en K classes ou chaque pixel x_j peut être représenté par un ensemble de D attributs : $X_j = (x_{j1}, x_{j2}, \dots, x_{jD})^T$

Les K classes $(c_1, c_2, \dots, c_K)$ peuvent être représenté par leurs centres $\{\bar{X}/i = 1, \dots, K\}$ tel que $\bar{X}_i = (\bar{x}_{i1}, \bar{x}_{i2}, \dots, \bar{x}_{iD})^T$.

La fonction objective de FCM pour partitionner $\{x_j / j=1 \dots N\}$ en K classes est donnée par :

$$J = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \|x_j - \bar{x}_i\|^2 = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m d^2(x_j, \bar{x}_i) \quad (4.1)$$

$$\text{ou } d^2(x_j, \bar{x}_i) = d_i^2 = \|x_j - \bar{x}_i\|^2 = \sum_{k=1}^D (x_{kj} - \bar{x}_{ki})^2$$

μ_{ij} est le degré d'appartenance du $j^{ième}$ observation à la classe i , la matrice $\mu_{K \times N} [\mu_{ij}]$ est appelée matrice de K partitions floues. Les degrés d'appartenances doivent satisfaire les conditions suivantes :

$$\forall i \in \{1 \dots K\}, \forall j \in \{1 \dots N\}, \quad \left\{ \begin{array}{l} \mu_{ij} \in [0 \ 1] \\ 0 < \sum_{j=1}^N \mu_{ij} < N \end{array} \right. \quad (4.2)$$

$$\forall j \in \{1 \dots N\}, \sum_{i=1}^K \mu_{ij} = 1 \quad (4.3)$$

Le problème de partition de X en K classes floues peut être formulé comme la minimisation d'une fonctionnelle J (4.1) sous les contraintes (4.2) et (4.3).

$m > 1$ est un paramètre contrôlant le degré de flou de la partition résultante.

$d^2(x_j, \bar{x}_i)$ est une distance du vecteur x_j au centre de la classe i noté $\bar{x}_i$.

La solution au problème de minimisation précédent est obtenue par la méthode des multiplicateurs de Lagrange. Plus précisément, le Lagrangien $L(x_j)$ est défini pour chaque vecteur x_j par :

$$L(x_j) = \sum_{i=1}^K \mu_{ij}^m d^2(x_j, \bar{x}_i) - \alpha (\sum_{i=1}^K \mu_{ij} - 1) \quad (4.4)$$

α étant le multiplicateur de Lagrange tel que $\alpha > 0$.

L'annulation des dérivées partielles par rapport à μ_{ij} et α donne :

$$\frac{\partial L(x_j)}{\partial \alpha} = 0 \quad \text{et} \quad \frac{\partial L(x_j)}{\partial \alpha} = m \cdot \mu_{ij}^{m-1} d^2(x_j, \bar{x}_i) - \alpha = 0 \quad (4.5)$$

Si $d^2(x_j, \bar{x}_i) \neq 0$, nous obtenons :

$$\sum_{i=1}^K \mu_{ij} = 1$$

et

$$\mu_{ij} = \left(\frac{\alpha}{m d^2(x_j, \bar{x}_i)} \right)^{\frac{1}{m-1}} \quad (4.6)$$

$$\text{avec} \quad \sum_{i=1}^K \mu_{ij} = \left(\frac{\alpha}{m} \right)^{\frac{1}{m-1}} \sum_{i=1}^K \left(\frac{1}{d^2(x_j, \bar{x}_i)} \right)^{\frac{1}{m-1}} = 1 \quad (4.7)$$

De cette équation, on tire $\left(\frac{\alpha}{m} \right)$ qu'on remplace dans l'équation (4.6). On obtient alors :

$$\mu_{ij} = \left[\sum_{k=1}^K \left(\frac{d^2(x_j, \bar{x}_i)}{d^2(x_j, \bar{x}_k)} \right)^{\frac{1}{m-1}} \right]^{-1} \quad (4.8)$$

Si $d^2(x_j, \bar{x}_i) = 0$, x_j est un prototype $\bar{x}_i$ et $\mu_{ij} = 1$ avec $\mu_{kj} = 0$, $k \neq i$

L'annulation des dérivées partielles par rapport à $\bar{x}_i$ donne :

$$\bar{x}_i = \frac{\sum_{k=1}^N \mu_{ik}^m x_k}{\sum_{k=1}^N \mu_{ik}^m} \quad (4.9)$$

L'algorithme FCM peut être résumé comme suit :

Etape 1 : Fixer les paramètres ;

-le nombre de classes

-le seuil ε représentant l'erreur de convergence

-le degré flou généralement pris égal à 2.

Etape 2 : Initialiser la matrice degrés d'appartenances U par des valeurs aléatoires dans l'intervalle $[0,1]$.

Etape 3 : Mettre à jour la matrice prototype par la relation (4.9).

Etape 4 : Mettre à jour la matrice degrés d'appartenance U par la relation (4.8).

Etape 5 : Répéter les étapes 3 et 4 jusqu'à satisfaction du critère d'arrêt qui s'écrit :

$$\|U_{ancien} - U_{nouveau}\| < \varepsilon .$$

L'algorithme FCM n'est pas robuste face au bruit et est sensible à l'initialisation des centres des classes. Plusieurs extensions ont donc été développées pour surmonter ces inconvénients.

4.3 Variantes de la méthode FCM

Depuis son apparition en 1975, l'algorithme FCM continue de susciter un grand intérêt. De nombreuses variations ont été ainsi proposées. Celles-ci consistent soit à changer la fonctionnelle à minimiser, soit à définir une autre métrique d^2 ou soit encore à modifier l'influence du facteur flou m .

4.3.1 Variantes de la distance

La version standard de FCM utilise la distance euclidienne pour mesurer la distance entre un objet et le centre d'une classe. Les classes ainsi détectées auront seulement une forme sphérique. Or en réalité les classes peuvent avoir des formes géométriques quelconques.

Ainsi afin de détecter ce type de classes, plusieurs variantes de FCM utilisant d'autres types de distances ont été proposées.

4.3.1.1 Algorithme de Gustafson et Kessel [30]

Gustafson et **Kessel** en 1978 ont proposé une variante des FCM adaptée à des ensembles dont les classes présentent des formes différentes. A chaque classe K est associée une matrice carrée A_i à p lignes et p colonnes, p étant la dimension de l'espace des données.

Les matrices $A_1, A_2, \dots, A_k$, qui dépendent des x_k , sont définies positives et sont utilisés pour le calcul des distances . Leurs déterminants sont fixés telles que:

$$\det(A_i) = \rho_i$$

avec $\rho_i > 0$ est un facteur fixé à priori pour chaque classe.

$$d_{ij}^2 = \|x_j - \bar{x}_i\|_{A_i}^2 = (x_j - \bar{x}_i)^T A_i (x_j - \bar{x}_i) \quad (4.10)$$

La fonction objective utilisée par **Gustafson** et **Kessel** est donc la suivante :

$$J^{GK} = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \|x_j - \bar{x}_i\|_{A_i}^2 \quad (4.11)$$

La formule permettant le calcul des différents centres reste inchangée. Celle relative aux degrés d'appartenance est légèrement modifiée en faisant intervenir les matrices A_i pour le calcul des distances et non plus la matrice unité correspondant à la distance euclidienne.

Ces matrices sont également remises à jour à chaque itération :

$$A_i = [\rho_i (\det S_i)]^{\frac{1}{p}} S_i^{-1} \quad (4.12)$$

Où la matrice de covariance floue de la classe i est définie par :

$$S_i = \frac{\sum_{j=1}^N \mu_{ij}^m (x_j - \bar{x}_i)(x_j - \bar{x}_i)^T}{\sum_{j=1}^N \mu_{ij}^m} \quad (4.13)$$

La mise à jour des centres des classes et des degrés d'appartenance se fait d'une manière identique à celles de l'algorithme FCM.

Gustafson et **Kessel** ont en particulier utilisé leur approche pour discriminer deux classes allongées suivant deux directions orthogonales et ayant le même centre de gravité.

4.3.1.2 Algorithme de Gath et Geva [30]

Gath et **Geva** ont proposé une variante de FCM afin de détecter des classes hyper-ellipsoïdales ayant des densités différentes en utilisant une distance exponentielle, $d_e(x_j, \bar{x}_i)$, entre chaque objet x_j et le centre $\bar{x}_i$ de chaque classe. Cette distance est utilisée dans le calcul des degrés d'appartenance de la manière suivante :

$$\mu_{ij} = \frac{\frac{1}{d_e^2(x_j - \bar{x}_i)}}{\sum_{k=1}^K \frac{1}{d_e^2(x_j - \bar{x}_k)}} \quad (4.14)$$

La distance d_e est définie par la relation suivante :

$$d_e^2(x_j - \bar{x}_i) = \frac{[\det F_i]^{1/2}}{P_i} \exp \left[(x_j - \bar{x}_i)^T F_i^{-1} (x_j - \bar{x}_i) / 2 \right] \quad (4.15)$$

où F_i ; est la matrice floue de covariance de la $i^{ième}$ classe définie par :

$$F_i = \frac{\sum_{j=1}^N P(i/x_j) (x_j - \bar{x}_i)(x_j - \bar{x}_i)^T \sum_{j=1}^N \mu_{ij} (x_j - \bar{x}_i)(x_j - \bar{x}_i)^T}{\sum_{j=1}^N P(i/x_j) \sum_{j=1}^N \mu_{ij}} \quad (4.16)$$

P_i est la probabilité à priori de cette même classe

Les degrés d'appartenance sont assimilés à des probabilités conditionnelles :

$$\mu_{ij} = P(i/x_j) \quad (4.17)$$

C'est-à-dire que le degré d'appartenance du vecteur x_k à la classe i est égal à la probabilité conditionnelle de choisir la classe i étant donnée le vecteur x_k .

Les auteurs présentent le nouvel algorithme résultant comme étant la généralisation au cas flous de l'estimation par le maximum de vraisemblance.

La boucle de calcul comporte maintenant deux étapes supplémentaires pour le calcul de probabilité à priori P_i et des matrices floues de covariances F_i .

4.3.1.3 Algorithme de Shen et al [31]

Dans le cadre de la segmentation d'images, **Shen et al** (2005) ont proposé la fonction distance suivante :

$$d^2(x_i, c_k) = \|x_i, c_k\|^2 (1 - \lambda H_{ki} - \varepsilon F_{ki}) \quad (4.18)$$

où H_{ki} représente un paramètre d'attraction et F_{ki} une distance d'attraction telle que :

$$H_{ki} = \frac{\sum_{j=1}^{N_v} \mu_{kj} g_{ij}}{\sum_{j=1}^{N_v} g_{ij}} \quad (4.19)$$

$$F_{ki} = \frac{\sum_{j=1}^{N_v} \mu_{kj}^2 \mu_{ij}^2}{\sum_{j=1}^{N_v} g_{ij}} \quad (4.20)$$

avec $g_{ij} = |x_i - x_j|$ et $q_{ij} = (a_i - a_j)^2 + (b_i - b_j)^2$

(a_i, b_i) et (a_j, b_j) représentent les coordonnées du pixels i et ceux de son voisin j .

N_v étant le nombre de pixels voisins.

λ et ε sont des nombres réels qui permettent d'ajuster la contribution de chaque terme d'attraction $0 < \lambda < 1$ et $0 < \varepsilon < 1$. **Forouzanfar et al** (2010) [32] ont proposé récemment une manière de les ajuster automatiquement.

La mise à jour des centres de classes et des degrés d'appartenance se fait de la même manière que pour l'algorithme FCM mais en utilisant la distance définie précédemment.

4.3.1.4 Algorithme de Wang et al [33]

La distance proposée par **Wang et al** (2004) est :

$$d_{ij}^{(w)} = \sqrt{\sum_{k=1}^D w_k^2 (x_{ik} - x_{jk})^2} \quad (4.21)$$

où w_k est un poids qui module la contribution de chaque attribut D.

Pour déterminer w_k , il suffit de minimiser l'expression suivante :

$$E(w) = \frac{2}{w(w-1)} \sum_{i=1}^K \sum_{\substack{j=1 \\ i \neq j}}^N \frac{1}{2} \left[\rho_{ij}^{(w)} (1 - \rho_{ij}) + \rho_{ij} (1 - \rho_{ij}^{(w)}) \right] \quad (4.22)$$

où $\rho_{ij}^{(w)} = \frac{1}{1 + \beta d_{ij}^{(w)}}$ est une mesure de similarité entre deux observations i et j.

Pour $w = (1, 1, \dots, 1)$, $\rho_{ij}^{(w)} = \rho_{ij}$.

β étant une constante positive.

La minimisation de $E(w)$ peut se faire par la méthode de gradient

$$w_k^{(t)} = w_k^{(t-1)} + \Delta w_k \quad (4.23)$$

$$\Delta w_k = -\eta \frac{\partial E(w)}{\partial w_k} \quad (0 < k < D) \quad (4.24)$$

Notons que comme précédemment, la mise à jour des centres des classes et des degrés d'appartenance utilisent les mêmes formules de l'algorithme FCM mais avec la distance $d_{ij}^{(w)}$.

4.3.2 Variantes de fuzzification

D'autres variantes de FCM sont basées sur l'étude de la fuzzification ; tel que les degrés d'appartenances appartiennent à $[0, 1]$ et non pas $\{0, 1\}$.

4.3.2.1 Algorithme de Rousseeuw et al [30]

Rousseeuw, **Trauwaert** et **Kaufman** (1995) ont proposé une version des FCM en remplaçant la quantité μ_{ij} par une fonction $f(\mu_{ij})$, telle que :

$$f(\mu_{ij}) = \alpha\mu_{ij} + (1 - \alpha)\mu_{ij}^2 \quad (4.25)$$

α étant un nombre réel dans l'intervalle $[0,1]$. La fonction objective devient alors :

$$J = \sum_{i=1}^K \sum_{j=1}^N [\alpha\mu_{ij} + (1 - \alpha)\mu_{ij}^2] d_{ij}^2 \quad (4.26)$$

Lorsque α tend vers 1 ou vers 0, $f(\mu_{ij})$ se rapproche de μ_{ij} ou de μ_{ij}^2 respectivement :

$$f(\mu_{ij}) \rightarrow \mu_{ij} \text{ lorsque } \alpha \rightarrow 1$$

$$f(\mu_{ij}) \rightarrow \mu_{ij}^2 \text{ lorsque } \alpha \rightarrow 0$$

La valeur du paramètre α permet donc de moduler le contraste entre les degrés d'appartenance élevés et ceux qui sont faibles.

Les centres des classes sont mis à jour à partir de la formule suivante:

$$\bar{x}_i = \frac{\sum_{j=1}^N f(\mu_{ij})x_j}{\sum_{j=1}^N f(\mu_{ij})} \quad (4.27)$$

4.3.2.2 Algorithme de Klawonn et Hoppner [30]

Klawonn et **Hoppner** prennent également en tant que point de départ les degrés d'appartenance qui ne prennent réellement jamais les valeurs 0 et 1. Ils proposent comme fonction objective :

$$J = \sum_{i=1}^K \sum_{j=1}^N g(\mu_{ij}) d_{ij}^2 \quad (4.28)$$

$$\text{avec } g(\mu) = \alpha \mu^2 + (1 - \alpha) \mu, \quad 0 < \alpha < 1 \quad (4.29)$$

La minimisation de cette fonction conduit à :

$$\mu_{ij} = \frac{1}{1-\beta} \left[\frac{1+(\hat{c}-1)\beta}{\sum_{\substack{l=1 \\ \mu_{lj} \neq 0}}^K \frac{d_{lj}}{d_{ij}}} - \beta \right] \quad (4.30)$$

où $\hat{c}$ représente le nombre de classes pour un objet x_j ayant des degrés d'appartenance non nul.

$$\beta = \frac{g'(0)}{g'(1)} = \frac{1-\alpha}{1+\alpha} \quad (4.31)$$

La mise à jour des centres des classes reste la même que dans le cas du FCM.

Une autre fonction $g(\mu) = \frac{(e^{\alpha\mu}-1)}{(e^{\alpha}-1)}$ a été proposée.

Dans ce cas on a :

$$\mu_{ij} = \frac{1}{\alpha c} \left[\alpha + \sum_{\substack{l=1 \\ \mu_{lj} \neq 0}}^K \ln \left(\frac{d_{lj}}{d_{ij}} \right) \right] \quad (4.32)$$

4.3.3 Variantes de la fonction objective

D'autres variantes de FCM agissent directement sur la fonction objective afin d'améliorer les résultats de la classification notamment en présence de bruit. Celles-ci introduisent un autre terme dans la fonction objective initiale.

4.3.3.1 Algorithme de Frigui et krishnapuram[30]

Frigui et **Krishnapuram** (1997) proposent un algorithme CA (Compétitive Agglomération algorithme) qui combine les approches hiérarchiques et de partitionnement.

Dans un premier temps, un grand nombre de classes qui rentrent en compétition, puis progressivement les classes qui perdent sont fusionnées.

CA minimise la fonction objective suivante :

$$J_{CA} = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \|x_j - \bar{x}_i\|^2 - \alpha \sum_{i=1}^K \left[\sum_{j=1}^N \mu_{ij}^m \right]^2 \quad (4.33)$$

La minimisation de cette fonction par les multiplicateurs de Lagrange aboutit à :

$$\mu_{ij} = \alpha \frac{N_i}{d^2(x_j, \bar{x}_i)} + \frac{1 - \alpha \sum_{k=1}^K \frac{N_K}{d^2(x_j, \bar{x}_k)}}{\sum_{k=1}^K \frac{d^2(x_i, \bar{x}_j)}{d^2(x_i, \bar{x}_k)}} \quad (4.34)$$

$$\text{et} \quad \bar{x}_i = \frac{\sum_{j=1}^N \mu_{ij}^m x_j}{\sum_{j=1}^N \mu_{ij}^m} \quad (4.35)$$

où $N_i = \sum_{j=1}^N \mu_{ij}$ est le cardinal de la classe i .

α est un paramètre proportionnel à :

$$\beta = \frac{\sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^2 d^2(x_j - \bar{x}_i)}{\sum_{i=1}^K [\sum_{j=1}^N \mu_{ij}]} \quad (4.36)$$

L'algorithme CA est un algorithme itératif où initialement le nombre de classes est surestimé puis à chaque itération ce nombre décroît jusqu'à atteindre le nombre optimal. Dans ce cas $\alpha(t) = \eta(t)\beta$ avec $\eta(t) = \eta_0 e^{-t/\tau}$.

A chaque itération, les classes ayant un cardinal faible sont éliminées.

L'un des inconvénients de cette méthode est que le nombre exact de classe n'est pas requis à priori.

4.3.3.2 Algorithme Pham [30]

Dans le cadre de la segmentation d'image IRM, **Pham** opte pour un critère différent qui consiste à contraindre la classification par les appartenances des pixels voisins.

Il propose donc la fonction objective suivante :

$$J = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \|x_j - \bar{x}_i\|^2 + \frac{\beta}{2} \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \sum_{l \in N_j} \sum_{k \in M_i} \mu_{lk}^m \quad (4.37)$$

Où β est un paramètre qui pondère l'influence du deuxième terme de la fonction objective.

N_j représente un ensemble de points voisins du point j et $M_i = \{1, \dots, K\} \setminus \{i\}$ représente l'ensemble des classes à l'exception de la classe i .

L'appartenance d'un pixel à une classe est conditionnée par l'appartenance de ses voisins aux autres classes.

Cette méthode dépend donc des valeurs des autres classes et non pas uniquement de la classe envisagée. Comme il n'y a pas de contraintes sur les centres des classes l'équation reste identique à celle de FCM. En revanche la fonction de partition est différente mais obtenue également par optimisation Lagrangienne.

$$\mu_{ij} = \frac{[d^2(x_j, \bar{x}_i) + \beta \sum_{l \in N_j} \sum_{k \in M_i} \mu_{lk}^m]^{\frac{-1}{(m-1)}}}{\sum_{n=1}^K [d^2(x_j, \bar{x}_n) + \beta \sum_{l \in N_j} \sum_{k \in M_n} \mu_{lk}^m]^{\frac{-1}{(m-1)}}} \quad (4.38)$$

L'algorithme procède les mêmes étapes que l'algorithme FCM.

4.3.3.3 Possibilitic-C-Means (PCM) [34]

Krishnapuram et **Keller** ont introduit la notion de possibilité très voisine du flou ou d'autres extensions des probabilités comme l'évidence. Les auteurs considèrent le degré d'appartenance d'un point à une classe comme une possibilité d'appartenir à une classe. Contrairement au FCM, la somme des degrés d'appartenance de toutes les classes n'est pas égale à 1.

La fonction objectif ainsi modifiée est :

$$J_m = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m d_{ij}^2 + \sum_{i=1}^K \eta_i \sum_{j=1}^N (1 - \mu_{ij})^m \quad (4.39)$$

telle que $\sum_{j=1}^N \mu_{ij} > 0 \quad \forall i \in \{1, \dots, K\}$.

$d_{ij}^2 = \|x_j - \bar{x}_i\|^2$: est la distance d'un échantillon j au centre de classe i .

η_i est un nombre positif qui peut être interprété comme la distance à laquelle le degré d'appartenance est égale à 0.5.

Le premier terme est identique à celui de FCM, le second force μ_{ik} à la plus grande valeur possible.

En dérivant la formule J_m par rapport à μ_{ij} , et en annulant cette dérivée, on obtient :

$$\mu_{ij} = \left[1 + \frac{d_{ik}^2}{\eta_i} \right]^{-1} \quad (4.40)$$

En dérivant la formule J_m par rapport à $\bar{x}_i$ et en annulant cette dérivée on obtient :

$$\bar{x}_i = \frac{\sum_{j=1}^N \mu_{ij}^m x_j}{\sum_{j=1}^N \mu_{ij}^m} \quad (4.41)$$

η_i peut être estimé par la formule suivante:

$$\eta_i = \frac{\sum_{j=1}^N \mu_{ij}^m d_{ij}^2}{\sum_{j=1}^N \mu_{ij}^m} \quad (4.42)$$

Les centres de classes sont mis à jour comme dans FCM .Un bon choix de m pour le PCM est $\approx 1,5$ et non pas 2 comme dans FCM.

Cette algorithme est robuste au bruit mais il souffre du même inconvénient que FCM, c'est que le nombre de classe doit être spécifié et nécessite une bonne initialisation et un bon choix de η_i .

4.3.3.4 Algorithme de Wu et al [35]

Wu et al (2005) proposent la fonction objective suivante :

$$J = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \|x_j - \bar{x}_i\|^2 - \sum_{i=1}^K \eta_i \sum_{j=1}^N \mu_{ij}^m \|\bar{x}_i - \bar{x}\|^2 \quad (4.43)$$

ou $\eta_i \geq 0$, $\bar{x}$ est centre de tous les points : $\bar{x} = \frac{1}{N} \sum_{j=1}^N x_j$

Le deuxième terme de J indique la variance inter-classe.

En minimisant J par rapport à μ_{ij} , $\bar{x}_i$ et η_i , on obtient :

$$\mu_{ij} = \frac{(\|x_j - \bar{x}_i\|^2 - \eta_i \|\bar{x}_i - \bar{x}\|^2)^{-1/m-1}}{(\sum_{k=1}^K \|x_j - \bar{x}_k\|^2 - \eta_k \|\bar{x}_k - \bar{x}\|^2)^{-1/m-1}} \quad (4.44)$$

$$\bar{x}_i = \frac{\sum_{j=1}^N \mu_{ij}^m x_j - \eta_i \sum_{j=1}^N \mu_{ij}^m \bar{x}}{\sum_{j=1}^N \mu_{ij}^m - \eta_i \sum_{j=1}^N \mu_{ij}^m} \quad (4.45)$$

$$\eta_i = \frac{\beta / 4 \min_{i \neq i'} \|\bar{x}_i - \bar{x}_{i'}\|^2}{\max_k \|x_k - \bar{x}\|^2} \quad (4.46)$$

β est un paramètre qui contrôle la taille d'une classe telle que : $0 \leq \beta \leq 1$, une valeur de 0.01 a été prise par les auteurs.

4.3.3.5 Algorithme d'Ahmed et al [36]

Ahmed et **al** proposent une modification de la fonction objective de FCM en rajoutant un 2^{ième} terme qui représente la contrainte spatiale exercé par les voisins en chaque pixel d'après leurs distances à une même classe.

Si un pixel est proche d'une classe et ses voisins affichent également une proximité pour ce même centre, alors le degré d'appartenance à cette classe augmente.

Donc un point est contraint de suivre son voisinage, ce qui produit un effet de régularisation :

$$J_m = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \|x_j - \bar{x}_i\|^2 + \frac{\alpha}{N_R} \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \left(\sum_{x_r \in N_j} \|x_r - \bar{x}_i\|^2 \right) \quad (4.47)$$

Où N_k est l'ensemble des voisins

α : Contrôle l'influence de 2^{ième} terme et m : est le degré flou $m \in [1, +\infty[$

La minimisation de la fonction objective J_m sera résolue en utilisant le multiplicateur de Lagrange :

$$L_m = \sum_{i=1}^K \sum_{j=1}^N (\mu_{ij}^m d_{ij} + \frac{\alpha}{N_R} \mu_{ij}^m \gamma_i) + \lambda (1 - \sum_{i=1}^K \mu_{ij}) \quad (4.48)$$

avec : $d_{ij}^2 = \|x_j - \bar{x}_i\|^2$ et $\gamma_i = \left(\sum_{x_r \in N_j} \|x_r - \bar{x}_i\|^2 \right)$

λ étant un multiplicateur de Lagrange.

En dérivant la fonction L_m par rapport à μ_{ij} et en annulant cette dérivée, on obtient :

$$\mu_{ij} = \frac{1}{\sum_{k=1}^K \left(\frac{d_{ij} + \frac{\alpha}{N_R} \gamma_i}{d_{kj} + \frac{\alpha}{N_R} \gamma_k} \right)^{1/m-1}} \quad (4.49)$$

En dérivant L par rapport à $\bar{x}_i$ et en annulant cette dérivée, on obtient :

$$\bar{x}_i = \frac{\sum_{j=1}^N \mu_{ij}^m \left((x_j) + \frac{\alpha}{N_R} \sum_{x_r \in N_j} (x_r) \right)}{(1+\alpha) \sum_{j=1}^N \mu_{ij}^m} \quad (4.50)$$

4.3.3.6 Algorithme de Kang et al [37]

Dans le cadre de la segmentation d'image **Kang et al** (2008) ont apporté une légère modification à la fonction objective proposée par **Ahmed et al** afin de réduire le temps de calcul. Elle consiste à remplacer le terme $\frac{1}{N_R} \sum_{x_r \in N_i} \|x_r - \bar{x}_i\|^2$ par $\|\bar{x}_j^* - \bar{x}_i\|^2$ ou $\bar{x}_j^*$ est le niveau de gris moyen calculé pour chaque pixel j de coordonnées (p,q) .

$$\bar{x}_j^* = \frac{\sum_{s=-r}^r \sum_{t=-r}^r w_{st} x_{p+s, q+t}}{\sum_{s=-r}^r \sum_{t=-r}^r w_{st}} \quad (4.51)$$

avec

$$w_{st} = (x_{p+s, q+t}) = \begin{cases} k_0 & \text{si } |x_{pq} - x_{p+s, q+t}| > k_1 \\ 0 & \text{sinon} \end{cases} \quad (4.52)$$

k_1, k_0 sont des constantes positives telles que : $0 < k_0 < 0.5$ et $0.5 < k_1 < 1$

Notons que $w_{00} = 1$

W_{-1-1}	W_{-10}	W_{-11}
W_{0-1}	W_{00}	W_{01}
W_{1-1}	W_{10}	W_{11}

La fonction objective devient :

$$J = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \|x_j - \bar{x}_i\|^2 + \alpha \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^m \|\bar{x}_j^* - \bar{x}_i\|^2 \quad (4.53)$$

La minimisation de cette fonction objective conduit aux formules de mise à jour pour les centres des classes et les degrés d'appartenances :

$$\bar{x}_i = \frac{\sum_{j=1}^N \mu_{ij}^m (x_j - \alpha \bar{x}_j)}{(1 + \alpha) \sum_{j=1}^N \mu_{ij}^m} \quad (4.54)$$

$$\mu_{ij} = \frac{1}{\sum_{k=1}^K \left(\frac{\|x_j - \bar{x}_i\|^2 + \alpha \|\bar{x}_j - \bar{x}_i\|^2}{\|x_j - \bar{x}_k\|^2 + \alpha \|\bar{x}_j - \bar{x}_k\|^2} \right)^{1/m-1}} \quad (4.55)$$

Cette méthode a été particulièrement appliquée à la détection des plaques dentaires.

4.3.3.7 Algorithme de Li and Mukaidono [38]

Li and **Mukaidono** (2008) proposent de rajouter à la fonction objective de l'algorithme FCM un deuxième terme caractérisé par l'entropie du degré d'appartenance comme suit :

$$J = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij} \|x_j - \bar{x}_i\|^2 + \delta \sum_{i=1}^K \sum_{j=1}^N \mu_{ij} \log \mu_{ij} \quad , (\delta > 0) \quad (4.56)$$

La mise à jour des degrés d'appartenance est réalisée par l'expression suivante :

$$\mu_{ij} = \frac{\exp\left(-\frac{\|x_j - \bar{x}_i\|^2}{\delta}\right)}{\sum_{l=1}^K \left(-\frac{\|x_j - \bar{x}_l\|^2}{\delta}\right)} \quad (4.57)$$

Les centres des classes sont déterminés comme dans l'algorithme FCM:

$$\bar{x}_i = \frac{\sum_{j=1}^N \mu_{ij} x_j}{\sum_{j=1}^N \mu_{ij}} \quad (4.58)$$

Une autre fonction objective a été également proposée par **Li and Mukaidono** (2008) en rajoutant un terme quadratique à la fonction objective du FCM :

$$J = \sum_{i=1}^K \sum_{j=1}^N \mu_{ij} \|x_j - \bar{x}_i\|^2 + \frac{1}{2} \delta \sum_{i=1}^K \sum_{j=1}^N \mu_{ij}^2, \quad (\delta > 0) \quad (4.59)$$

4.3.3.8 Algorithme de Yang et al [39]

Yang et al (2009) proposent une variante de l'algorithme FCM ayant comme fonction objective suivante :

$$J = \sum_{i=1}^K \sum_j \mu_{ij}^2 d_{ij}^2 + \sum_{j=1}^N \alpha_j \sum_{i=1}^K \mu_{ij} (1 - \mu_{ij}) \quad (4.60)$$

$$\text{où } d_{ij}^2 = \|x_j - \bar{x}_i\|^2$$

Le facteur flou m est généralement fixé à 2.

α_j désigne un paramètre de pondération de chaque pixel. Il peut être déterminé par la relation suivante :

$$\alpha_j = \min \{d_{ij}^2 / i \in \{1, 2, \dots, k\}\} - \eta \quad (4.61)$$

η est un paramètre strictement positif ($\eta > 0$), les auteurs ont choisi la valeurs $\eta = 0,01$ pour leurs tests.

La minimisation de J par rapport à μ_{ij} conduit à :

$$\mu_{ij} = \frac{1}{\sum_{k=1}^K (d_{ij}^2 - \min d_{*j}^2) / (d_{kj}^2 - \min d_{*j}^2)} \quad (4.62)$$

La formule du centre d'une classe $\bar{x}_i$ reste identique à celle de l'algorithme FCM.

4.3.3.9 Algorithme de Tang et al [34]

Très récemment **Tang et al** (2010) proposent de minimiser la fonction objective suivante :

$$J = \sum_{j=1}^N e^{tj} \sum_{i=1}^K \mu_{ij}^m d_{ij}^2 \quad (4.63)$$

où e^{tj} est appelé exposant du facteur d'impact du $j^{ième}$ point et t le facteur poids de x_j tel que :

$$\prod_{j=1}^N e^{tj} = 1 \quad (4.64)$$

La minimisation de la fonction objective J représente un problème d'optimisation non linéaire qui peut être résolu par la méthode des multiplicateurs de Lagrange. On aboutit alors à :

$$\mu_{ij} = \frac{d_{ij}^{\frac{-2}{m-1}}}{\sum_{i=1}^K d_{ij}^{-2/m-1}} \quad (4.65)$$

$$e^{tj} = \frac{\prod_{l=1}^N [(\sum_{i=1}^K \mu_{il} d_{il}^2)^{\frac{1}{N}}]}{\sum_{i=1}^K \mu_{ij}^m d_{ij}^2} \quad (4.66)$$

$$\bar{x}_i = \frac{\sum_{j=1}^N e^{\frac{t_j}{s}} \mu_{ij}^m x_j}{\sum_{j=1}^N e^{\frac{t_j}{s}} \mu_{ij}^m} \quad (4.67)$$

s est un paramètre réel qui joue le même rôle que le facteur flou m. Une valeur égale à -1 est préconisée par les auteurs.

4.4 Conclusion

Dans ce chapitre, nous avons présenté un algorithme de classification non supervisé qui fait intervenir la notion de flou. Cet algorithme appelé FCM (Fuzzy C-Means) est très populaire, il est de nos jours le plus utilisé en pratique. Il a l'avantage d'être à la fois simple et efficace. Cependant cet algorithme souffre de plusieurs inconvénients, et pour s'acquiescer de ces derniers, plusieurs variantes ont été présentées dans ce chapitre. Ces variantes tendent à changer soit la fonctionnelle à minimiser, soit à définir une autre métrique d^2 ou encore à modifier l'influence du facteur flou m.

Il existe aussi dans la littérature d'autres variantes qui combinent l'algorithme FCM avec celui du PCM [34] ou avec les réseaux de neurones [40,41, 42] ou encore avec les méthodes d'optimisation et le classifieur Support Vector Machine SVM [43]. Comme il existe des versions de FCM qui utilisent des fonctions noyaux [44,45].

L'algorithme FCM et quelques unes de ses variantes seront testés et évalués dans le chapitre suivant dans le cadre de la segmentation.

Chapitre 5 : Tests et résultats

5.1 Introduction

Dans ce chapitre nous allons présenter les résultats de l'algorithme FCM et quelques unes de ses variantes dans le cadre de la segmentation d'images. Cinq variantes de la méthode FCM sont considérées. Il s'agit de l'algorithme PCM, Ahmed, Wu, Tang et celui de Li. Ces méthodes ont été implémentées sous MATLAB version 7.8.

La méthode de segmentation sera appliquée sur des images synthétiques de texture et sur plusieurs images médicales.

5.2 Méthode de segmentation

La méthode de segmentation que nous proposons se déroule en deux phases. Dans la première phase, chaque pixel de l'image est caractérisé par un ensemble d'attributs de texture. Il s'agit plus particulièrement de calculer une matrice de cooccurrence pour chaque pixel de l'image à partir de ses pixels voisins situés à l'intérieur d'une fenêtre de voisinage centrée sur le pixel à traiter. La taille de cette fenêtre est de taille $(2w+1)^2$. A partir de cette matrice, on extrait les attributs de **Haralick** décrits au chapitre 2. Ces attributs correspondent à la corrélation, l'énergie, l'entropie et le contraste. Dans la deuxième phase, on regroupe les pixels ayant des attributs identiques en un ensemble de K (connu) classes en utilisant l'algorithme FCM ou l'une de ses variantes. La méthode de segmentation ainsi proposée est conditionnée d'une part par la taille de la fenêtre de voisinage (paramètre w) et par les paramètres propres à chaque algorithme de classification à savoir:

- Le facteur flou m pour l'algorithme FCM
- Le facteur flou m pour l'algorithme PCM
- Le facteur flou m et les paramètres α et N_r pour l'algorithme d'Ahmed.
- Le facteur flou m et le paramètre δ pour l'algorithme de Li.
- Le facteur flou m et le paramètre β pour l'algorithme de Wu.
- Le facteur flou m et le paramètre s pour l'algorithme de Tang.

5.2.1 Influence des paramètres de la méthode

Pour étudier l'influence des paramètres de la méthode de segmentation proposée, nous avons utilisé une image synthétique composée de 4 textures (**Fig 5.1**).

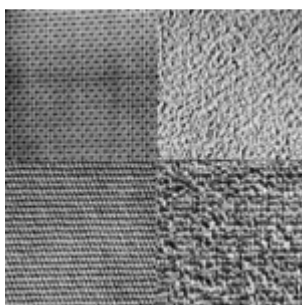


Fig 5.1 : Image test

Les figures (5.2) à (5.7) montrent respectivement les images segmentées obtenus par chaque algorithme de classification avec plusieurs valeurs de w , du facteur flou m et des paramètres propres à chaque algorithme.

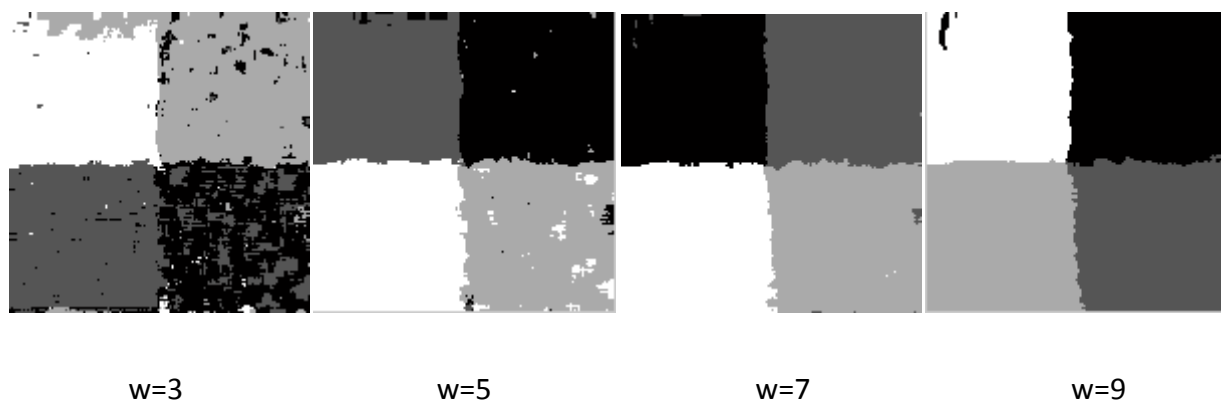
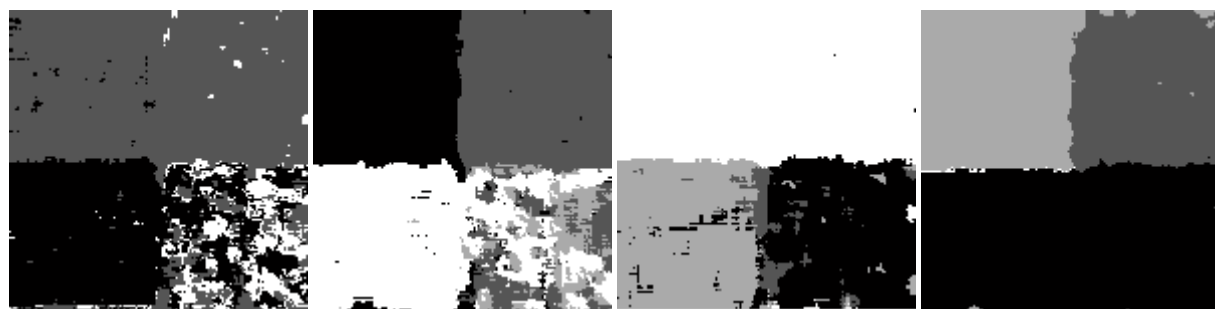


Fig 5.2 : Segmentation de l'image test par la méthode FCM



$w=3$

$w=5$

$w=7$

$w=9$

$m=1.5$



$m=0.5$

$m=1$

$m=1.5$

$m=2$

$w=5$

Fig 5.3 : Segmentation de l'image test par la méthode PCM



$W=3$

$W=5$

$W=7$

$W=9$

$b=0.01/ m=2$



$m=0.5$

$m=1$

$m=1.5$

$m=2$

$\beta = 0.01 / w=9$



$\beta=0.01$

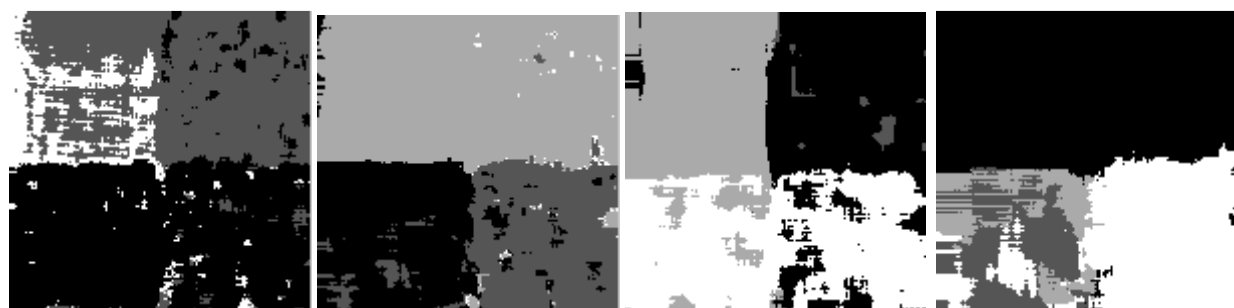
$\beta = 0.05$

$\beta = 0.09$

$\beta = 0.1$

$m=2 / w=9$

Fig 5.4 : Segmentation de l'image test par la méthode Wu



$W=3$

$W=5$

$W=7$

$W=9$

$m=2 / s=-1$

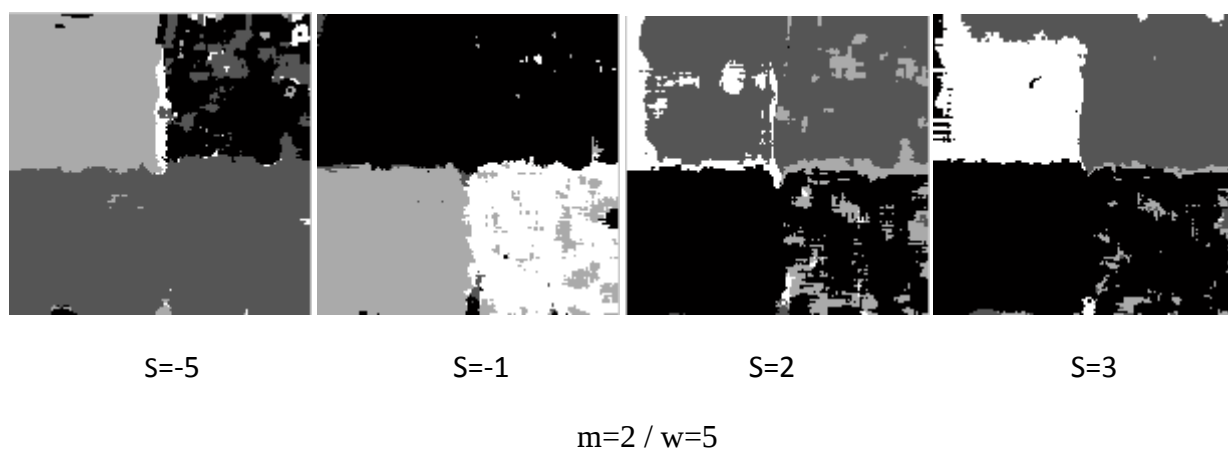
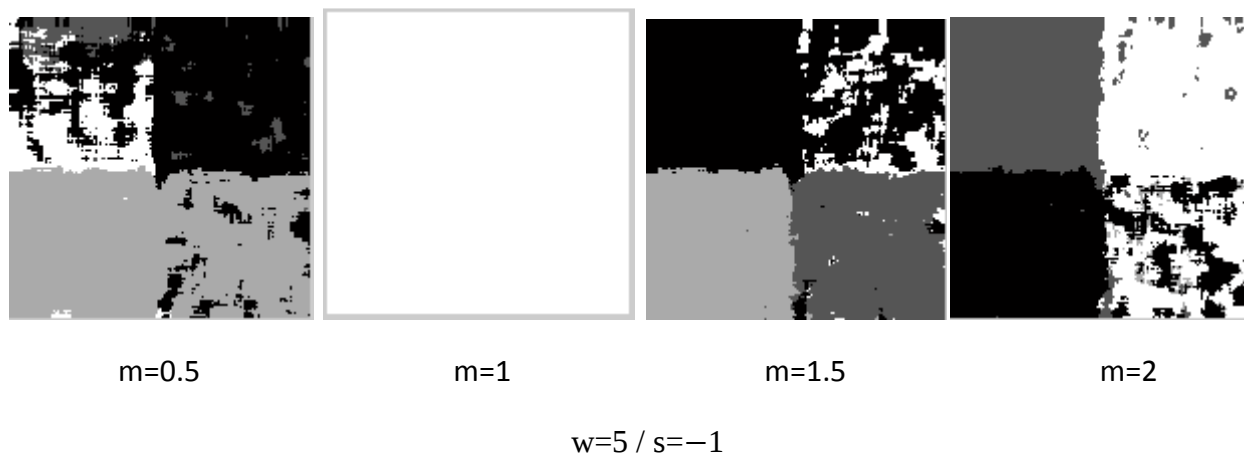
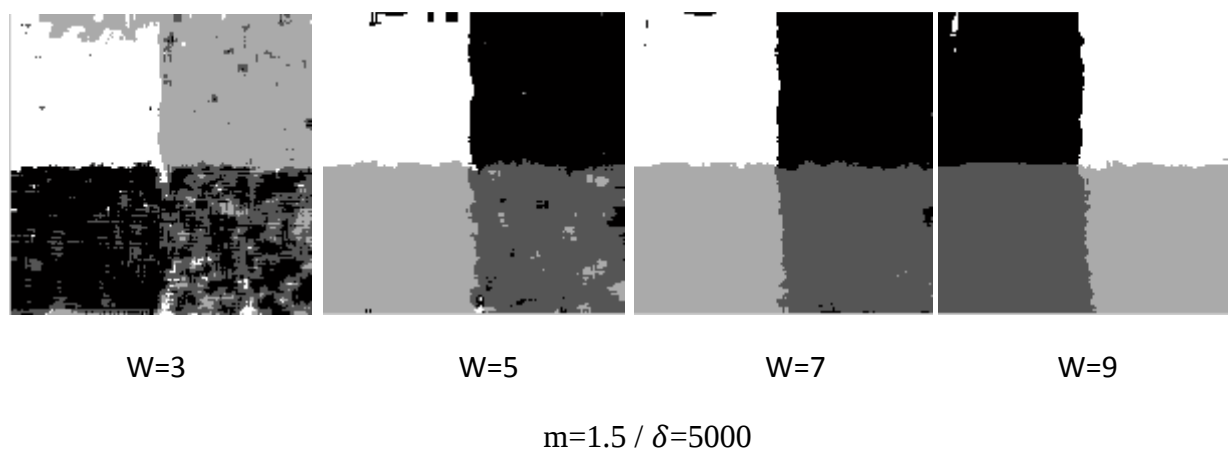


Fig 5.5 : Segmentation de l'image test par la méthode de Tang



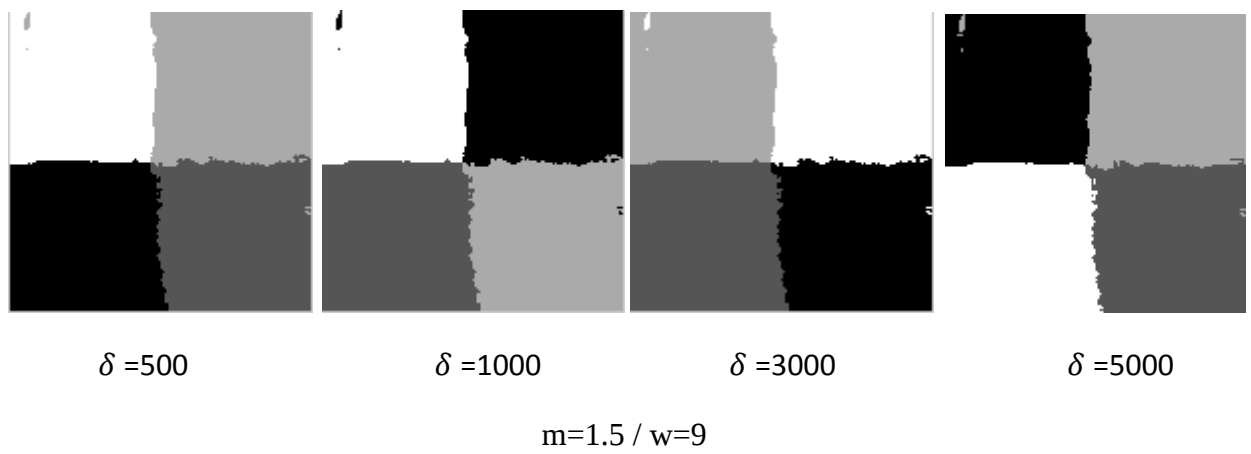
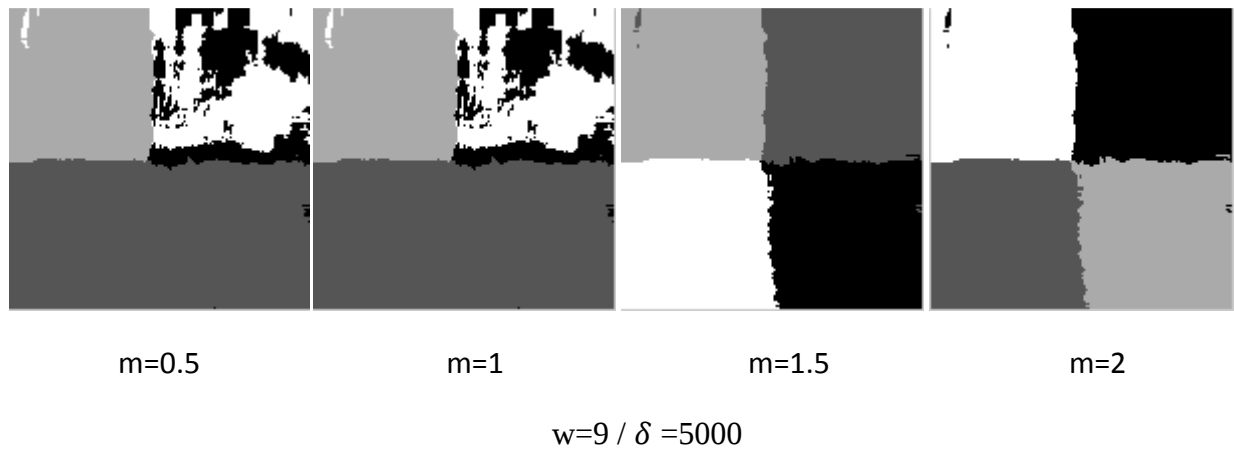
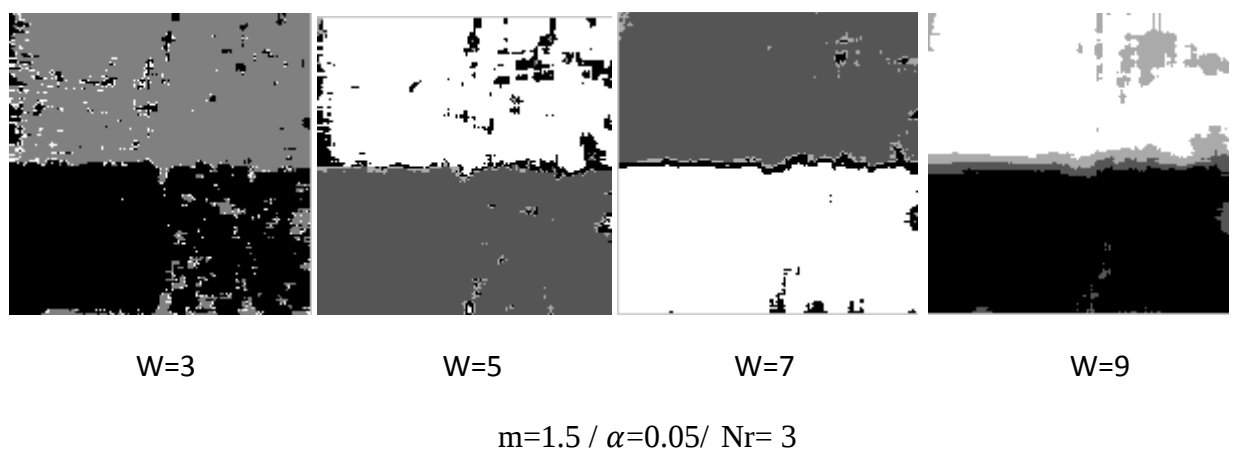


Fig.5.6 : Segmentation de l'image test par la méthode Li



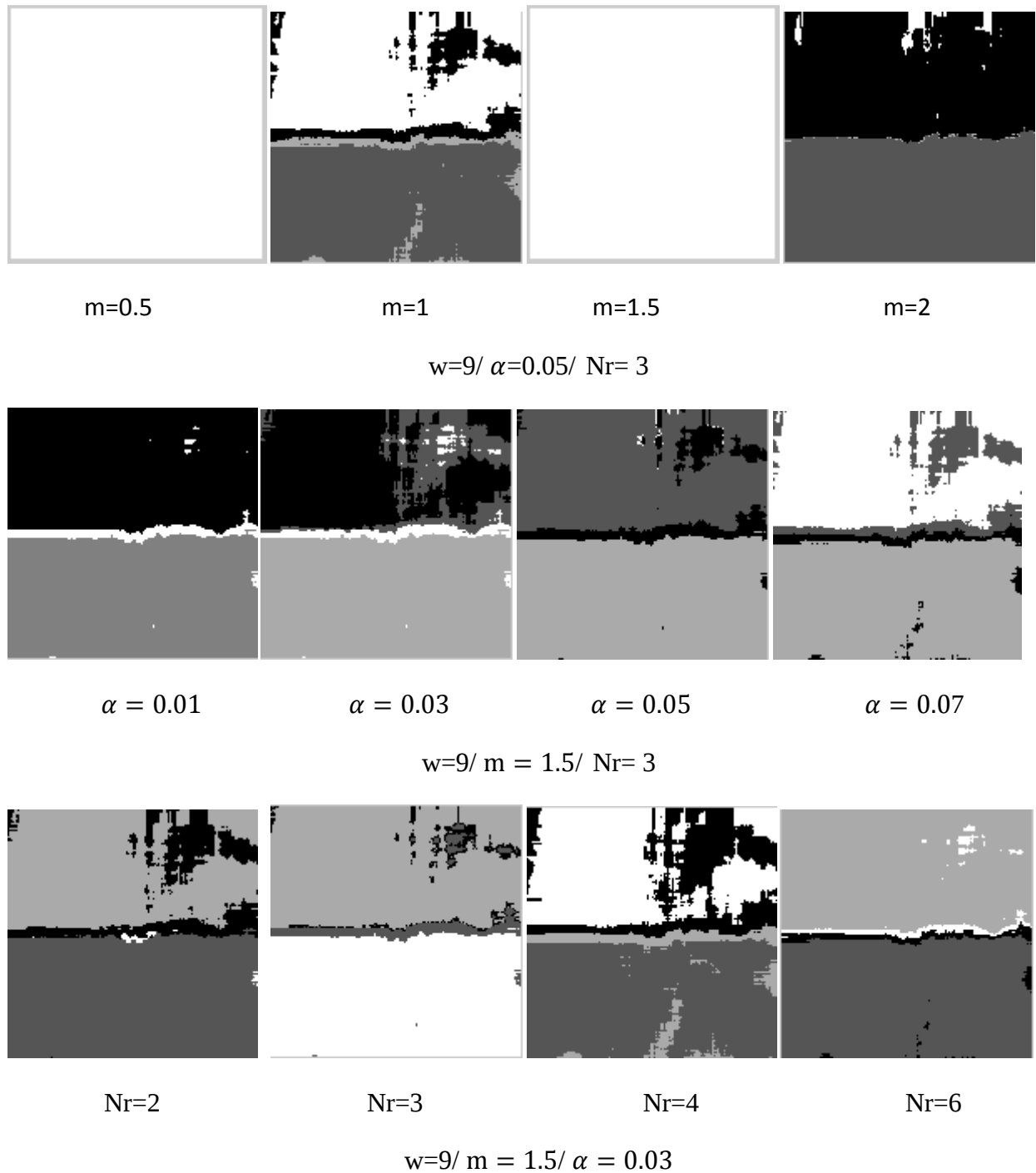


Fig.5.7: Segmentation de l'image test par la méthode d'Ahmed

Interprétation :

Notre objectif par l'illustration de ces résultats est de montrer l'influence des paramètres propres de chaque algorithme de classification. En utilisant une image synthétique composée de 4

textures telles que les attributs de cette dernière sont l'entropie, l'énergie, la corrélation et le contraste.

En analysant les résultats obtenus nous remarquons que les paramètres propres et la taille de la fenêtre de voisinage jouent un rôle prépondérant lors de la classification de textures.

Dans la méthode FCM, la fenêtre de voisinage influe sur l'image, car plus le w augmente plus la segmentation est bonne.

La méthode PCM donne de meilleur résultats pour $m=1.5$ et $w=5$, plus la fenêtre de voisinage w augmente, le nombre de classe diminue.

Pour une fenêtre de voisinage $w>7$, $m=2$ et $0<b<1$ Wu donne de bon résultats.

La méthode Tang donne de meilleur résultat pour $m=2$ et $w=5$ et $a=-1$.

Pour une grande valeur de v et un $m=1.5$. $w=9$; Li a donné un bon résultat.

Quelque soit les paramètres choisis, Ahmed ne donne pas de bon résultats car les régions ne sont pas bien identifiées.

5.3 Segmentation d'images médicales

Nous avons également appliqué la méthode proposée pour segmenter quelques images médicales.

Image radiographique :

La figure 5.8 montre les résultats de la segmentation d'une image radiographique (radio du poignet gauche après un traumatisme d'un homme de 38 ans) par les différentes méthodes de classification. Le nombre de classes de texture est fixé à 3 correspondant respectivement au fond, la chaire et les os.



Image originale



FCM



PCM



Ahmed



Li



Wu



Tang

Fig 5.8: Segmentation d'une image radiographique

Les méthodes FCM, Ahmed, Li, et Wu donnent presque les mêmes résultats

Pour cette image, les 4 méthodes FCM, Ahmed, Li, et Wu donnent de meilleurs résultats par rapport aux autres méthodes car les frontières séparant les régions sont plus claires et les objets de l'image sont bien détectés.

Image IRM :

La figure 5.9 montre les résultats de la segmentation d'une image IRM de coupe sagittale du cerveau par les différentes méthodes de classification. Le nombre de classes de texture est fixé à 5 correspondant respectivement au fond, la voûte crânienne, le parenchyme, la chaire et les os.

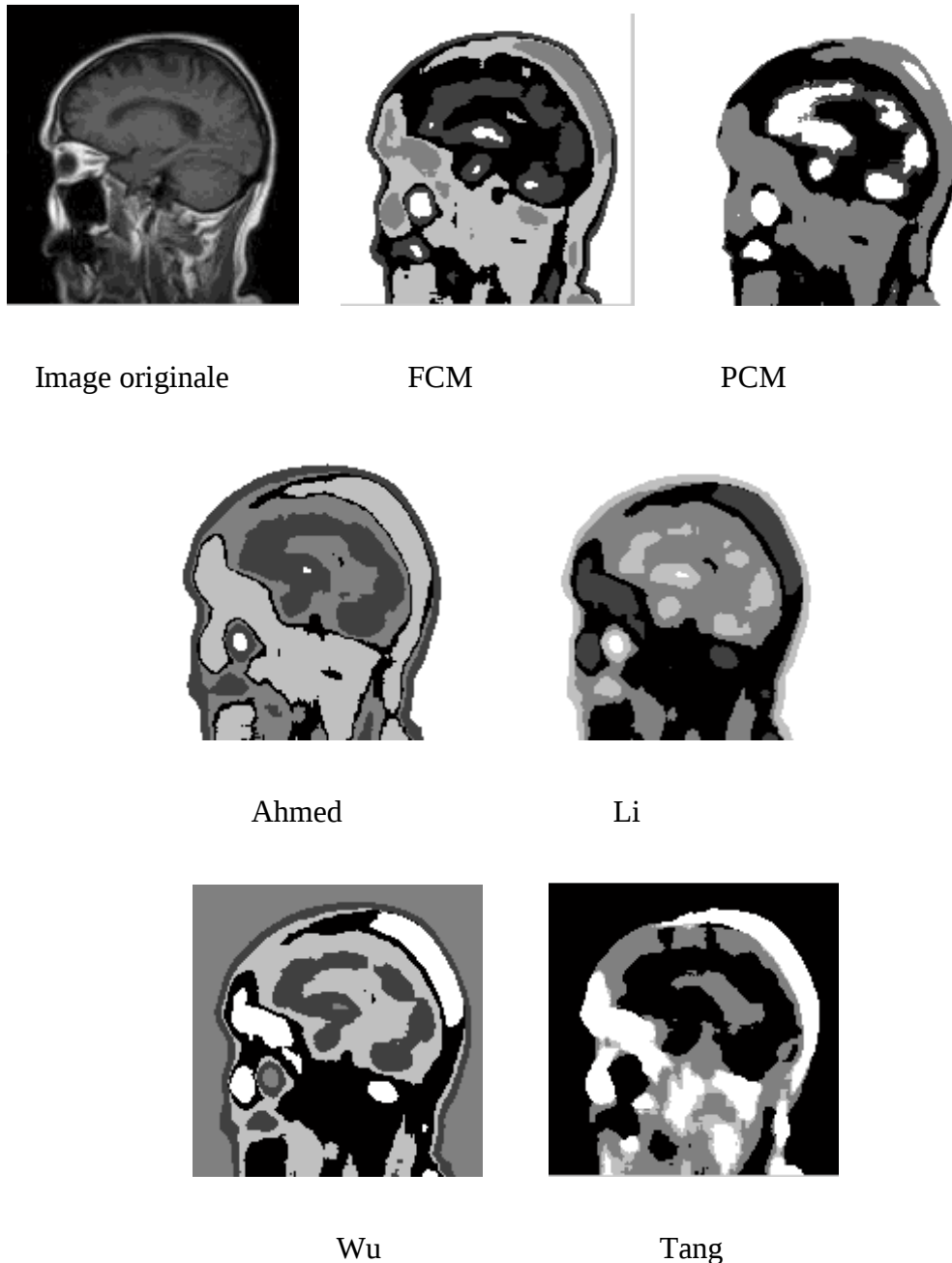


Fig 5.9: Segmentation d'une image IRM

Pour cette image, les méthodes de Ahmed et Tang donnent de meilleurs résultats par rapport aux autres méthodes car la région parenchymateuse est plus claire.

Image endoscopique :

La figure 5.10 montre les résultats de la segmentation d'une image endoscopique qui montre un ulcère d'estomac par les différentes méthodes de classification. Le nombre de classes de texture est fixé à 3

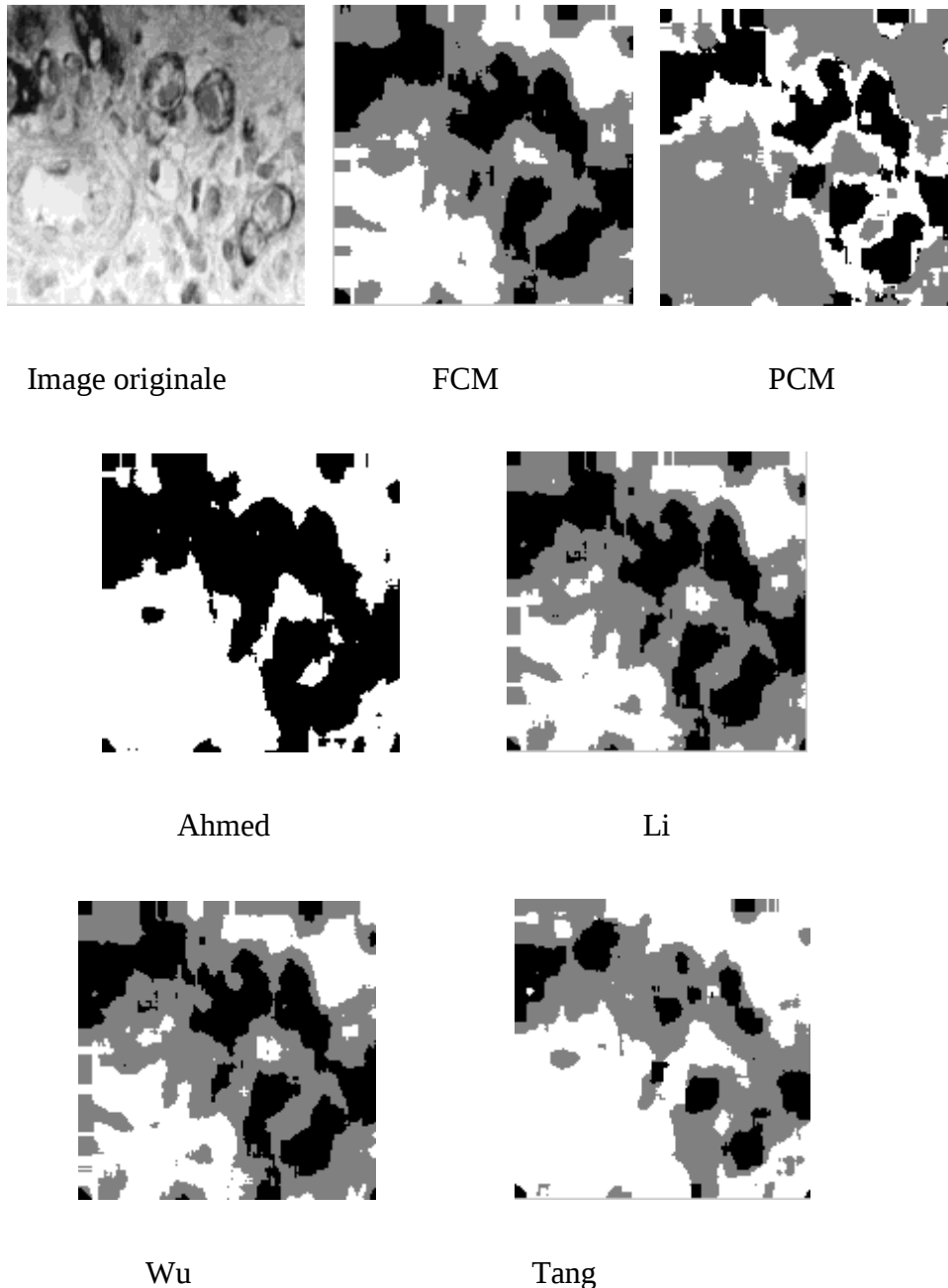


Fig 5.10: Segmentation d'une image endoscopique

La méthode de Ahmed fournit 2 classes seulement.

Pour cette image, les méthodes de FCM, PCM, Li et Wu donnent de meilleurs résultats par rapport aux autres méthodes car toutes les régions sont bien détecté.

5.4 Temps de calcul

Le tableau 5.1 donne le temps de calcul consommé par chaque méthode de classification. Ces résultats sont obtenus sur l'image de la figure 5.9. Le nombre d'itération pour chaque méthode est le même.

Méthodes	Temps d'exécution(s)
FCM	1.232334
PCM	1.552791
Ahmed	316.546013
Li	1.501927
Wu	0.958647
Tang	1.493971

Tableau 5.1 : Temps d'exécution de chaque méthode de classification

On remarque que la méthode Wu reste la méthode la plus rapide car les autres méthodes font tous appel à des calculs supplémentaires. Le temps d'exécution de Ahmed est plus lent comparant aux autres méthodes. Cependant, le temps de calcul des méthodes FCM, PCM, Li, Tang restent presque les mêmes.

5.5 Conclusion

Nous avons présenté dans ce chapitre les résultats de la segmentation par la méthode de classification floue.

Une série de tests sont d'abord effectués sur une image synthétique pour étudier l'influence des paramètres de la méthode proposée. En suite nous avons également appliquée la méthode proposée pour segmenter quelques images médicales.

Après une série de tests on ne peut pas dire qu'une méthode est plus efficace que l'autre car chaque méthode est efficace pour une image donnée.

Conclusion générale

Le travail présenté dans ce mémoire concerne le domaine du traitement d'images et plus précisément celui de la segmentation.

Parmi les nombreuses techniques de segmentation, nous nous sommes intéressés particulièrement à la méthode de classification floue qui est basée sur l'algorithme Fuzzy –C-Means. Cette méthode est formalisée comme un problème de minimisation d'un critère de classification sous certaines contraintes. Son inconvénient principal réside dans sa sensibilité au bruit. Nous avons alors étudié certaines variantes proposées dans la littérature pour remédier à ce problème. Celles-ci tentent soit à changer la fonctionnelle à minimiser, soit à définir une autre métrique ou soit encore à modifier l'influence du facteur flou m . Parmi toutes ces méthodes, nous avons implémenté 5 d'entre elles à savoir les algorithmes PCM, de Ahmed, de Wu, de Tang et celui de Li.

L'application de ces méthodes dans le cadre de la segmentation d'images texturées de synthèse et des images médicales montre que les méthodes de Li, Wu et FCM donnent de meilleurs résultats que les autres méthodes. D'un point de vue rapidité la méthode de Wu reste la plus rapide et la méthode d'Ahmed demeure la plus lente.

Ce travail reste tout de même perfectible. Nous pensons notamment à introduire d'autres attributs de texture.

Bibliographie

- [1] S.HORWITZ et T.Pavlidis: Picture segmentation by a directed split and merge procedure. Rapport de Recherche, Departement of Electrical Engineering, Princeton University. 1975.
- [2] S.Horowitz et T.Pavlidis: Picture Segmentation by a tree traversal algorithm,journal of the Association for Computing Machinery,23(3):368-388.avril 1976.
- [3] S.Zucker : Region growing:clihood and adolescence.Computer Graphics,Image Processing 5(3): 382-390,septembre 1976.
- [4] Henri Maitre – Le traitement des images , Hermes-science, Lavoisier, Paris, 2003
- [5] R.M. Haralick, K. Shanmugam and I.H. Dinstein. Textural features for image classification. IEEE Transactions on Systems Man and cybernitics, Vol. SMC-3, n°6, 1973.
- [6] Jean_Loic Rose . Croissance de région variationnelles et contraintes géométriques tridimensionnelles pour la segmentation d’images .Thèse doctorat Institut Nationaldes sciences Appliquées de Lyon .2008.
- [7] S.L.Horowitz and T.Pavlidis.Picture . Segmentation by a directed split and merge procedure.In CMetImAly77, pages 101.11.1977.
- [8] Yassine Boukir et Nadjib Benkhellat . Segmentation d’image en niveau de gris par des méthodes évolutives (AG, AGQ, ANTSEG) Mémoire d’ingénieur.INI 2005.
- [9] M.Dempster,N.M.Laird,D.B,Jain . Maximum likelihood for incomplete data via the EM algorithm,J.Royal Statistics Society,Series B.Vol.39.1977,pp 1_38.
- [10] K.Fukunaga et L.Hostetler .The estimation of the gradient of a density function whith IEEE Transaction on information Théory , 6:721-741.1984.
- [11] D.Comaniciu et P.MEER:Mean Shift:a Robust approach Toward feature space analysis.IEEE Transactions on Pattern Analysis and Machine Intelligence,24(5):603_619.mai 2002.
- [12] C. Marisala. Logique floue, principes, aide à la décision.
- [13] Ensfloue. Pdf. [www.ensiie. Fr/r gacogne/ensfloue.pdf](http://www.ensiie.fr/r_gacogne/ensfloue.pdf).

- [14] Méchel Ménard, 1998: The fuzzy c+2 Means: Solving The Extended Ambiguity Rejet in clustering. In IEEE Transactions on fuzzy systems, vol.1, N° 2, pages 195-203.
- [15] S.A. Mingoti et J.O. Lima. Comparing SOM neural network with Fuzzy c-means, K-means and traditional hierarchical clustering algorithms. European Journal of Operational Research. Vol. 174, pp. 1742–1759, 2006.
- [16] F. Masulli, Andrea Schenone. A fuzzy clustering based segmentation system as support to diagnosis in medical imaging. Artificial Intelligence in Medicine. Vol.16, pp. 129–147, 1999.
- [17] R. Babuslka, L. Alica, M.S. Lourens, A.F.M. Verbraak, J. Bogaard. Estimation of respiratory parameters via fuzzy clustering. Artificial Intelligence in Medicine. Vol. 21, pp. 91-105, 2001.
- [18] T. Chaira. A novel intuitionistic fuzzy Cmeans clustering algorithm and its application to medical images. Applied Soft Computing (2010) doi:10.1016/j.asoc.2010.05.005.
- [19] L.-H. Chen, C. Kao, S. Kuo, T.-Y. Wang and Y.-C. Jang. Productivity Diagnosis via Fuzzy Clustering and Classification" An Application to Machinery Industry. Omega, Int. J. Mgmt Sci. Vol. 24, No. 3, pp. 309-319, 1996.
- [20] J. Soto, M. I. V. Aguiar, A. Flores-Sintas. A fuzzy clustering application to precise orbit determination. Journal of Computational and Applied Mathematics. Vol. 204, pp. 137 – 143, 2007.
- [21] V. Barra, V. Delouille, J.-F. Hochedez. Segmentation of extreme ultraviolet solar images via multichannel fuzzy clustering. Advances in Space Research. Vol. 42, pp. 917–925, 2008.
- [22] J. Fan, M. Han, J.Wang. Single point iterative weighted fuzzy C-means clustering algorithm for remote sensing image segmentation. Pattern Recognition. Vol. 42, pp. 2527-2540, 2009.
- [23] A.B. Goktepe, S. Altun, A. Sezer. Soil clustering by fuzzy c-means algorithm. Advances in Engineering Software. Vol. 36, pp. 691–698, 2005.
- [24] F. Naghdy. Fuzzy clustering of human motor motion. Applied Soft Computing (2010) doi:10.1016/j.asoc.2010.01.013.
- [25] C.-C. Lo, S.-J. Wang. Video segmentation using a histogram-based fuzzy c-means clustering algorithm. Computer Standards & Interfaces. Vol. 23, pp. 429–438, 2001.
- [26] R. Hasanzadeh P.R., A.H. Rezaiea, S.H.H. Sadeghi, M.H. Moradi, M. Ahmadi. A density-based fuzzy clustering technique for non-destructive detection of defects in materials. NDT&E International. Vol. 40, pp. 337–346, 2007.
- [27] A. Mukhopadhyaya, U. Maulik. Towards improving fuzzy clustering using support vector machine: Application to gene expression data. Pattern Recognition. Vol. 42, pp. 2744-2763, 2009.

- [28] G. Naadimuthu, P. Gultom, E.S. Lee. Fuzzy clustering in cell formation with multiple attributes. *Computers and Mathematics with Applications*. Vol. 59, pp. 3137-3147, 2010.
- [29] A. Rick, *Représentation de la variabilité dans le traitement d'image floue application à la mammographie numérique*, Thèse de doctorat de l'université Paris 6. 1999.
- [30] R. Kruse, C. Döring, and M.-J. Lesot. *Fundamentals of Fuzzy Clustering. Advances in Fuzzy Clustering and its Applications* Edited by J. Valente de Oliveira and W. Pedrycz. 2007 John Wiley & Sons.
- [31] S. Shen, W. Sandham, M. Granat and A. Sterr. MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. *IEEE Trans. Int. Techrol. Biomed.* Vol.9, pp.459-467, 2005.
- [32] M. Forouzanfar, N. Forghani and M. Teshehla. Parameter optimization of improved fuzzy c-means clustering algorithm for brain MR image segmentation. *Engineering Application of Artificial Intelligence*. 2010.
- [33] X. Wang, Y. Wang and L. Yang. Improving fuzzy c-means clustering based on feature – weight learning. *Pattern Recognition*, vol. 25, pp. 1123-1132, 2004.
- [34] C. Tang, S. Wang, W. Xu. New fuzzy c-means clustering model based on the data weighted approach. *Data & Knowledge Engineering*. Vol. 69, pp. 881–900, 2010.
- [35] K.-L. Wu, J. Yu, M.-S. Yang. A novel fuzzy clustering algorithm based on a scatter matrix with optimality tests. *Pattern Recognition Letters* 26 (2005) 639–652
- [36] M.N. Ahmed, S.M. Yamany, N. Mohamed, A.A. Farag, T. Moriarty, A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data, *IEEE Trans. Med. Imaging* 21 (2002) 193–199.
- [37] J. Kang, L. Min, Q. Luan, X. L. Li and J. Liu. *Novel modified fuzzy c-means algorithm with applications*. Digital Image Processing, 2008.
- [38] S. Miyamoto, H. Ichihashi, K. Honda. *Algorithms for Fuzzy Clustering : Methods in c-Means Clustering with Applications*. Springer, 2008.
- [39] Z. Yang, F.-L. Chung and W. Shitong. Robust fuzzy clustering-based image segmentation. *Applied Soft Computing*. Vol. 9, pp. 80–84, 2009.
- [40] A. Mukhopadhyaya, U. Maulik. Towards improving fuzzy clustering using support vector machine: Application to gene expression data. *Pattern Recognition*. Vol. 42, pp. 2744-2763, 2009.
- [41] U. Maulika, I. Sahab. Modified differential evolution based fuzzy clustering for pixel classification in remote sensing imagery. *Pattern Recognition*. Vol. 42, pp. 2135-2149, 2009.

- [42] H. Izakian, A. Abraham. Fuzzy C-Means and Fuzzy Swarm for Fuzzy Clustering Problem. Expert Systems with Applications. (2010) Doi 10.1016/j.eswa.2010.07.112
- [43] A. Mukhopadhyaya, U. Maulik. Towards improving fuzzy clustering using support vector machine: Application to gene expression data. Pattern Recognition. Vol. 42, pp. 2744-2763, 2009.
- [44] C.G. Looney. Fuzzy connectivity clustering with radial basis kernel functions. Fuzzy Sets and Systems. Vol. 160, pp. 1868–1885, 2009.
- [45] D. Graves, W. Pedrycz. Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study. Fuzzy Sets and Systems. Vol. 161, pp. 522–543, 2010.
- [Web1] <http://www.centre-irm.com/images/favicon.ico>
- [Web2] <http://sante-medecine.commentcamarche.net/rss/actualite.php3>
- [Web3] http://pagesperso-orange.fr/chimie.sup/IRM_fichiers/filelist.xml