



République Algérienne Démocratique et Populaire

Université Abou Bakr Belkaid– Tlemcen

Faculté des Sciences

Département d'Informatique

MÉMOIRE DE MAGISTER

En Informatique

Option: Système d'information et de communication

Thème

CMBIR (Content Medical Based Image Retrieval)

développement d'outil logiciel d'annotation

d'images médicales, utilisant les méthodes

d'indexation par descripteurs invariants de

contenus

Présenté par : M^{elle} Imane NEDJAR

Les membres de jurys sont :

M^r Mohammed Amine CHIKH Professeur à université de Tlemcen *directeur de mémoire*

M^r Said MAHMOUDI M.C à université de Mont Belgique *co-directeur de mémoire*

M^r Fethi BEREKSI REGUIG Professeur à université de Tlemcen *président du jury*

M^r Abdelhafid BESSAID Professeur à université de Tlemcen *examinateur*

M^r Amine ABDERRAHIM M.C.B à université de Tlemcen *invité*

Année universitaire : 2011-2012

Remerciements

Tout d'abord, je tiens à remercier mes directeurs de thèse, Mr Mohammed Amine CHIKH et Mr Said MAHMOUDI, pour leur patience, leur aide, la pertinence de leurs conseils. Le mémoire est une expérience un peu aventureuse, à l'issue incertaine ; et c'est avant tout à eux que je dois d'avoir mené la mienne à son terme. Travailler à leurs côtés a été un réel plaisir, et j'espère en avoir de nouveau l'occasion par la suite.

Nous exprimons notre reconnaissance à Mr Fethi BEREKSI REGUIG, pour m'avoir fait honneur en acceptant de présider le jury de ce mémoire.

Je suis également reconnaissant à Mr Abdelhafid BESSAID et Mr Amine ABDERRAHIM pour m'avoir fait l'honneur de rapporter sur ce mémoire.

Je remercie enfin les étudiants de la promotion informatique 2012, pour leur solidarité et leur esprit de groupe.

A mon père et ma mère

à mes frères et mes sœurs,

à mes amies Assia ,Fatiha, Zineb ,Khadidja ,Farah,Hanane.

Je dédie ce travail

TABLE DES MATIÈRES

TABLE DES MATIÈRES.....	4
LISTE DES FIGURES.....	7
LISTE DES TABLEAUX.....	9
Introduction générale.....	10
Chapitre 1 : Indexation d’images médicales par le contenu	12
Introduction.....	13
I. Systèmes médicaux existants CBIR.....	13
I.1.ASSERT (Automatic Search and Selection Engine with Retrieval Tools)	13
I.2.CasImage	14
I.3. IRMA (Image Retrieval in Medical Applications).....	14
I.4. NHANES II (The Second National Health And Nutrition Examination Survey)	14
II .Les descripteurs d’image.....	15
II .1.Les descripteurs de texture	16
II .1.1 .Taxonomie des modèles de textures.....	17
II .1.1.1 L’approche statistique.....	17
1. La matrice de co-occurrence (Gay Level Co-Occurrence Matrix GLCM).....	19
2. Les caractéristique d’auto corrélation (Autocorrelation features).....	19
II.1.1.2.Les méthodes géométrique (Geometrical methods).....	20
1. Les Caractéristiques de Voronoi tessellation	21
2. L’approche structurale (Structural methods).....	21
II .1.1.3.Les méthodes fondées sur un modèle (<i>Model based methods</i>).....	22
II .1.1.4.Les méthodes de traitement de signal (<i>Signal processing methods</i>)	22
1. Filtres spatiaux (Spatial domain filters)	22
2. Le domaine de Fourier (Fourier domain filtering)	23
3. Le model de Gabor et Ondelette (Gabor and Wavelet models)	24
II.2. Les descripteurs de formes	27
II.2 .1.Classification des formes.....	27
II.2.1.1.La représentation de la forme basée sur le contour.....	28
II.2.1.2. Représentation de forme basée sur la région et les techniques de description (Region-based shape representation and description Techniques).....	33

II.3.Les descripteurs de couleur :	39
II.3.1. Les espaces de couleurs	39
II.3.1.1.Espace couleur HSI	40
II.3.1.2.Espace couleur YCbCr	41
II.3.1.3.Espace couleur I1I2I3	42
II.3.1.4.Espace couleur La*b*	43
II.3.2.L'histogramme	44
II.3.3.Histogramme de couleur-structure CS	44
II.3.4.Le Corrélogramme	44
II.3.5.Vecteur de cohérence de couleurs	45
II.3.6. Descripteur par disposition de couleurs"	46
II.3.7.Descripteur par couleurs dominantes	46
Conclusion	47
Chapitre 2 :L'annotation d'images médicales	48
Introduction	49
I. Les ontologies	49
I.1.Les ontologies formelles (Formal ontologies)	50
I.2.Les ontologies biomédicales (Biomedical ontology)	51
I.2.1. Le modèle fondamental de l'ontologie anatomie (The Foundational Model of Anatomy ontology)	51
I.2.2.Les ontologies des maladies (Disease Ontology)	53
I.2.3. Unified Medical Language System (UMLS)	53
I.2.4. La rubrique de sujet médical (Medical Subject Headings)	55
I.2. 5. GALEN	55
I.2.6. Nomenclature systématisée de médecine - Conditions cliniques (Systematized Nomenclature of Medicine--Clinical Terms)	57
I.2.7. L'ontologie des gènes (The Gene Ontology)	58
II. Le système d'archivage et de communication (Picture Archiving and Communication System)	59
III. Les classifieurs utilisés dans la recherche d'image médical par le contenu	60
Conclusion	62
Chapitre 3 :Les méthodes utilisées	63
Introduction	64
I.L'extraction des caractéristiques	64

I.1 Prétraitement par segmentation d'images :	64
I.1.1 .Le modèle de contour actif :	65
I.2. les descripteurs visuels utilisés :	69
I.2.1 .Les ondelettes de Gabor.....	69
I.2.2. La matrice de co-occurrence (Gay Level Co-Occurrence Matrix GLCM).....	73
II. Classification	77
II.1.Les Machines à Vecteurs Supports (SVM)	77
II.2.Les méthodes de décomposition	88
II.2.1.L'approche "un contre tous"	88
II.2.2.L'approche "un contre un".....	88
II.2.3.L'approche de codes correcteurs d'erreurs (ECOC).....	88
Conclusion	89
Chapitre 4 : Architecture du système de recherche par le contenu proposé.....	90
Introduction.....	91
I. L'architecture générale d'un système d'indexation et de recherche d'images :	91
I.1.La recherche d'image par contenu dans le domaine médicale.....	92
I.2.Les modalités d'imagerie et CBIR.....	93
II. Catégorisation d'images.....	95
II.1.Base d'images utilisée	95
II.2 .Architecture du système proposé.....	99
2.1. Normalisation	101
2.2. L'extraction des caractéristiques	101
2.3. La classification	105
2.4. La distance Euclidienne	106
II.3.L'annotation.....	107
Conclusion	110
Chapitre 5 : Validation et analyse des résultats	111
Introduction.....	112
I .Analyse des résultats	112
I.1.Première expérimentation.....	112
I.1.1.Interprétation	118
I.1.2.Comparaison.....	119
I.2.Deuxième expérimentation :	123

I .2.1. Interprétation	125
I .2.2. Mesure de la qualité des réponses	126
I.3. Troisième expérimentation :	132
I.3.1. Interprétation	133
I.3.2. Comparaison.....	134
II .Mesure de la qualité des descripteurs	134
Conclusion	137
Conclusion générale	138
Bibliographie.....	140
Résumé	143

LISTE DES FIGURES

Figure 1 les interactions entre les méthodes et l'image d'origine [1].....	18
Figure 2 les Caractéristiques de texture provenant du spectre de puissance (a) la texture d'une image, et (b) spectre de sa puissance.....	20
Figure 3 Classification des techniques de la représentation de forme [21]	28
Figure 4 (a) Convexe et ses concavités; (b) l'arbre de concavité [21]	38
Figure 5 Construction de l'axe médian d'une forme rectangulaire [21].....	38
Figure 6 Espace couleur RGB.....	40
Figure 7 Espace couleur HSI	41
Figure 8 Espace couleur YCbCr.....	42
Figure 9 Espace couleur I1I2I3	43
Figure 10 Espace couleur La*b*.....	44
Figure 11 Les différents sous-domaines intégrés dans l'UMLS [31]	54
Figure 12 Les concepts actifs dans SNOMED au cours de six ans.....	58
Figure 13 les cas possibles de la position de la courbe. Le terme est minimisé seulement dans le cas où la courbe est à la limite de l'objet.....	66
Figure 14 la courbe se propage dans une direction	67
Figure 15 Deux régularisations différentes de la fonction Heaviside (en haut) et la fonction Dirac (en bas)	68
Figure 16 La partie réelle	71
Figure 17 La partie imaginaire	71
Figure 18 Phase angle of the oscillation function	71
Figure 19 the Gauss function	71
Figure 20 Gabor function Amplitude part	72
Figure 21 Gabor function - phase part.....	72
Figure 22 Morphologie des 40 filtres de Gabor	72
Figure 23 (a) La partie réelle du noyau de Gabor (b) Cinq grandeurs de la Gabor	73

Figure 24 la matrice de co-occurrence avec D=1 et PH (0°), PV (90°), PRD (45°), PLD (135°).....	74
Figure 25 Des classifieurs linéaires (hyperplan) pour un espace de deux dimensions	79
Figure 26 Fonction de classification SVM: la maximisation de la marge d'hyperplan dans un espace bidimensionnel.....	80
Figure 27 Les relations entre ξ , ξ_i , et C	85
Figure 28 Principe de la technique SVM.....	86
Figure 29 un classifieur non linéaires (hyperplan) pour un espace plus de deux dimensions.	86
Figure 30 Architecture générale d'un système CBIR	92
Figure 31 Exemples d'images de la base de données d'annotation ImageCLEF 2005.....	97
Figure 32 Deux images pour les classes n°1 et 2 respectivement.....	98
Figure 33 Deux images de la classe n°8.	98
Figure 34 le diagramme de framework pour la première phase.....	99
Figure 35 le diagramme de framework pour la deuxième phase	100
Figure 36 La croissance de la courbe autour de l'objet.....	102
Figure 37 le centroïde de la forme.....	102
Figure 38 (A) Représentation en ondelettes de Gabor de l'image 10001 de la classe n°12 ImageCLEFmed 2005 (B) grandeurs de la Gabor (c) La partie réelle de noyaux de Gabor..	103
Figure 39 l'image n° 2975 de la classe n°1 (B) La partie réelle de noyaux de Gabor (C) le descripteur de l'image	104
Figure 40 l'image n° 2975 de la classe n°1 (B) le descripteur de l'image	105
Figure 41 La fonction à base radiale	106
Figure 42 L'annotation d'une image	107
Figure 43 les images les plus proches à la requête.....	109
Figure 44 Exemple d'images des 52 classes utilisées.....	117
Figure 45 Les résultats de classification de 52 catégories.....	118
Figure 46 Exemple d'images des 21 classes utilisées.....	124
Figure 47 Les résultats de classification de 21 catégories.....	124
Figure 48 Deux images de la base de test pour la classe n°9.....	125
Figure 49 Deux images de la base d'apprentissage pour la classe n°20	125
Figure 50 Deux images de la base d'apprentissage pour les classes n°19 et n°20	126
Figure 51 Allure d'une courbe de rappel-précision.....	128
Figure 52 Courbe de rappel et précision pour une image requête de la classe 15	129
Figure 53 Courbe de rappel et précision pour une image requête de la classe 19	130
Figure 54 la courbe moyenne de rappel et précision.....	131
Figure 55 Exemple d'images des 15 classes utilisées.....	133
Figure 56 Les résultats de classification de 15 catégories par la matrice de co-occurrence ..	133

LISTE DES TABLEAUX

Tableau 1 Les modalités d'imagerie médicale et biologique et les systèmes existants CBIR [39]	95
Tableau 2 Les classes d'images de rayons X: le numéro de la classe dans image Clef anatomique, l'angle	117
Tableau 3 Les classes d'images de rayons X: anatomique, l'angle	120
Tableau 4 Les résultats de classification basé sur la base de test.....	121
Tableau 5 Un tableau comparatif entre les deux méthodes	122
Tableau 6 les numéros des classes	123
Tableau 7 Des valeurs de rappel-précision pour une image de classe 15.....	129
Tableau 8 Des valeurs de rappel-précision pour une image de classe 19.....	130
Tableau 9 la moyenne de la précision et le rappel pour les deux requêtes.....	131
Tableau 10 les numéros des classes	132
Tableau 11 le rappel et la précision pour les descripteurs de 1 cas.....	135
Tableau 12 le rappel et la précision pour les descripteurs de 2 cas.....	136
Tableau 13 le rappel et la précision pour les descripteurs de 3 cas.....	137
Tableau 14 Le taux de classification pour chaque descripteur	137

Introduction générale

Grâce aux avancées récentes de la technologie ces dernières années, l'informatique numérique est devenu le cœur de tous les secteurs d'activités : dans le monde industriel, scientifique, juridique, médical, etc.

Ces progrès se sont accompagnés d'une baisse des coûts des équipements informatiques qui a facilité la diffusion et l'échange des données multimédia numérisées vers le grand public. Particulièrement, l'image a envahi tous les supports d'informations grâce à son pouvoir expressif et la facilité de son transfert via les réseaux informatiques. Il en résulte une expansion continue des bases de données d'images. De ce fait, l'accès aux informations inhérentes à ces dernières devient de plus en plus difficile et coûteux.

Les premières solutions adoptées pour la recherche d'image sont inspirées des méthodes de recherche dans les documents textuels. En fait, un progrès acceptable a été réalisé dans la recherche de ce type de documents visuels.

Dans ces systèmes de recherche d'image, l'image est indexée par un ensemble d'annotations textuelles qui vont décrire son contenu visuel. De là, il suffit d'utiliser un moteur de recherche textuel pour retrouver les documents demandés. Cette solution se heurte à plusieurs problèmes comme la subjectivité d'interprétation du contenu visuel, les coûts élevés d'annotation manuelle, etc. Une autre solution au problème de recherche d'images consiste à effectuer une navigation libre « Free browsing » dans la base de données (BDD) jusqu'à ce qu'on trouve l'image candidate. Cette solution ne peut être appliquée que pour une recherche occasionnelle et pour des BDD de petites tailles. En plus, elle n'est pas adéquate pour des applications professionnelles.

Pour combler l'inefficacité et les limites des méthodes traditionnelles de la recherche d'images, les chercheurs ont orienté leurs travaux vers l'exploration des voies possibles de l'utilisation d'une description inhérente au contenu visuel de l'image. Cette approche est connue sous le nom de la recherche d'images par le contenu « Content Based Image Retrieval ou CBIR ». Dans ce paradigme, l'image est décrite par un ensemble de descripteurs numériques qui essaient de caractériser des aspects différents du contenu visuel comme la texture, la forme et la couleur. En effet, les systèmes de recherche d'image actuels ne peuvent pas effectuer la recherche en exploitant directement le contenu brut des images. Donc la

recherche d'images similaires se réduit à une simple recherche d'image présentant des descripteurs similaires.

La recherche d'images par le contenu concerne tous les domaines où l'image constitue un support essentiel d'informations et plus particulièrement, la radiologie. En fait, dans les hôpitaux sont produites, quotidiennement, des milliers d'images qui sont sauvegardées dans leurs archives. L'exploitation de ces données sera fonction des techniques, d'archivages et de recherche mises en place. C'est dans ce contexte restreint de la recherche d'images médicales par le contenu que se situera notre travail.

Notre manuscrit comportera les parties suivantes :

Dans les deux premiers chapitres, nous ferons un tour d'horizon des méthodes d'indexations et d'annotations des images, existants dans la littérature. Le chapitre 3 donne une vue plus détaillée des techniques utilisées. Nous présenterons dans le chapitre 4 le système d'indexation et de recherche d'image par le contenu mise en œuvre. Ensuite une analyse des résultats du système est décrite dans le dernier chapitre.

Chapitre 1 : Indexation d'images médicales par le contenu

Dans ce chapitre, nous allons présenter les systèmes médicaux existants dans la recherche des images par contenu, puis nous aborderons les méthodes d'indexations rencontrées dans la littérature.

Introduction

La recherche d'images par le contenu « Content-based visual information retrieval (CBVIR) » ou « Content-Based Image Retrieval (CBIR) » a été l'un des domaines de recherche les plus vives dans le champ de la vision par ordinateur au cours de ces vingt dernières années. Cette nouvelle technique pallie aux problèmes posés par la recherche textuelle, et permet d'améliorer des applications interrompues et contribue aussi à faire émerger des nouvelles applications dans divers domaines, notamment dans le domaine médical. En effet dans ce domaine les images et en particulier les images numériques utilisées pour le diagnostic et la thérapie sont produites en quantités toujours croissantes ; ce qui rend l'utilisation d'un système de recherche et de stockage automatique nécessaire.

I. Systèmes médicaux existants CBIR

Bien que la recherche d'images par le contenu ait souvent été proposée pour une utilisation dans la gestion des images médicales, seuls quelques systèmes de recherche basés sur le contenu ont été développés spécifiquement pour les images médicales. Ces axes de recherche sont généralement construits dans des instituts de recherche et continuent d'être améliorés, développés et évalués au fil du temps.

Cette section présentera plusieurs grands systèmes médicaux d'indexation et de recherche d'images médicales par le contenu.

I.1. ASSERT (Automatic Search and Selection Engine with Retrieval Tools)

- Le système ASSERT utilise l'approche « physician-in-the-loop » pour la recherche des images HRCT du poumon. Cette approche oblige les médecins à délimiter les régions portant la pathologie et d'identifier certains repères anatomiques pour chaque image.

- Ce système extrait 255 caractéristiques de la texture, de la forme, des bords, et d'échelle de niveau gris des régions portant la pathologie.
- Une table multi-dimensionnelle de hachage est construite pour indexer les images HRCT.

Cet outil a été développé dans l'université de Purdue, l'université d'Indiana, et l'hôpital universitaire de Wisconsin, USA.

I.2. CasImage

- Le système CasImage, a été intégré dans un environnement PACS (Picture Archiving And Communication System). Il contient une base de données pour l'enseignement ainsi que le moteur de recherche medGIFT.
- Le système de recherche medGIFT extrait les couleurs (globales et régionales) et les caractéristiques de texture, y compris les 166 couleurs dans l'espace de couleur HSV, et les réponses des filtres de Gabor dans quatre directions chacune à trois échelles différentes.
- Les combinaisons d'étiquettes textuelles et les caractéristiques visuelles sont utilisées pour l'indexation et la recherche des images médicales.

Ce système est développé par l'hôpital universitaire de suisse à Genève.

I.3. IRMA (Image Retrieval in Medical Applications)

- Le système IRMA est implémenté comme une plate-forme pour la recherche d'images par le contenu dans des applications médicales.
- Ce système divise le processus de recherche d'images en sept étapes consécutives, y compris la catégorisation, l'enregistrement, l'extraction des caractéristiques, la sélection des caractéristiques, l'indexation, l'identification et la récupération.

Ce système est développé par: l'université de technologie de Aachen, Allemagne.

I.4. NHANES II (The Second National Health And Nutrition Examination Survey)

- Ce système contient l'outil « the Active Contour Segmentation (ACS) », qui permet aux utilisateurs de créer un modèle par les points de marquage autour des vertèbres. Si la segmentation d'un modèle est acceptée, l'outil ACS estime la localisation de la vertèbre, et place le gabarit sur l'image, puis la segmente.
- Dans la représentation des données, un processus d'approximation polygonale est appliqué pour éliminer les caractéristiques de formes insignifiantes et réduire le nombre de points de données. Les données obtenues dans le processus d'interpolation polygonale représentent la forme d'une vertèbre. Puis, la courbe approchée de la vertèbre est converti en espace tangent à la mesure de similarité.

Ce système est développé par: la librairie national de médecine, USA.

II .Les descripteurs d'image

Le contenu de l'image désigne à la fois son contenu sémantique et son contenu visuel. Le contenu visuel est décrit par la texture, la forme et la couleur.

Le contenu sémantique est souvent associé à l'image manuellement ou automatiquement. L'utilisation du contenu sémantique dans la description des images rencontre plusieurs problèmes. D'une part, la méthode manuelle est limitée par la subjectivité inhérente de l'interprétation de l'image. En fait, la même image peut être interprétée différemment par deux individus différents .D'autre part, la précision des méthodes automatiques sera en fonction de la précision des méthodes de reconnaissance des formes, et les techniques d'apprentissage adoptées, dont les performances ne sont pas garanties.

Un descripteur est défini comme la connaissance utilisée pour caractériser l'information contenue dans les images .De nombreux descripteurs sont utilisés dans les systèmes de recherche pour décrire les images. Ceux-ci peuvent être différenciés selon deux niveaux :

1-les descripteurs bas niveau : ils décrivent le contenu bas niveau de l'image, principalement en termes de couleurs, textures et de formes. Ce sont les descripteurs les plus utilisés dans les systèmes actuels, car ils sont plus simples à mettre en place.

2-les descripteurs haut niveau : ils décrivent le contenu sémantique de l'image, et sont principalement des mots clés fournis par l'utilisateur lors de l'indexation. Ces attributs

sémantiques peuvent être le résultat d'un processus d'analyse sémantique automatisé utilisant des approches de classification ou des ontologies spécialisées.

Un descripteur visuel peut être soit local, ou global. Un descripteur global utilise les caractéristiques visuelles de toute l'image, tandis qu'un descripteur local utilise les caractéristiques des régions ou des objets pour décrire le contenu de l'image.

II .1.Les descripteurs de texture

La texture est une notion très générale qui peut être attribuée à presque tout dans la nature. Pour un être humain, la texture se rapporte principalement à une répétition spatiale particulière, en générale la répétition implique des variations locales de l'échelle, de l'orientation, ou tout autre caractéristique géométriques et optique.

Dans le dictionnaire, la texture est la répétition spatiale d'un motif de base dans différentes directions de l'espace. Cette définition indique la nature structurée de la surface considérée ; par exemple la texture d'un tissu, d'un mur de briques. Les partisans d'une telle définition s'orientent généralement vers une approche spectrale, structurelle ou statistique de la texture.

Plusieurs auteurs considèrent cependant qu'une texture ne possède pas de contours francs, mais plutôt un certain désordre, c'est-à-dire une disposition aléatoire que l'on pourrait considérer (avec une certaine approximation) comme visuellement homogène. On aurait alors une surface fermée, sans motifs isolables ou fréquences de répétition discernables : c'est le cas d'une photographie à distance d'herbes, de sable, de graviers etc. Les partisans d'une telle définition s'orientent généralement vers une approche probabiliste.

La texture est le second attribut visuel largement utilisé dans la recherche d'image par le contenu. Elle permet de combler un vide que la couleur est incapable de le faire, notamment lorsque les distributions des couleurs sont très proches. Elle est vue comme un ensemble de pixels spatialement agencés selon un certain nombre de relations spatiales. Les recherches sur la modélisation des textures se sont portées sur la caractérisation de ces relations spatiales. De nombreuses approches et modèles ont été proposées pour la caractérisation de la texture ; La section suivante traitera les modèles de texture.

II .1.1 .Taxonomie des modèles de textures

Les variations d'intensité qui caractérisent la texture dans une image sont généralement dues à une certaine variation physique sous-jacente dans la scène ; ce qui explique le fait qu'il n'existe pas une définition précise de la texture dans la littérature de vision par ordinateur [1].

En dépit de cela, il y a un certain nombre des propriétés intuitives de texture qui sont généralement pris en charge tel que :

- La texture d'un point n'est pas définie, donc la texture est une propriété contextuelle et sa définition doit impliquer les valeurs de niveaux de gris dans un voisinage spatial. La taille de ce voisinage dépend du type de texture, ou de la taille des primitives définissant la texture.
- La texture d'une image peut être perçue à différentes échelles ou niveaux de résolution. Par exemple, la texture représentée dans un mur de briques dans une résolution grossière, la texture est perçue comme formée par les différentes briques dans le mur; les détails à l'intérieur de la brique sont perdus. À une résolution plus élevée, seulement quelques briques sont dans le champ de vision, et la texture perçue montre les détails dans la brique.
- Une région est perçue comme ayant de la texture lorsque le nombre d'objets primitifs perçus dans la région est grand.
- Les propriétés suivantes jouant un rôle important dans la texture : l'uniformité, la densité, la grossièreté, la rugosité, la régularité, la linéarité, la directivité, la fréquence. Certaines de ces qualités ne sont pas indépendantes. Par exemple, la fréquence n'est pas indépendante de la densité et la propriété de la direction ne s'applique qu'aux textures directionnelles.
- Le fait que la perception de la texture a donc de nombreuses dimensions différentes est une raison importante pour laquelle il n'existe aucune méthode unique de la représentation de texture qui est suffisante pour une variété de textures.

II .1.1.1 L'approche statistique

Une des qualités définissant la texture est la distribution spatiale des valeurs de niveau de gris. La méthode statistique a été l'une des premières méthodes proposées dans la littérature. Et un grand nombre des caractéristiques de texture ont été proposées, mais ces caractéristiques ne sont pas indépendantes; voir la figure suivante:

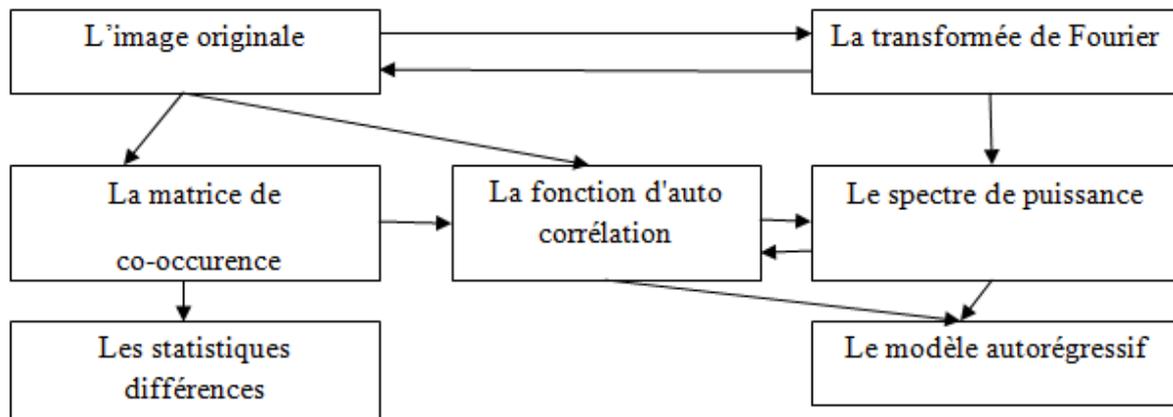


Figure 1 les interactions entre les méthodes et l'image d'origine [1]

L'approche statistique inclue les statistiques de premier ordre (l'ordre détermine le nombre de pixels mis en jeu lors de la caractérisation de la distribution statistique des primitives de la texture), qui peuvent être calculées à partir de l'histogramme (moyenne, variance, écart-type, etc.) et les statistiques de second ordre (matrice de cooccurrence).

La méthode statistique modélise les notions qualitatives usuelles de texture: granularité, contraste, homogénéité, répétitivité, fragmentation, orientation, etc.

La granularité est un trait dominant de la texture et même parfois par abus de langage synonyme de texture. Le grain qui donne la granularité est lui-même constitué des pixels voisins possédant le même niveau de gris. La taille et la densité des grains déterminent le niveau de finesse de la texture.

Le contraste est basé sur le nombre de niveaux de gris et leur taux de variation. Changer le contraste c'est modifier ces paramètres, ce qui modifie la qualité de l'image mais pas sa structure.

L'orientation est une propriété globale pour une région et traduit la direction générale prise par les motifs ou grains d'une texture.

1. La matrice de co-occurrence (Gay Level Co-Occurrence Matrix GLCM)

Beaucoup des caractéristiques statistiques de textures sont basé sur la matrice de co-occurrence introduite initialement par Haralick [2], GLCM donne une description de la structure locale ainsi qu'une information sur le champ aléatoire qui contrôle les détails à petite échelle. Cette approche est basée sur la probabilité jointe de la distribution des pixels dans l'image. l'élément $P_{d\theta}(i,j)$ de la matrice de co-occurrence définit la fréquence d'apparition des couples de niveaux de gris i,j pour les couples de pixels séparés par une distance d selon la direction θ . Donc le calcul de la matrice nécessite le choix d'une distance et d'un angle de déplacement. Aujourd'hui il ya beaucoup de progrès dans les techniques d'imagerie, de plus en plus les images sont en trois dimension (ou plus). La matrice de co-occurrence n'échappe pas cette progression avec son modèle en de 3D. En forme tridimensionnelle, la matrice GLCM peut être étendu en spécifiant un déplacement en utilisant $d = (dx, dy, dz)$, où dx et dy sont les mêmes que décrit pour la matrices de co-occurrences en deux dimension, et dz représentant le nombre de voxels déplacés tout au long de l'axe z d'une image 3D, [3]. Nous détaillerons la matrice de co-occurrence GLCM dans le chapitre suivant. La matrice de co-occurrence a été largement utilisé dans la recherche d'images médicales par le contenu, elle fournit des indications scéniquement approfondie sur la distribution spatiale des variations tonale, et l'orientation de la surface utilisée pour l'extraction des caractéristique, que ce soit en 2D [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] et même pour la 3D [3].

2. Les caractéristique d'auto corrélation (Autocorrelation features)

La fonction d'auto corrélation d'une image peut être utilisée pour évaluer le montant de la régularité ainsi que la finesse / grossièreté de la texture dans l'image. Formellement, la fonction d'autocorrélation d'une image $I(x, y)$ est définie comme suit :

$$p(x, y) = \frac{\sum_{u=0}^N \sum_{v=0}^N I(u, v) I(u + x, v + y)}{\sum_{u=0}^N \sum_{v=0}^N I^2(u, v)} \quad \text{Équation 1}$$

Cette fonction est en relation avec la taille de la texture primitive (c-à-d, la finesse de la texture). Si la texture est grossière, alors la fonction d'autocorrélation va baisser lentement sinon, elle va baisser très rapidement.

La fonction d'autocorrélation est également lié au spectre de puissance de la transformée de Fourier (the power spectrum of the Fourier transform) voir la figure suivante.

Considérons la fonction d'image dans le domaine spatial $I(x,y)$ et sa transformé de Fourier $F(u,v)$. La quantité $|F(u,v)|^2$ est défini comme le spectre de puissance ou $|\cdot|$ est le module d'un nombre complexe. L'exemple de la figure 02 illustre l'effet de la direction d'une texture sur la distribution de l'énergie dans le spectre de puissance.

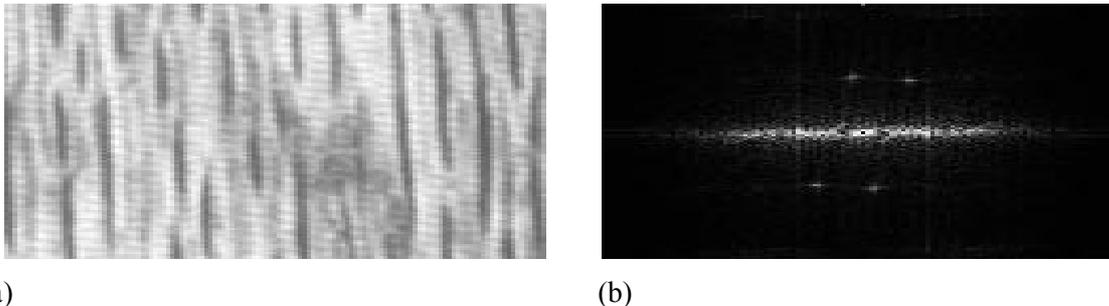


Figure 2 les Caractéristiques de texture provenant du spectre de puissance (a) la texture d'une image, et (b) spectre de sa puissance.

II.1.1.2. Les méthodes géométrique (Geometrical methods)

Les méthodes d'analyse de texture dans basé sur les approches géométriques sont caractérisé par leur définition de la texture comme étant composée « d'éléments de texture » ou de primitives. La méthode d'analyse dépend généralement des propriétés géométriques de ces éléments de texture.

Il y a deux grandes approches d'analyse de la texture, une fois les éléments de texture sont identifiés dans l'image

- le calcule des propriétés statistiques extraites à partir des éléments de la texture pour les utiliser comme des caractéristiques de la texture.
- Extraction des règles de placement pour décrire la texture.

1. Les Caractéristiques de Voronoï tessellation

Tuceryan et Jain [1] ont proposé l'extraction des jetons de texture en utilisant les propriétés de la tessellation de Voronoï de l'image donnée.

Supposons un ensemble S de trois ou plusieurs jetons (pour plus de simplicité, nous supposons que le jeton est un point) dans un plan euclidien, et ces points ne sont pas tous alignés. Considérons une paire arbitraire de points P et Q , la bissectrice de la ligne de jonction PQ est le lieu des points équidistants pour P et Q . Le demi-plan H_P^Q (H_Q^P) est le lieu des points les plus proches de P (Q) que Q (P). $\bigcap_{Q \neq P} H_P^Q$ définit une région polygonale constituée par des points plus proches de P que les autres ; Cette région est appelée polygone de Voronoï et l'ensemble des polygones complets est appelé le diagramme de Voronoï de S . Le diagramme de Voronoï avec les polygones incomplets dans l'enveloppe convexe définit une tessellation de Voronoï de l'ensemble du plan [1].

Tessellation de Voronoï a été proposé en raison de ses propriétés souhaitables dans la définition des voisinages spatiaux et parce que les distributions spatiales locales des jetons se retrouvent dans les formes des polygones de Voronoï.

Les jetons de texture sont extraits puis la tessellation est construite. Les jetons peuvent être aussi simples comme les points de fort gradient dans les structures de l'image ou complexes tels que des segments de ligne.

2. L'approche structurale (Structural methods)

Cette catégorie d'approches modélise les relations spatiales entre éléments primitifs constituant l'image. Les règles de placement ou d'agencement spatial déterminent l'existence et la nature de la texture.

Cette classe d'algorithmes, en général, est limitée en puissance sauf s'il s'agit des textures très réguliers.

L'analyse de la texture structurale se compose de deux étapes majeures:

- (a) L'extraction des éléments de texture.
- (b) les règles de placement.

Il ya un certain nombre de façons pour extraire les éléments de texture dans les images. Habituellement les éléments de texture comprennent les régions de l'image avec des niveaux de gris uniformes.

II .1.1.3.Les méthodes fondées sur un modèle (*Model based methods*)

Ces méthodes d'analyse de texture sont basées sur la construction d'un modèle d'image qui peut être utilisé non seulement pour décrire la texture, mais aussi pour faire la synthèse. Les paramètres du modèle capturent les qualités essentielles perçues de la texture. Par exemple les champs de Markov aléatoires- **Markov random fields** (MRFs)- ont été très populaires pour la modélisation des images. Ils sont capables de capturer l'information contextuelle locale dans une image. Ces modèles supposent que l'intensité de chaque pixel de l'image dépend seulement de l'intensité des pixels voisins. Les modèles MRF ont été utilisés dans différentes applications de traitement d'image comme la classification de texture, la segmentation et la compression [1].

II .1.1.4.Les méthodes de traitement de signal (*Signal processing methods*)

Des recherches en psychophysique ont fourni la preuve que le cerveau humain effectue une analyse en fréquence de l'image. La texture est particulièrement adapté pour ce type d'analyse en raison de ses propriétés .Cette section passera en revue les différentes techniques d'analyse de texture qui s'appuient sur les techniques de traitement du signal.

La plupart de ces techniques calculent des caractéristiques à partir des filtres de l' image qui sont ensuite utilisés dans les tâches de classification ou de segmentation.

1. Filtres spatiaux (*Spatial domain filters*)

Les filtres spatiaux sont la façon la plus directe utilisée pour capturer les propriétés d'une texture d'image.

La mesure de bord est généralement calculée par des masques de bord simples, telles que celui de

Robert ou l'opérateur Laplacien .

- Filtre de Roberts

Les filtres de Roberts ont des masques de taille 2×2 , qui sont les suivantes [14]:

Filtre diagonal : a =

1	0
0	-1

Filtre anti-diagonal : b =

0	1
-1	0

L'image du contour est obtenue par la racine de la somme de convolution de l'image originale avec les deux fenêtres a et b.

➤ Filtre du Laplacien :

Le Laplacien mesure la rapidité des changements d'intensité sur le voisinage. Il approxime de façon très simple la dérivée seconde de l'image. Le noyau 3×3 de ce filtre est [14] :

0	-1	0
-1	4	-1
0	-1	0

2. Le domaine de Fourier (Fourier domain filtering)

La transformée de Fourier est une analyse globale des fréquences d'un signal. Les canaux multiples accordés à des fréquences différentes sont également dénommé multi-résolution dans la littérature. Le concept de la transformation multi-résolution est affiné et développé dans le modèle d'ondelettes décrit ci-dessous. Dans la recherche des images par contenu le filtre de Fourier est utilisé pour analyser l'image [8].

3. Le model de Gabor et Ondelette (Gabor and Wavelet models)

De nombreuses applications nécessitent l'analyse dans le domaine fréquentiel pour être localisée dans le domaine spatial. Ceci est habituellement traité par l'introduction de la dépendance spatiale dans l'analyse de Fourier. La manière classique de faire ceci est par ce qu'on appelle la transformée de Fourier à fenêtre. the window Fourier transform (or short-time Fourier Transform) d'un signal à une dimension $f(x)$ et qui est définie comme suit:

$$F_W(u, \zeta) = \int_{-\infty}^{\infty} f(x)W(x, \zeta)e^{-j2\pi u x} dx \quad \text{Équation 2}$$

Lorsque la fonction de fenêtre $W(x)$ est gaussienne, la transformée devient une transformée de Gabor.

Les limites de la résolution dans le temps et le domaine fréquentiel de la transformée de Fourier à fenêtre sont déterminés par le produit temps-fréquence- la bande passante.

$$\Delta t \Delta u \geq \frac{1}{4\pi}$$

Une fois la fenêtre de la transformée de Fourier choisie, la résolution de temps-fréquence est fixée sur l'ensemble du plan temps-fréquence. Pour surmonter la limitation de la résolution de la fenêtre nous laissons Δt et Δu varier dans le domaine temps-fréquence. Intuitivement, la résolution temporelle doit augmenter à mesure que la fréquence centrale du filtre d'analyse est augmentée.

Une fonction de Gabor bidimensionnels comprend une onde sinusoïdale d'une certaine fréquence et l'orientation modulé par une enveloppe gaussienne, et est donnée par :

$$f(x, y) = \exp\left(-\frac{1}{2}\left[\frac{X^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right]\right) \cos(2\pi u_0 + \phi) \quad \text{Équation 3}$$

Où u_0 et ϕ sont la fréquence et la phase de l'onde sinusoïdale.

Et les valeurs σ_x et σ_y sont les tailles de l'enveloppe de la Gaussienne dans les directions x et y.

Les filtres de Gabor

Il s'agit des filtres extrayant une information fréquentielle localisée et orientée.

Les filtres de Gabor sont le produit de Gaussiennes par des sinus ou des cosinus

Sinus / cosinus : analyse en fréquences

Gaussienne : différents niveaux de lissage (échelles)

$$G(x, y) = \cos(w_x x + w_y y) e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad \text{Équation 4}$$

$$G(x, y) = \sin(w_x x + w_y y) e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad \text{Équation 5}$$

Pour l'extraction des caractéristiques de texture le filtre de Gabor a été utilisé [6] [15].

La transformé de Gabor est utilisé dans les systèmes de recherche d'images médicales avec la 3D Gabor Transform (3D GT) [3].

Les ondelettes (Wavelet)

La notion d' " ondelettes" ou "Wavelets" a été utilisé pour la première fois au début des années 80 par le géophysicien français J.Morlet pour désigner des fonctions mathématiques utilisées dans la représentation des données sismiques.

Selon Yves Meyer, "c'est le signal transitoire le plus simple que l'on puisse imaginer". Une ondelette est un objet mathématique : il s'agit d'une fonction, localisée (par exemple, nulle hors d'un certain intervalle) et oscillante.

Il existe de nombreuses ondelettes. Chacune possède des qualités bien spécifiques, comme par exemple :

– La symétrie : utile pour éviter le déphasage,

– La régularité : utile pour obtenir des signaux reconstruits lisses et réguliers.

L'ondelette mère est un prototype de base pour générer toute une famille d'ondelettes (Les *ondelettes filles*), aussi appelées atomes de base. Ce sont des copies dilatées, comprimées et translatées dans le temps de l'ondelette mère. D'où la formule suivante d'un atome de base

$$\psi_{a,b} = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), a \in \mathbb{R}^{*+}, b \in \mathbb{R} \quad \text{Équation 6}$$

Le paramètre $a > 0$ est le paramètre d'échelle, c'est lui qui va agir sur la compression ou la dilatation de l'atome de base. Lorsque $0 < a < 1$, l'atome de base est *comprimé* et lorsque $a > 1$, l'atome de base est *dilaté*. Le paramètre b est le paramètre de position, il permet de faire translater l'atome de base à gauche ($b < 0$) ou à droite ($b > 0$), le long du signal à analyser.

La transformée en ondelettes

Une manière pratique d'introduire la transformée en ondelettes consiste à établir une analogie avec la musique. De la même manière que les notes de musique possèdent une certaine hauteur et une durée, l'ondelette a une certaine fréquence et pendant un temps limité. Prenons l'exemple d'un extrait d'une partition musicale, l'expérience consiste à comparer la transformée de Fourier et la transformée en ondelettes de l'extrait musical, afin de mettre en avant quelques remarques préliminaires concernant l'analyse par ondelettes. Les transformées de Fourier nous indiquent quelles sont les notes qui sont présentes dans l'extrait musical, en nous renseignant précisément sur leurs fréquences ; on ne sait cependant pas l'ordre dans lequel les notes sont jouées. Dans la transformée en ondelettes, on possède une bonne résolution fréquentielle aux basses fréquences (au détriment d'une perte de localisation en temps) et qu'à l'inverse, on possède une bonne résolution temporelle aux hautes fréquences (au détriment d'une perte de localisation en fréquence).

Les ondelettes ont été largement utilisées dans les systèmes de recherche des images médicales par le contenu comme par exemple [16], [8], les ondelettes de Gabor [13] [17], et dans [18], pour la base de donnée (ImageCLFE Medical Image Dataset) les ondelettes 3D [3].

L'ondelette discrète de (Haar) a été utilisée dans [19], et la transformé discrète d'ondelette (DWT) pour la base de donnée (COREL Photo Image Dataset) dans [18]. Le descripteur d'ondelette pour le moment de Zernike (WZMD) [20] a été également utilisé pour la récupération et la classification des objets.

II.2. Les descripteurs de formes

La forme est l'un des attributs de bas niveau les plus utilisés pour décrire la structure géométrique du contenu visuel. Même si la caractérisation du contenu géométrique s'est avérée complexe plusieurs primitives géométriques sont montrées leurs intérêts dans les systèmes de recherche.

Contrairement aux descripteurs de texture et de couleur, qui peuvent être utilisés pour décrire globalement l'image, les descripteurs de forme décrivent des propriétés locales aux régions composant l'image. Donc, leur robustesse sera fonction de l'exactitude de la préalable segmentation et elle est, généralement, mesurée par le degré d'invariance à la translation, rotation et au facteur d'échelle.

La représentation de la forme est une tâche difficile. C'est parce que quand un objet en 3-D du monde réel est projeté sur un plan d'image en 2-D, une dimension de l'information de l'objet est perdue. En conséquence, la forme extraite de l'image ne représente que partiellement l'objet projeté.

Pour rendre le problème encore plus complexe, la forme est souvent corrompu avec le bruit, la distorsion arbitraire et l'occlusion.

Nous présentons dans ce qui suit quelques méthodes de description de la forme.

II.2 .1. Classification des formes

La représentation de la forme et ces techniques de description peuvent être classés en deux classes de méthodes: **Les méthodes de contour** et **les méthodes basées sur la région** [21]. Sous chaque catégorie, les différentes méthodes sont divisées en deux approches *structurelles* et *globales*. Les caractéristiques de forme sont dérivées du domaine spatial ou à partir d'une transformé de domaine.

Ces techniques sont discutées en détail dans les sections suivantes.

II.2.1.1. La représentation de la forme basée sur le contour

Il ya deux types d'approches pour la modélisation du contour de la forme: l'approche continue (globale) et l'approche discrète (structurale). Les approches continues ne divisent pas la forme en sous-parties, mais généralement un vecteur d'objet est utilisé pour décrire la forme.

Les approches discrètes brisent la frontière de la forme en segments dites primitives avec l'utilisation d'un critère particulier. La représentation finale est généralement une chaîne, un graphe ou un arbre.

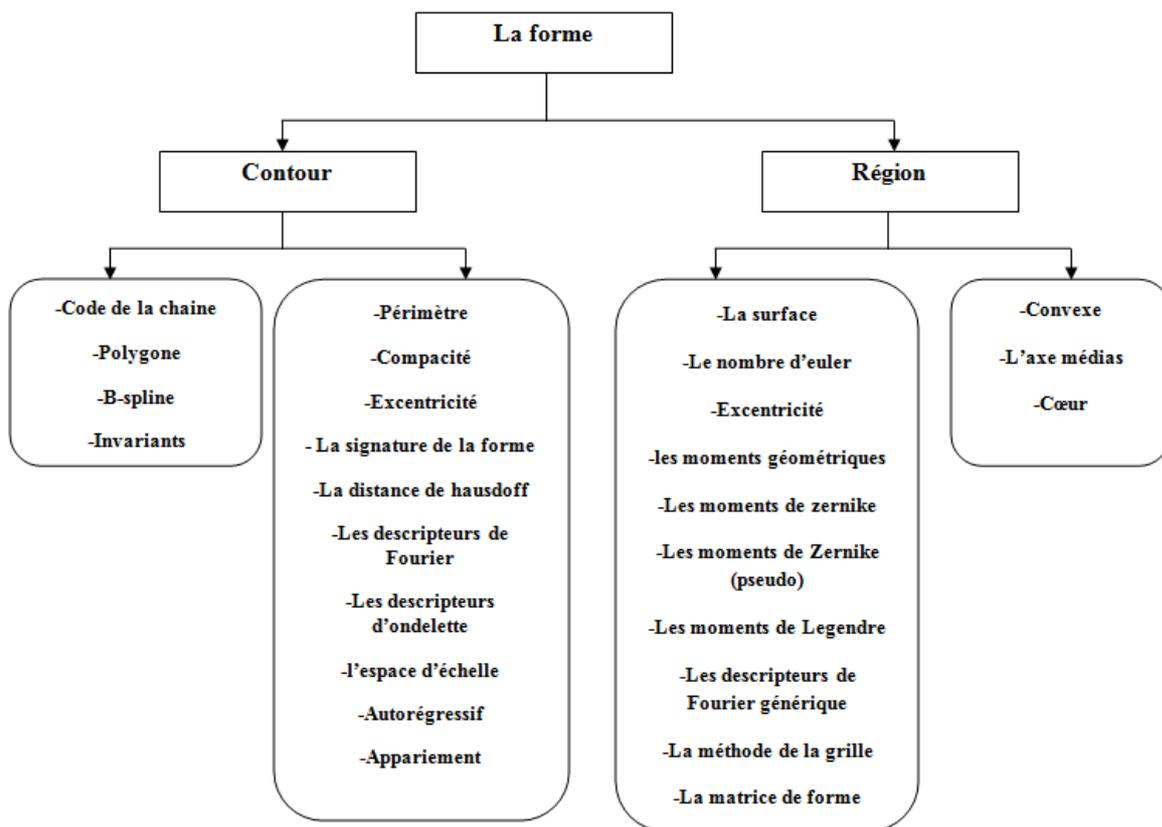


Figure 3 Classification des techniques de la représentation de forme [21]

II.2.1.1.1. Les méthodes globales

Les techniques globales prendront le contour entier pour représenter la forme ; La correspondance entre les formes peut être dans le domaine spatial ou dans le domaine fréquentiel. Pour la description d'une forme, il ya toujours un compromis entre l'exactitude et l'efficacité.

D'une part, la forme devrait être décrite aussi précisément que possible; d'autre part, un descripteur de forme doit être aussi compact que possible afin de simplifier l'indexation et la récupération.

Les descripteurs globaux d'une forme sont compact, cependant, ils sont des descripteurs de forme très inexacts ; Ils doivent être combinés avec d'autres descripteurs de forme pour créer des descripteurs plus exacts et complets.

II.2 .1.1.1.1. Les simples descripteurs de forme (Simple shape descriptors)

Les simples descripteurs globaux de la forme sont : la surface (area), la circularité ($\text{perimeter}^2 = \text{area}$), l'excentricité (eccentricity) (longueur de grand axe / longueur du petit axe).

La surface et l'excentricité sont utilisées comme des descripteurs de forme dans la récupération des images [19], et les images médicales [7] [8].

II .2.1.1.1.2. La distance de Hausdorff (Hausdorff distance)

Est une méthode classique de la mesure de la correspondance entre les formes avec l'utilisation d'une similitude point -à-point d'objet. Elle est souvent utilisée pour localiser les objets dans une image et pour mesurer la similitude entre les formes.

Etant donnés deux formes représentées par deux ensemble de points

$A = \{a_1; a_2; \dots; a_p\}$ and $B = \{b_1; b_2; \dots; b_q\}$, la distance de Hausdorff entre A et B est défini par

$H(A; B) = \max(h(A; B); h(B; A))$ [21]

Ou $h(A; B) = \max_{a \in A} \min_{b \in B} \|a-b\|$.

S'il y a un seul point dans A qui est loin de tout point dans B, h (A, B) sera grande.

L'avantage de cette approche est que la forme peut être adaptée en partie ; par contre la distance de Hausdorff varie selon l'échelle et la rotation.

II .2.1.1.1.3. Une signature de forme (Shape signature)

Représente la forme par une fonction unidimensionnelle extraite à partir des points de contour de la forme ; Notons que de nombreuses signatures de forme existent. Elles comprennent le profil centroïde, les coordonnées complexes, et la distance centroïde.

La signature de forme est sensible au bruit, et un changement léger dans la bordure de la forme peut provoquer des erreurs importantes dans la comparaison. Par conséquent, il n'est pas souhaitable de décrire directement une forme avec l'utilisation d'une signature de forme.

II.2 .1.1.1.4.Les moments de contours (Boundary moments)

Peuvent être utilisés pour réduire les dimensions de la représentation des bordures de la forme. En supposant que la bordure de la forme est représentée comme par z (i).

$$m_r = \frac{1}{N} \sum_{i=1}^N [Z(i)]^r \quad \text{Équation 7} \quad \mu_r = \frac{1}{N} \sum_{i=1}^N [Z(i) - m_1]^r \quad \text{Équation 8}$$

μ_r Le moment central.

m_r r moment.

N est le nombre de points de bordure.

Les moments normalisés sont invariants à la rotation et mise à l'échelle.

$$\mu_r = \sum_{i=1}^K (v_i - m)^2 p(v_i) \quad \text{Équation 9}$$

$$m = \sum_{i=1}^K v_i p(v_i) \quad \text{Équation 10}$$

L'avantage de descripteurs basés sur les moments de bordures, c'est qu'ils sont faciles à mettre en œuvre.

II .2.1.1.1.5. Les méthodes stochastique (Stochastic method)

Les séries chronologiques des modèles et surtout le modèle autorégressif (AR) ont été utilisés pour le calcul des descripteurs de forme.

Les méthodes de cette classe sont basées sur une approche stochastique. Un modèle linéaire autorégressif exprime la valeur d'une fonction par la combinaison linéaire d'un certain nombre de valeurs précédentes [21].

$$f_t = \alpha + \sum_{j=1}^m \theta_j f_{t-j} + \sqrt{\beta} w_t \quad \text{Équation 11}$$

θ_j Sont les coefficients du modèle AR.

m est l'ordre du modèle, c'est à dire, il indique combien de valeurs des fonction précédentes le modèle utilise.

$\sqrt{\beta} w_t$ le terme d'erreur courant ou résiduel.

α est proportionnelle à la moyenne des valeurs de fonction.

L'inconvénient de la méthode AR, c'est que dans le cas des bordures complexes, un petit nombre de paramètres AR ne suffit pas pour une description adéquate. Le choix de m est un problème complexe et il est pris généralement de manière empirique.

II .2.1.1.1.6. Les méthodes de l'espace échelle (Scale space method)

Pour corriger les problèmes de la sensibilité au bruit et aux variations de la limite de forme, la plupart des méthodes spatiales s'inspire de l'utilisation de l'analyse d'espace d'échelle.

La représentation de l'espace échelle d'une forme est créée par le suivi de la position des points d'infection dans une limite de forme filtrés par des filtres gaussiens passe-bas avec des largeurs variables.

Quand la largeur (σ) de filtre gaussien augmente, signifie les infections de bordure sont éliminés et la limite et la forme devient plus lisse.

II.2.1.1.2. Les méthodes structurel

II.2.1.1.2.1 . La représentation par un code de la chaîne

Dans cette approche, une courbe arbitraire est représentée par une séquence de vecteurs de petites unités de longueur et un nombre limité de directions possibles. Elle est appelée la méthode des unités de vecteur.

Dans la mise en œuvre, la limite numérique d'une image est superposée avec une grille, et les points de frontière sont approximés au point le plus proche la grille, puis une image échantillonnée est obtenue.

Avec un point de départ sélectionné, un code de la chaîne peut être généré en utilisant la chaîne de 4 directions ou 8-directions.

Le code de la chaîne a généralement une grande dimension et est sensible au bruit. Il est souvent utilisé comme une entrée à un niveau supérieur d'analyse.

II.2.1.1.2.2. La décomposition polygonale (Polygon decomposition)

La bordure de la forme est décomposée en segments de ligne par approximation polygonale. Les sommets du polygone sont utilisés en tant que primitives.

La caractéristique de chaque primitive est exprimée en une chaîne à quatre éléments qui se compose des angles interne, de la distance à partir du sommet suivant, et ses coordonnées x et y.

La similitude entre les deux formes est mesurée par la distance entre les deux chaînes de caractéristique.

Pour plus d'efficacité et pour des raisons de robustesse, seul un nombre fixe (5) de sommets sont choisis dans chaque forme [21].

II.2.1.2. Représentation de forme basée sur la région et les techniques de description (Region-based shape representation and description Techniques)

Dans cette approche tous les pixels à l'intérieur de la forme sont pris en compte ; pour obtenir la représentation de la forme, plutôt que seulement les informations du contour.

II.2.1.2.1. Les méthodes globales

II.2.1.2.1.1. Les moments invariants géométriques (Geometric moment invariants)

Les moments géométriques permettent de décrire une forme à l'aide des propriétés statistiques. Ils représentent les propriétés spatiales de la distribution des pixels dans l'image. Ils sont facilement calculés et implémentés, par contre, cette approche est très sensible au bruit et aux déformations.

La formule générale des moments géométriques est donnée par la relation suivante :

$$m_{p,q} = \sum_{p=0}^m \sum_{q=0}^n x^p y^q f(x, y) \quad \text{Équation 12}$$

$P+q$ est l'ordre du moment. le moment d'ordre 0 $m_{0,0}$ représente l'aire de la forme de l'objet.

Les deux moments d'ordre 1 $m_{0,1}$ et $m_{1,0}$, associés au moment d'ordre 0 $m_{0,0}$ permettent de calculer le centre de gravité de l'objet. les coordonnées de ce centre sont :

$$X_c = \frac{m_{1,0}}{m_{0,0}} \quad Y_c = \frac{m_{0,1}}{m_{0,0}}$$

Il est possible de calculer à partir de ces moments l'ellipse équivalente à l'objet. Afin de calculer les axes de l'ellipse, il faut ramener les moments d'ordre 2 au centre de gravité :

$$m_{2,0}^g = m_{2,0} - m_{0,0} x_c^2 \quad \text{Équation 13}$$

$$m_{1,1}^g = m_{1,1} - m_{0,0}x_c y_c \quad \text{Équation 14}$$

$$m_{0,2}^g = m_{0,2} - m_{0,0}y_c^2 \quad \text{Équation 15}$$

Puis on détermine l'angle d'inclinaison de l'ellipse α :

$$\alpha = \frac{1}{2} \arctan \frac{2m_{1,1}^g}{m_{2,0}^g - m_{0,2}^g} \quad \text{Équation 16}$$

Le principal problème avec des moments géométriques est que seuls quelques invariants dérivés à partir des moments d'ordre inférieur ne sont pas suffisants pour décrire avec précision la forme. Et les invariants d'ordre supérieur sont difficiles à tirer.

Les moments invariants ont été utilisés pour l'indexation d'images [22], et les images médicales [8] ; et dans le système de recherche des images Endoscopie [7] ou les trois moments centraux ont été utilisé pour extraction des caractéristiques.

Les moments invariants de Hu

HU a proposé un ensemble de sept moments invariants aux translations, rotation et changement d'échelle. ces moments sont très utilisés dans la littérature pour la description des formes en vue d'une classification ou d'une indexation, mais sont assez sensibles aux bruits.

II .2.1.2.1.2. Les moments orthogonaux (Orthogonal moments)

Par opposition aux moments géométriques qui sont définis par rapport à une base quelconque $(x^p y^q)$, les moments orthogonaux, comme leur nom l'indique, sont définis dans une base orthogonale, ce qui évite la redondance des informations portées par chacun des moments.

Les deux types de moments orthogonaux les plus utilisés sont : les moments de Legendre et les moments de Zernike, dont nous donnons les définitions ci-dessous.

Le moment de Legendre

Les moments de Legendre sont définis à partir des polynômes du même nom. Ils sont définis dans le carré unité $[-1,1]$ et $[-1,1]$, ce qui oblige à normaliser l'objet dont on veut calculer les moments.

Le polynôme de Legendre d'ordre n est donné par

$$\forall x \in [-1,1], \forall n \in \mathbb{N}, P_n(x) = \frac{1}{2^n n!} \frac{d^n (x^2 - 1)^n}{dx^n} \quad \text{Équation 17}$$

Les polynômes de Legendre $\{P_n(x)\}$ forment une base complète et orthogonale sur le domaine de définition $[-1,1]$:

$$\forall (x, y) \in [-1,1]^2, \forall (m, n) \in \mathbb{N}^2, \int \int_{-1}^1 P_m(x) P_n(y) dx dy = \frac{2}{2m+1} \delta_{mn}$$

Équation 18

δ_{mn} Représente la fonction de Kronecker.

Les moments de Legendre d'ordre N sont donc donnés par :

$$\forall (x, y) \in [-1,1]^2, \forall (p, q) \in \mathbb{N}^2, N = p + q, \quad L_{pq} = \frac{(2p+1) \times (2q+1)}{4} \int \int_{-1}^1 P_p(x) P_q(y) f(x, y) dx, dy \quad \text{Équation 19}$$

Ou, $f(x,y)$ est le niveau de gris d'un pixel de l'image I sur laquelle on calcule le moment. À partir de cette équation, nous pouvons générer une infinité de moments de Legendre. Plusieurs études sur la reconnaissance des formes ont démontré que l'utilisation des moments de Legendre d'ordres inférieurs (jusqu'à l'ordre 3) est suffisante pour représenter la forme globale de l'entité donnée.

Les moments de Zernike

Ce type des moments a été initialement introduit par Teague et sont construits à partir des polynômes complexes et forment un ensemble orthogonal complet définie sur le disque unité. Ils sont invariants par rapport à la rotation et aux changements d'échelles et présentent des

capacités intéressantes en termes de résistance aux bruits, efficacité informative et possibilité de reconstruction.

Les moments orthogonaux de Zernike d'ordre p sont définis de la manière suivante :

$$A_{mn} = \frac{m+1}{\pi} \iint I(x, y) [V_{m,n}(x, y)] dx dy \quad \text{Équation 20}$$

Où m et n définissent l'ordre du moment et $I(x,y)$ le niveau de gris d'un pixel de l'image I sur laquelle on calcule le moment.

Les polynômes de Zernike $V_{m,n}(x,y)$ sont exprimés en coordonnées polaires :

$$V_{m,n}(r, \theta) = R_{m,n}(r) e^{-jn\theta} \quad \text{Équation 21}$$

Où $R_{m,n}(r)$ est le polynôme radial orthogonal :

$$R_{m,n}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s \frac{(m-s)!}{s! \left(\frac{m+|n|}{2} - s\right)! \left(\frac{m-|n|}{2} - s\right)!} r^{m-2s} \quad \text{Équation 22}$$

Avec $n=0, 1, 2, 3 \dots \infty \leq |m| \leq n$ et $n-|m|$ un entier pair.

Les polynômes de Zernike sont orthogonaux, et donc les moments correspondants le sont également. Cette propriété d'orthogonalité annule l'effet de redondance de l'information portée par chaque moment.

Les moments de Zernike ont été utilisés dans les systèmes de recherche et d'indexation d'images [22] ; et même dans CBMIR [23] .

II.2 .1.2.1.4. Le descripteur de Fourier générique (Generic Fourier descriptor)

Bien que le descripteur Zernike a une solide performance, il présente plusieurs tout de même quelques lacunes. Tout d'abord, le noyau de moments de Zernike est complexe à calculer.

Les caractéristiques radiales et les caractéristiques circulaires capturées par les moments de Zernike ne sont pas compatibles. En effet, un est dans le domaine spatial et l'autre est dans le domaine spectral. Pour surmonter ces lacunes, un descripteur de Fourier générique (GFD) a

été proposé. GFD est construit par l'application d'une transformée de Fourier 2-D sur une trame polaire.

Le descripteur GFD est plus simple à calculer. En effet, les caractéristiques sont dans le domaine spectral et ont des meilleures performances d'indexation en raison de son analyse multi-résolution dans les deux directions radiales et circulaire de la forme.

Dans [8] le descripteur de Fourier a été utilisé pour extraire les caractéristiques de la forme.

Le descripteur de Fourier générique (GFD), est utilisé pour la recherche et la classification des objets [20].

II.2.1.2.1.5. La méthode basé sur la grille (Grid based method)

Le principe de cette méthode est qu'une grille de cellules est superposée sur la forme, la grille est ensuite balayée de gauche à droite et de haut en bas. Le résultat est une image bitmap.

Les cellules couvertes par la forme sont attribués à 1, et ceux qui ne sont pas couvertes par la forme sont attribués à 0. La forme peut alors être représentée comme un vecteur de caractéristique binaire. La distance de Hamming est utilisée pour mesurer la similitude entre les deux formes.

Les avantages du descripteur de la grille sont sa simplicité. Le principal problème avec cette méthode est la normalisation de l'axe principal de rotation. Le grand axe est sensible au bruit et peu fiables.

II.2.1.2.2. Les méthodes structurels

Les méthodes structurelles décomposent la forme en parties qui sont ensuite utilisées pour représenter la forme et la description de l'objet.

II.2.1.2.2.1. Convexe (Convex hull)

Une région R est convexe si et seulement si pour deux points quelconques $x_1; x_2 \in R$, la ligne entière du segment x_1x_2 est à l'intérieur de la région. [21]

convex hull d'une région est la région convexe H la plus petite qui satisfait la condition $R \subset H$. la différence $H-R$ est appelé « carence convexe » D d'une région R .

Une meilleure représentation de la forme peut être obtenue par un procédé récursif, ce qui conduit à un arbre de concavités.

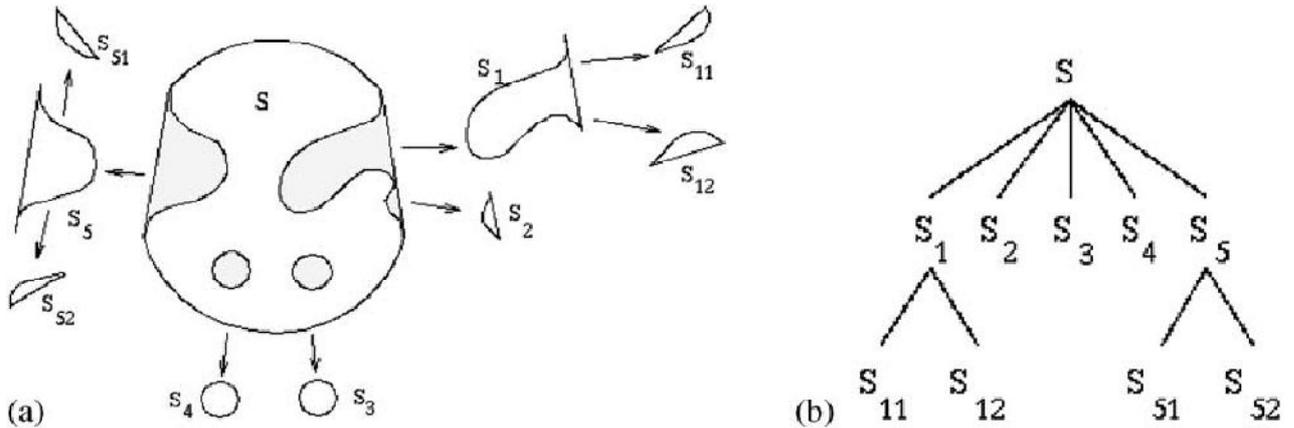


Figure 4 (a) Convexe et ses concavités; (b) l'arbre de concavité [21]

II .2.1.2.2.2. L'axe médian (Medial axis)

Comme convexe, un squelette d'une région peut également être utilisé pour la description et la représentation d'une forme.

Un squelette peut être défini comme un ensemble connexe de lignes médianes. Par exemple dans le cas des caractères dessinés à la main, le squelette peut être considéré comme le chemin effectivement parcouru par la plume. L'idée de base du squelette est de supprimer les informations redondantes tout en conservant uniquement les informations concernant la topologie de la structure d'objet qui peut aider à la reconnaissance.

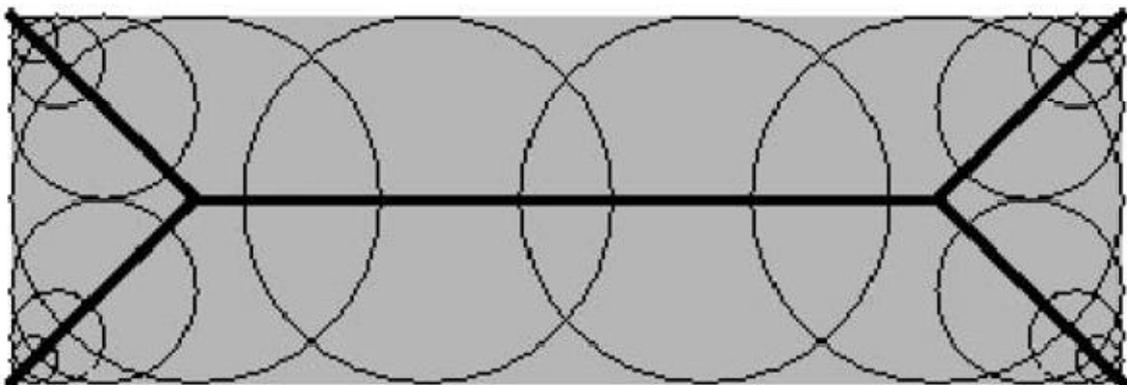


Figure 5 Construction de l'axe médian d'une forme rectangulaire [21]

L'axe médian est le lieu des centres des disques maximaux qui correspondent à la forme ; La ligne en gras dans la figure est le squelette de la forme rectangulaire ombragée. Le squelette peut alors être décomposé en segments et représenté par un graphe selon certains critères. La concordance entre les formes devient un appariement de graphe.

Le calcul de l'axe médian est un problème assez difficile. L'axe médian acquit de cette manière est appelé *le noyau de la forme*.

II.3. Les descripteurs de couleur :

La couleur est une caractéristique riche d'information et très utilisée pour la représentation des images. Elle forme une partie significative de la vision humaine. La couleur est devenue la première signature employée pour la recherche d'image par le contenu en raison de son invariance par rapport à l'échelle, la translation et la rotation. Ces valeurs tridimensionnelles font que son potentiel discriminatoire soit supérieur à la valeur en niveaux de gris des images. Une indexation couleur repose sur deux principaux choix : l'espace colorimétrique et le mode de représentation de la couleur dans cet espace.

II.3.1. Les espaces de couleurs

Avant de sélectionner un type de description du contenu couleur, il convient de choisir un espace de couleurs. Une couleur est généralement représentée par trois composantes. Ces composantes définissent un espace de couleurs. Plusieurs études ont été réalisées sur l'identification d'espaces colorimétriques les plus discriminants mais sans succès car il n'existe pas d'espace de couleur idéal. Il existe plusieurs espaces colorimétriques qui ont chacun certaines caractéristiques intéressantes.

L'espace RGB est très simple à utiliser, car c'est celui employé par de nombreux appareils de capture d'image qui effectuent leurs échanges d'information uniquement en utilisant les triplets (Rouge, Vert, Bleu). On parle d'espace colorimétrique orienté matériel. Cette manière de représenter la couleur est extrêmement basique, puisqu'aucun traitement n'est nécessaire. Cependant, ces trois composantes sont fortement corrélées, et cet espace est sensible aux changements d'illumination, et ne correspond pas au processus de perception humaine. La représentation des couleurs dans cet espace donne un cube de Maxwell comme illustré dans la

figure. En vision par ordinateur, les trois composantes sont représentées par trois octets (24 bit). On obtient avec ce choix une palette de 256^3 couleurs possibles. Les systèmes de recherche d'image par le contenu utilisent une quantification pour réduire le nombre de couleurs.

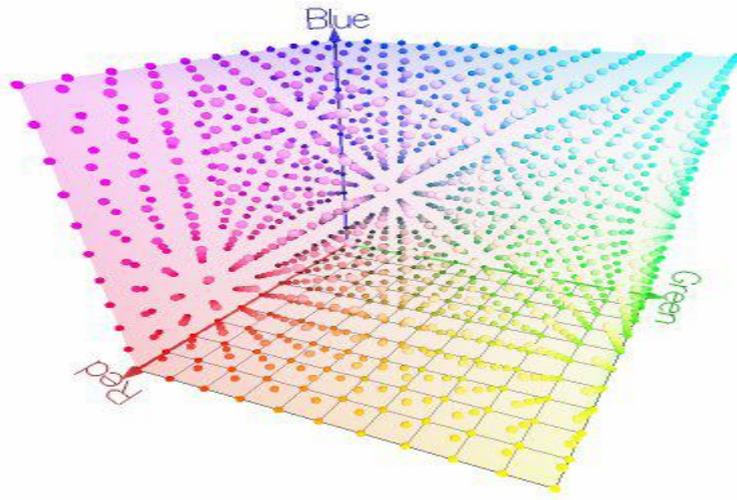


Figure 6 Espace couleur RGB

De la représentation RGB, nous pouvons dériver d'autres genres de représentation de couleurs en employant des transformations linéaires ou non linéaires.

II.3.1.1. Espace couleur HSI

Pour les applications qui requièrent des manipulations de couleur comme le traitement d'images, le modèle **HSI** (teinte saturation intensité) est couramment utilisé.

La luminosité, la teinte et la saturation sont des caractéristiques qui sont généralement utilisées pour distinguer une couleur d'une autre. La luminosité embarque la notion chromatique d'intensité (l'amplitude de la puissance de la lumière). L'ensemble teinte saturation est appelé chrominance.

L'espace couleur **HSI** a été développé pour positionner les couleurs qui sont plus faciles à quantifier pour les humains.

Deux facteurs principaux font du modèle **HSI** l'un des modèles idéals pour développer des applications de vision :

- Le découplage de la composante intensité de l'information couleur.
- La relation étroite entre la chrominance et la manière dont les humains perçoit la couleur.

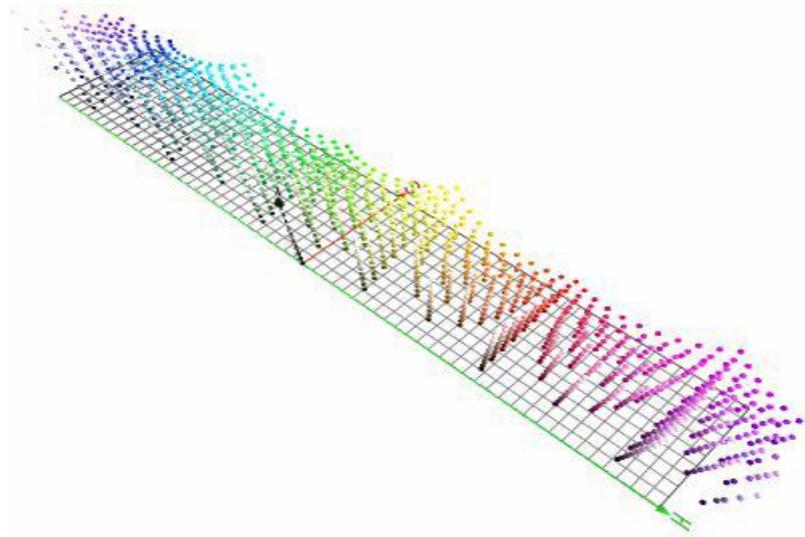


Figure 7 Espace couleur HSI

Où

$$H = \arctan\left(\frac{\beta}{\alpha}\right)$$

$$S = \sqrt{\alpha^2 + \beta^2}$$

$$I = (R + G + B) / 3$$

Avec **Équation 23**

$$\alpha = R - \frac{1}{2}(G + B)$$

$$\beta = \frac{\sqrt{3}}{2}(G - B)$$

II.3.1.2. Espace couleur YCbCr

L'œil humain est plus sensible à la luminance (où luminosité) qu'à la chrominance d'une image. Or la luminance est présente dans les trois couleurs rouge, vert, bleu. Ainsi on transforme les composantes RGB de l'image en une composante de luminance notée **Y** et en deux composantes de chrominance notées **Cb** et **Cr**. Pour cela on utilise des formules de passage (pour un pixel) :

$$\begin{aligned}
 Y &= 0.2989 \times R + 0.5866 \times G + 0.1145 \times B \\
 Cb &= -0.1688 \times R - 0.3312 \times G + 0.5000 \times B \\
 Cr &= 0.5000 \times R - 0.4184 \times G - 0.0816 \times B
 \end{aligned}
 \quad \text{Équation 24}$$

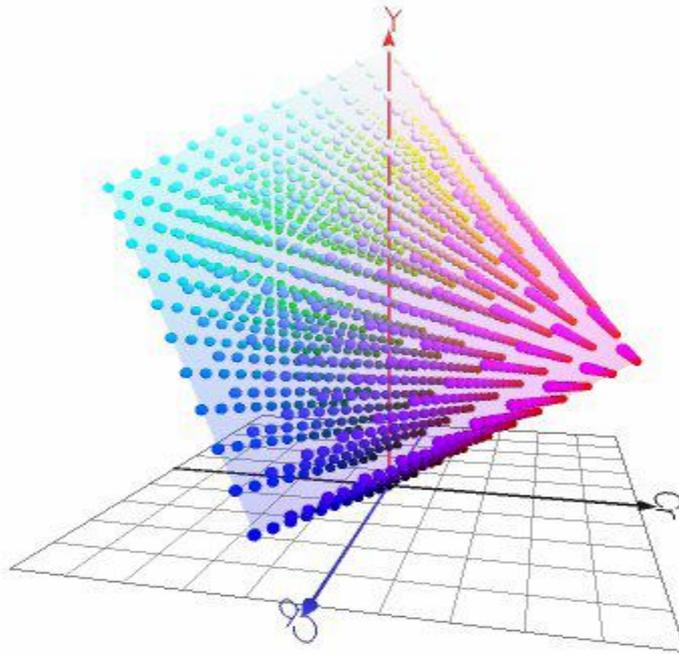


Figure 8 Espace couleur YCbCr

Cette transformation des composantes RGB en composantes YCbCr est linéaire et bijective. Ainsi il n'y a pas de pertes de données pendant cette étape.

II.3.1.3. Espace couleur I1I2I3

Plusieurs expériences systématiques de segmentation de région ont été exécutées pour dériver un ensemble de composantes couleurs efficaces. À chaque étape de division, de nouvelles composantes couleurs sont calculées par la transformation karhunen-loeve des composantes R, G et B. Ceci est appliqué sur huit types d'images couleur et analysé avec plus de 100 composantes couleurs, ce offre un ensemble de composante de couleurs comme suit :

$$\begin{aligned}
 I_1 &= \frac{1}{3}(R + G + B) \\
 I_2 &= \frac{1}{2}(R - B) \\
 I_3 &= \frac{1}{4}(2G - R - B)
 \end{aligned}
 \quad \text{Équation 25}$$

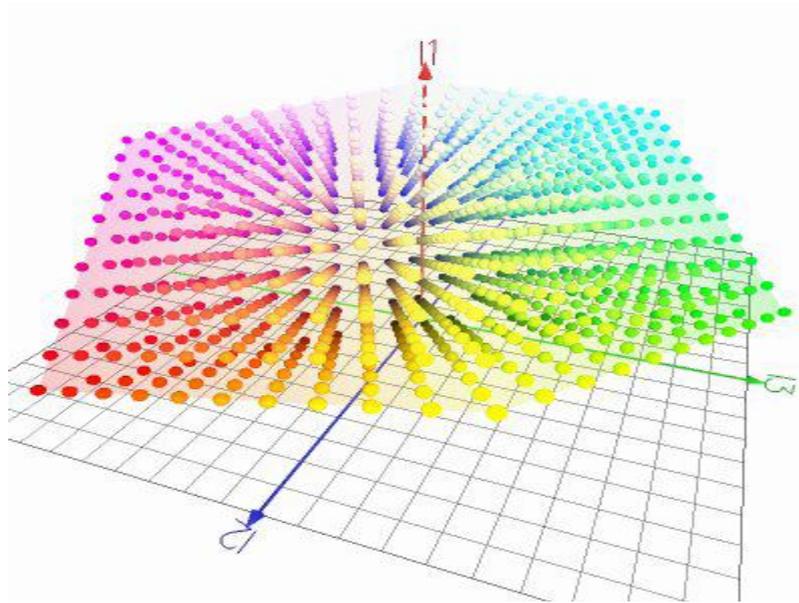


Figure 9 Espace couleur I1I2I3

En comparaison avec d'autres espaces couleur I1I2I3 est le plus efficace en termes de qualité de segmentation.

II.3.1.4. Espace couleur La^*b^*

La^*b^* (aussi connu sous le nom de *CIELab*), dans lequel une couleur est repérée par trois valeurs :

- **L**, la luminance, exprimée en pourcentage (0 pour le noir à 100 pour le blanc)
- **a** et **b** deux gammes de couleur allant respectivement du vert au rouge et du bleu au jaune avec des valeurs allant de -120 à +120.

Le mode *Lab* couvre ainsi l'intégralité du spectre visible par l'œil humain et le représente de manière uniforme. Il permet donc de décrire l'ensemble des couleurs visibles indépendamment de toute technologie graphique. De cette façon il comprend la totalité des couleurs RGB et CMYK, c'est la raison pour laquelle des logiciels tels que PhotoShop utilisent ce mode pour passer d'un modèle de représentation à un autre. Il s'agit d'un mode utilisé dans la recherche des images par contenu [19].

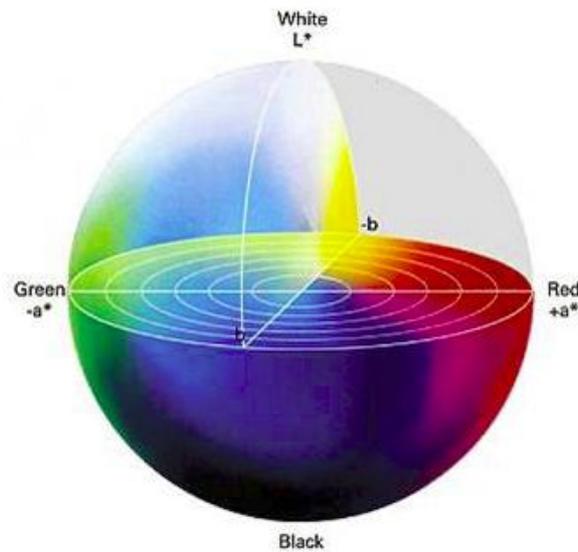


Figure 10 Espace couleur $L^*a^*b^*$

II.3.2.L'histogramme

L'histogramme de couleurs est le descripteur de couleurs le plus courant et le plus rencontré dans la littérature. L'histogramme représente la répartition des pixels en fonction de leur niveau de gris.

II.3.3.Histogramme de couleur-structure CS

Le descripteur de couleurs structure CS étend et enrichit la notion d'histogramme en introduisant dans la représentation un minimum d'information spatiale locale. Un élément structurant, définie un masque binaire, est translaté en chaque pixel de l'image. Tous les intervalles de l'histogramme correspondant aux couleurs présentent à l'intérieur du masque sont alors incrémentés. Ainsi, l'histogramme CS représente la fréquence relative des éléments structurants contenant une couleur donnée. Pour assurer une certaine invariance par rapport aux homothéties, les images sont normalisées à une taille fixe avant l'extraction du descripteur.

II.3.4.Le Corrélogramme

La représentation d'images la plus commune du domaine du traitement d'images dans les bases de données est l'histogramme de couleur. Il est raisonnable de penser que les images qui se ressemblent aient des compositions de couleur similaires et donc il est envisageable que les histogrammes soient intéressants. Les histogrammes de couleur ont d'autres avantages : ils

sont invariants par rapport à la taille, à l'échelle et à la rotation des images. De plus ils sont compacts si nous les comparons avec la taille des images. Une fois que les images sont représentées par les histogrammes dans un espace de couleurs, nous pouvons choisir une parmi les différentes mesures de similarité existantes. Un des problèmes de l'utilisation directe des histogrammes est qu'un petit décalage des valeurs des vecteurs caractéristiques implique une grande différence entre des deux histogrammes en utilisant une des mesures de similarité communes. Les histogrammes de couleurs souffrent d'un manque d'information locale parce que ces derniers n'enregistrent pas la structure locale des couleurs mais seulement la composition globale. Le corrélogramme de couleurs enregistre la probabilité de présence d'une couleur dans un voisinage prédéfini d'une autre couleur. Une nouvelle représentation des informations locales et globales conjointement est nommée « histogrammes de blobs ».

Un corrélogramme présente les corrélations spatiales entre les niveaux de gris en fonction de la distance interpixel. Pour une image A de N pixels, quantifiée sur n niveaux de gris $\{c_1, c_2, \dots, c_n\}$.

Chaque pixel p en (x, y) est doté d'un niveau de gris $A(p)$:

$$p \in A_c = A(p) = c$$

Pour une distance $d \in \{1, 2, \dots, N\}$ fixée à priori, le corrélogramme du niveau de gris i par rapport à j est défini par :

$$\gamma_{C_i, C_j}^{(K)} = \text{prob} \left| p_1 \in A_{c_j} / |p_1 - p_2| = k \right|$$

$$\forall k \in \{1, 2, \dots, d\}$$

Équation 26

Pour tout pixel de niveaux de gris c_i dans l'image, la valeur $\gamma_{C_i, C_j}^{(K)}$ donne la probabilité de trouver à une distance k un pixel de niveau de gris c_j

Le corrélogramme couleur exprime comment la corrélation spatiale de couples de couleurs évolue en fonction de la distance dans l'image. Il prend simultanément en compte la distribution globale des couleurs des pixels ainsi que la corrélation des couleurs entre pixels voisins.

II.3.5. Vecteur de cohérence de couleurs

Le CCV représente une autre variance, plus de détaille. Dans cette technique, chaque rang de l'histogramme peut être partitionné en deux catégories :

- Cohérent, s'il appartient à une région de couleur uniforme
- Sinon incohérent.

II.3.6. Descripteur par disposition de couleurs"

Le CLD représente la disposition spatiale de l'image dans un format très compact. Il est obtenu par application d'une transformation discrète en cosinus (DCT) dans un espace de couleur YCbCr. Il se compose de quatre étapes, le partitionnement d'image, détection de couleur représentative, la transformation DCT, et une quantification non linéaire du zigzag des coefficients numérisée. Ce descripteur a été utilisé dans la classification des images biomédical [12].

II.3.7. Descripteur par couleurs dominantes

Ce descripteur fournit une description compacte des couleurs représentatives dans une image. Il est défini par :

$$F = \left\{ (C_i, p_i, v_i, s) \right\}, (i=1, 2, \dots, N) \text{ Équation 27}$$

Où N désigne le nombre de couleurs dominantes, C_i la valeur de la $i^{\text{ème}}$ couleur dominante, P_i un poids exprimant sa fréquence relative d'apparition dans l'image et V_i sa variance. La cohérence spatiale S représente l'homogénéité spatial des couleurs dominantes dans l'image. Considérons les deux descripteurs ci-dessous :

$$\begin{aligned} F_1 &= \left\{ (C_{1i}, p_{1i}, v_{1i}, s_1) \right\}, (i=1, 2, \dots, N_1) \\ F_2 &= \left\{ (C_{2i}, p_{2i}, v_{2i}, s_2) \right\}, (i=1, 2, \dots, N_2) \end{aligned} \text{ Équation 28}$$

Le degré de similarité entre deux couleurs C_{1k} et C_{2l} , noté est défini par:

$$\alpha_{1k, 2l} = \begin{cases} 1 - \frac{\delta(c_{1k}, c_{2l})}{\delta_{\max}} \\ 0 \end{cases} \text{ si } \delta(c_{1k}, c_{2l}) \leq T_d \text{ Équation 29}$$

Où $\delta(c_{1k}, c_{2l})$ désigne la distance euclidienne entre deux couleurs C_{1k} et C_{2l} , T_d est un seuil prédéfini.

Conclusion

Le premier chapitre de ce mémoire avait pour but de présenter les différentes méthodes d'indexation existant ainsi que les descripteurs d'image par contenu de bas niveau tel que la couleur la forme et la texture.

Chapitre 2 : L'annotation d'images médicales

Dans ce chapitre, nous allons présenter les méthodes d'annotation, puis nous aborderons les classifieurs utilisés dans le domaine médical

Introduction

L'annotation peut être définie comme le processus d'interprétation explicite du document. La création des métadonnées à partir des documents annotés est l'une des principales techniques permettant aux données d'être compréhensible par la machine. Les métadonnées associées aux images peuvent être classés comme suit :

- i) *les métadonnées de contenu indépendant*, où les métadonnées sont liées à l'image, comme les noms des auteurs, les dates, le lieu.
- ii) *le contenu en fonction de métadonnées*, où les métadonnées sont liés aux caractéristiques de bas niveau et / ou niveau intermédiaire, par exemple: la couleur, la texture, la forme.
- iii) *les métadonnées descriptives de contenu*, où les métadonnées sont liées au contenu sémantique. [24]

Le contenu descriptif des métadonnées peut être fourni à deux niveaux de spécificité:

- i) un contenu descriptif associé à l'image complète.
- ii) la segmentation d'images avec des liens vers le contenu descriptif de chaque segment. Les systèmes d'annotations des images par un contenu descriptif basées sur les ontologies, utilisent généralement deux types d'ontologies : une permettant de définir un schéma d'annotation et une autre pour définir les concepts de domaine [24] .

Dans la section suivante nous allons rapidement présenter les ontologies ainsi les ontologies utilisé dans la recherche des images médicales.

I. Les ontologies

L'ontologie du domaine, est un modèle formel décrivant les concepts du domaine et leurs relations. Une ontologie peut prendre des formes diverses, mais comprendra nécessairement un vocabulaire de termes, et une certain spécification de leur signification. Gruber [25] a défini une ontologie comme **« la spécification des conceptualisations, utilisé pour aider les programmes et les humains pour partager les connaissances »**

La conceptualisation est l'abaissement de la connaissance sur les termes d'entités. La spécification est la représentation de cette conceptualisation dans une forme concrète. Une étape dans cette spécification est le codage de la conceptualisation dans un langage de représentation des connaissances.

Guarino [26] a défini une ontologie comme **« Est une structure sémantique intensionnelle qui code les règles implicites qui limitent la structure d'un morceau de réalité »**

Une ontologie définit les termes utilisés pour décrire et représenter un domaine de connaissance.

I.1. Les ontologies formelles (Formal ontologies)

L'ontologie formelle, exige une sémantique claire pour la langue, des motivations claires pour les distinctions adoptées ainsi que des règles strictes sur la façon de spécifier les termes et les relations.

Ceci est obtenu en s'appuyant sur l'analyse ontologique (au sens philosophique) et en utilisant la logique formelle où la signification des termes est garantie par la sémantique formelle.

La complexité d'un système de représentation se divise en deux aspects distincts:

- l'organisation de la structure des connaissances.
- l'information spécifique à un domaine d'application.

Une ontologie formelle a une certaine structure logique sous-jacente qui nous permet de raisonner sur les concepts d'ontologie et aussi de les réutiliser.

Les ontologies sont classiquement découpées en 3 niveaux :

Les ontologies de domaine (Domain ontologies): ce sont des ontologies formelles qui se concentrent sur un domaine d'application (c.-à- d la modélisation d'entreprise, de l'anatomie, de l'astrophysique, etc.) Son but est de fournir une description des concepts de base qui stable et sans ambiguïté, ainsi que les entités et les relations utilisées dans un domaine spécifique.

Les ontologies noyaux (Core ontologies): Ce sont des ontologies formelles qui fournissent l'organisation de haut niveau (général) des concepts utilisés dans (ou à travers) certaines

communautés et domaines d'application. Le but est de faciliter l'échange d'information fiable au sein de ces groupes.

Les ontologies fondamentales (Foundational ontologies) : Ce sont des ontologies formelles plus générales. Le but de ces ontologies est de caractériser les entités et les relations qui sont communes dans tous les domaines et à fournir une vue cohérente et fédératrice.

Dans la section on va voir les différentes ontologies biomédicales existantes dans la littérature.

I.2. Les ontologies biomédicales (Biomedical ontology)

Les ontologies biomédicales sont maintenant un élément important dans la recherche biomédicale, les soins de santé et la santé publique [27].

Au sens large, les ontologies biomédicales comprennent les thésaurus, les terminologies, les classifications et d'autres systèmes de codage utilisées dans le domaine biomédical [27]. Quelques ontologies biomédicales seront décrites dans la section suivante.

I.2.1. Le modèle fondamental de l'ontologie anatomie (The Foundational Model of Anatomy ontology)

FMA est une source de connaissances évolutive pour l'informatique biomédicale qui a pour objectif ; la représentation des catégories et les relations nécessaires à la représentation symbolique de la structure phénotypique du corps humain sous une forme qui soit compréhensible pour les humains et également navigable.

FMA est l'une des ressources d'informations intégrées dans le cadre du « système anatomie distribuée » développée et maintenue par le groupe structurel d'informatique à l'Université de Washington.

FMA est une ontologie de domaine qui représente un ensemble cohérent de connaissances déclaratives et explicites sur l'anatomie humaine.

Le modèle fondamental d'ontologie de l'anatomie FMA met à disposition des informations anatomiques sous une forme symbolique (non graphique) aux modélisateurs de connaissances et d'autres développeurs d'applications pour l'éducation, la médecine clinique, pour le dossier

électronique de santé, la recherche biomédicale et tous les domaines liés aux prestations des soins de santé et de la gestion.

Le modèle fondamental de l'ontologie d'anatomie a quatre composantes interdépendantes [28]:

1. La taxonomie d'anatomie At (Anatomy taxonomy)

La taxonomie d'anatomie At classe les entités anatomiques en fonction des caractéristiques qu'elles partagent (genre) et par laquelle elles peuvent être distinguées les unes des autres (différenciation).

2. Abstraction structurelle d'anatomie ASA (Anatomical Structural Abstraction)

Précise les relations partie-tout et spatiale qui existent entre les entités représentées dans AT.

3. Transformation anatomique ATA (Anatomical Transformation Abstraction)

Spécifie la transformation morphologique des entités représentées dans AT pendant le développement prénatal et dans le cycle de vie postnatale.

4. Métaconnaissances Mc (Metaknowledge)

Précise les principes, les règles et les définitions selon laquelle les classes et les relations dans les trois autres composantes de la FMA sont représentés.

Donc FMA est représenté $FMA = (At, ASA, ATA, Mk)$

L'ontologie FMA est implémenté par Protégé (L'environnement d'édition des bases de connaissances (ontologies)). Les utilisateurs de FMA peuvent modifier les données dans une base de données MySQL avec un schéma base de données relationnelle spécifique.

Le contenu de FMA:

FMA est l'une des plus grandes sources informatiques à base des connaissances dans les sciences biomédicales. Elle contient environ 75.000 classes et plus de 120.000 termes et plus de 2,1 millions des relations d'instances [28].

La composante la plus compréhensive de FMA est la taxonomie Anatomy (AT). La classe dominante dans AT est *Anatomical Structure*.

Les structures anatomiques comprennent des macromolécules biologiques, les cellules et leurs parties, des portions de tissus, les organes et leurs parties, ainsi que des systèmes d'organes et les régions du corps.

Les structures anatomiques macroscopiques sont plus globalement représentées, tandis que les molécules biologiques ont été saisis principalement pour illustrer le continuum structurel des

parties importantes du corps, tels que le thorax, et les macromolécules biologiques, telles que la myosine.

Des portions de substances corporelles, comme le sang, la matrice intercellulaire, et le cytoplasme, sont définis en termes de leur relation avec les structures anatomiques. Les auteurs dans [29] ont intégré des connaissances externe qui sont encodées en varie lexiques, thésaurus et les ontologies comme FMA et ceci pour supporter des requêtes sémantiques, un apprentissage basé sur la détection de l'anatomie et une segmentation par l'utilisation de la distribution priors gratuite ont été utilisé, afin que le système s'adapte rapidement à de nouveaux domaines et de nouvelles maladies.

1.2.2. Les ontologies des maladies (Disease Ontology)

Les ontologies des maladies (DO) a été développée par le centre de la médecine génétique de l'université Northwestern depuis 2003.

DO a été conçu comme une ontologie open source basé sur les principes de l'ontologie des gènes (GO), en prenant en compte des concepts de maladie, organisés en un graphe acyclique dirigé (Directed Acyclic Graph DAG).

La construction initiale de l'ontologie des maladies en 2003 et 2004 a utilisées le modèle d'informations commune (the Common Information Model CIM-9) comme un vocabulaire fondamental, et a été réorganisé par processus, et la cause des maladies (maladies génétiques, maladies infectieuses, troubles métaboliques). Dans la version actuelle de la DO, il ya 341,850 références externes et 12,564 termes [30].

1.2.3. Unified Medical Language System (UMLS)

Le système Unified Medical Language (UMLS) est un référentiel des vocabulaires biomédicaux développés par la librairie national de médecine des États-Unis¹.

UMLS contient plus de 1 million concepts biomédicaux et 2,8 millions de noms de concepts et, plus de 100 vocabulaires contrôlés ; de plus de 60 familles de vocabulaires biomédicaux, ainsi que 12 millions de relations entre ces concepts.

Les vocabulaires intégrés dans le Metathesaurus UMLS comprennent la taxonomie du centre national pour l'information biotechnologie (NCBI), l'ontologie de gène (GO), La rubrique de

¹<http://www.nlm.nih.gov/research/umls/>

sujet médical (MeSH), l'hérédité mendélienne chez l'homme en ligne (OMIM) et La base de connaissances des symboles numériques d'anatomies.

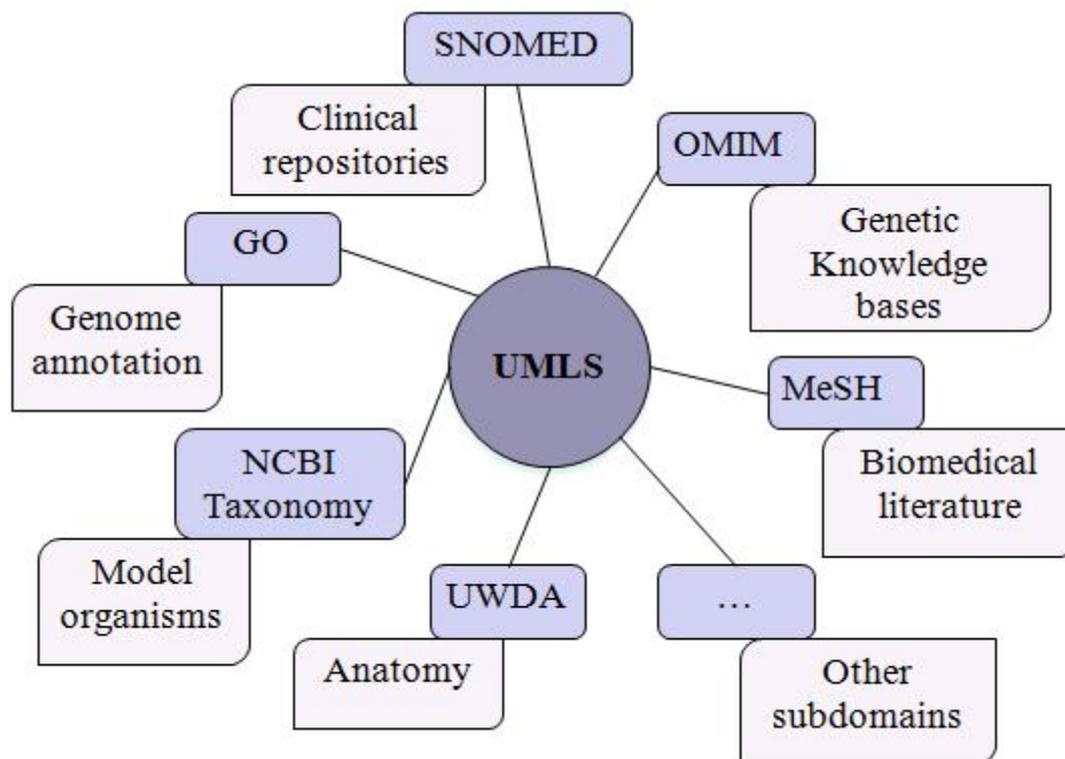


Figure 11 Les différents sous-domaines intégrés dans l'UMLS [31]

Les concepts UMLS ne sont pas seulement liés entre eux, mais peuvent aussi être liés à des ressources externes telles que la base de données banque de gènes (GenBank).

Dans UMLS, la connaissance est organisée par concepts (c-à-d sens). Les termes synonymes sont regroupés pour former un concept, et les concepts sont liés à d'autres concepts par le biais de divers types de relations, ce qui entraîne un graphe riche. Les relations Inter-Concept sont soit héritées de la structure des vocabulaires sources ou générées spécifiquement par les rédacteurs des Metathesaurus. Les relations symboliques peuvent être hiérarchiques (par exemple « is a kind of » ou « is a », « part of ») ou associative (« location of », « caused by »).

Une telle structure, permet aux utilisateurs d'exécuter des tâches telles que:

- (i) La collecte des divers termes utilisés pour nommer un concept.
- (ii) L'extraction des relations d'un concept avec d'autres concepts, que ce soit hiérarchiques ou associatives, symbolique ou statistique.

(iii) L'obtention d'un ensemble de concepts pour une catégorie donnée, en utilisant la liste des concepts qui ont été affectés à un type sémantique.

Les sources de connaissances UMLS sont multi-usages. Ils ne sont pas optimisés pour des applications particulières. Mais peuvent être appliquée aux les systèmes qui exécutent une série de fonctions impliquant un ou plusieurs types d'informations .Par exemple, les dossiers des patients, de la littérature scientifique, des lignes directrices, et les données de santé publique.

Les sources de connaissances UMLS sont mises à jour trimestriellement. Tous les vocabulaires sont disponibles, sans frais à des fins de recherche au sein d'une institution. Les auteurs dans [24] proposent une architecture visant à soutenir le processus de récupération des connaissances à partir de bases de données en utilisant des images annotées dans de domaine anatomique. Et parmi les ontologies biomédicales qui ont été utilisé il y'avait : **FMA**, **UMLS** et **DO**.

I.2.4. La rubrique de sujet médical (Medical Subject Headings)

MeSH est une hiérarchie taxinomique des termes médicaux et biologiques proposés par la bibliothèque nationale de médecine des États-Unis.NLM a adopté le langage XML (Extensible Markup Language) comme langage de description pour MeSH.

Les descripteurs MeSH sont représentés en deux formes alphabétiques et une structure hiérarchique. Au niveau le plus général de la structure hiérarchique, il y a les rubriques très larges telles que «anatomie» ou «troubles mentaux».

Les rubriques les plus spécifiques se trouvent à des niveaux plus étroits d'une hiérarchie de douze niveaux, tels que "la cheville" et "les troubles de conduite ".

Il ya 26,142 descripteurs MeSH (2011)².Il ya également plus de 177.000 termes d'entrée qui aident à trouver les entête MeSH les plus appropriés, par exemple, "vitamine C" est un terme d'entrée à "acide ascorbique".

I.2. 5. GALEN

² <http://www.nlm.nih.gov/pubs/factsheets/mesh.html>

GALEN vise à fournir des ressources terminologiques réutilisables pour les systèmes cliniques. Le cœur de Galien est l'utilisation de l'ontologie : modèle de référence commun (the Common Reference Model), formulée dans une logique de description spécialisés : GRAIL.

L'objectif global des ressources terminologique de GALEN est de soutenir les systèmes d'information cliniques, d'aide à la décision, et les systèmes bibliographique.

Plus généralement, l'ontologie GALEN peut être décrits en termes de [32]:

Expressivité : la capacité de représenter formellement les notions nécessaires à ses usagers. Pour les ontologies médicales, cela signifie tous les symptômes et les maladies concernées.

Classification : la capacité d'inférer une classification correcte (indexation) des expressions représentées, et cela peut se réaliser de deux manières :

- a) profondément : toutes les inférences faites sont correctes.
- b) complètement : toutes les inférences possibles sont réalisés.

Parcimonie : - GALEN a été spécialement conçu pour une utilisation en tant que «post-coordonnée du système », dans lequel la classification des nouvelles expressions est déduite et dynamiquement maintenue.

Les concepts du modèle GALEN sont représentés par des catégories. Les catégories peuvent être abstraite, tels que «phénomène» ou «maladie», comme «le diabète» ou très spécifiques tels que «l'aspirine». Les catégories peuvent être spécialisés par exemple : «aspirine» peut être spécialisé pour «children's aspirin», et «foot» peut être spécialisé à «left foot», «deformed foot», «deformed left foot».

Le modèle GALEN fournit les taxonomies qui contiennent des milliers de catégories dans une complexe hiérarchie.

I.2.6. Nomenclature systématisée de médecine - Conditions cliniques (Systematized Nomenclature of Medicine--Clinical Terms)

SNOMED CT a été acquise en Avril 2007 par l'International Standards Organisation de la Santé de la terminologie (International Health Terminology Standards Organisation (IHTSDO)).

SNOMED CT est la plus complète, terminologie clinique multilingue dans les soins de santé dans le monde.³

SNOMED CT contribue à l'amélioration des soins en soutenant le développement des dossiers de santé électroniques qui enregistrent des informations cliniques de manière à permettre la récupération basée sur le sens. Cela permet un accès efficace à l'information nécessaire pour l'aide à la décision et le reporting et l'analyse.

SNOMED CT est :

- Une terminologie clinique.
- Une ressource complète avec un contenu scientifiquement validé, essentielle pour les dossiers de santé électroniques.
- Déjà utilisé dans plus de cinquante pays.

SNOMED CT fournit une terminologie générale de base pour le dossier de santé électronique (the electronic health record (EHR)) et contient plus de 311.000 concepts actifs avec des significations uniques et formels basés sur la logique organisées en hiérarchies.

³ <http://www.ihtsdo.org/snomed-ct/>



Figure 12 Les concepts actifs dans SNOMED au cours de six ans

SNOMED CT peut être utilisé pour représenter les informations cliniquement pertinentes de manière cohérente, et fiable.

1.2.7. L'ontologie des gènes (The Gene Ontology)

GO est un effort collaboratif permettant de répondre au besoin de description cohérente des gènes dans différentes bases de données. Le projet a commencé par une collaboration entre trois bases de données : FlyBase (*Drosophila*), the *Saccharomyces* Genome Database (SGD), et Mouse Genome Database (MGD).

Le projet de Gene Ontology fournit une ontologie qui définit des termes représentant les propriétés des gènes produits. L'ontologie couvre trois domaines⁴:

- **Les composants cellulaires** (cellular component), les parties d'une cellule ou de son environnement extracellulaire.
- **Les fonctions moléculaires** (molecular function), les activités élémentaires des gènes au niveau moléculaire, telles que la liaison ou la catalyse.
- **Le processus biologique** (biological process), des opérations ou des ensembles d'événements moléculaires, avec un début et une fin, pertinentes pour le

⁴ <http://www.geneontology.org/GO.doc.shtml>

fonctionnement des unités de vie intégrées tel que : les cellules, tissus, organes, et les organismes.

Par exemple, le produit de gène du cytochrome « C » peut être décrit par le terme de la *fonction moléculaire* « l'activité oxydoréductase », et décrit dans *le processus biologique* par « phosphorylation oxydative » et « l'induction de la mort cellulaire », dans les termes *des composants cellulaires* correspondent à « la matrice mitochondriale » et « la membrane interne mitochondriale ».

L'ontologie GO est structurée sous forme d'un graphe acyclique dirigé, et chaque terme a des relations avec un ou plusieurs autres termes dans le même domaine, et parfois dans d'autres domaines.

II. Le système d'archivage et de communication (Picture Archiving and Communication System)

Le PACS, est un système d'archivage et de communication des images médicales. Ce type de système informatique a pu se développer grâce à la numérisation de toutes les images produites dans les services de radiologie (radiographies standards, angiographies), d'imagerie en coupe (scanner, IRM), d'échographie ou de médecine nucléaire (scintigraphie, TEP scanner).

Ce système PACS devient de fait la « mémoire » des services d'imagerie et des services où sont hospitalisés les patients et permet la communication des images entre les professionnels de santé concernés par cet élément du dossier patient.

Concrètement, le PACS se traduit à terme par la disparition du film et par la visualisation des images médicales sur une station informatique.

Les principaux objectifs d'un projet PACS peuvent se résumer aux points suivants :

1) Améliorer l'accessibilité aux images médicales : pour les médecins radiologues et les médecins cliniciens. En effet, dès leur production, les images sont envoyées sur le PACS. Elles sont donc disponibles immédiatement pour les imageurs et pour les cliniciens. Elles sont accessibles simultanément en plusieurs endroits du CHU.

L'accès aux examens antérieurs du patient, présents sur le PACS, est facilité. Le compte rendu de l'examen d'imagerie devient accessible dans les mêmes conditions que l'image.

2) Permettre l'accès à l'imagerie médicale numérique : à tous les médecins du centre hospitalier. Cela permet de bénéficier de toute la richesse de l'image et donc de pouvoir appliquer certains traitements à cette image.

3) Permettre la disparition du support actuel de l'image : (film, papier). Le film est un élément polluant, coûteux, gourmand en mètre linéaire d'archive.

4) Améliorer la sécurité des dossiers d'imagerie : Le risque de perdre des images est très faible au regard des différentes sécurités mises en place dans le système.

III. Les classifieurs utilisés dans la recherche d'image médical par le contenu

Dans [18], les auteurs ont proposé une nouvelle méthode d'apprentissage semi-supervisé par SVM, pour résoudre le retour de pertinence dans la recherche d'images par le contenu, le système palie à deux inconvénients principaux de l'apprentissage SVM qui sont :

- les performances des SVM sont généralement limitées par le nombre des exemples étiquetés.
- les approches classiques ne prennent pas en compte la redondance des exemples, et pourraient sélectionner de multiples exemples qui sont similaires (ou même identiques).

Deux nouveaux algorithmes sont proposés pour résoudre l'optimisation combinatoire. La première approche du problème se rapproche à un programme quadratique, et la seconde permet de résoudre l'optimisation combinatoire par un algorithme glouton.

Dans [12] un système de recherche des images médicales de différentes modalités a été défini. Les auteurs ont proposé une approche de fusion de similarité pour faire la combinaison linéaire de différentes caractéristiques (la couleur, la forme et la texture). Cette approche utilise une technique combinant le théorème de Bayes et un classifieur supervisé SVM (multi-classe) afin de déterminer l'importance d'une caractéristique.

Dans [6] un algorithme pour la classification d'images médicale en fonction de leur contenu visuel a été proposé. Il est basé sur l'apprentissage du noyau multiple (Multiple Kernel Learning). MKL est une méthode utilisée pour apprendre la combinaison optimale d'un ensemble de noyaux pour une classification par SVM.

Trois caractéristiques globales ont été utilisées : la matrice de cooccurrence, les caractéristiques de Gabor, les caractéristiques de la texture Tamura; et une caractéristique locale est ModSIFT. La combinaison des caractéristiques de Gabor et modSIFT par MKL a donné le meilleur résultat.

Dans [33] une annotation automatique est proposée basée sur les attributs visuels pour étiqueter les images dans le cadre du processus d'indexation. Au cours de la recherche d'images ces étiquettes sont utilisées pour filtrer ou réorganiser les résultats. Les auteurs ont utilisé les textures et les descripteurs de brassage comme caractéristiques visuelles et le classifieur Naive Bayes NearestNeighbor (NBNN) pour classer les images selon leurs modalités.

Dans [23] une nouvelle approche est proposée pour soutenir l'aide au diagnostic médical (CAD), visant à la recherche basée sur la similitude des lésions de masse par mammographie.

Cette méthode comprend les techniques de moments de Zernike pour récupérer les caractéristiques de la forme, les règles d'association statistiques pour identifier les attributs pertinents et de la théorie fractale pour réduire la dimension des vecteurs de caractéristique ; ensuite le classifieur K plus proche voisin a été utilisé.

Dans [8] un classificateur à base de fusion hiérarchique a été proposé. Les performances de classification ont été améliorées par l'application d'une fusion des résultats de la classification. Les auteurs ont proposé les caractéristiques de tessellation basées sur le spectre, et l'histogramme de direction pour l'extraction des caractéristiques. La classification hiérarchique comporte deux niveaux et dans chaque niveau un perceptron multicouche (MLP) a été utilisé.

Dans [10] les auteurs ont proposé l'utilisation de la fonction Ripley K et la matrice de cooccurrence comme des signatures de texture, appliquées pour la recherche des images mammographiques par le contenu. Pour chaque image deux vecteurs sont obtenus, l'un contient les mesures de la fonction Ripley K, et l'autre les mesures de texture de Haralick, et enfin la distance euclidienne a été utilisée.

Dans [34] la contribution principale des auteurs est l'introduction d'une nouvelle méthode pour la recherche des images médicales. C'est la *récupération déductible*, et elle est réalisée par une nouvelle méthode d'annotation SEMI-SECC, qui présente la combinaison de modèle codes correcteurs sémantique d'erreurs et une classification semi-supervisée. Trois

différents types de caractéristiques ont été utilisés et qui sont : l'intensité, les ondelettes de Haar, et l'ondelette Gabor.

Conclusion

Dans ce chapitre, nous avons vu les descripteurs de contenu de haut niveau comme les ontologies et le système d'archivage et de communication, et nous avons conclut ce chapitre par une recherche bibliographique sur l'utilisation des classifieurs dans la recherche d'images médicales par contenu.

Chapitre 3 : Les méthodes utilisées

Nous allons présenter dans ce chapitre les méthodes d'analyse d'image permettant de décrire le contenu visuel que nous avons retenues pour construire le vecteur de caractéristiques .En suite nous détaillons le système de classification proposé afin d'associer chaque images de la base de données à une classe spécifique .

Introduction

Ce chapitre donne une vue plus détaillée des techniques utilisées dans notre travail. Nous abordons plus précisément deux aspects du problème. Tout d'abord l'extraction des caractéristiques, et dans ce contexte nous allons présenter les méthodes d'analyse d'image permettant de décrire le contenu visuel que nous avons retenues pour cette application. Ensuite nous nous attacherons à utiliser un système de classification SVM pour classifier les images et ceci par l'utilisation de leur descripteurs.

I.L'extraction des caractéristiques

La première étape dans le processus de compréhension d'une image est le plus souvent l'extraction des objets qui la composent. Aussi, la segmentation est l'une des étapes importantes conduisant à l'analyse de données des images traitées. Il s'agit principalement de diviser une image en parties ou ensembles qui ont une forte corrélation avec les objets contenus dans l'image. Nous avons utilisé la segmentation pour l'extraction des caractéristiques d'objet. Dans la section suivante nous allons présenter les grandes approches existantes dans la segmentation.

I.1 Prétraitement par segmentation d'images :

En général, la segmentation totale d'une image passe par une segmentation partielle a priori de celle-ci, et elle est suivie de traitements utilisant des informations à un niveau supérieur. Les deux grandes familles de segmentation en imagerie faisant intervenir des équations aux dérivées partielles sont les suivantes :

Les modèles déformables : C'est une approche analytique, qui permet d'agir interactivement sur la modélisation. Ce la permet de faire évoluer (en temps et en espace) la représentation du modèle vers la solution du problème de minimisation introduit dans la modélisation. Concrètement, cela revient à introduire un terme d'évolution en temps dans le critère de minimisation, ce qui permet à chaque pas de temps d'influencer le modèle a priori si nécessaire et de se recalculer à une meilleure solution. Les premiers travaux en la matière sont ceux de Kass, Terzopoulos et Witkin en 1987.

Les surfaces de niveaux ou "Level set methods" : C'est une approche géométrique. Dans le cas bidimensionnel, la courbe définissant le contour d'un objet est définie comme la ligne de niveau zéro d'une hyper-surface ($u(x,y,t)$) qui vérifie une équation aux dérivées partielles. Cette méthode a été introduite par Stanley Osher (UC Los Angeles) et James A. Sethian (UC Berkeley & Lawrence Berkeley Nat. Lab.) en 1988.

Nous avons utilisé la méthode des contours actifs pour segmenter l'image. L'idée de base dans les modèles de contours actifs (*active contour ou snakes*) est d'évoluer une courbe soumise à des contraintes d'une image donnée. Nous aborderons cette méthode dans la section suivante.

1.1.1 .Le modèle de contour actif :

Dans les modèles classiques de contours actifs, le gradient est utilisé pour arrêter l'évolution de la courbe à la limite de l'objet désiré. Ces modèles ne peuvent détecter que des objets avec des bords définis par gradient. Dans la pratique, les gradients discrets sont limités et la fonction d'arrêt n'est jamais nulle sur les bords, ainsi que la courbe peut passer à travers la frontière [35], et si l'image est très bruitée, le lissage gaussien doit être fort, ce qui peut lisser les bords aussi.

Dans notre cas nous avons utilisé le modèle de contours actifs proposés par F. Chan et A. Vese. Ce modèle peut détecter des objets dont les limites ne sont pas nécessairement définies par gradient, et le terme d'arrêt est basé sur des techniques de segmentation de Mumford-Shah [35]. Ce modèle se base aussi sur l'utilisation des surfaces de niveaux (level sets). Les contours intérieurs sont détectés automatiquement et la courbe initiale peut être n'importe où dans l'image.

Expliquons d'abord l'idée de base du modèle dans un cas simple. Soit une image u_0 formée par deux régions avec des intensités constantes et des valeurs distinctes u_0^i et u_0^o . Supposons en outre que l'objet à détecter est représentée par la région de la valeur u_0^o . Sa limite est désignée par C_0 .

$$\text{Soit } F_1(C) + F_2(C) = \int_{\text{inside}(c)} |u_0(x,y) - c_1|^2 dx dy + \int_{\text{outside}(c)} |u_0(x,y) - c_2|^2 dx dy \quad \text{Équation 30}$$

Où C est une autre courbe variable, et les constantes c_1, c_2 sont les moyennes à l'intérieur et à l'extérieur de C , respectivement. Dans ce cas, il est évident que la limite de l'objet C_0 , est déterminée par la minimisation de ces forces.

$$\text{Inf}_c \{F_1(c) + F_2(c)\} \approx 0 \approx F_1(c_0) + F_2(c_0).$$

Ceci peut être vu facilement. Par exemple, si la courbe C est à l'extérieur de l'objet, alors $F_1(c) > 0$ et $F_2(c) \approx 0$; et si la courbe C est à l'intérieur de l'objet, alors $F_2(c) > 0$ et $F_1(c) \approx 0$ et si la courbe est à la fois à l'intérieur et à l'extérieur de l'objet alors $F_2(c) > 0$ et $F_1(c) > 0$. Enfin, l'énergie est minimisée si $C = C_0$, c a d, la courbe C est à la limite de l'objet. Ces principes de base sont illustrés dans la figure suivante.

	
$F_1(c) > 0$ et $F_2(c) \approx 0$	$F_2(c) > 0$ et $F_1(c) \approx 0$
	
$F_2(c) > 0$ et $F_1(c) > 0$	$F_2(c) \approx 0$ et $F_1(c) \approx 0$

Figure 13 les cas possibles de la position de la courbe. Le terme est minimisé seulement dans le cas où la courbe est à la limite de l'objet

F. Chan et A. Vese ont ajouté quelques termes de régularisation pour le modèle de contours actifs comme la longueur de la courbe C , et (ou) la surface de la région à l'intérieur de C . L'énergie fonctionnelle devient donc :

$$F(C_1, C_2, C) = \mu \cdot \text{length}(c) + \nu \cdot \text{Area}(\text{inside}(c)) + \lambda_1 \int_{\text{inside}(c)} |u_0(x, y) - c_1|^2 dx dy + \lambda_2 \int_{\text{outside}(c)} |u_0(x, y) - c_2|^2 dx dy$$

Équation 31

Avec $\mu, \nu \geq 0, \lambda_1 \lambda_2 > 0$

Ce modèle de contour actif est un cas particulier du problème de partition minimale, dans le quelle on cherche la meilleure approximation u de u_0 , une fonction prenant seulement deux valeurs :

$$u = \begin{cases} \text{average}(u_0) \text{ inside } C \\ \text{average}(u_0) \text{ outside } C \end{cases}$$

Ce cas particulier du problème de partition minimale peut être formulé et résolu en utilisant la méthode des ensembles de niveau (the level set method).

$c \subset \Omega$ est représentée par l'ensemble de niveau zéro d'une fonction Lipschitz de telle sorte que

$$\begin{aligned} c &= \{(x, y) \in \Omega; \phi(x, y) = 0\} \\ \text{inside}(c) &= \{(x, y) \in \Omega; \phi(x, y) > 0\} \\ \text{outside}(c) &= \{(x, y) \in \Omega; \phi(x, y) < 0\} \end{aligned} \quad \text{Équation 32}$$

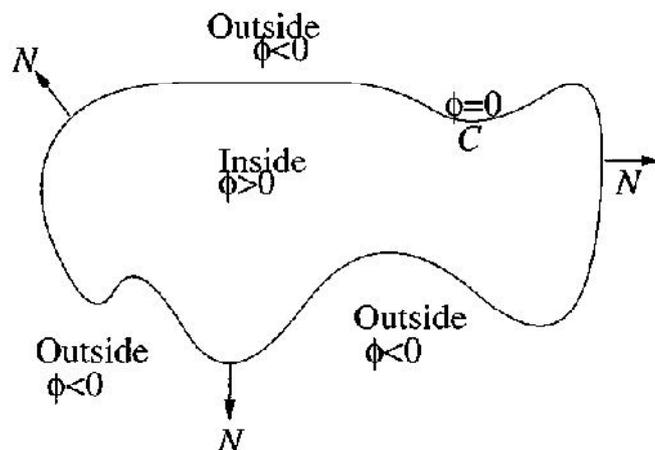


Figure 14 la courbe se propage dans une direction

La variable C est remplacée par la variable ϕ , et avec l'utilisation de la fonction de Heaviside, et la unidimensionnelle mesure de Dirac δ_0 .

$$H(z) = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{if } z < 0 \end{cases}$$

$$\delta_0(z) = \frac{d}{dx} H(z) \quad \text{Équation 33}$$

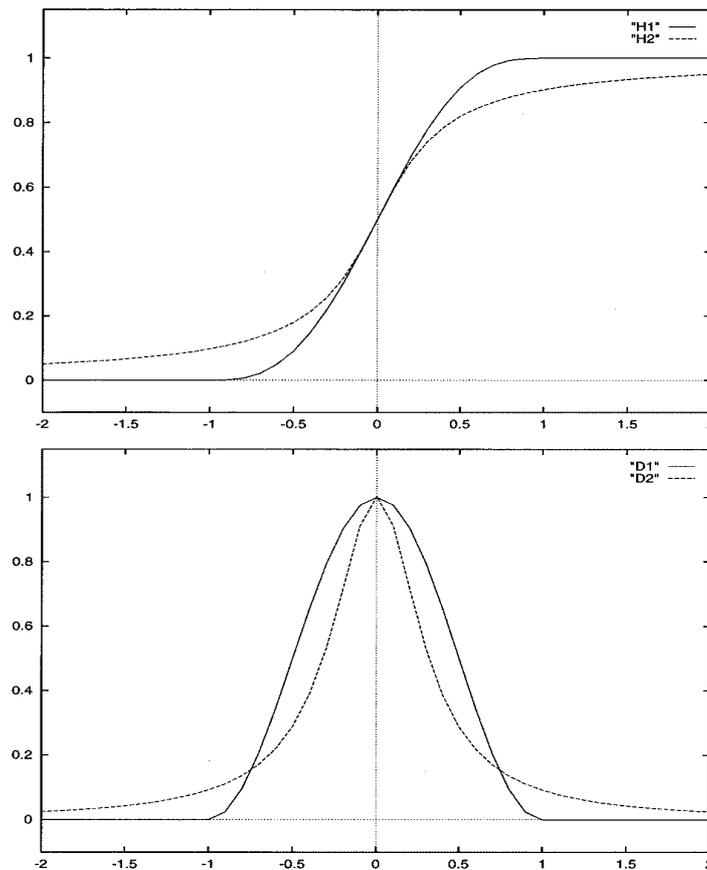


Figure 15 Deux régularisations différentes de la fonction Heaviside (en haut) et la fonction Dirac (en bas)

Les termes de l'énergie sont exprimés de la façon suivante :

$$\begin{aligned}
length\{\phi = 0\} &= \int_{\Omega} \nabla H(\phi(x, y)) dx dy \\
&= \int_{\Omega} \delta_0(\phi(x, y)) \nabla \phi(x, y) dx dy && \text{Équation 34} \\
Area\{\phi \geq 0\} &= \int_{\Omega} H(\phi(x, y)) dx dy
\end{aligned}$$

$$\begin{aligned}
&\int_{\phi > 0} |u_0(x, y) - c_1|^2 dx dy \\
&= \int_{\Omega} |u_0(x, y) - c_1|^2 H(\phi(x, y)) dx dy \\
&\int_{\phi < 0} |u_0(x, y) - c_2|^2 dx dy && \text{Équation 35} \\
&= \int_{\Omega} |u_0(x, y) - c_2|^2 (1 - H(\phi(x, y))) dx dy
\end{aligned}$$

Donc l'énergie devient

$$\begin{aligned}
F(c_1, c_2, \phi) &= \mu \int_{\Omega} \delta(\phi(x, y)) \nabla \phi(x, y) dx dy \\
&+ \nu \int_{\Omega} H(\phi(x, y)) dx dy + \lambda_1 \int_{\Omega} |u_0(x, y) - c_1|^2 H(\phi(x, y)) dx dy \\
&+ \lambda_2 \int_{\Omega} |u_0(x, y) - c_2|^2 (1 - H(\phi(x, y))) dx dy
\end{aligned}$$

Équation 36

Dans la section suivante nous allons expliquer les descripteurs utilisés pour le système proposé.

I.2. les descripteurs visuels utilisés :

La comparaison directe des images entre elles n'est pas envisageable. Il est donc nécessaire d'en extraire au préalable des informations représentatives qui décrivent les images. Nous avons choisi les descripteurs de texture que nous allons les détaillés dans la section suivante.

I.2.1 .Les ondelettes de Gabor

Les fonctions Complexes de Gabor ont été introduites par Gabor. Elles présentent des exponentielles complexes avec une enveloppe gaussienne, ou des gaussiennes qui sont modulés par des harmoniques complexes.

La transformé d'ondelette de Gabor (GWT) est un outil d'analyse en temps-fréquence efficace et puissant.

L'efficacité de la GWT est fortement influencée par la forme d'ondelette qui contrôle la propriété de localisation temps-fréquence.

Nous avons utilisé la forme GWT [36] suivante

$$\psi_{u,v}(z) = \frac{\|K_{u,v}\|^2}{\sigma^2} e^{(-\|K_{u,v}\|^2 \|z\|^2 / 2\sigma^2)} \left[e^{ik_{u,v}z} - e^{-\sigma^2/2} \right] \text{Équation 37}$$

Ou $Z=(x,y)$

$$\text{Et } K_{\mu\nu} = K_v e^{-i\theta\mu} \text{ Équation 38}$$

$$K_v = K_{\max}/f^v \text{ Équation 39}$$

K_{\max} est la fréquence maximale et f est facteur d'espacement compris entre les noyaux dans le domaine fréquentiel; avec une orientation u et une échelle v .

$$\text{Et } \theta\mu = \pi\mu / 8$$

$e^{-\sigma^2/2}$ Est une DC composition ; σ est une constante.

e est une fonction de d'oscillation, dont la partie réelle est la fonction cosinus et la partie imaginaire est une fonction sinusoïdale comme le montre les figures suivantes :

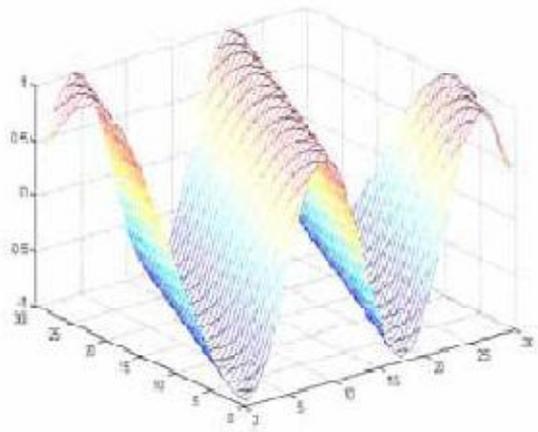


Figure 16 La partie réelle

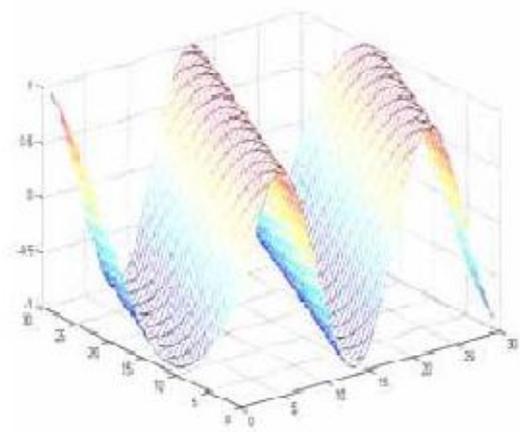


Figure 17 La partie imaginaire

La fenêtre de Gauss reflète la localisation du filtre de Gabor à la fois dans le domaine temporel et fréquentiel, et elle limite la plage de la fonction d'oscillation.

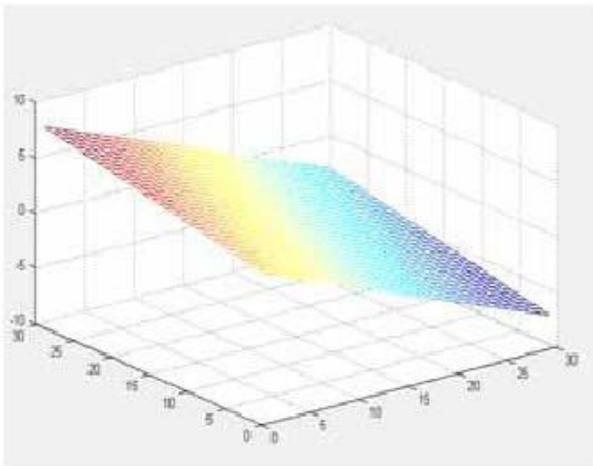


Figure 18 Phase angle of the oscillation function

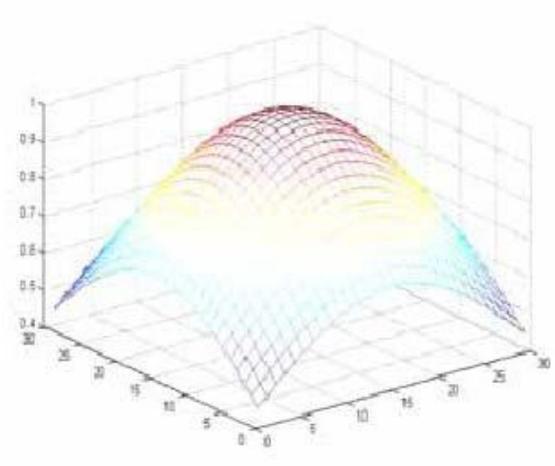
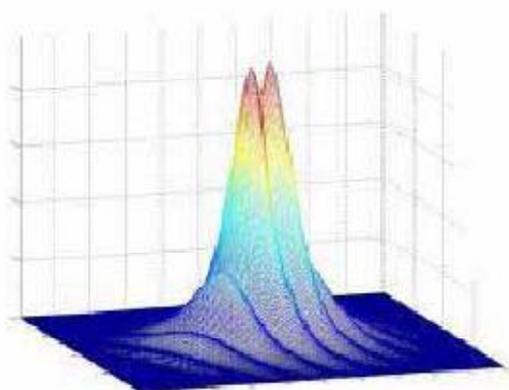
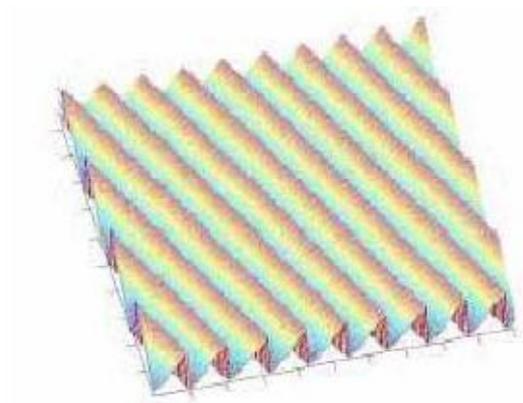
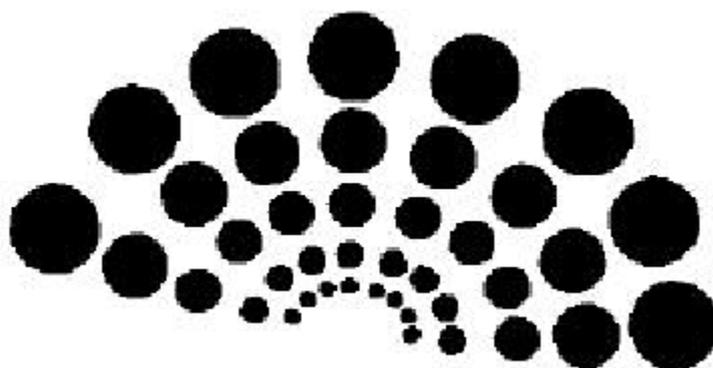


Figure 19 the Gauss function

**Figure 20 Gabor function Amplitude part****Figure 21 Gabor function - phase part**

La morphologie des filtres de Gabor est représentée dans la figure suivante

**Figure 22 Morphologie des 40 filtres de Gabor**

Les parties réelles du filtre de Gabor, et les grandeurs sont représenté dans la figure suivante :

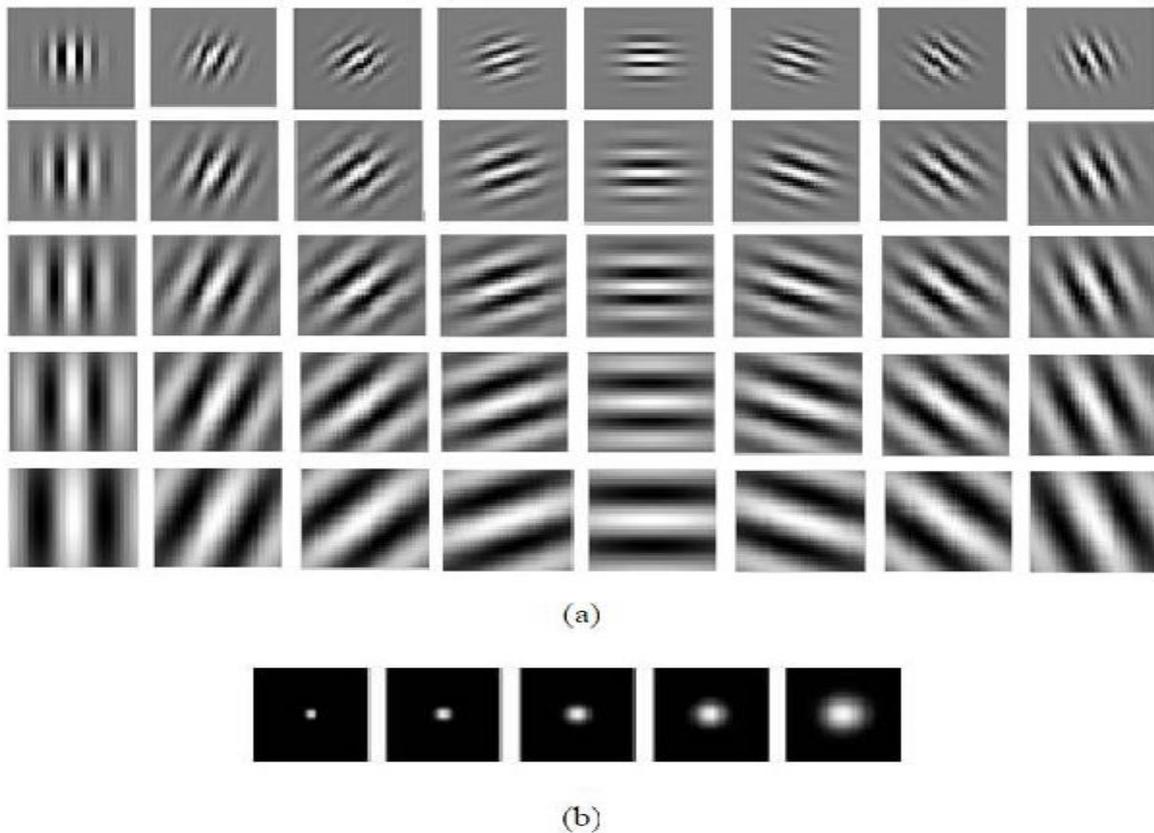


Figure 23 (a) La partie réelle du noyau de Gabor (b) Cinq grandeurs de la Gabor

L'énergie d'information pour une image $f(x,y)$ pour chaque échelle p et orientation q est définis comme suit [37]

$$E(p, q) = \sum_x \sum_y |O(x, y)| \quad \text{Équation 40}$$

Le deuxième descripteur d'image utilisé est la matrice de co-occurrence ; Nous allons le présenter dans la section suivante.

I.2.2. La matrice de co-occurrence (Gay Level Co-Occurrence Matrix GLCM)

Le deuxième descripteur d'image utilisé est la matrice de co-occurrence introduit par Haralick [2]. Il permet de décrire la fréquence relative avec laquelle deux pixels séparés par une distance D sous angle θ se produisent dans une image en niveaux de gris.

Pour chaque (i, j) de P_d est le nombre d'occurrences des paires de niveaux de gris et qui sont séparé par une certaine distance. Formellement :

$$P_d(i, j) = \{((r, s), (t, v)): I(r, s) = i, I(t, v) = j\} \text{ Équation 41}$$

Les directions d'analyse dans GLCM sont généralement :

- Horizontale (0°).
- Vertical (90°).
- Diagonale (45°, 135°)

	0	1	2	3
0	0	3	4	2
1	2	1	3	4
2	0	3	2	1
3	2	1	0	3

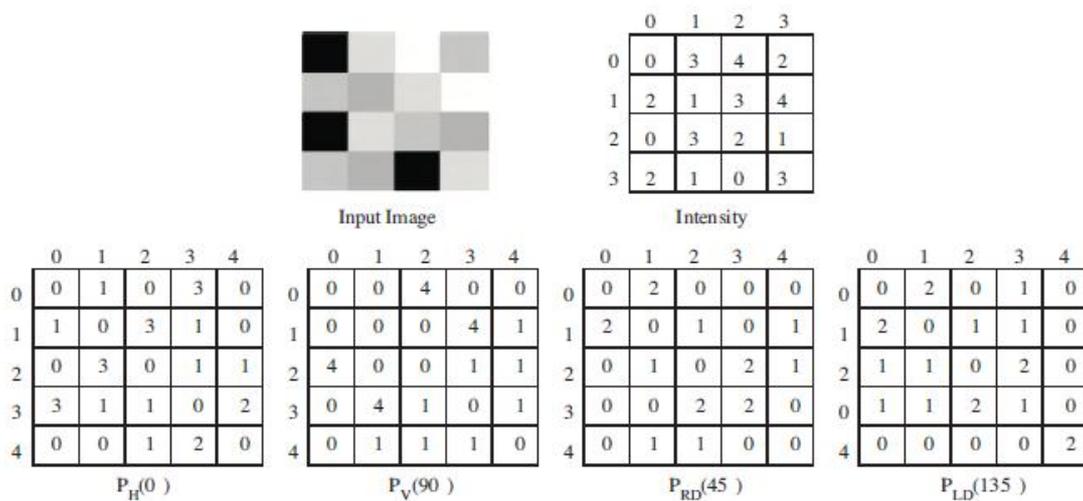


Figure 24 la matrice de co-occurrence avec D=1 et PH (0°), PV (90°), PRD (45°), PLD (135°).

L'utilisation d'une approche statistique comme la matrice de co-occurrence aidera à fournir des informations précieuses sur la position relative des pixels voisins dans une image.

Haralick a défini 14 paramètres caractéristiques de texture [2]

1. Les moments d'ordre deux (Angular Second Moment)

$$F1 = \sum_i \sum_j \{p(i, j)^2\}. \text{ Équation 42}$$

2. Le contraste (Contrast)

L'indice de contraste exprime le passage fréquent d'un pixel clair à un pixel foncé ou inversement. Si l'image contient peu de régions homogènes, le contraste sera élevé. Ce paramètre permet aussi de caractériser la dispersion des valeurs de la matrice par rapport à sa diagonale principale.

$$F2 = \sum_{u=0}^{Ng-1} n^2 \left\{ \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j) \right\} \quad \text{Équation 43}$$

3. Corrélation (Correlation)

$$F3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad \text{Équation 44}$$

Avec μ_x , μ_y , σ_x , σ_y sont les moyens et l'écart type de p_x et p_y .

4. La somme des variance carrés (Sum of squares Variance)

$$F4 = \sum_i \sum_j (i - \mu)^2 p(i, j) \quad \text{Équation 45}$$

5. Le moment de différence inverse (Inverse Différence Moment)

$$F5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad \text{Équation 46}$$

6. La somme de moyenne (sum Average)

$$F6 = \sum_{i=2}^{2Ng} i p_{x+y}(i) \quad \text{Équation 47}$$

7. La somme de variance (Sum variance)

$$F7 = \sum_{i=2}^{2Ng} (i - f_8)^2 p_{x+y}(i) \quad \text{Équation 48}$$

8. La somme entropie (Sum Entropy)

$$F8 = - \sum_{i=2}^{2Ng} p_{x+y}(i) \log \{ p_{x+y}(i) \} \quad \text{Équation 49}$$

9. L'entropie

L'indice d'entropie exprime le degré de désordre dans la texture. Il est d'autant plus faible qu'on a souvent le même couple de niveaux de gris. Mais, lorsque les valeurs de la matrice sont presque toutes égales, l'entropie est élevée. Ceci permet de caractériser le degré de granulation de l'image. Plus l'entropie est élevée, plus la granulation est grossière.

$$F9 = - \sum_i \sum_j p(i, j) \log(p(i, j)) \quad \text{Équation 50}$$

10. La différence de variance (différence Variance)

$$F10 = \text{variance of } p_{x-y} \quad \text{Équation 51}$$

11. La différence d'entropie (Différence entropy)

$$F11 = - \sum_{i=0}^{Ng-1} p_{x-y}(i) \log\{p_{x-y}(i)\} \quad \text{Équation 52}$$

12, 13) Les mesures d'information de corrélation (Information Measures of correlation)

$$F12 = \frac{HXY - HXY1}{\max\{HX, HY\}} \quad \text{Équation 53}$$

$$F13 = (1 - \exp[-2.0(HYX2 - HXY)])^{1/2} \quad \text{Équation 54}$$

$$HXY = - \sum_i \sum_j p(i, j) \log(p(i, j)) \quad \text{Équation 55}$$

HX ET HY sont les entropies of p_x, p_y

$$HXY1 = - \sum_i \sum_j p(i, j) \log\{p_x(i)p_y(j)\} \quad \text{Équation 56}$$

$$XHY2 = - \sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\} \quad \text{Équation 57}$$

14. Les coefficients de corrélation maximale (Maximal Correlation Coefficient)

F14 = (Second largest eigenvalue of Q)^{1/2}

$$\text{Avec } Q(i, j) = \sum_k \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)} \quad \text{Équation 58}$$

La réussite de la matrice de co-occurrence repose sur le bon choix des paramètres qui sont :

La taille de la matrice sur laquelle s'effectue la mesure, et la distance D qui sépare les deux pixels du motif, et l'angle θ .

Après l'étape d'extraction des caractéristiques nous procédons à une classification permettant de catégoriser les images requêtes du système.

II. Classification

Les méthodes de classification ont pour but d'identifier les classes auxquelles appartiennent des objets à partir de certains traits descriptifs. Elles s'appliquent à un grand nombre d'activités humaines et conviennent en particulier au problème de la prise de décision automatisée. La procédure de classification sera extraite automatiquement à partir d'un ensemble d'exemples. Un exemple consiste en la description d'un cas avec la classification correspondante. Un système d'apprentissage doit alors, à partir de cet ensemble d'exemples, extraire une procédure de classification. Il s'agit en effet d'extraire une règle générale à partir des données observées. La procédure générée devra classer correctement les exemples de l'échantillon et avoir un bon pouvoir prédictif pour classer correctement de nouvelles descriptions.

Les méthodes utilisées pour la classification sont nombreuses, citons : la méthode des Machines à Vecteurs Supports (SVM), les Réseaux de Neurones, etc. Nous présentons dans cette mémoire une étude détaillée de la technique SVM. Cette méthode a montré son efficacité dans de nombreux domaines d'applications tels que le traitement d'image, la catégorisation de textes ou les diagnostics médicaux, et ce même sur des ensembles de données de très grandes dimensions.

II.1. Les Machines à Vecteurs Supports (SVM)

Les machines à vecteurs supports (**SVM**) ont été mises en place par Vladimir Vapnik et ses collègues.

La plus ancienne version était celle de (Vapnik, 1979), mais le papier principal était dans

(Vapnik, 1995). Le succès de cette méthode est justifié par les solides bases théoriques qui la soutiennent. Elles permettent d'aborder des problèmes très divers dont la classification. SVM est une méthode particulièrement bien adaptée pour traiter des données de très haute dimension telles que les textes et les images. Depuis leur introduction dans le domaine de la reconnaissance de formes, plusieurs travaux ont pu montrer l'efficacité de ces techniques principalement en traitement d'image.

Cette technique est une méthode de classification à deux classes qui tente de séparer linéairement les exemples positifs des exemples négatifs dans l'ensemble des exemples.

Chaque exemple doit être représenté par un vecteur de dimension n . La méthode cherche alors l'hyperplan qui sépare les exemples positifs des exemples négatifs, en garantissant que la marge entre le plus proche des positifs et des négatifs soit maximale. Intuitivement, cela garantit un bon niveau de généralisation car de nouveaux exemples pourront ne pas être trop similaires à ceux utilisés pour trouver l'hyperplan mais être tout de même situés franchement d'un côté ou l'autre de la frontière. L'efficacité des SVM est très bonne pour la reconnaissance de formes. Un autre intérêt est la sélection de Vecteurs Supports qui représentent les vecteurs discriminant grâce auxquels est déterminé l'hyperplan. Les exemples utilisés lors de la recherche de l'hyperplan ne sont alors plus utiles et seuls ces vecteurs supports sont utilisés pour classer un nouveau cas. Cela en fait une méthode très rapide. Dans ce chapitre nous présentons les aspects théoriques de la méthode SVM.

Les machines à vecteurs supports sont une classe d'algorithmes d'apprentissage utilisée pour la classification binaire. L'idée de base est de trouver un hyperplan qui sépare les données parfaitement dans ses deux catégories.

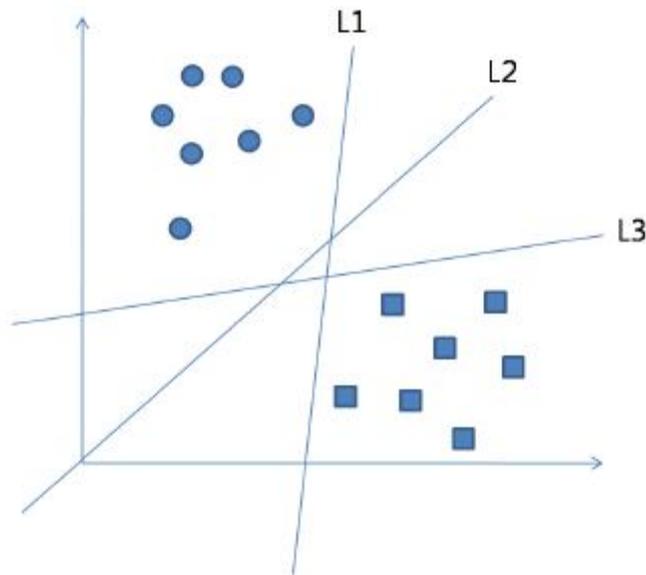


Figure 25 Des classifieurs linéaires (hyperplan) pour un espace de deux dimensions

Les points de données D de la figure 25. (ou ensemble d'apprentissage) peuvent être exprimée mathématiquement comme suit :

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$$

Où x_i est un vecteur à n dimensions, y_i est soit 1 soit -1 désignant la classe dont le point x_i appartient.

La fonction de classification SVM $F(x)$ prend la forme

$$F(x) = w \cdot x - b \quad \text{Équation 59}$$

w est le vecteur de pondération et b est la polarisation, sont calculée par la méthode SVM dans le processus d'apprentissage.

Tout d'abord, pour un classement correcte de l'ensemble d'apprentissage, $F(\cdot)$ (ou w et b) doit retourner un nombre positif pour des points de données positives et des nombres négatifs si les données sont négatifs. Et ceci pour chaque point x_i dans D .

$$w \cdot x_i - b > 0 \text{ if } y_i = 1$$

$$w \cdot x_i - b < 0 \text{ if } y_i = -1$$

Ces conditions peuvent être écrites comme suit :

$$y_i(w \cdot x_i - b) > 0, \forall (y_i, x_i) \in D$$

S'il existe une telle fonction F linéaire qui permet de classer correctement tous les points de D alors D est appelé linéairement séparable.

F (ou l'hyperplan) doit maximiser la marge. La marge est la distance entre l'hyperplan et les plus proches points de données. Un exemple d'un tel hyperplan est illustré sur la figure 26.

La distance entre l'hyperplan et un vecteur x_i est formulée sous forme $\frac{|F(x_i)|}{\|w\|}$ Ainsi, la

marge devient : $\frac{1}{\|w\|}$.

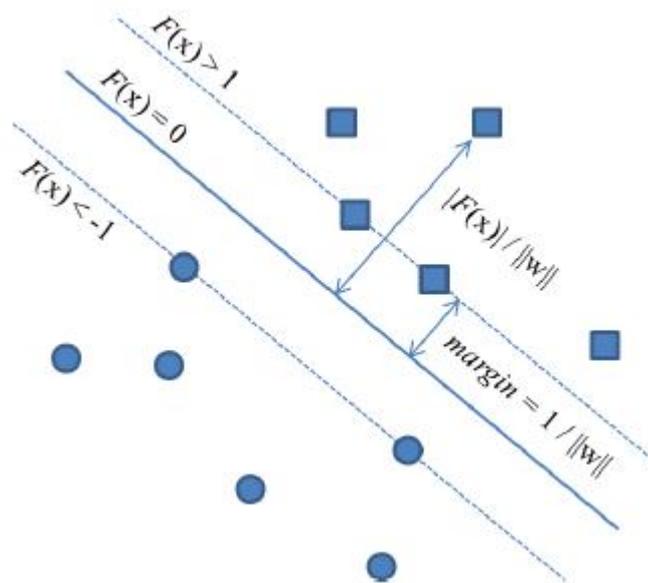


Figure 26 Fonction de classification SVM: la maximisation de la marge d'hyperplan dans un espace bidimensionnel

Les plus proches vecteurs, qui satisfont l'équation $y_i(w \cdot x_i - b) \geq 1, \forall (y_i, x_i) \in D$, avec un signe d'égalité, sont appelés vecteurs de support.

Maximiser la marge revient à minimiser $\|w\|$. Ainsi, le problème d'apprentissage dans SVM devient un problème d'optimisation sous contrainte comme suit :

Minimiser : $Q(w) = \frac{1}{2} \|w\|^2$ Équation 60

Avec : $y_i(w.x_i - b) \geq 1, \forall (x_i, y_i) \in D$ Équation 61

Le facteur 1/2 est utilisé pour plus de commodité mathématique.

Le problème d'optimisation contraint (60) et (61) est appelé problème primal. Il est caractérisé comme suit:

- La fonction objective (60) est une fonction convexe de w .
- Les contraintes sont linéaires dans w .

Le problème d'optimisation sous contrainte peut être résolu avec utilisation de la méthode des multiplicateurs Lagrange. Tout d'abord, par la construction de la fonction Lagrange:

$$J(w, b, \alpha) = \frac{1}{2} w \cdot w - \sum_{i=1}^m \alpha_i \{ y_i (w \cdot x_i - b) - 1 \} \quad \text{Équation 62}$$

Où les variables non négatives α sont appelés multiplicateurs de Lagrange. La solution du problème d'optimisation est déterminée par la fonction de Lagrange $J(w, b, \alpha)$. Elle doit être réduite au minimum par rapport à w et b et doit également être maximisée par rapport à α

nous obtenons ce qui suit deux conditions d'optimalité:

$$\text{Condition 1: } \frac{\partial J(w, b, \alpha)}{\partial w} = 0$$

$$\text{Condition 2: } \frac{\partial J(w, b, \alpha)}{\partial b} = 0$$

Après réarrangement des termes, la condition 1 donne $w = \sum_{i=1}^m \alpha_i y_i x_i$

Et la condition 2 donne $\sum_{i=1}^m \alpha_i y_i = 0$

Devant un tel problème d'optimisation de contrainte, il est possible de construire un autre problème appelé problème dual. Le problème dual a la même valeur optimale que le problème primal, mais avec les multiplicateurs de Lagrange fournissant une solution optimal.

On développe terme à terme équation (62), comme suit:

$$J(w, b, \alpha) = \frac{1}{2} w \cdot w - \sum_{i=1}^m \alpha_i y_i w \cdot x_i - b \sum_{i=1}^m \alpha_i y_i + \sum_{i=1}^m \alpha_i \quad \text{Équation 63}$$

Le troisième terme sur le côté droit de l'équation (63) est égal à zéro selon la condition 2.

Et à partir condition 1 on a $w \cdot w = \sum_{i=1}^m \alpha_i y_i w \cdot x = \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \cdot x_i x_j$

En conséquence, suivant la fonction objective $J(w, b, \alpha) = Q(\alpha)$ nous pouvons reformuler(63)

$$Q(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i x_j$$

Où les α_i sont positifs.

Nous allons maintenant énoncer le problème dual:

Maximiser

$$Q(\alpha) = \sum_{i=1}^m \alpha - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i x_j \quad \text{Équation 64}$$

Avec $\sum_i \alpha_j y_i = 0$

$$\alpha \geq 0$$

La fonction $Q(\alpha)$ à maximiser ne dépend que des données d'entrée sous la forme d'un ensemble de produit scalaire $\{x_i \cdot x_j\}_{(i,j)=1}^m$.

Après avoir déterminé les multiplicateurs optimal de Lagrange, noté par un α^* , on peut calculer le vecteur de poids optimum w^* .

$$w^* = \sum_{i=1}^m \alpha_i^* y_i x_i \quad \text{Équation 65}$$

La solution du problème dual α^* doit satisfaire la condition suivante :

Pour $i=1,2,\dots,m$

$$\alpha_i^* \{y_i (w^* \cdot x_i - b) - 1\} = 0$$

Cette condition implique que c'est seulement quand x_i est un vecteur de support que $y_i(w^* \cdot x_i - b) = 1$, et leurs coefficients correspondant α^* sont différents de zéro. En d'autres termes, les x_i dont les coefficients correspondant α^* sont nuls n'affectera pas le vecteur de pondération optimal w^* en raison de l'équation (65). Ainsi, le vecteur optimal w^* dépendra uniquement des vecteurs de support, dont les coefficients sont positifs. Une fois quand on a les α^* positifs calculés et leurs vecteurs supports correspondants, nous pouvons calculer le biais b

$$b^* = 1 - w^* \cdot x_i \quad \text{Équation 66}$$

Donc $F(x)$ devient comme suit

$$F(x) = \sum_i \alpha_i y_i x_i \cdot x - b \quad \text{Équation 67}$$

La discussion jusqu'ici a porté sur les cas linéairement séparables ; Cependant le problème d'optimisation (60) et (61) n'aura pas une solution si D n'est pas linéairement séparables.

Pour traiter un tels cas, SVM à marge souple (soft margin SVM) permet d'étiqueter ces points de données, tout en maximisons la marge.

La méthode introduit des variables d'écart ξ_i qui mesurent le degré de classification erronée.

Ce qui suit est un problème d'optimisation pour soft margin SVM.

$$\text{Minimiser } Q_1(w, v, \xi_i) = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i \quad \text{Équation 68}$$

$$\begin{aligned} & y_i(w \cdot x_i - b) \geq 1 - \xi_i \\ \text{Avec : } & \forall (x_i, y_i) \in D \\ & \xi_i \geq 0 \end{aligned} \quad \text{Équation 69}$$

En raison de ξ_i dans l'équation (69) les points de données sont autorisés à être mal classés, et les erreurs de classification seront réduites au minimum tout en maximisant la marge suivant la fonction d'objectif (68).

C est un paramètre qui détermine le compromis entre la taille et la marge la quantité d'erreur dans l'apprentissage.

Cette forme primitive peut être transformée à la forme suivante en utilisant les multiplicateurs de Lagrange.

Maximiser :

$$Q_2(\alpha) = \sum_i \alpha_i - \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i x_j \quad \text{Équation 70}$$

Avec : $\sum_i \alpha_i y_i = 0$
 $C \geq \alpha \geq 0$

Notez que ni les variables d'écart ξ_i ni leurs multiplicateurs de Lagrange apparaissent dans le problème dual.

Le cas non séparable diffère du cas séparable dans la contrainte $\alpha_i \geq 0$ qui est remplacée par la contrainte plus stricte $C \geq \alpha_i \geq 0$.

L'optimisation sous contrainte pour le cas non séparable et le calcul des valeurs optimales des vecteurs de pondération w et b est procédée de la même façon que dans le cas linéairement séparable.

$$w^* = \sum_{i=1}^{m_s} \alpha_i^* y_i x_i \quad \text{Équation 71}$$

Où m_s est le nombre de vecteurs de support dont coefficient correspondant $\alpha_i > 0$.

La détermination des valeurs optimales de la polarisation suit également le mode opératoire similaire à celui décrit précédemment.

Une fois que α et b sont calculés, $F(x)$ est utilisé pour classer un nouveau objet.

les relation entre ξ , α , et C sont définies par

$$\alpha_i \{y_i (w \cdot x_i - b) - 1 + \zeta_i\} = 0 \quad \text{Équation 72}$$

$$i = 1, 2, 3, \dots, m$$

Et

$$\mu_i \zeta_i = 0$$

$$i = 1, 2, \dots, m \quad \dots \text{Équation 73}$$

$$\alpha_i + \mu_i = c \dots\dots \text{Équation 74}$$

μ_i Sont des multiplicateurs de Lagrange qui ont été introduits pour appliquer la positivité sur les variables d'écart pour tout i .

En combinant (73), (74) nous voyons que

$$\begin{aligned} \zeta_i &= 0, \text{ if } \alpha_i < c \\ \zeta_i &\geq 0, \text{ if } \alpha_i = c \end{aligned} \text{Équation 75}$$

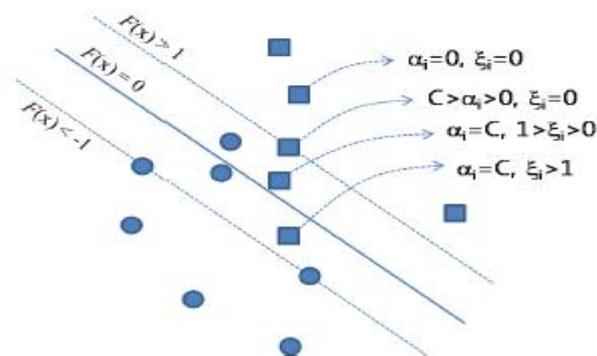


Figure 27 Les relations entre ξ , x_i , et C

Toutefois, les exemples de données ne sont pas souvent linéairement séparables, et il n'y a aucun hyperplan linéaire qui peut séparer les classes. Afin d'apprendre une fonction non linéaire dans ce cas, l'espace des données peut toujours être plongé dans un espace de plus grande dimension appelé **espace de caractéristiques** de façon à ce que les données deviennent linéairement séparables. Dans cet espace, on construit un hyperplan optimal séparant les classes tel que :

- Les vecteurs appartenant aux différentes classes se trouvent de différents côtés de l'hyperplan.
- La plus petite distance entre les vecteurs et l'hyperplan (la marge) soit maximale.

Le principe de cette technique est représenté dans la figure ci-dessous:

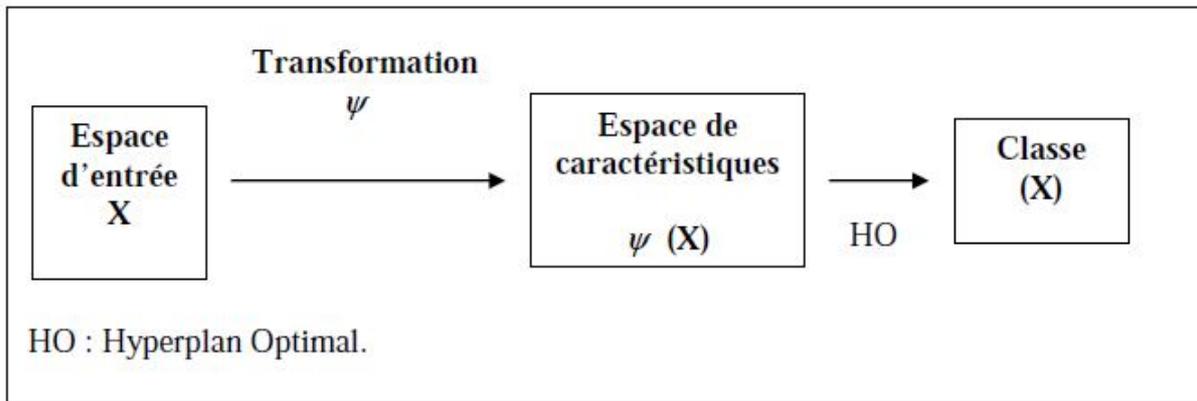


Figure 28 Principe de la technique SVM.

Le processus de recherche d'une fonction de classification en utilisant les SVMs non linéaires compose en de deux étapes :

Tout d'abord, les vecteurs d'entrée sont transformés en vecteurs de caractéristiques de grande dimension où les données d'apprentissage peuvent être linéairement séparées.

Puis, les SVMs sont utilisés pour trouver l'hyperplan de marge maximale dans le nouvel espace de caractéristique.

L'hyperplan de séparation devient une fonction linéaire dans l'espace de caractéristiques transformé, mais une fonction non linéaire dans l'espace d'entrée d'origine.

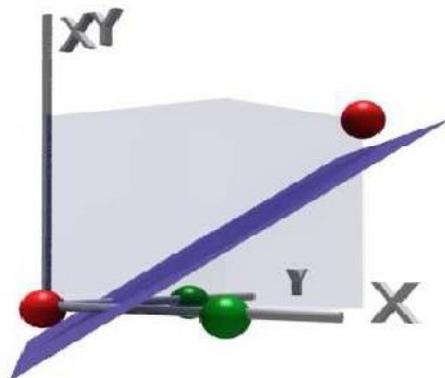


Figure 29 un classifieur non linéaires (hyperplan) pour un espace plus de deux dimensions

Soit x un vecteur d'entrée dans l'espace à n dimensions, et $\varphi(\cdot)$ une fonction non linéaire de l'espace de caractéristiques de grande dimension. L'hyperplan qui représente la frontière de décision dans l'espace des attributs est défini comme suit :

$$w \cdot \varphi(x) - b = 0 \quad \text{Équation 76} \quad \text{Avec} \quad w = \sum \alpha_i y_i \varphi(x_i) \quad \text{Équation 77}$$

La fonction de décision devient

$$F(x) = \sum_i^m \alpha_i y_i \varphi(x_i) \cdot \varphi(x) - b \quad \text{Équation 78}$$

Et

$$Q(\alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \varphi(x_i) \cdot \varphi(x_j) \quad \text{Équation 79}$$

Le calcul du produit intérieur dans l'espace transformé semble être tout à fait complexes et souffrent du problème de dimensionnalité. Pour éviter ce problème, l'astuce du noyau est utilisée. L'astuce de noyau remplace le produit scalaire dans l'espace de caractéristiques avec une fonction noyau K dans l'espace d'entrée d'origine comme suit.

$$K(u, v) = \varphi(u) \cdot \varphi(v) \quad \text{Équation 80}$$

Le problème dual est maintenant défini en utilisant la fonction noyau comme suit.

Maximiser

$$Q_2(\alpha) = \sum_i \alpha_i - \sum_i \sum_j \alpha_i \alpha_j y_i y_j k(x_i, x_j) \quad \text{Équation 81}$$

$$\text{Avec : } \begin{cases} \sum_i \alpha_i y_i = 0 \\ C \geq \alpha \geq 0 \end{cases} \quad \text{Équation 82}$$

Et la fonction de classification devient :

$$F(x) = \sum_i \alpha_i y_i k(x_i, x) - b \quad \text{Équation 83}$$

La fonction du noyau est une sorte de fonction de similarité entre deux vecteurs où la fonction de sortie est maximisée lorsque les deux vecteurs deviennent équivalents.

Pour cette raison, SVM peut apprendre une fonction de toutes les formes de données au-delà de vecteurs (tels que des arbres ou des graphes), tant que nous pouvons calculer une fonction de similarité entre toutes les paires d'objets de données.

Les fonctions noyaux suivants sont couramment utilisées :

$$\text{Polynomial: } K(a, b) = (a \cdot b + 1)^d \quad \text{Équation 84}$$

$$\text{Radial Basis Function (RBF): } k(a, b) = \exp(-\gamma \|a - b\|^2) \quad \text{Équation 85}$$

Sigmoid : $\tanh(ka \cdot b + c)$.

Les Séparateurs à vaste marge ont été développés pour traiter des problèmes binaires mais ils peuvent être adaptés pour traiter les problèmes multi-classes.

Une classification multi-classe peut être mise en œuvre en utilisant la méthode de couplage par paire et ce ci par la combinaison de multiple classificateurs binaire.

II.2. Les méthodes de décomposition

Les méthodes de décomposition permettent d'aborder un problème de discrimination à catégories multiples comme une combinaison de problèmes de calcul de dichotomies.

II.2.1. L'approche "un contre tous"

L'approche "un contre tous" est la plus simple et la plus ancienne des méthodes de décomposition. Elle consiste à utiliser un classifieur binaire (à valeurs réelles) par catégorie. Le k -ième classifieur est destiné à distinguer la catégorie d'indice k de toutes les autres. Pour affecter un exemple, on le présente donc à Q classifieurs, et la décision s'obtient en application du principe " le gagnant prend tout " : l'étiquette retenue est celle associée au classifieur ayant renvoyé la valeur la plus élevée.

II.2.2. L'approche "un contre un"

Une autre méthode de décomposition très naturelle est la méthode "un contre un". Elle consiste à utiliser un classifieur par couple de catégories. Le classifieur indicé par le couple (K, L) (avec $1 \leq K < L \leq Q$), est destiné à distinguer la catégorie d'indice K de celle d'indice L . Pour affecter un exemple, on le présente donc à C_Q^2 classifieurs, et la décision s'obtient habituellement en effectuant un vote majoritaire.

II.2.3. L'approche de codes correcteurs d'erreurs (ECOC)

Les codes correcteurs d'erreurs (Error-Correcting Output Codes) sont utilisés depuis quelques années pour concevoir des ensembles de classifieurs divers. L'idée de base derrière l'utilisation de ces méthodes est d'éviter de concevoir directement un classifieur multi-classes en décomposant le problème de classification en un ensemble de classifieurs binaires (fonctions discriminantes). Chaque classifieur dans l'ensemble est utilisé pour discriminer deux classes, ces classes pouvant être des regroupements des classes d'origine.

On définit ainsi une matrice binaire $M = (m_{cn})$ de $M_{C, N}(\{-1, 1\})$ contenant C lignes (où C est le nombre de classes) et N colonnes (N étant le nombre de classifieurs binaires f_n impliqués).

Chaque vecteur ligne M_c de la matrice M est un mot-code pour une classe donnée.

Chaque exemple x , après classification par les N classifieurs, se voit attribuer un vecteur $f = [f_1(x); \dots; f_N(x)]$ regroupant les résultats des fonctions de décision.

Le principe de la méthode ECOC (Error Correcting Output Codes) consiste à attribuer à l'exemple x la classe minimisant la distance L (distance de Manhattan) entre son mot-code M_c et le mot-code f de l'exemple :

$$Y = \arg \min_{1 \leq c \leq C} L(f, M_c) = \arg \min_{1 \leq c \leq C} \sum_{n=1}^N |f_n(x) - m_{cn}| \dots \text{Équation 86}$$

Conclusion

Nous avons présenté dans ce chapitre les descripteurs de texture et de forme utilisés dans notre travail.

La matrice de co-occurrence a été utilisée pour caractériser la texture. Nous proposons également d'utiliser la moyenne et la variance d'une image après l'application d'ondelettes de Gabor. Et pour décrire la forme nous avons utilisé les contours actifs pour segmenter l'objet.

Dans ce chapitre nous avons aussi présenté les concepts de la méthode SVM, pour laquelle nous avons expliqué la démarche de construction de l'hyperplan optimal dans les cas des données linéairement séparables et non linéairement séparables, et cela en se basant sur les fondements mathématiques pour le cas binaire (à deux classes). Nous avons ensuite décrit la démarche suivie pour généraliser cette technique pour le cas de multi-classes à partir du cas binaire.

Chapitre 4 : Architecture du système de recherche par le contenu proposé

L'objectif de ce chapitre est de définir et de mettre en place un système d'indexation et de recherche d'image par le contenu.

Introduction

Dans le domaine de l'imagerie médicale, un nombre croissant d'images sont traitées tout les jours. Ainsi, dans la littérature des efforts d'investigation de nouvelles méthodes de classification et d'archivage de ces données sont menées par plusieurs chercheurs.

Les bases de données classiques permettent des recherches uniquement textuelles à partir les métas-données. Souvent, le système de base de données ne contient que des références pour les données d'image qui sont stockées sous forme de fichiers individuels. En particulier, les images peuvent contenir des informations sémantiques qui ne sont pas transmises par une description textuelle. D'où, l'intérêt de la recherche d'images par le contenu. Les systèmes CBIR permettent une indexation et une recherche d'images portant sur les caractéristiques de l'image. Ils sont basées sur la recherche des images spécifiques qui sont semblable à une image requête donnée.

Les systèmes CBIR sont utilisés dans plusieurs domaines d'application. Par exemple, dans les grands hôpitaux les images médicales sont de plus en plus acquises, transférés et stockés. En conséquent des données équivalentes à plusieurs milliers de téra octets doivent être gérées chaque année.

I. L'architecture générale d'un système d'indexation et de recherche d'images :

La recherche d'images par le contenu se décompose en deux phases de traitement :
Une phase d'indexation dite hors ligne qui est une étape de caractérisation où les attributs sont automatiquement extraits à partir des images de la base, et stockés dans le vecteur de caractéristiques. Ensuite, le descripteur est stocké dans une base de données.
Une phase en ligne consiste à extraire le vecteur descripteur de l'image requête et le comparer avec les descripteurs de la base de données en utilisant une mesure de distance. Le système renvoi ensuite le résultat de la recherche dans une liste d'images ordonnées en fonction de la similarité entre leurs descripteurs visuels et le descripteur visuel de l'image requête.

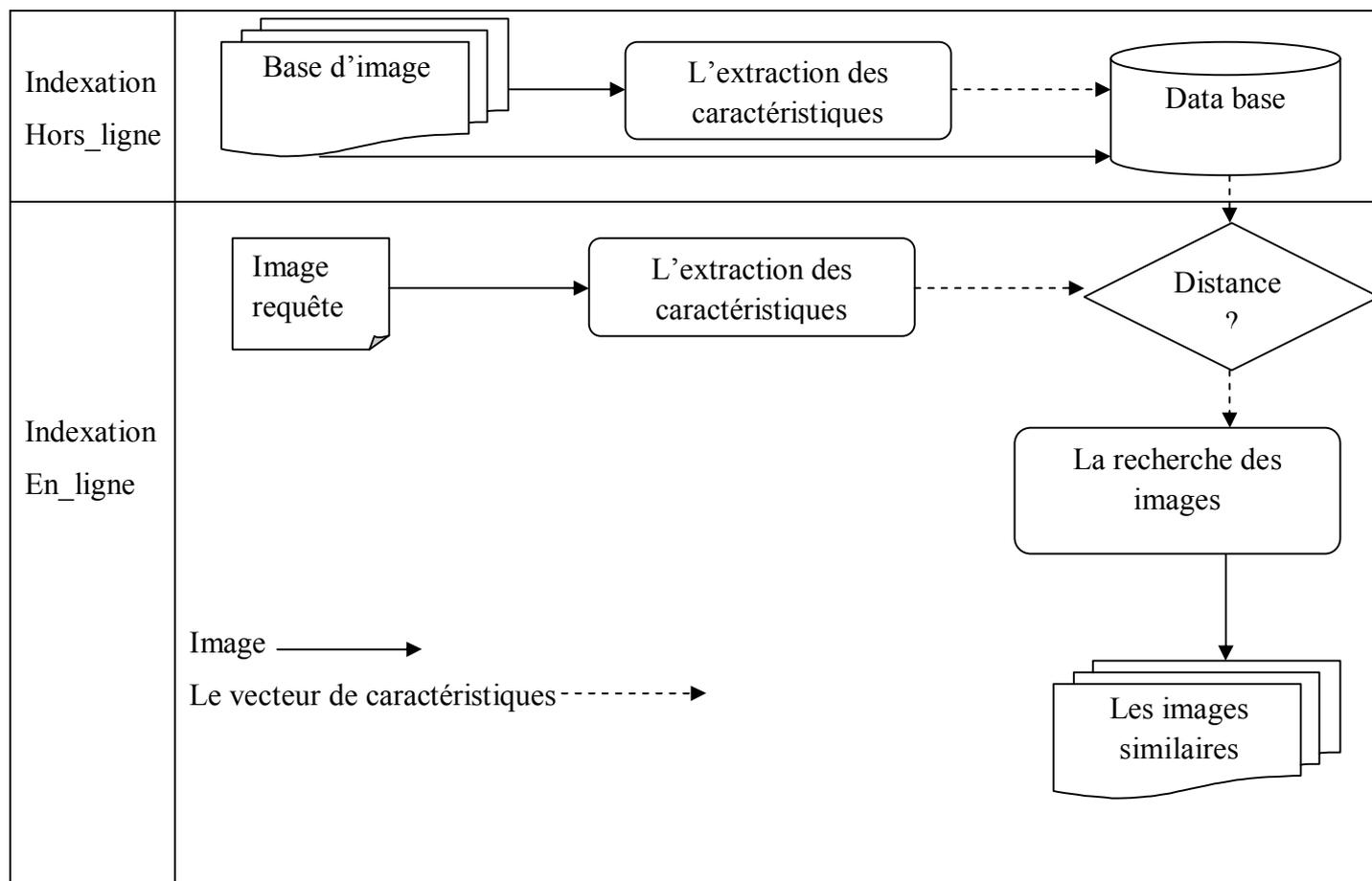


Figure 30 Architecture générale d'un système CBIR

La figure ci-dessus montre l'architecture générale d'un système CBIR. Les caractéristiques numériques sont calculées pour chaque image en entrée. En utilisant l'approche d'interrogation par exemple (QBE), les mêmes caractéristiques sont extraites à partir de l'image d'interrogation. Les images ayant les caractéristiques les plus similaires à l'image requête sont alors extraites et présentées à l'utilisateur comme réponse à sa requête.

I.1. La recherche d'image par contenu dans le domaine médicale

Les systèmes d'informations médicaux permettent souvent de fournir des informations nécessaires au bon moment et au bon endroit dans le but d'améliorer la qualité et l'efficacité du processus de soins, et cela par la recherche des images de la même modalité, et la même région anatomique de la même maladie. La plupart des systèmes appliquent plusieurs restrictions aux images d'entrées (par exemple sur type de la modalité). Généralement ces systèmes sont incapables de faire face à des images de plus d'une région du corps ou d'une

modalité, car ils utilisent implicitement ou explicitement des propriétés particulières des données au cours du traitement de l'image. Les systèmes médicaux diffèrent dans la formulation de la requête [38].

La difficulté trouvée pour développer un système d'indexation et de recherche par le contenu des images médicales est due à plusieurs facteurs [15] :

- Dans la plupart des cas, les images médicales sont des images d'intensités qui transportent moins d'informations que les images couleur.
- Les images médicales sont généralement de basse résolution et avec beaucoup de bruit. C'est difficile de les analyser automatiquement pour extraire des caractéristiques. Les images médicales acquises avec différents appareils, même en utilisant la même modalité peuvent avoir des propriétés différentes.
- Le diagnostic automatique des images médicales est le plus souvent impossible aujourd'hui (sauf dans certain cas rares et particuliers), car l'interprétation des images médicales est souvent difficile, même pour les radiologues.

I.2. Les modalités d'imagerie et CBIR

Les images médicales couvrent un large éventail de modalités. Voici quelques exemples:

- Images 2D (2D ou en temps) : Ce sont des images projetées ou des tranches du corps, par exemple, la radiographie du thorax, l'échocardiogramme.
- Images 3D volumétriques, par exemple, l'imagerie par résonance magnétique du cerveau (IRM) ou la tomodensitométrie abdominale (TDM).
- 4D (3D en temps ou le long des différentes modalités) des images, par exemple, Cardiac CT / MR / échographie.
- 5D (3D en temps et le long de modalités) des images d'un organe mobile, par exemple, asservie à la respiration PET-CT étude du poumon.

Le tableau suivant regroupe la plupart des modalités d'imagerie médicale et biologique existant d'aujourd'hui, ainsi que certaines modalités émergence.

<i>Technologie</i>	<i>Modalité d'imagerie</i>	<i>Dimension</i>	<i>Systèmes CBIR, et les modalités spécifiques</i>
Radiographie	Planaire aux rayons X (par exemple, fluoroscopie, mammographie, radiographie thoracique)	2D	Seins: Korn, Giger Spine: CBIR2, MIRS
	Tomodensitométrie (TDM)	3D	Cerveau: MIMS, Liu Poumon: ASSERT, Liu
	Angioscanner (CTA)	3D	
	Microtomographie	3D	
	Dynamique tomodensitométrie	3D+temps	
Résonance magnétique	Imagerie par résonance magnétique (IRM)	3D	Cerveau: Chu, iLive Cœur: Robinson
	Diffusion MRI (DTI)	3D	
	Angiographie par résonance magnétique (ARM)	3D	
	L'IRM fonctionnelle (IRMf)	3D/4D	Cerveau: Bai
	La spectroscopie par résonance magnétique (SRM)	3D	
	Contraste Dynamique améliorée IRM (DCE-MRI)	3D+ temps	Cœur: Glatard
ULTRASONS	En mode M	1D+ temps	
	En mode B	2D	
	Doppler	2D	
	Echographie 3D	3D	
	Imagerie échographie dynamique 3D	3D+ temps	
NUCLÉAIRE	Planar imagerie nucléaire	2D	
	Positron Emission Tomography (PET)	3D	
	Émission à photon unique Comp. Tomo. (SPECT)	3D	
	PET dynamique	3D+ temps	Oncology: FICBDS
OPTIQUE	Tomographique d'imagerie Optique	2D	
	Tomographie par cohérence optique (OCT)	3D	
	L'imagerie optique diffuse (TOD)	2D	
	Topographique d'imagerie fonctionnelle	2D+ temps	
FLUORESCENCE	Imagerie de vie fluorescence (FLIM)	3D	
	Fluorescence spectroscopie de corrélation (FCS)	3D	
	Fluorescence tomographie	3D	

	moléculaire (FMT)		
MICROSCOPIE	La lumière microscopie	2D	BASS, IDEM, I-Browse
	L'électron microscopie	2D/3D	
	microscopie confocale	2D/3D	
AUTRES	L'endoscopie	2D	
	Imagerie par bioluminescence (BLI)	2D	
	photo	2D	Peau: MELDOQ, Schmidt-Saugeon
Des modalités fusionnées	PET/CT, SPECT/CT, PET/MR	3D+mod.	
	PET/CT	(+ temps)	

Tableau 1 Les modalités d'imagerie médicale et biologique et les systèmes existants CBIR [39]

Avec l'évolution de la technologie, nous aurons plus de modalités, plus d'images par patient, ce qui demande encore plus d'efficacité de traitement des systèmes CBIR.

II. Catégorisation d'images

En général, la catégorisation automatique d'images comprend trois principes :

La représentation: l'extraction des caractéristiques pour décrire le contenu d'une image.

L'adaptation: à savoir la sélection des meilleurs caractéristiques suivant l'information discriminante.

La généralisation: l'apprentissage et l'évaluation d'un classificateur.

Jusqu'à présent, la catégorisation automatique des images médicales est limitée à un petit nombre de catégories [38].

Un framework pour la recherche des images médicale par contenu est proposé. Dans l'étape de représentation nous utiliserons les descripteurs de texture (la matrice de co-occurrence et les ondelettes de Gabor) et aussi les descripteurs de forme (la surface et le controïde) ensuite un certain nombre des caractéristiques de ces descripteurs seront adaptés pour décrire les images ; dans la partie de généralisation nous avons utilisé le classifieur SVM et la distance euclidienne.

La section suivante décrit la base utilisée.

II.1. Base d'images utilisée

Nous avons utilisé la base d'images « ImageCLEFmed 2005 ». Cette collection rassemble des radiographies anonymes, qui ont été arbitrairement choisis au département de radiologie diagnostique, Aachen University of Technology (RWTH), Aachen, Allemagne.

L'imagerie représente différents âges, sexes, vues et les pathologies. Par conséquent, la qualité d'images est très variable. Toutes les images ont été revues à la baisse pour tenir dans une dimension vers 512 x 512 en conservant les proportions d'origine⁵.

ImageCLEF vise à fournir un forum d'évaluation pour l'annotation inter-langues et la recherche d'images médicales par le contenu.

L'objectif principal de ImageCLEF est de soutenir le progrès dans le domaine de l'analyse des médias visuels, de l'indexation, la classification et la recherche par le contenu, par le développement d'une infrastructure nécessaire pour l'évaluation des systèmes de recherche d'information visuels, opérant à la fois en monolingue, inter-langues et indépendamment du contexte de la langue.

ImageCLEF a été lancé en 2003 dans le cadre du [Cross Language Evaluation Forum \(CLEF\)](#) dans le but d'apporter un soutien à l'évaluation des systèmes de recherche par le contenu pour⁶:

1. Les méthodes indépendantes de la langue pour l'annotation automatique d'images avec des concepts.
2. Les méthodes multimodales de recherche d'information basée sur la combinaison de caractéristiques visuelles et textuelles.
3. Les méthodes multilingues de recherche d'images, afin de comparer l'effet de la recherche d'images et des formulations de requêtes dans plusieurs langues.

⁵ <https://ganymed.imib.rwth-aachen.de>

⁶ <http://www.imageclef.org/>



Figure 31 Exemples d'images de la base de données d'annotation ImageCLEF 2005

Les problèmes majeurs liés aux images médicales X-Ray sont :

- Le chevauchement élevé entre les images qui appartiennent à des classes différentes.
- Une grande variabilité intra-classe.

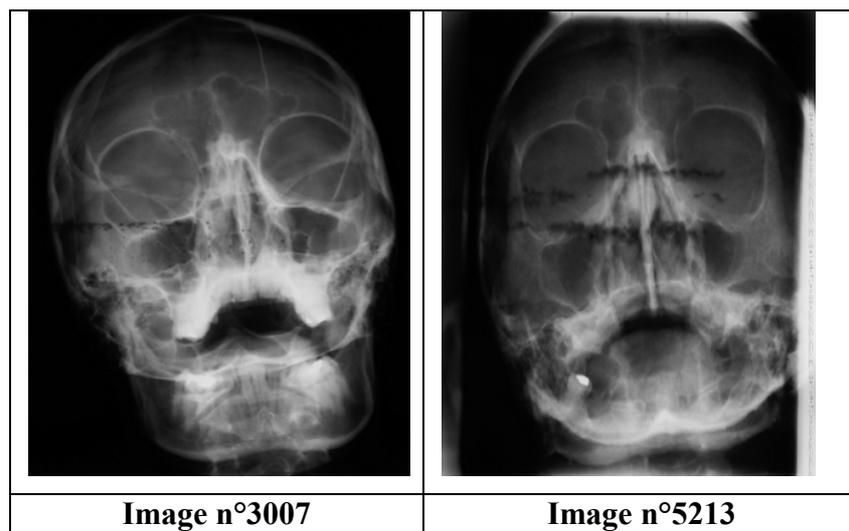


Figure 32 Deux images pour les classes n°1 et 2 respectivement.

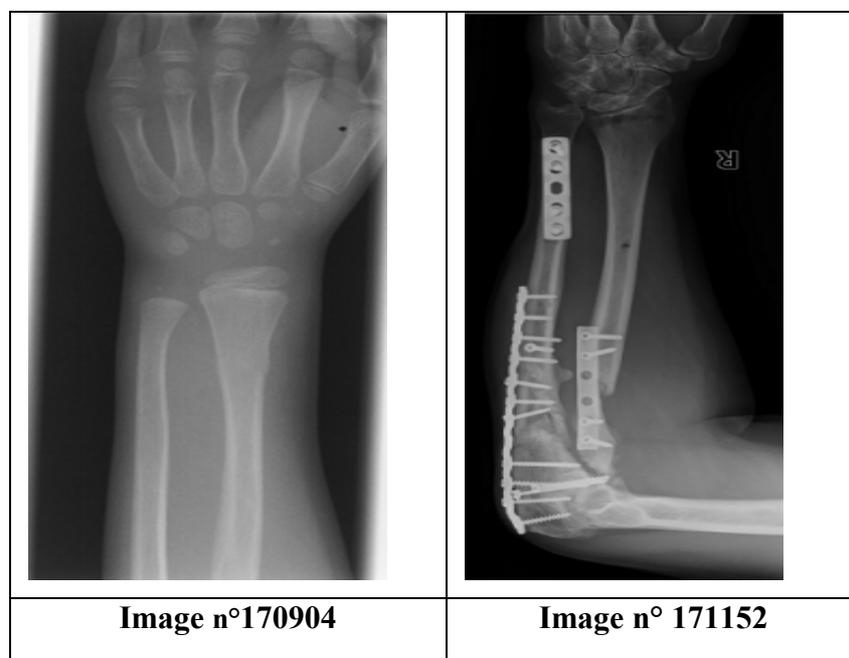


Figure 33 Deux images de la classe n°8.

II.2 .Architecture du système proposé

Notre système de recherche est composé de deux phases :

- la première phase de classification : consiste à retrouver les images correspondant à la même classe de l'image requête, la figure 34 montre l'architecture de moteur de recherche.

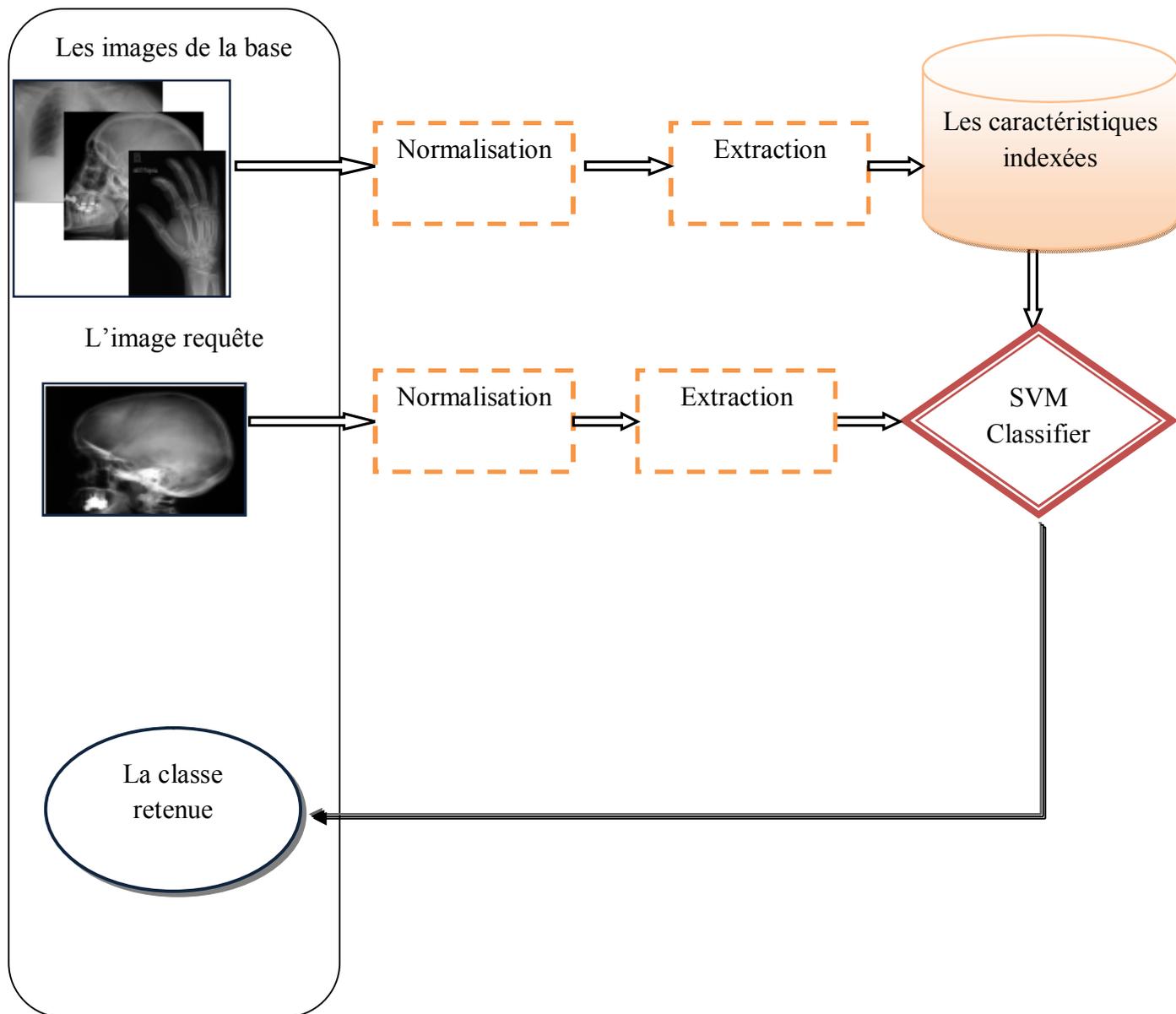


Figure 34 le diagramme de framework pour la première phase

- La deuxième phase de sélection : fait une recherche de similarité que dans les éléments de la classe retrouvée en phase 1, afin d'afficher les images les plus similaires à l'image requête par exemple : une pathologie.

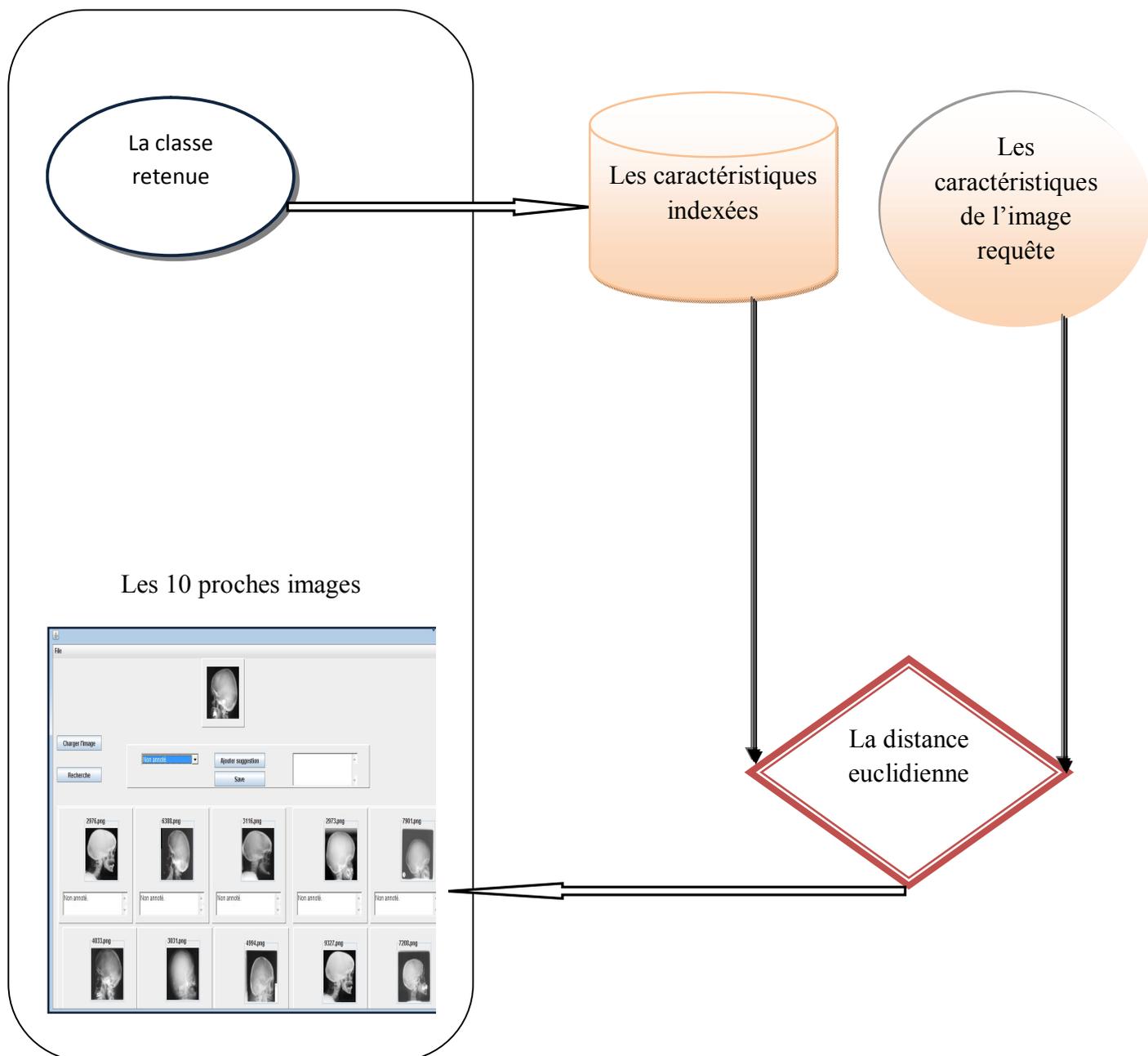


Figure 35 le diagramme de framework pour la deuxième phase

2.1. Normalisation

Il y a des incohérences entre les dimensions des images de la base. Et pour cela l'étape de normalisation est nécessaire, nous avons pris la dimension de 160×160.

2.2. L'extraction des caractéristiques

La deuxième étape après la normalisation est la construction d'un descripteur. Nous rappelons que le but de l'indexation est de fournir une représentation d'image permettant des recherches efficace. Il ne s'agit pas de coder toute l'information portée par l'image mais de se concentrer sur l'information qui permet de traduire efficacement une similarité proche des besoins exprimés par un utilisateur.

L'une des étapes clés d'une indexation efficace est l'extraction des caractéristiques primaires en accord avec le type et le but des recherches visées par le système. Ces caractéristiques sont généralement simples, intuitives et génériques.

2.2.1. Prétraitement par segmentation

Comme prétraitement, nous avons utilisé la méthode des contours actifs pour segmenter l'image et extraire les caractéristiques de la forme telles que les coordonnées du centroïde et la surface de la forme. Les pixels qui n'appartiennent pas à la forme sont mis à 0. La figure 36 montre l'état de l'évolution de la courbe.

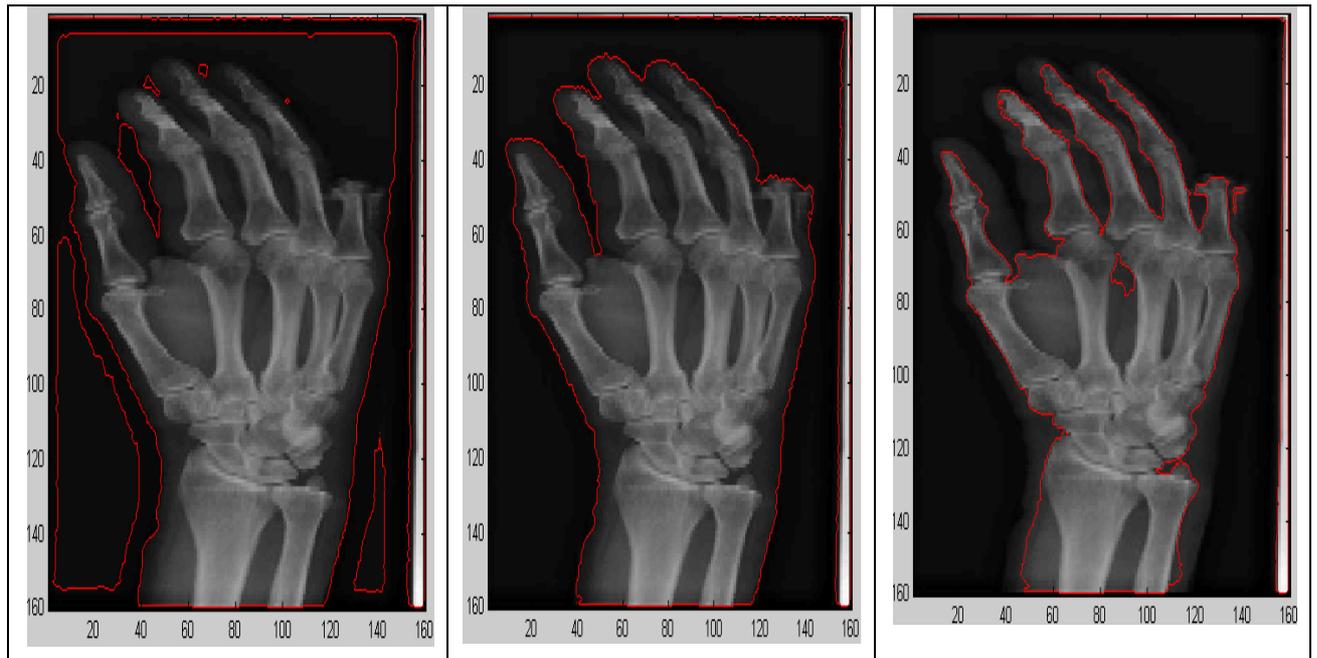


Figure 36 La croissance de la courbe autour de l'objet



Figure 37 le centroïde de la forme

Le descripteur de forme nous a permis d'avoir une bonne estimation de la surface de l'objet dans l'image. Après l'utilisation des descripteurs de forme nous avons utilisé des descripteurs de texture.

2.2.2. Les ondelettes de Gabor

Trois valeurs d'échelles sont choisies pour décrire la longueur d'onde de la fenêtre de Gauss, et pour contrôler l'échelle de l'échantillonnage.

Cinq orientations sont choisies pour décrire la fonction d'oscillation avec différentes directions.

Nous avons choisi les paramètres suivants : $\sigma = 2\pi$, $k_{max} = \pi / 2$ and $f = 2^{1/2}$; la partie réel du filtre de Gabor et ces trois grandeurs sont représentés dans la figure 38.

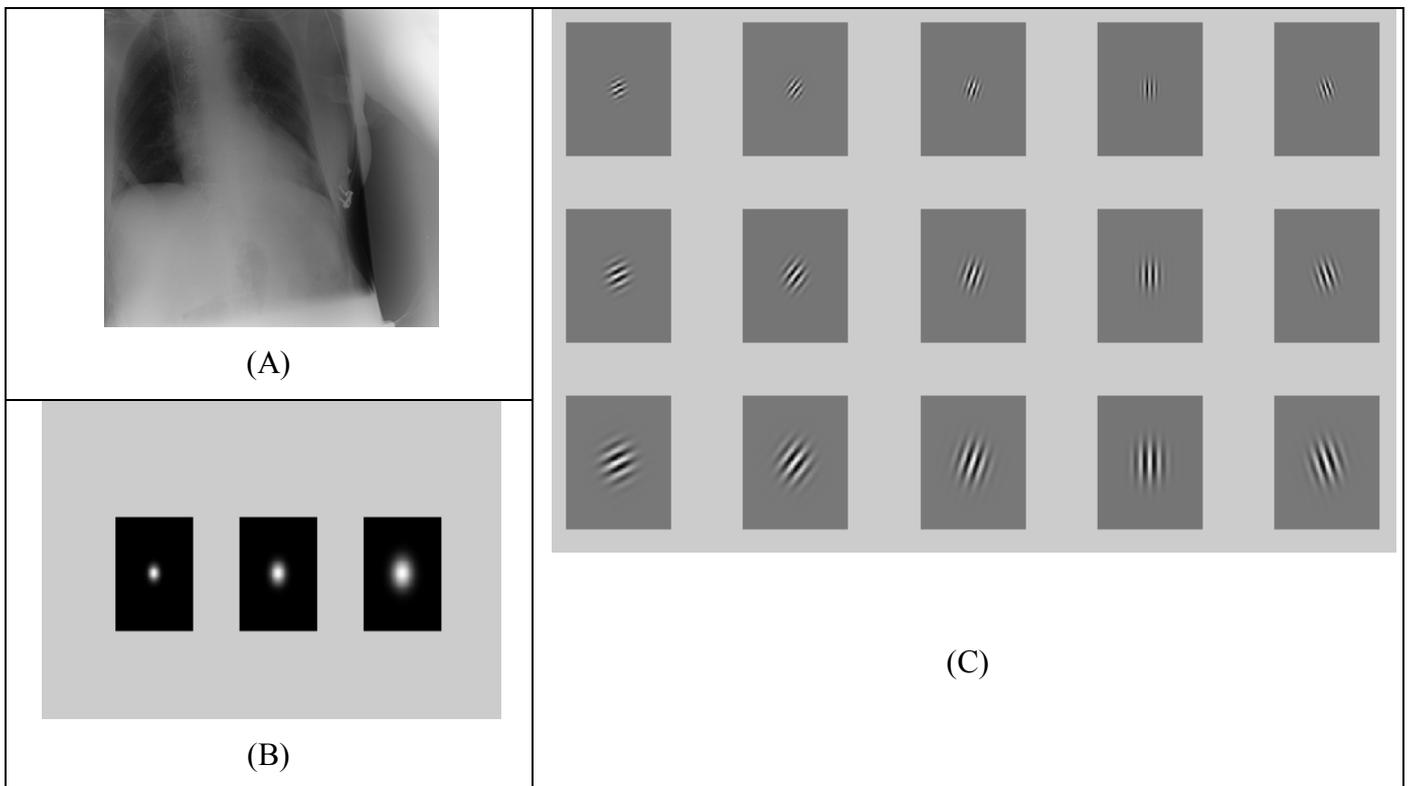


Figure 38 (A) Représentation en ondelettes de Gabor de l'image 10001 de la classe n°12 ImageCLEFmed 2005 (B) grandeurs de la Gabor (c) La partie réelle de noyaux de Gabor

Nous avons utilisé la moyenne et la variance comme des caractéristiques extraites à partir les quinze filtres de Gabor.

La moyenne est en fonction de l'énergie d'information (l'équation 40), et de la taille d'image. Supposons une image de $M \times N$.

$$\mu_{pq} = E_{pq}(x) / MN \quad \text{Équation 87}$$

La variance s'écrit sous la forme

$$\sigma_{pq} = \sqrt{\frac{\sum_x \sum_y (|G(x,y) - \mu_{pq}|)^2}{MN}} \quad \text{Équation 88}$$

Le vecteur de caractéristiques de Gabor est constitué est de dimension= 30. Les premières valeurs représentent la moyenne et les 15 derniers représentent la variance.

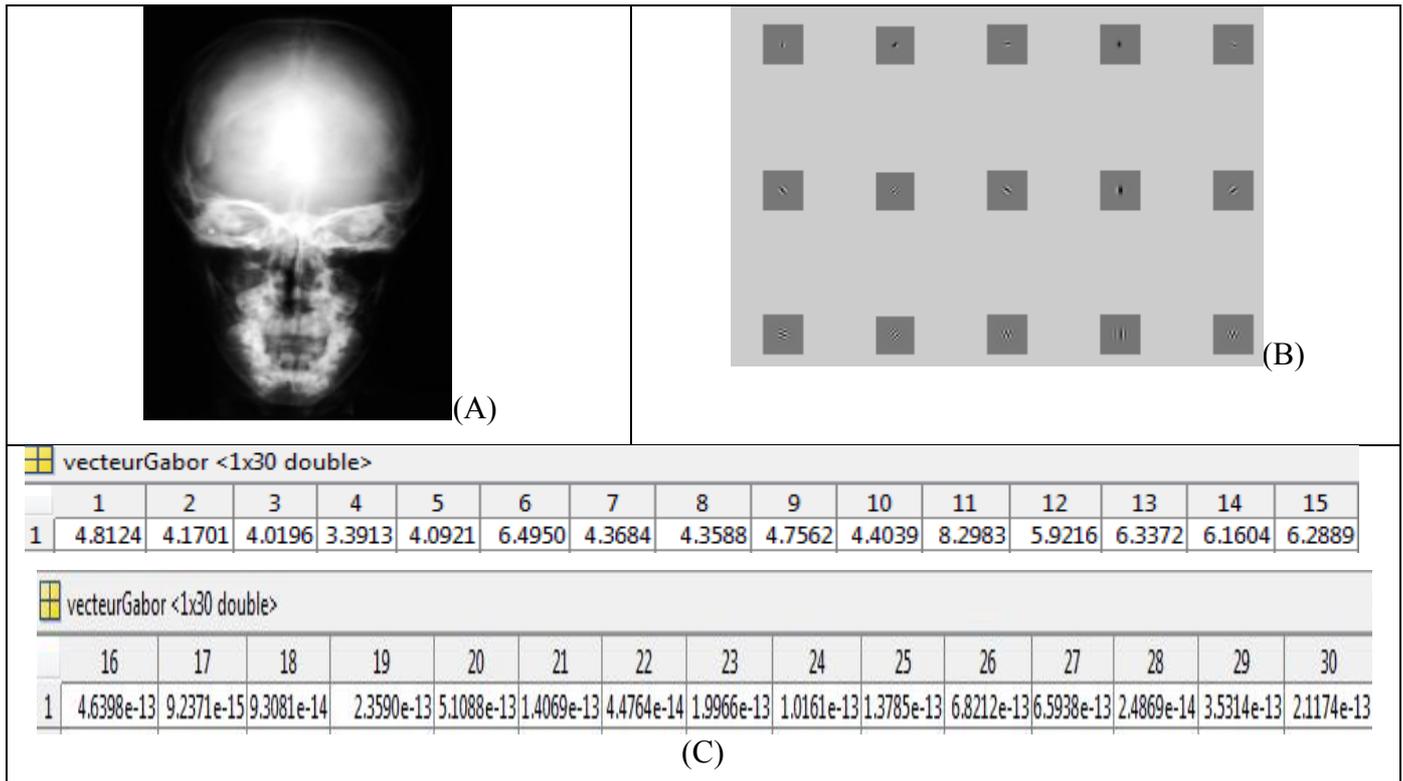


Figure 39 l'image n° 2975 de la classe n°1 (B) La partie réelle de noyaux de Gabor (C) le descripteur de l'image

2.2.3. La matrice de co-occurrence

C'est le deuxième descripteur de texture utilisé, nous avons choisi les caractéristiques suivantes :

Le contraste, corrélation, l'entropie et la somme carré des variances [2] (ont été décrit dans le chapitre précédant). En plus l'énergie, auto- corrélation, la dissemblance, groupe d'ombre, groupe d'importance et l'homogénéité [40].

$$L'energie f1 = \sum_i \sum_j p(i, j)^2 \dots \text{Équation 89}$$

$$\text{Auto- corrélation } f2 = \sum_i \sum_j (ij) p(i, j) \dots \text{Équation 90}$$

$$\text{La dissemblance } f3 = \sum_i \sum_j |i - j| p(i, j) \dots \text{Équation 91}$$

$$\text{Groupe d'ombre (Cluster Shade) } f4 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 p(i, j) \dots \text{Équation 92}$$

$$\text{Groupe d'importance (Cluster Prominence) } f5 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 p(i, j) \text{ Équation 93}$$

$$\text{Avec } \mu_x = \sum_i \sum_j i \cdot p(i, j) \dots \text{Équation 94}$$

$$\mu_y = \sum_i \sum_j j \cdot p(i, j) \dots \text{Équation 95}$$

$$\text{L'homogénéité } f_6 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i, j) \dots \text{Équation 96}$$

L'ordre des caractéristiques est comme suite :

Auto- corrélation, le contraste, corrélation, groupe d'ombre, groupe d'importance, la dissemblance, l'énergie, l'entropie, l'homogénéité, la somme carré des variances.

Nous avons utilisé les angles de 0°, 45°, 90°, 135° et une distance d=1.

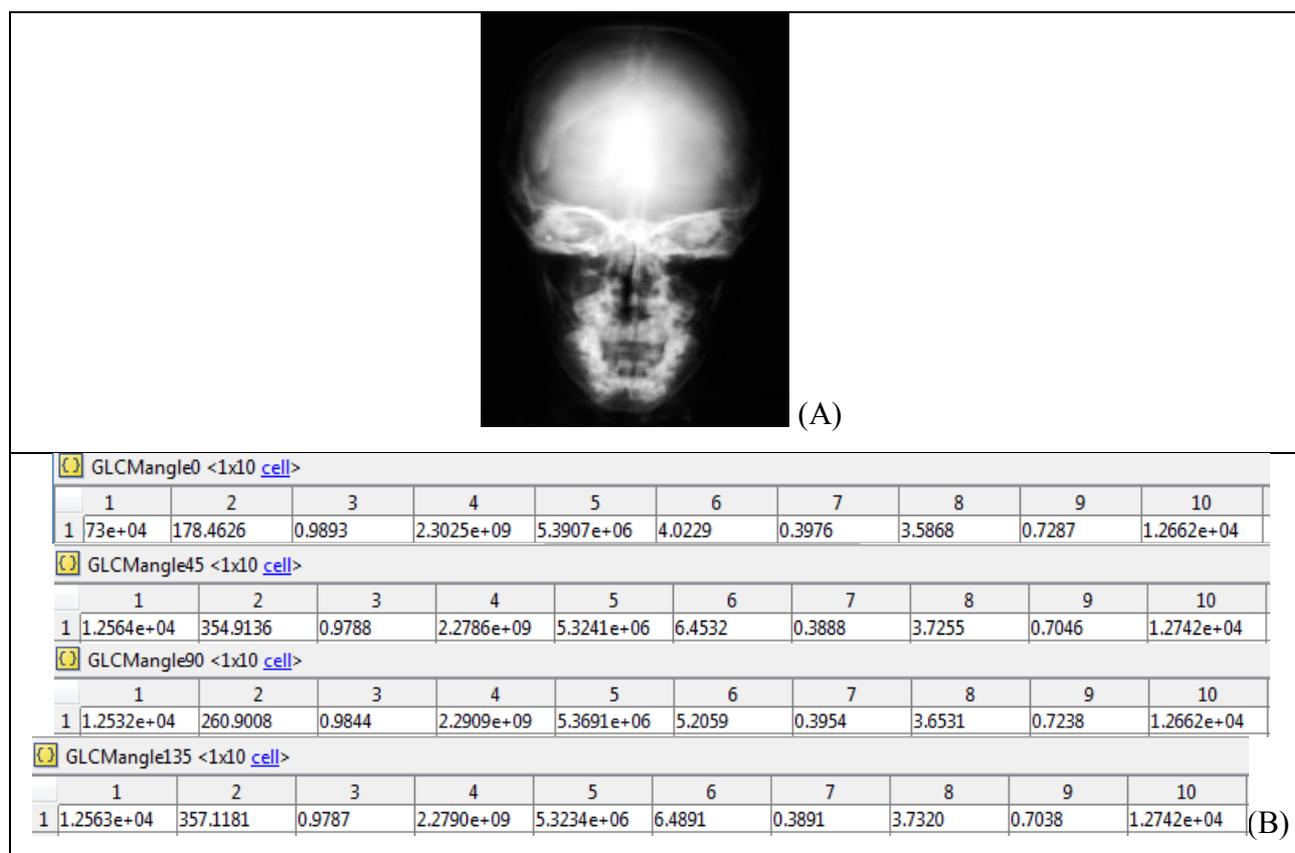


Figure 40 l'image n° 2975 de la classe n°1 (B) le descripteur de l'image

Après l'étape d'extraction des caractéristiques des images, l'étape de classification est présentée dans la suite :

2.3. La classification

La méthode SVM est bien adaptée pour le traitement des données de très haute dimension telles que les images ; pour cela nous l'avons choisi dans notre application.

Nous avons utilisé SVM multiclasse avec l'approche un contre tous pour classer les descripteurs d'images ; et nous avons utilisé la fonction noyau « la fonction à base radiale ».

$$\mathbf{K}(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2) \dots \text{Équation 97}$$

RBF est l'un des fonctions du noyau les plus populaires .Elle ajoute une «bosse» autour de chaque point de données.

$$F(\mathbf{x}) = \sum_{i=1}^m \alpha_i \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}\|^2) + \mathbf{b} \dots \text{Équation 98}$$

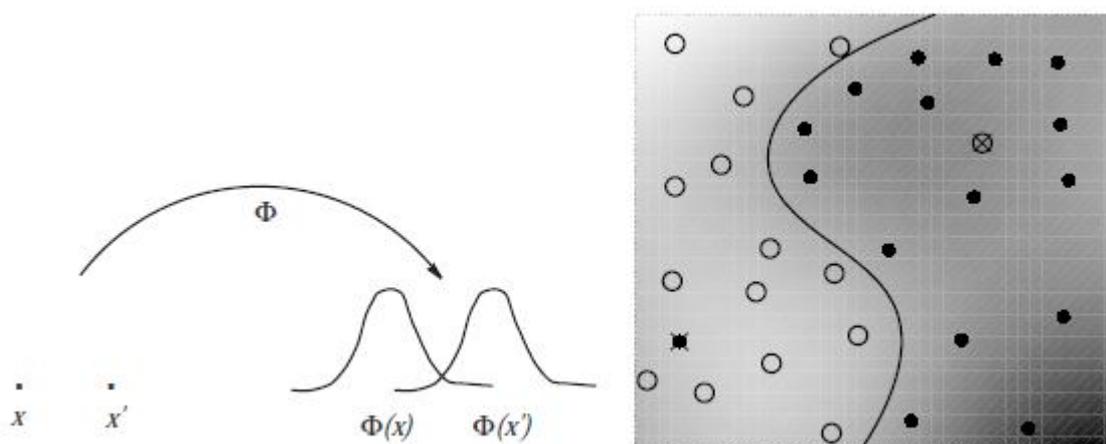


Figure 41 La fonction à base radiale

Ensuite à partir de la classe retenue nous avons utilisé le calcul de distance pour récupérer les dix images les plus proches à la requête.

2.4. La distance Euclidienne

Notre choix est porté sur la distance Euclidienne pour récupérer les dix images les plus proches et les représenter à l'utilisateur. L'idée de cette distance est que chaque image qui a N caractéristiques est un point dans l'espace de N dimensions. Chaque caractéristique est un vecteur de $V(K)$ de cet espace.

La distance entre deux images, c'est la distance entre deux points de cet espace, elle est donnée par la formule suivante :

$$\text{Dist}(\mathbf{Im}_{\text{req}}, \mathbf{Im}_{\text{cour}}) = (\sum_{k=1}^K |V(k)_{\text{imreq}} - V(k)_{\text{imcour}}|^2)^{1/2} \quad \dots \text{Équation 99}$$

Où im_{req} est l'image requête et im_{cour} est l'image courante.

$V(k)_{\text{imreq}}$ le vecteur descripteur de l'image requête.

$V(k)_{\text{imcour}}$ le vecteur descripteur de l'image courante.

II.3.L'annotation

Nous avons utilisé une simple annotation pour annoter l'image requête et même les images de la base. Cette annotation permet une aide précieuse dans le processus de diagnostic en permettant au médecin de voir des cas similaires à l'image requête. L'annotation est faite manuellement pour une image d'entrée. Il y a une liste de suggestions qui contiennent les noms des classes qui existent dans la base, l'utilisateur peut ajouter une suggestion dans l'éditeur adroite ensuite il peut ajouter d'autres informations concernant l'image.

La figure 42 montre l'annotation d'une image.

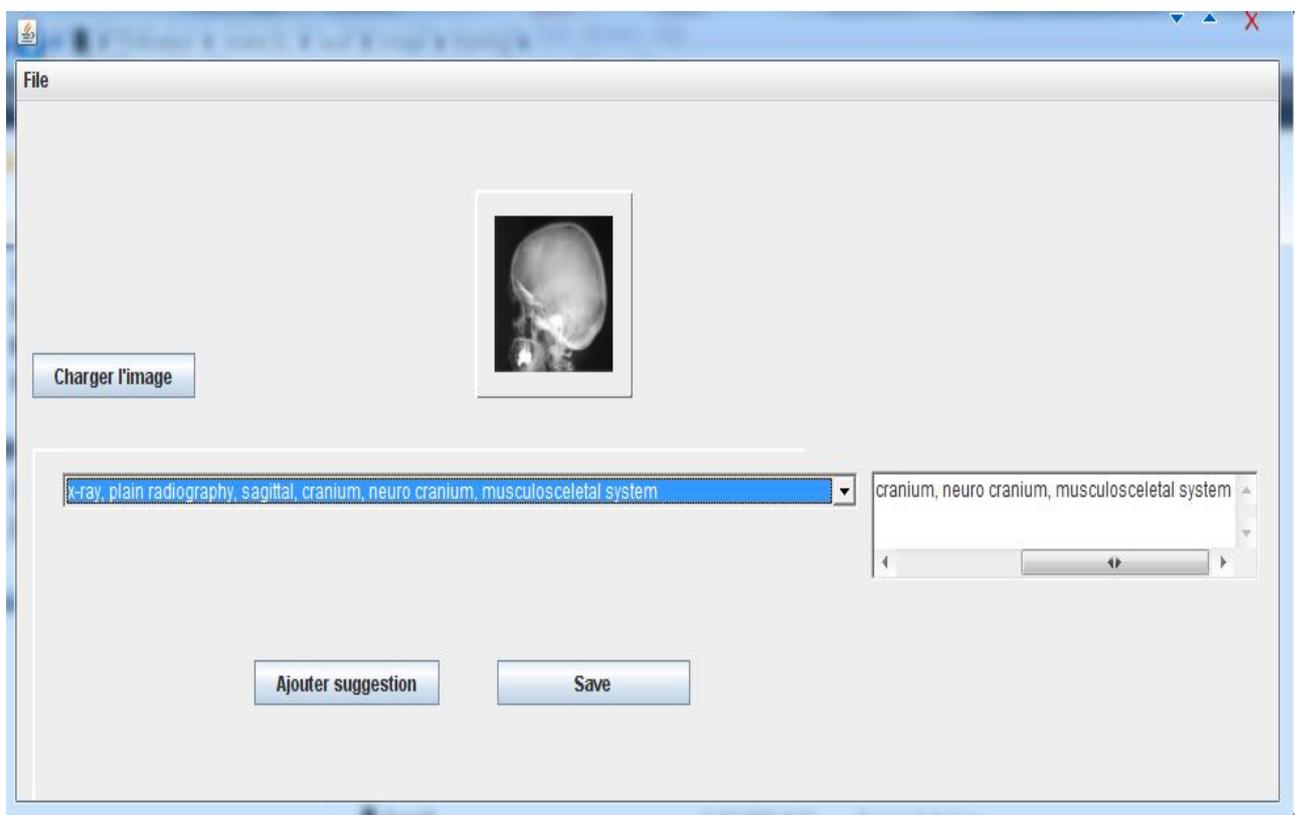


Figure 42 L'annotation d'une image

La figure 43 montre le cas d'annotation d'image requête, après la recherche et l'affichage des images les plus similaires à l'image requête. Cette dernière peut être annotée par l'utilisateur à l'aide d'une liste de suggestions, qui a été construite à partir des annotations des images résultante.

L'image requête est une image d'un téléthorax (la classe n°12), les images résultats correspondent aux 10 images les plus similaires à la requête, qui sont triées et affichées, selon le score, de gauche à droite et de haut en bas.

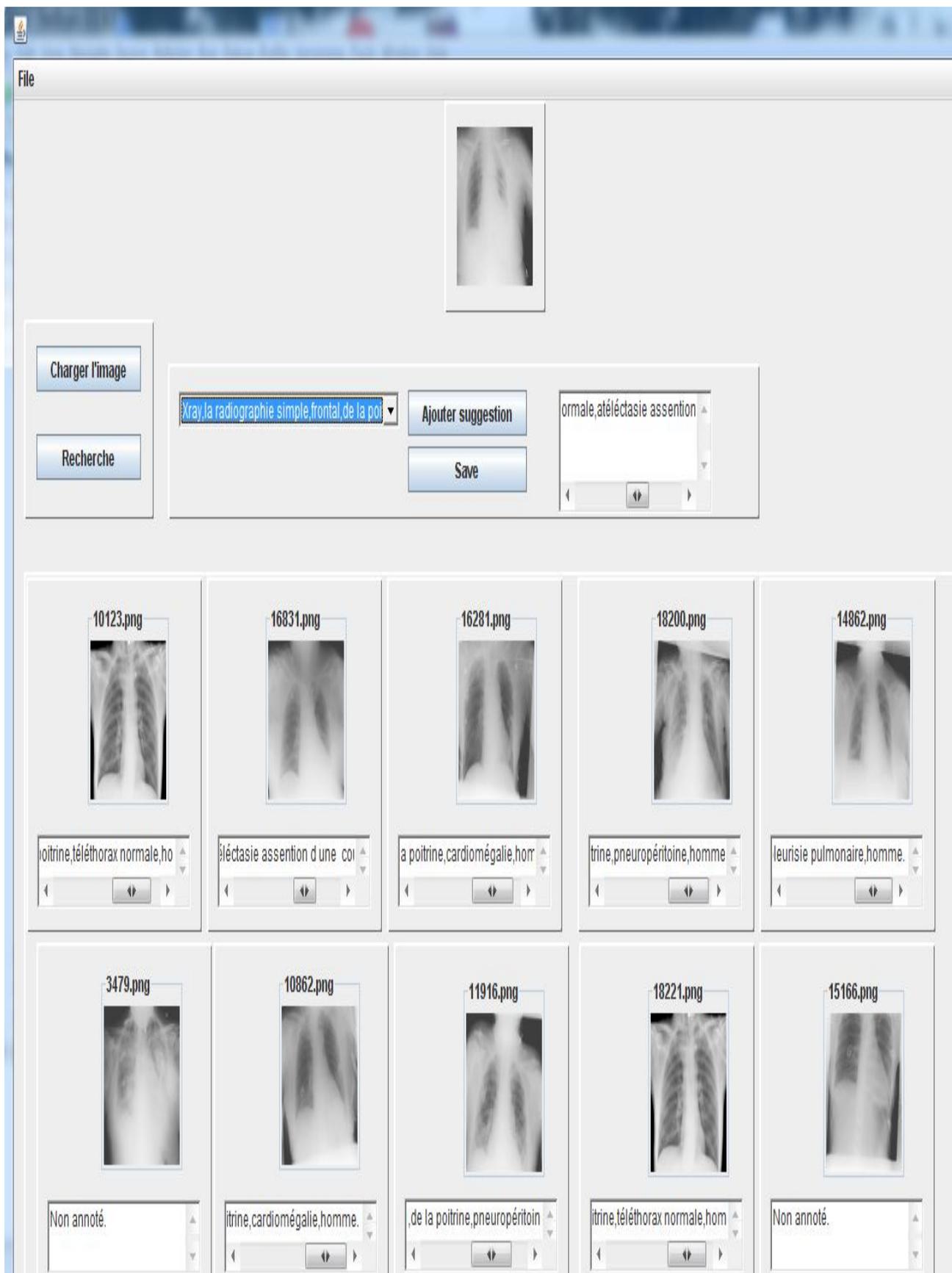


Figure 43 les images les plus proches à la requête

Conclusion

Nous avons présenté dans ce chapitre notre contribution à la recherche d'image médical par le contenu. Ce travail est appliqué à la base d'image ImageCLEFmed 2005. Nous y avons développé les principales étapes d'un processus de recherche d'image, à savoir l'extraction des caractéristiques et la comparaison des images et ensuite l'annotation manuelle. La première étape a été de définir des descripteurs adaptés aux images testées. Pour cela, nous avons utilisé les descripteurs de texture et de forme. Ensuite, le système de classification SVM a été utilisé afin d'associer l'image requête à une classe spécifique. Et enfin une simple annotation a été ajoutée à l'image pour minimiser l'écart sémantique entre le contenu formel et informel de l'image.

Chapitre 5 : Validation et analyse des résultats

Dans ce chapitre nous allons tester le système de recherche que nous avons proposé en utilisant la base ImageCLEF2005, et cela afin de mettre en évidence le degré de sa pertinence et son efficacité.

Introduction

Nous avons vu dans le chapitre précédent les détails des méthodes utilisées dans notre système. Dans ce chapitre, nous allons les tester afin de mettre en évidence leurs degrés d'efficacité, et la pertinence des descripteurs utilisés. 52 catégories sont testés, puis 21 catégories et en fin 15 catégories.

I .Analyse des résultats

I.1.Première expérimentation

En premier lieu nous avons choisi 52 classes différentes (les détails des catégories sont décrits dans le tableau suivant), avec une base d'apprentissage de 8569 images et une base de test de 995 images.

Nous avons choisi les caractéristiques de la matrice de cooccurrence (décrites dans le chapitre précédent), et aussi la surface de la forme, et le contour extraits à partir d'une image segmentée par le contour actif. Ce qui donne un vecteur de caractéristique de taille (43) éléments .Nous avons obtenus un taux de classification de 55.6783%.

N° classe	Nom de la classe dans Image CLEF 2005	Anatomique	Direction
1	1	crâne, le système musculo-squelettique	Coronaire
2	2	crâne, crâne facial, le système musculo-squelettique	Coronaire
3	3	la colonne vertébrale, du rachis cervical, système musculo-squelettique	Coronaire
4	4	colonne vertébrale, épine dorsale, système musculo-squelettique	Coronaire
5	5	la colonne vertébrale, du	Coronaire

		rachis lombaire, système musculo-squelettique	
6	6	membre supérieur (bras), de la main, système musculo-squelettique	Coronaire
7	7	membre supérieur (bras), articulation radio carpien, système musculo-squelettique	Coronaire
8	8	membre supérieur (bras), l'avant-bras à la main, système musculo-squelettique	Coronaire
9	9	membre supérieur (bras), le coude, système musculo-squelettique	coronaire
10	10	membre supérieur (bras), le bras, système musculo-squelettique	Coronaire
11	11	membre supérieur (bras), de l'épaule, système musculo-squelettique	Coronaire
12	12	poitrine	Coronaire
13	13	poitrine, des os, système musculo-squelettique	Coronaire
14	14	abdomen, système gastro-intestinal	Coronaire
15	15	abdomen, système uropoietic	Coronaire
16	16	l'abdomen, l'abdomen supérieur, système gastro-intestinal	Coronaire
17	17	bassin, système musculo-squelettique	Coronaire
18	18	membres inférieurs (jambes),	Coronaire

		le pied, système musculo-squelettique	
19	19	membres inférieurs (jambes), articulation de la cheville, système musculo-squelettique	Coronaire
20	20	membres inférieurs (jambes), inférieure de la jambe, système musculo-squelettique	Coronaire
21	21	membres inférieurs (jambes), du genou, système musculo-squelettique	Coronaire
22	22	membres inférieurs (jambes), la jambe supérieure, système musculo-squelettique	Coronaire
23	23	membres inférieurs (jambes), de la hanche, système musculo-squelettique	Coronaire
24	24	crâne, crâne facial, système musculo-squelettique	Sagittal
25	25	crâne, crâne neuro, système musculo-squelettique	Sagittal
26	26	la colonne vertébrale, du rachis cervical, système musculo-squelettique	Sagittal
27	27	colonne vertébrale, épine dorsale, système musculo-squelettique	Sagittal
28	28	la colonne vertébrale, du rachis lombaire, système musculo-squelettique	Sagittal
29	29	membre supérieur (bras), de la main, système musculo-	Sagittal

		squelettique	
30	30	membre supérieur (bras), articulation radio carpien, système musculo- squelettique	Sagittal
31	31	membre supérieur (bras), avant-bras, système musculo- squelettique	Sagittal
32	32	membre supérieur (bras), le coude, système musculo- squelettique	Sagittal
33	33	membre supérieur (bras), de l'épaule, système musculo- squelettique	Sagittal
34	34	poitrine	Sagittal
35	35	membres inférieurs (jambes), le pied, système musculo- squelettique	Sagittal
36	36	membres inférieurs (jambes), articulation de la cheville, système musculo-squelettique	Sagittal
37	37	membres inférieurs (jambes), inférieure de la jambe, système musculo-squelettique	Sagittal
38	38	membres inférieurs (jambes), du genou, système musculo- squelettique	Sagittal
39	39	membres inférieurs (jambes), la jambe supérieure, système musculo-squelettique	Sagittal
40	41	du sein (maman), poitrine droite, le système	Axial+ Autre orientation

		reproducteur	
41	42	du sein (maman), le sein gauche, le système reproducteur	Axial+ Autre orientation
42	43	membres inférieurs (jambes), du genou, système musculo-squelettique	Axial
43	45	crâne, crâne neuro, système musculo-squelettique	Autre orientation
44	47	membre supérieur (bras), de la main, système musculo-squelettique	Autre orientation
45	50	membres inférieurs (jambes), le pied, système musculo-squelettique	Autre orientation
46	51	thorax, hile, les voies respiratoires	Coronaire
47	52	l'abdomen, l'abdomen supérieur, système gastro-intestinal	Coronaire
48	53	bassin, du système cardiovasculaire	Coronaire
49	54	membres inférieurs (jambes), la jambe, le système cardiovasculaire	Coronaire
50	55	membres inférieurs (jambes), le genou, le système cardiovasculaire	Coronaire
51	56	membres inférieurs (jambes), la cuisse, le système cardiovasculaire	Coronaire
52	57	bassin, du système	Coronaire

		cardiovasculaire	
--	--	------------------	--

Tableau 2 Les classes d'images de rayons X: le numéro de la classe dans image Clef anatomique, l'angle

Les 52 catégories sont représentées dans la figure 44



Figure 44 Exemple d'images des 52 classes utilisées

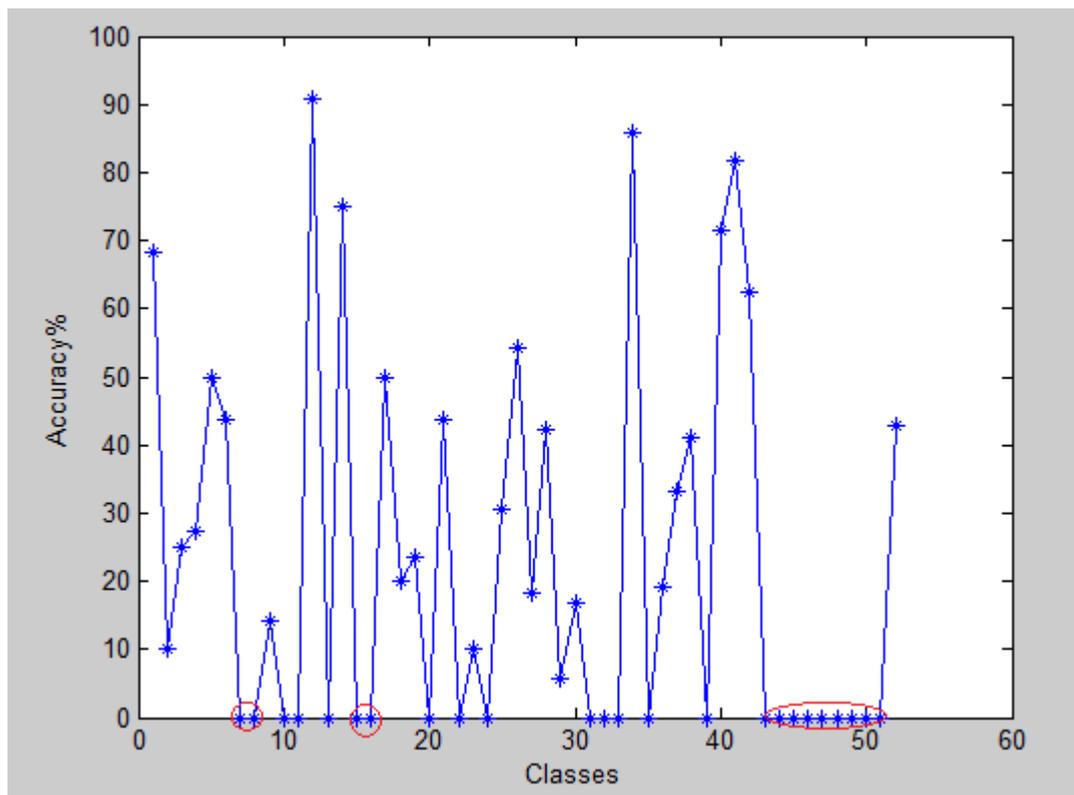


Figure 45 Les résultats de classification de 52 catégories.

I.1.1. Interprétation

- Les classes ayant un ensemble d'apprentissage grand, ont eu un taux de classification élevé.

Exemples :

Pour la classe n°1, l'ensemble d'apprentissage est composé de 336 images, et le taux de classification obtenu est de 68.421%.

Pour la classe n°34, l'ensemble d'apprentissage est composé de 880 images, et le taux de classification obtenu est de 85.8974%.

Pour la classe n°12, l'ensemble d'apprentissage est composé de 2563, et le taux de classification obtenu est 90.8784%.

- Les classes ayant un ensemble d'apprentissage petit, n'ont pas été détectées.

Exemple :

La classe n°43, l'ensemble d'apprentissage est composé de 9 images, et le taux de classification obtenu est de 0%.

La classe n°47, l'ensemble d'apprentissage est composé de 9 images, et le taux de classification obtenu est de 0%.

La classe n°50, l'ensemble d'apprentissage est composé de 10 images, et le taux de classification obtenu est de 0%.

- Les classes ayant un ensemble de test petit, n'ont pas été détectées.

Exemple

Les classes n°46, n°47, n°48, n°49, n°50, n°51 l'ensemble de test contient moins de 7 images.

- Des ambiguïtés dans la classification pour des classes ayant la même région de corps comme la classe crâne de système musculo-squelettique avec une vue coronaire et la classe de crâne facial de système musculo-squelettique d'une vue coronaire.

I.1.2.Comparaison

Dans [8] les auteurs proposent une méthode de classification hiérarchique comportant deux niveaux ; en utilisant un ensemble de caractéristiques de formes et de texture.

La base de données utilisée pour tester cette méthode est une collection de plus de 9100 images radiologiques (X-ray) comprenant 40 différentes classes (les détails des catégories sont décrits dans le tableau suivant).

Elle a été divisée en deux ensembles : pour l'apprentissage 7861 images et le test 1239 images.

N° classe	Anatomique	Direction
1	Crâne	Coronaire
2	Rachis cervical	Coronaire
3	Rachis lombaire	Coronaire
4	Main	Coronaire
5	Articulation radio carpien	Coronaire
6	Coude	Coronaire
7	Epaule	Coronaire
8	Poitrine	Coronaire
9	Os poitrine	Coronaire
10	Abdomen gastro-système	Coronaire

11	Abdomen-uropoietic système	Coronaire
12	Abdomen supérieur	Coronaire
13	Bassin	Coronaire
14	Pied	Coronaire
15	Articulation de la cheville	Coronaire
16	Inférieure de la jambe	Coronaire
17	Genou	Coronal posterior
18	Cuisse	Coronaire
19	Hanche	Coronaire
20	Crâne	Sagittal
21	Rachis cervical	Sagittal
22	Colonne thoracique	Sagittal
23	Articulation radio carpien	Sagittal
24	Avant bras	Sagittal
25	Coude	Sagittal
26	Poitrine	Sagittal
27	Articulation de la cheville	Sagittal
28	Genou	Sagittal
29	Sein droit	Axial
30	Sein gauche	Axial
31	Genou	Axial
32	Crâne facial	Autre orientation
33	Crâne	Autre orientation
34	Main	Autre orientation
35	Sein droit	Autre orientation
36	Sein gauche	Autre orientation
37	Abdomen supérieur	Coronaire
38	Inférieure de la jambe	Coronaire
39	Genou	Coronaire
40	Bassin du système cardiovasculaire	Coronaire

Tableau 3 Les classes d'images de rayons X: anatomique, l'angle

Les caractéristiques suivantes ont utilisé : Excentricité, les longueurs des axes majeurs et mineurs (O), le descripteur de la distance Fourier pour la représentation de centroïde (FD), le descripteur de Fourier complexe (FZ), les moments invariants (I), tessellation à base spectrale (S), l'histogramme directionnelle (D) et le contraste, la corrélation, l'énergie et l'homogénéité (T).

Ensuite un perceptron multicouche (MLP) a été utilisé. Les résultats de classification obtenus par les auteurs de cette méthode par l'utilisation de 40-classe et les différentes caractéristiques sont montrés dans le tableau suivant :

Espace des caractéristiques	La taille de vecteur	Précision
O	4	29
FD	127	18
FZ	254	42
I	5	4
S	16	39
D	36	36
T	16	22
FZ, O	258	57
FZ, S	270	51
FZ, D	290	52
FZ, T	270	44
FZ, I, O, S	279	58
FZ, I, O, T	279	54
FZ, I, O, D	299	61
FZ, I, O, D, S	331	64

Tableau 4 Les résultats de classification basé sur la base de test

Les auteurs ont utilisés une seule caractéristique à la fois et puisque le taux de classification de l'espace des caractéristiques FZ est supérieure au reste des caractéristiques, FZ a été est inclut dans le vecteur de caractéristique optimale.

L'espace des caractéristiques (FZ, I, O, D, S) offre un grand taux de classification 64%.

Un tableau comparatif est représenté par la suite :

Anatomique	Notre méthode	La méthode [8]
Crâne(Coronaire)	68.421	80
Rachis cervical	25	98
Rachis lombaire	50	90
Main	43.6364	45
Articulation radio carpien	0	32
Coude	14.2857	15
Epaule	0	21
Poitrine	90.8784	90
Os poitrine	0	75

Abdomen gastro-système	75	50
Abdomen-uropoietic système	0	32
Abdomen supérieur	0	60
Bassin (coronaire)	50	80
Pied	20	59
Articulation de la cheville	23.5294	72
Inférieure de la jambe	10	55
Genou (sagittal)	43.7500	72
Cuisse	0	21
Hanche	10	8
Crâne (autre orientation)	30.5556	62
Rachis cervical	54.1667	65
Colonne thoracique	85.8974	50
Articulation radio carpien	16.6667	40
Avant bras	0	41
Coude	0	51
Poitrine	85.8974	80
Sein droit	71.4286	75
Sein gauche	81.8182	80
Genou (coronaire)	62.5000	0
Abdomen supérieur (sagittal)	33.3333	100
Bassin du système cardiovasculaire	42.8571	91

Tableau 5 Un tableau comparatif entre les deux méthodes

Dans ImageCLEF 2005 l'ensemble de test ne dépasse pas 1000 images et la base d'apprentissage de 9000 images ; les auteurs dans [8] ont utilisés une autre décomposition de la base pour faire l'apprentissage et le test ; par contre nous avons utilisé la taille des ensembles originaux.

La taille de descripteur dans [8] est de 331 ; alors que la taille de descripteur utilisé dans notre méthode est de 43, ce descripteur a donné des bons résultats dans certain classe par contre il n'a pas détecté autre ; le taux maximum été de 90.87%. Le descripteur proposé par les auteurs [8] n'a pas pu détecter qu'une seule classe ce qui montre la qualité des caractéristiques utilisées et même le choix de la taille de l'ensemble d'apprentissage et de teste.

I.2. Deuxième expérimentation :

Nous avons choisi 21 classes différentes (voir le tableau 6). Avec une base d'apprentissage de 6916 images et une base de test de 782 images et un vecteur de caractéristique de taille (43) contenant les caractéristiques de la matrice de cooccurrence et de la forme.

N° Classe	N° classe dans Image CLEF 2005
1	1
2	3
3	4
4	5
5	6
6	12
7	14
8	17
9	18
10	19
11	21
12	25
13	26
14	28
15	34
16	37
17	38
18	41
19	42
20	43
21	57

Tableau 6 les numéros des classes



Figure 46 Exemple d'images des 21 classes utilisées

Nous avons obtenu lors de cette expérimentation un taux de classification moyen de 69.8210%.

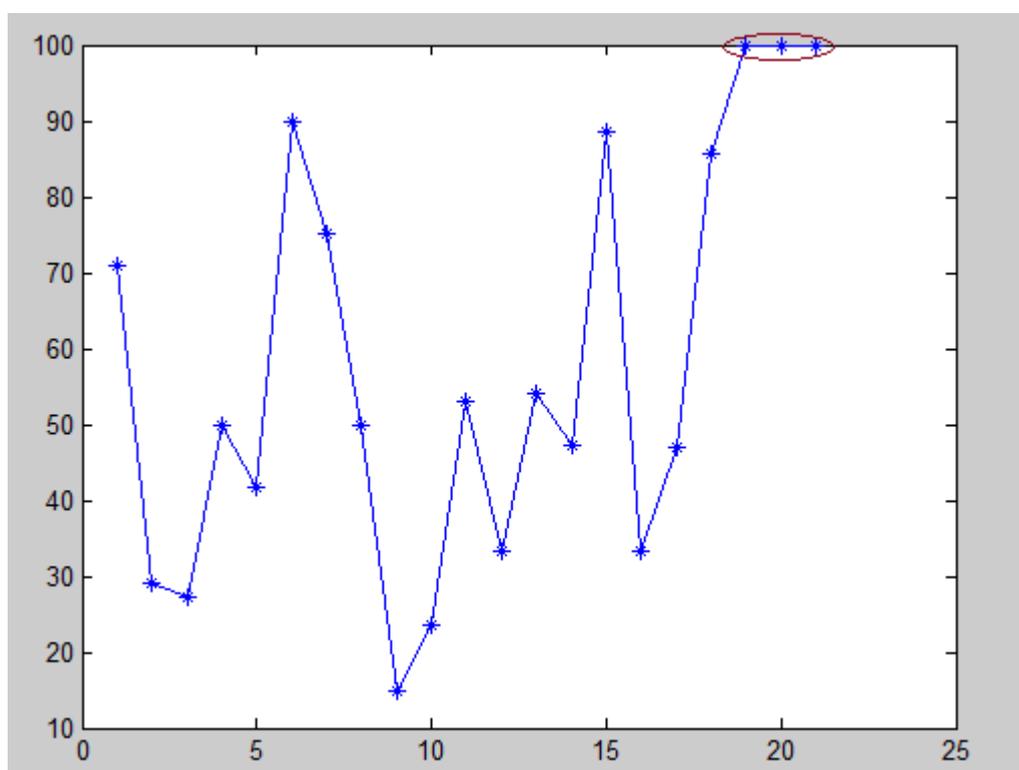


Figure 47 Les résultats de classification de 21 catégories.

I.2.1. Interprétation

- Le faible taux de classification a été pour la classe n°9, la taille de l'ensemble de test pour cette classe est de 20 images.

45% des images de cette classe ont été classées comme la classe n°20.



Figure 48 Deux images de la base de test pour la classe n°9

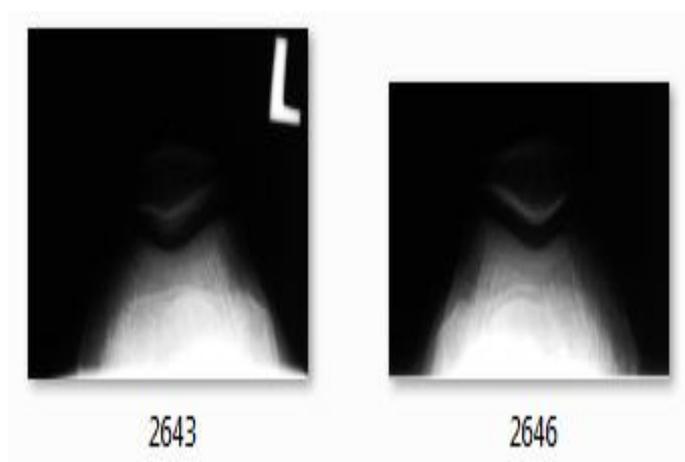


Figure 49 Deux images de la base d'apprentissage pour la classe n°20

Nous remarquons que le contraste de l'image a beaucoup influencé sur le taux de classification.

- Deux classes ayant la même texture et la même forme ont un taux de classification de 100% ce qui montre l'efficacité de la caractéristique de controiide.

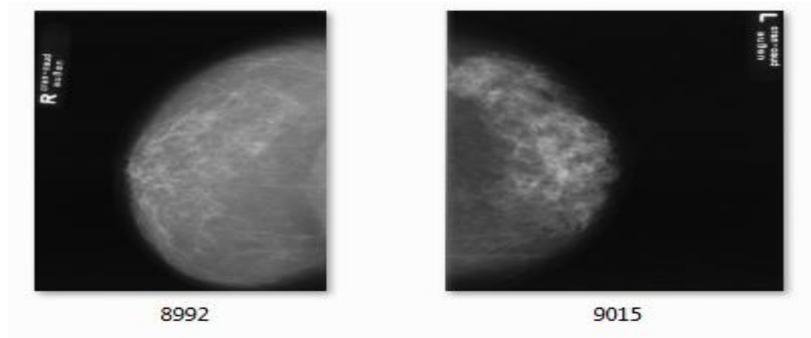


Figure 50 Deux images de la base d'apprentissage pour les classes n°19 et n°20

I.2.2.Mesure de la qualité des réponses

I.2.2.1.Le rappel (Recall)

Le rappel est le rapport entre le nombre d'images pertinentes dans l'ensemble des images trouvées et le nombre d'images pertinentes dans la base d'image.

$$\text{Rappel} = \frac{|Ra|}{|R|} \quad \text{Équation 100}$$

I.2.2.2.La précision (precision)

La précision est le rapport entre le nombre d'images pertinentes dans l'ensemble des images trouvées et le nombre d'images trouvées.

$$\text{Précision} = \frac{|Ra|}{|A|} \quad \text{Équation 101}$$

Où R est l'ensemble d'images pertinentes dans la base d'images utilisée (R est l'ensemble des images similaires à la requête), $|R|$ est le nombre d'images pertinentes dans la base d'images. A est l'ensemble des réponses, et $|R_a|$ est le nombre d'images pertinentes dans l'ensemble des réponses.

Par exemple, si la base contient 50 image d'une catégorie jardin, et si le système renvoie 5 bonne réponses parmi les 10 images qu'il peut afficher, le rappel vaut $(5/50)=0,1$ et la précision $(5/10)=0,5$.

Les deux métriques rappel et précision s'utilisent conjointement pour l'évaluation des performances des systèmes de recherche d'information et varient inversement : lorsque la précision diminue, le rappel augmente et réciproquement. Les valeurs de ces deux métriques reflètent le point de vue de l'utilisateur :

- Si le rappel est faible, une partie de l'information pertinente ne lui sera pas accessible.
- Si la précision est faible, l'utilisateur ne sera pas satisfait à cause de la forte concentration des informations non-pertinentes fournies dans les résultats.

Dans les deux cas, le système ne répond pas aux attentes des utilisateurs à retourner l'information utile et pertinente et par la suite, il est non-performant. Le cas idéal est d'avoir la valeur de précision et rappel respectivement égale à un, chose qui n'est pas atteinte en réalité.

Une des manières de tenir compte à la fois du rappel et de la précision d'un système est d'exprimer les valeurs de précision en fonction des différents niveaux de rappel selon la courbe rappel/précision.

1.2.2.3. La courbe rappel/précision

L'évaluation des performances d'un système CBIR est une tâche ardue, car elle dépend étroitement de la perception visuelle de l'humain, du domaine d'application et du contenu de la base elle-même. Nous proposons d'utiliser la courbe de précision/rappel, pour évaluer la performance de notre système. Il s'agit d'un calcul statistique sur la base de données mettant aussi en évidence la capacité de la méthode étudiée à retrouver les classes d'images.

Cependant, cet outil nécessite un grand nombre de requêtes pour évaluer un système. La courbe rappel/précision permet de suivre la qualité du résultat en fonction du nombre d'images retournées par le système en réponse à une requête. En pratique, cette courbe correspondant à une recherche idéale a l'allure générale de la figure 51.

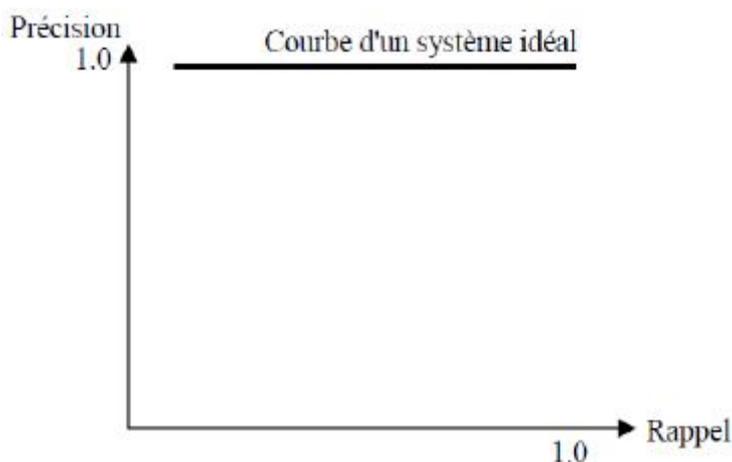


Figure 51 Allure d'une courbe de rappel-précision

Nous présentons dans cette partie quelques courbes de rappel/précision, afin d'avoir une représentation visuelle globale de la qualité de descripteur utilisé.

Nous avons construit une base de tests utilisant 20 images de chaque classe.

Considérons :

- Un vecteur de caractéristique de taille (43) avec le descripteur de texture est la matrice de co-occurrence.
- La distance euclidienne pour calculer la similarité entre les descripteurs.

Pour l'image requête de la classe n°15 ; la liste des réponses du système est :

{image1, image2, image3, image4, image5, image6, image7, image8, image9, image10}

Image	La classe	Pertinent	Précision	Rappel
Image1	n°13	NP	0	0
Image2	n°15	P	0.5	0.05
Image3	n°06	NP	0.33	0.05

Image4	n°15	P	0.5	0.1
Image5	n°15	P	0.6	0.15
Image6	n°20	NP	0.5	0.15
Image7	n°15	P	0.57	0.2
Image8	n°15	P	0.625	0.25
Image9	n°15	P	0.666	0.3
Image10	n°15	P	0.7	0.35

Tableau 7 Des valeurs de rappel-précision pour une image de classe 15

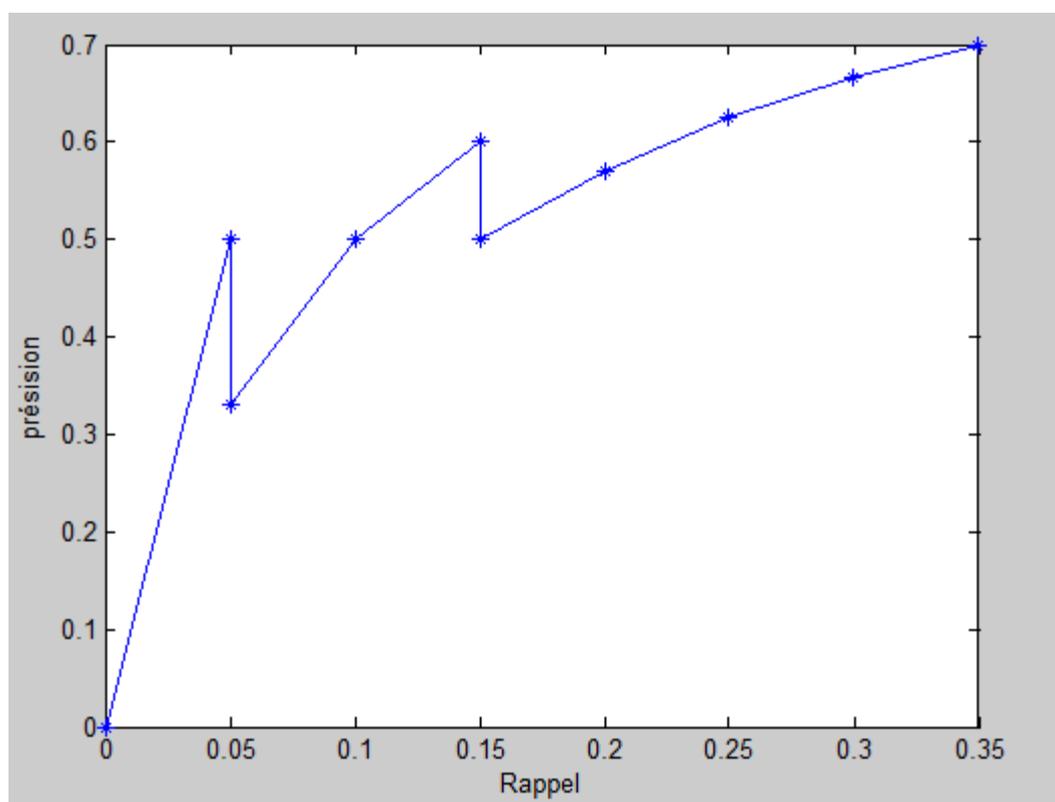


Figure 52 Courbe de rappel et précision pour une image requête de la classe 15

Maintenant une autre image requête de la classe 19. Avec les mêmes conditions.

Soit la liste des réponses du système :

{image1, image2, image3, image4, image5, image6, image7, image8, image9, image10}

Image	La classe	Pertinent	Précision	Rappel
Image1	n°19	P	1	0.05
Image2	n°19	P	1	0.1
Image3	n°19	P	1	0.15
Image4	n°19	P	1	0.2
Image5	n°19	P	1	0.25
Image6	n°18	NP	0.83	0.25
Image7	n°04	NP	0.71	0.25
Image8	n°19	P	0.75	0.3
Image9	n°08	NP	0.66	0.3
Image10	n°04	NP	0.6	0.3

Tableau 8 Des valeurs de rappel-précision pour une image de classe 19

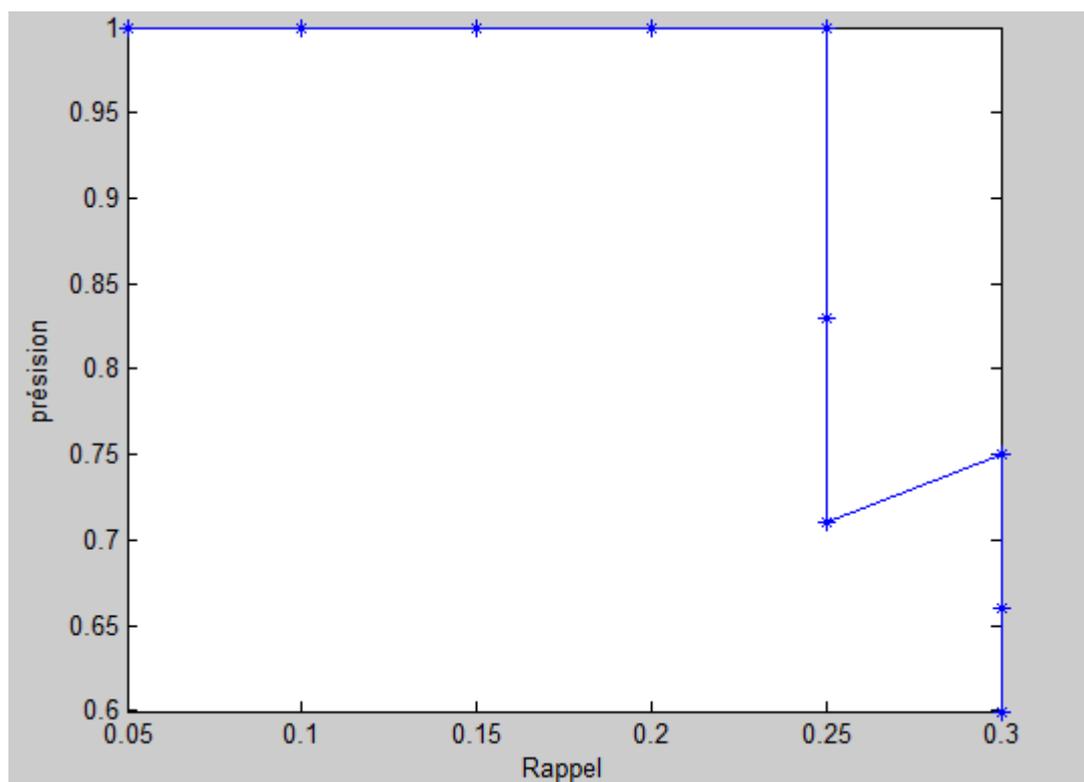


Figure 53 Courbe de rappel et précision pour une image requête de la classe 19

Le tableau suivant montre la moyenne des deux requêtes :

Image	Précision	Rappel
Image1	0.5	0.025
Image2	0.75	0.075
Image3	0.66	0.1
Image4	0.75	0.15
Image5	0.8	0.2
Image6	0.66	0.2
Image7	0.64	0.22
Image8	0.68	0.27
Image9	0.66	0.3
Image10	0.65	0.32

Tableau 9 la moyenne de la précision et le rappel pour les deux requêtes

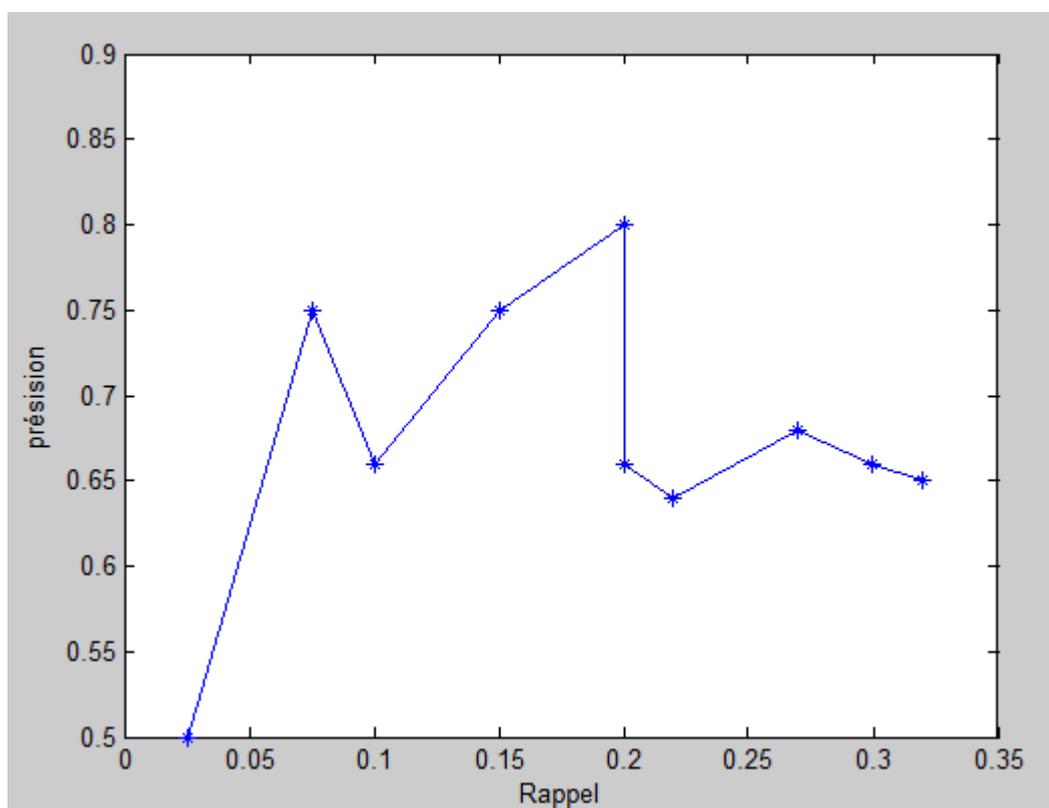


Figure 54 la courbre moyenne de rappel et précision

I.3. Troisième expérimentation :

Nous avons choisi pour cette expérimentation 15 classes différentes (voir le tableau 10). Avec une base d'apprentissage de 5951 images et une base de test de 671 images et un vecteur de caractéristique de taille (43) contenant les caractéristiques de la matrice de cooccurrence et de la forme.

N° Classe	N° classe dans Image CLEF 2005
1	1
2	5
3	6
4	12
5	14
6	17
7	21
8	26
9	28
10	34
11	38
12	41
13	42
14	43
15	57

Tableau 10 les numéros des classes

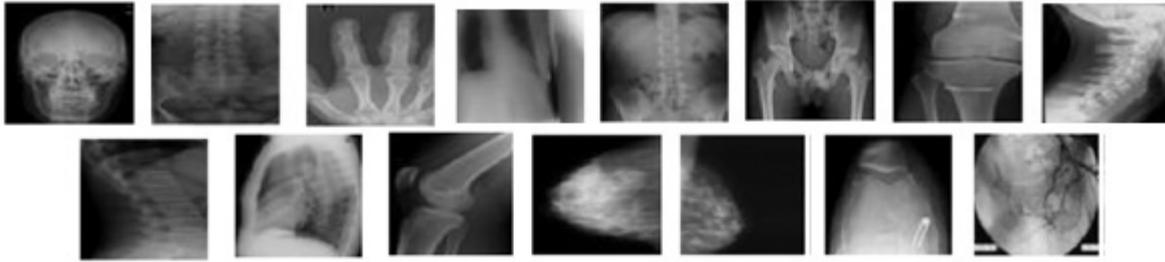


Figure 55 Exemple d'images des 15 classes utilisées

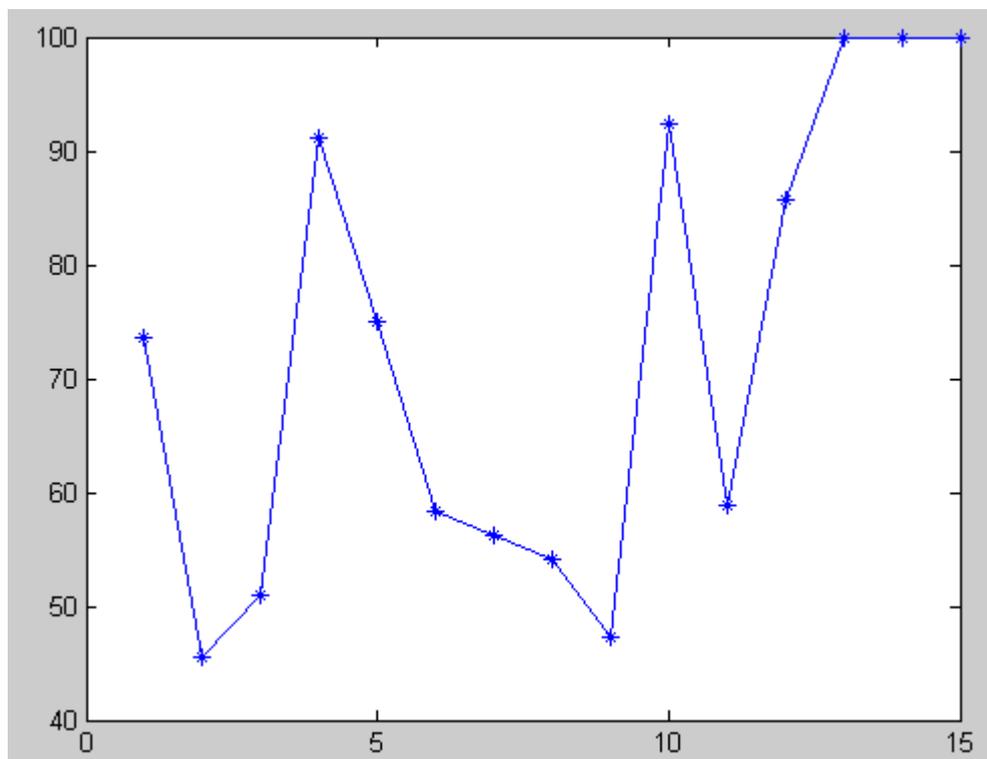


Figure 56 Les résultats de classification de 15 catégories par la matrice de co-occurrence

On a obtenu un taux de classification de 79.4337%.

I.3.1. Interprétation

D'après les résultats obtenues nous pouvons constater que :

- 53.33% des classes ont un taux de classification >70%.
- 33.33% des classes ont un taux de classification >50%.
- 13.33% des classes ayant un taux de classification <50%.
- On remarque que la classe n°2 a le plus faible taux de classification : 45.4545%.
- Le meilleur taux a été atteint pour les classes n°13, 14,15 de 100%.

- La base d'apprentissage pour 15 catégories représente la majorité des images de la base (5951/9000). Ce qui est le cas aussi pour la base de test (671/1000).

I.3.2. Comparaison

Dans [6] les auteurs ont utilisé un apprentissage à noyau multiple : (Multiple Kernel Learning « MKL ») qui est une méthode permettant d'apprendre la combinaison optimale d'un ensemble de noyaux pour une classification en fonction de SVM.

Trois caractéristiques globales ont été utilisées : la matrice de cooccurrence, les caractéristiques de Gabor, les caractéristiques de la texture Tamura; et une caractéristique locale est ModSIFT. La méthode proposée a été testée sur des images radiologiques pour 14 classes. 100 images ont été sélectionnées de façon aléatoire pour chaque catégorie afin de construire un ensemble de données de 1400 images.

La combinaison des caractéristiques de Gabor et modSIFT avec MKL a donné la meilleure performance parmi toutes les combinaisons. Pour un vecteur de 620 la précision a été 96.68%.

II .Mesure de la qualité des descripteurs

Afin de mesurer la qualité des descripteurs par l'utilisation des métriques de rappel et de précision, nous avons construit une base d'exemple de 20 images pour 15 classes ; et à partir de l'ensemble de test de chaque classe on cherche l'image ayant le taux le plus élevé de précision.

Considérons :

- La distance euclidienne pour calculer la similarité entre les descripteurs.
- 10 images comme des réponses de système.

Pour un vecteur de caractéristiques qui contient la combinaison suivante :

1. la matrice de cooccurrence, l'ondelette de Gabor et les caractéristiques de forme.
2. la matrice de cooccurrence et l'ondelette de Gabor.
3. la matrice de cooccurrence.

Le même tableau est obtenu pour les trois descripteurs :

Image requête de Classe n°	précision	rappel
1	0.6	0.3
2	0.5	0.25
3	0.2	0.1
4	0.4	0.2
5	0.3	0.15
6	0.4	0.2
7	0.2	0.1
8	0.7	0.35
9	0.4	0.2
10	0.7	0.35
11	0.2	0.1
12	0.1	0.05
13	0.6	0.3
14	0.2	0.1
15	0.3	0.15

Tableau 11 le rappel et la précision pour les descripteurs de 1cas

On remarque que les mêmes images de chaque classe ont été sélectionné par les différents descripteurs ce qui implique que les valeurs des caractéristiques de la matrice de cooccurrence domine dans la discrimination, c'est parce que la métrique de mesure de similarité été la distance euclidienne.

Pour un vecteur de caractéristiques qui contient la combinaison suivante

1. L'ondelette de Gabor et les caractéristiques de forme.
2. Les descripteurs de forme.

Le même tableau est obtenu pour les deux descripteurs :

Image requête de Classe n°	précision	rappel
1	0.40	0.2
2	0.3	0.15

3	0.4	0.2
4	0.6	0.3
5	0.4	0.2
6	0.3	0.15
7	0.5	0.25
8	0.4	0.2
9	0.4	0.2
10	0.5	0.25
11	0.3	0.15
12	0.6	0.3
13	0.8	0.4
14	0.5	0.25
15	0.7	0.35

Tableau 12 le rappel et la précision pour les descripteurs de 2 cas

On remarque que les descripteurs de forme permettent une bonne discrimination pour les classes n° 13 et 15 ce qui est pas le cas pour la matrice de coocurrence ; et aussi les caractéristiques de Gabor ont les plus petites valeurs.

Pour un vecteur de caractéristiques qui contient juste les caractéristiques de l'odelette de Gabor, on a le résultat suivant :

Image requête de Classe n°	précision	rappel
1	1	0.5
2	1	0.5
3	0.6	0.3
4	0.9	0.45
5	0.9	0.45
6	1	0.5
7	0.6	0.3
8	0.6	0.3
9	0.7	0.35

10	0.7	0.35
11	0.5	0.25
12	0.5	0.25
13	0.5	0.25
14	0.5	0.25
15	0.6	0.3

Tableau 13 le rappel et la précision pour les descripteurs de 3 cas

Dans ce descripteur on a obtenu les images ayant la moyenne de précision la plus élevée.

Le tableau 14 monte le taux de classification avec les différents descripteurs utilisés :

Nombre de catégorie	Matrice de co-occurrence		Gabor		Combinaison	
	Taux de classification	Taux d'erreur	Taux de classification	Taux d'erreur	Taux de classification	Taux d'erreur
14	81.7073%	18.2927%	63.8211%	36.1789%	60.7724%	39.2276%
15	79.4337%	20.5663%	39.0462%	60.9538%	47.2429%	52.7571%

Tableau 14 Le taux de classification pour chaque descripteur

A partir de ce tableau on remarque que la matrice de co-occurrence donne des bons résultats par rapport au descripteur de Gabor et la combinaison des deux.

Conclusion

Dans ce chapitre, trois expérimentations sont décrites. La première pour un ensemble de 52 catégories. La deuxième pour 21, et la dernière pour 15 catégories. Pour les trois expérimentations nous avons utilisé le même descripteur qui regroupe les caractéristiques de la matrice de cooccurrence et de forme. Cependant, les résultats de la classification ont montré qu'un descripteur contenant les caractéristiques de la matrice de cooccurrence, l'ondelette de Gabor et de la forme d'une taille 73, n'a pas donnée des bons résultats, et que la discrimination d'un descripteur de taille 43 contenant juste les caractéristiques de la matrice de cooccurrence et de la forme a été plus efficace.

Conclusion générale

Dans ce mémoire, nous avons abordé le problème de la recherche d'images médicale. Plus précisément, nous nous sommes focalisés sur la recherche d'images basée sur le contenu visuel. Notre choix a été motivé par la quantité d'images radiologique disponible aujourd'hui, qui ne cesse de croître.

Le but de ce travail n'a pas été de proposer une nouvelle définition, néanmoins l'idée a été d'extraire à partir des images, une certaine information pertinente de la texture et de la forme. Ces informations permettent une recherche efficace. Pour atteindre cet objectif, un système d'indexation et de recherche d'image médicale par contenu a été proposé. Ce système fonctionne avec des descripteurs visuels qui sont extraits d'une image, telle que les caractéristiques de la matrice de co-occurrence, et les caractéristiques de la forme. Nous avons aussi développé une interface qui permet à l'utilisateur de proposer facilement une requête et de visualiser les résultats. Un module d'annotation manuelle de l'image requête a été également développé pour combler le fossé sémantique qui sépare les descripteurs de bas niveau extraits du contenu des images et la sémantique que l'utilisateur place dans sa requête.

Notre système a été testé par l'utilisation de la base ImageCLEF2005. Pour évaluer la qualité des descripteurs, nous avons calculé les courbes de rappel et précision pour des images de test. Nous avons également proposé et testé un système de classification par le classifieur SVM. Le taux de classification obtenu par la combinaison des caractéristiques de la matrice de co-occurrence et la forme a été satisfaisant comparé à la combinaison de tous les descripteurs. Néanmoins, les caractéristiques d'ondelette de Gabor avec utilisation de la distance euclidienne comme métrique pour la recherche d'image ont donné de bons résultats. A ce stade, plusieurs problèmes d'actualité complexes et incontournables ont été soulevés :

- Le premier se pose sur l'extraction et le choix des descripteurs visuels.
- Le deuxième se porte sur la malédiction de la dimension de l'espace de vecteurs de caractéristiques souvent de grande taille.

A l'issue des travaux menés dans le cadre de ce mémoire, nous dégagons quelques perspectives :

- Nous envisageons d'enrichir notre système en utilisant d'autres caractéristiques visuelles afin d'augmenter la collection des descripteurs utilisables.
- La seconde amélioration intervient dans la phase d'annotation où nous prévoyons de combiner les approches d'annotation automatique afin d'améliorer l'interprétation sémantique et pour que le système devienne plus indépendant.
- Nous proposons également d'étudier et de tester les différentes mesures de similarité afin de les comparer et de pouvoir sélectionner celles qui ont la meilleure performance.

Bibliographie

- [1] Anil K. Jain Tuceryan M, «The Handbook of Pattern Recognition and Computer Vision», 1998.
- [2] Shanmugam K, Dinstein I Haralick R.M, « Textural Features for Image Classification », November 1973.
- [3] Qian Y , Hui , Loomes M , Comley R , Barn B , Chapman A and Rix J Gao X, «TEXTURE-BASED 3D IMAGE RETRIEVAL FOR MEDICAL APPLICATIONSIADIS », 2010.
- [4] Zhu H Zhang P, « Medical Image Retrieval Based on Co-Occurrence Martix and Edge Histogram», 2011.
- [5] Chandran K. R, Aishwarya S., Janaranjani P Ramamurthy B, «CBMIR: Content Based Image Retrieval using Invariant Moments, GLCM and Grayscale Resolution for Medical », 2011.
- [6] Hao Z, Chao L Hong W, «Medical Image Classification with Multiple Kernel Learning» , 2010.
- [7] Weirong M, Zanchao Z Shunren X, «A Content-Based Retrieval System for Endoscopic Images», 2005.
- [8] Ghassemian H Pourghassem H, « Content-based medical image classification using a new hierarchical merging scheme », 2008.
- [9] William H Jayashree Kalpathy C, «Effectiveness of global features for automatic medical image classification and retrieval – The experiences of OHSU at ImageCLEFmed », 2008.
- [10] Cardoso de Paiva A, de Souza Baptista C. Leonardo S. Kury A. Mendes, «Content Based Mammography Images Retrieval using Ripley's K Function», 2009.
- [11] Fatemizadeh E Rafi Nazari M, «A CBIR System for Human Brain Magnetic Resonance Image Indexing », 2010.
- [12] Sameer K. Antani , George R. Thoma Rahman M, «A Classification-Driven Similarity Matching Framework for Retrieval of Biomedical Images» , 2010.
- [13] Dacher J , Darmoni S Bensrhair A, « Recherche des images médicales par contenu et type de ressource dans un catalogue de santé en ligne» , mai 12-13, 2005.

- [14] Burel G, « Introduction au traitement d'image simulation sous matlab », 2001.
- [15] Montagnat J, Magnin E Glatard T, «Texture Based Medical Image Indexing and Retrieval: Application to Cardiac Imaging», 2004.
- [16] Darwish A , Mohamed R.A Youssif A, «Content based medical image retrieval based on pyramid structure wavelet », 2010.
- [17] Ma Z.M, Deng L Zhang G, « Texture Feature Extraction and Description Using Fuzzy Set of Main Dominant Directions of Variable Scales in Content-Based Medical Image Retrieval».
- [18] JIANKE Z, MICHAEL R. LYU STEVEN C. H. HOI, «Semisupervised SVM Batch Mode Active Learning with Applications to Image Retrieval ».
- [19] Kompatsiaris L, Strintzis M.G Mezaris V, « Region-Based Image Retrieval Using an Object Ontology and Relevance Feedback ».
- [20] Naveen K. Nishchal, Arun K. Gupta, Vinod K. Rastogi Yadav B R, «Retrieval and classification of objects using generic Fourier,Legendre moment, and wavelet Zernike moment descriptors and recognition using joint transform correlator».
- [21] Guojun L Dengsheng Z, Review of shape representation and description techniques, 2003.
- [22] Imran Shafiq A Nan X, «Shape-Based Image Retrieval» MoMM December 14-16, Kuala Lumpur, Malaysia, 2009.
- [23] Marcela X. Ribeiro, Elaine P. M. Sousa, Agma J. M. Traina, Caetano Traina Jr Joaquim C. Felipe, « Effective Shape-based Retrieval and Classification of Mammograms».
- [24] Fabiano D. Beppler, Roberto C.S. Pacheco H. Andres Melgar S., « Knowledge Retrieval in the Anatomical Domain».
- [25] Gruber.T.R, «Towards Principles for the Design of Ontologies used for Knowledge Sharing» , 1993.
- [26] Guarino.N, «Formal Ontology and Information Systems », 1998.
- [27] Bodenreider.O, «Biomedical ontology in action ».
- [28] <http://sig.biostr.washington.edu/projects/fm/AboutFM.html> , visité 18/02/2012.
- [29] Zillner S, Moeller M, Sintek M, Zhan Y, Krishnan A, Gupta A Zhou X S, «Semantics and CBIR: A Medical Imaging Perspective».

- [30] <http://do-wiki.nubic.northwestern.edu/index.php/History#>, , visité 18/02/2012.
- [31] Bodenreider.O, The Unified Medical Language System (UMLS): integrating biomedical terminology.
- [32] Rogers JE I Rector AL, Ontological & Practical Issues in using a Description Logic to Represent Medical Concepts: Experience from GALEN, 2005.
- [33] Cramer Jayashree K Hersh W, « Multimodal Medical Image Retrieval -Image Categorization to Improve Search Precision».
- [34] Zhanga M, Antanib S, Longb R, Thoma G Yaoa J, "Automatic medical image annotation and retrieval", 2008.
- [35] Luminita A. Vese Tony F. Chan, "Active Contours Without Edges", FEBRUARY 2001.
- [36] Harry Wechsler Chengjun Liu, "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition".
- [37] Z.M. Ma,Liguo Deng Gang Zhang, "Texture Feature Extraction and Description Using Fuzzy Set of Main Dominant Directions of Variable Scales in Content-Based Medical Image Retrieval".
- [38] Mark O , Deselaers T, Keysers D, Schubert H, Spitzer K, Ney H, Berthold B. Wein Lehmann M, «Automatic categorization of medical images for content-based retrieval and data mining».
- [39] Zillner S, Moeller M, Sintek M, Zhan Y, Krishnan A, Gupta A Zhou X, "Semantics and CBIR: A Medical Imaging Perspective", 2008.
- [40] Tsatsoulis c Soh LK, Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-Occurrence Matrices.
- [41] Borgo.S, «Classifying (medical) Ontologies».
- [42] Agma J. M. Traina ,Caetano Traina Jr Joaquim C. Felipe, «A New Similarity Measure for Histograms Applied to Content-based Retrieval of Medical Images».
- [43] Thies C, Fischer B, Lehmann M Mark O, «Content-Based Retrieval of Medical Images by Combining Global Features».

ظهرت في العقدين الأخيرين من الزمن، العديد من أنظمة البحث عن الصور باستعمال المحتوى المرئي بخلاف الأنظمة القديمة التي كانت تعتمد على الطريقة النصية للبحث. في هذا العمل، نقترح نظاما للبحث عن الصور الطبية عن طريق المحتوى. يستند هذا النظام على تطبيق أساليب الفهرسة الهندسية مثل الشكل و الملمس. ويعمل النظام أيضا على تقديم شروح للصور الطبية المكتسبة حديثا من خلال استغلال نتائج البحوث.

كلمات البحث:

الصور الطبية، المحتوى المرئي، أساليب الفهرسة، الشرح.

Abstract:

During the two decades, several content-based image retrieval systems (CBIR) have emerged to fill the limits of the textual image retrieval. In this work, we propose a content-medical based image retrieval system. This system is based on the application of low-level indexing methods such as texture and shape. It will also provide annotations for newly acquired medical images by exploiting the results of research.

Keywords:

Medical image, indexing methods, ontology, annotation.

Résumé

Pendant les deux décennies, plusieurs systèmes d'indexation et de recherche d'images par le contenu visuel ont émergé afin de combler les limites de l'indexation textuelle des images. Dans ce travail, nous proposons un système de recherche d'images médicale par contenu. Ce système est basé sur l'application des méthodes d'indexation de bas niveau tel que la texture et la forme. Il permettra également de proposer des annotations pour les images médicales nouvellement acquises en exploitant les résultats de la recherche.

Mots-clés:

Images médicales, contenu visuel, méthodes d'indexation, ontologies, annotation.