MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE MOULOUD MAMMERI, TIZI-OUZOU

FACULTE DE GENIE ELECTRIQUE ET D'INFORMATIQUE
DEPARTEMENT AUTOMATIQUE

MEMOIRE DE MAGISTER

en ***Automatique***

Option : **Traitement d'Images et Reconnaissance de Formes**

Présenté par

AMROUNI Samia

Ingénieur U.M.M.T.O.

Thème

Recherche et classement d'images dans une base de données de grande taille à partir de requêtes relatives à leurs descripteurs

Mémoire soutenu le : devant le jury d'examen composé de :

HAMMOUCHE Kamal, Professeur à l'UMMTO	Président.
DIAF Moussa, Professeur à l'UMMTO	Rapporteur.
AMIROU Zahia, Maître de Conférences A à l'UMMTO	Examinatrice.
ALKAMA Sadia, Maître de Conférences B à l'UMMTO	Examinatrice.

Avant-propos

J'aimerai tout d'abord citer que l'élaboration de ce mémoire a été effectuée au laboratoire LVAAS (Laboratoire Vison Artificielle et Automatique des Systèmes » du département Automatique, FGEI, UMMTO.

J'aimerais, avant tout, exprimer mes remerciements à mon Directeur de mémoire, Monsieur **DIAF Moussa**, Professeur à l'UMMTO, pour m'avoir encadrée et donné la chance de me rapprocher du domaine du traitement d'images. Je le remercie pour ses encouragements pour surmonter les moments difficiles tout le long de ce travail.

Nous adressons nos remerciements les plus sincères à Monsieur **HAMMOUCHE Kamal**, Professeur de l'UMMTO, pour avoir accepté de juger ce travail en tant que président de jury.

Nos sincères remerciements s'adressent également à Madame **AMIROU Zahia,** Maître de Conférences classe A, à l'UMMTO, pour avoir accepté de juger notre travail en tant que examinatrice.

Nous ne manquerons pas de remercier Madame **ALKAMA Sadia**, Maître de Conférences classe B, à l'UMMTO, pour nous avoir fait l'honneur d'accepter de juger notre présent travail.

Nous tenons aussi à exprimer notre gratitude à toutes les personnes qui ont contribués de près ou de loin pour nous avoir guidés et encouragés à l'élaboration de ce travail.

Sommaire

Introduction générale 1

Chapitre I: Généralités sur la recherche d'images par le contenu

1.1. Introduction.....3
1.2. Etat de l'art.....4
1.3. La recherche d'images par le contenu.....6
1.4. Systèmes de recherche d'images par le contenu.....7
1.4.1. Principales composantes d'un système RIC.....8
La requête.....8
Les bases d'image 9
1.4.2. L'indexation 9
1.4.3. La recherche 10
1.5. Attributs visuels 11
Les attributs globaux 11
Les attributs locaux.....11
1.5.1. La couleur 12
1.5.2 La texture 12
1.5.3 La forme 12
1.6. Stratégies de la recherche 12
1.7. Mesure de similarité entre images..... 13
1.7.1 La distance de Minkowski 14
1.7.2 La distance quadratique 15
1.7.3 Distance de Mahalanobis..... 15
1.7.4 Intersection d'histogrammes 16
1.7.5 Distance Earth Mover Distance (EMD) 16
1.7.6 La distance de Kolmogorov Smirnov..... 17
1.7.7 La distance de Cramer Von Mises 17
1.7.8 Distance de Kullback-Leibler.....17
1.7.9 La Divergence de Jeffrey (JD)..... 18

1.8. Quelques systèmes RIC.....18
1.9. Domaines d'application.....20
1.10. Conclusion.....21

Chapitre II : Méthodes d'indexation d'images

2.1. Introduction.....22
2.2 La description de la couleur 22
2.2.1. L'espace couleur.....23
2.2.2. Histogramme couleur 26
2.2.3. Les moments statistiques.....27
2.2.4. Cohérence spatiale.....27
2.2.5. Les couleurs moyennes 28
2.2.6. Le corrélogramme.....28
2.3.7. Le vecteur de cohérence de couleur 29
2.3.8. Les couleurs dominantes 29
2.3. La description de la texture 30
2.3.1. Les Méthodes statistiques 30
2.3.1.1. Les statistiques du premier ordre.....31
2.3.1.2. Les statistiques du second ordre 31
1) Les matrices de cooccurrence 32
2) La matrice de longueur de plages.....34
2.3.2. Les méthodes fréquentielles.....35
2.3.2.1. La transformée de Fourier 35
2.3.2.2. Les filtres de Gabor 36
2.3.2.3. La transformée en ondelettes 37
2.4 La description de la forme... 38
2.4.1. Les descripteurs géométriques.....38
2.4.2. Les moments géométriques.....39
- Les moments de Hu 40
2.4.3. Les moments orthogonaux 40
2.4.3.1 Les moments de Legendre 40
2.4.3.2 Les moments de Zernike 41
2.4.4. La matrice norme de gradient 42
2.5. Conclusion.....43

Chapitre III: Tests et résultats

3.1. Introduction.....44
3.2. Indexation de la couleur.....45
3.2.1. Histogramme couleur 45
3.2.2. Mesure de similarité entre histogramme.....46
3.3. Indexation de La texture 46
3.3.1. La transformée en ondelettes continue.....47
3.3.2. Principe de la décomposition en ondelettes discrètes 47
3.3.3 Analyse Multirésolution 48
3.3.4. Mesure de similarité entre les vecteurs de texture.....52
3.4. Indexation de la forme.....52
3.4.1. Les moments de Hu 52
3.4.2. Mesure de similarité entre les vecteurs de forme.....54
3.5. Protocole d'évaluation d'un système de RIC.....54
3.5.1. Le rappel(Recall).....54
3.5.2. La précision 55
3.6 Les bases d'images utilisées.....56
La base ORL 56
La base COIL 100.....56
La base cérébrale standard MR.....57
La base d'images de texture.....57
3.7. Tests et résultats 58
3.7.1 La recherche d'images par histogramme 58
3.7.2 La recherche d'images par les ondelettes.....60
3.7.3 La recherche d'images par la méthode des moments de Hu.....62
3.8 La qualité des réponses 64
3.8.1 Principe de la méthode de TREC[63].....65
3.8.1 Interprétation des résultats.....68
3.9 Le bruitage d'images.....69
3.9.1 Interprétation des résultats.....72
3.10 Conclusion 72
Conclusion générale.....74

Introduction générale

Suite au développement de la technologie et la démocratisation de l'informatique, les bases de données images représentent des volumes d'informations de plus en plus considérables. Des outils de compression, archivage et recherche sont nécessaires afin de gérer ces stocks numériques. La recherche d'images est une technique qui consiste à identifier, parmi les images de la base, des images désirées par un utilisateur. Elle représente un domaine très vaste qui recouvre de nombreuses thématiques : le traitement d'images, les bases de données, la reconnaissance de formes, l'intelligence artificielle…

Le début était une problématique d'annotation textuelle par des mots clés des images respectant une forte dépendance entre le contenu de ces images et les étiquettes textuelles associées [1]. Le meilleur exemple est la recherche d'images sur internet. La recherche revient donc à une recherche textuelle sur ces mots-clés. Cette technique nécessite à l'utilisateur une entrée manuelle des mots définissants les images. Malgré ses avantages, ce type de caractérisation comporte des inconvénients, d'une part, la difficulté de l'annotation textuelle pour le cas d'une base d'image de grande taille et d'autre part le problème de la langue qui peut offrir des résultats qui sont complètement hors sujet.

Pour éviter les problèmes de l'annotation textuelle, une autre technique a été proposée au début des années 90, qui prend en considération l'image et uniquement l'image pour effectuer les recherches. Cette technique est connue sous le nom de la recherche d'image par le contenu (Content Based Image Retrieval CBIR).

Dans les systèmes de recherche d'images par le contenu [2], le principe général consiste à extraire de chaque image de la base un vecteur descripteur stocké dans une base de données, qui représentera fidèlement cette image lors de la recherche.

Dans ce travail nous nous intéressons de la recherche d'images par le contenu visuel à l'aide d'une description bas niveau contenant la couleur, la texture et la forme. Le contenu de ce mémoire comporte les étapes suivantes :

- Dans le premier chapitre on a donnée quelques généralités sur la recherche d'image par le contenu et les principales composantes constituant un système de recherche d'image par le contenu visuel ainsi les différentes distances de similarité à laquelle, un système de recherche d'images, peut faire appel pour comparer des images.
- Au deuxième chapitre, nous avons cité les trois attributs bas niveau utilisés dans notre travail qui sont la couleur, la texture et la forme, ainsi les différentes techniques d'indexation pour chaque attribut et l'influence que porte chaque attribut sur l'efficacité d'un système de recherche d'image par le contenu.
- Le troisième chapitre est consacré pour les tests et les résultats concernant la capacité des différentes méthodes choisies dans la récupération d'images similaires. Ainsi nous avons bruité notre image requête ensuite nous avons appliqué un filtre gaussien pour avoir une image requête dégradée, dont l'objectif est de tester la sensibilité de ces méthodes à la présence du bruit.
- Et à la fin, nous terminons par une conclusion générale, qui est consacrée à une évaluation des résultats des méthodes choisies de notre travail et leur sensibilité au bruit.

Chapitre I

Généralités sur la recherche d'images par le contenu

1.1 Introduction

Le premier moyen utilisé dans le domaine de la recherche d'images était l'utilisation de mots clés qui sont associés aux images pour les caractériser. Pour la recherche, il suffit de taper un mot-clé et l'image correspondante va immédiatement apparaitre. Grâce à la facilité de son utilisation, ce mode de recherche a été très appliqué. Désormais, malgré ses avantages, ce type de caractérisation comporte un certain nombre d'inconvénients. En effet, pour le cas des bases d'images de grandes tailles, la description textuelle pour toutes les images de la base est une opération longue, coûteuse et pénible pour l'utilisateur. De plus, le principal inconvénient de cette recherche est que le mot clés dépend de la langue et le résultat peut être complètement hors sujet. Comme exemple le mot 'avocat' qui désigne un homme de loi et en même temps un fruit. Afin d'échapper à de telles situations, une solution consiste à éviter la procédure de l'utilisation des mots-clés et donc de passer à une autre technique qui prend en considération l'image et uniquement l'image pour effectuer les recherches. Cette méthode est dite la Recherche d'Images par le Contenu (RIC ou CBIR, *Content Based Image Retrieval*).

Dans la suite de ce chapitre, nous présentons un état de l'art sur les systèmes de recherche d'images par le contenu ainsi que les principales composantes d'une telle recherche. Nous décrivons ensuite, les attributs visuels et les différentes distances de similarité les plus utilisées par les systèmes de recherche d'images. Après le rappel de quelques systèmes de recherche d'image par le contenu et les domaines d'application, nous terminons ce chapitre par une conclusion.

1.2 Etat de l'art

La recherche d'images est une technique qui consiste à identifier, parmi les images d'une base, des images désirées par un utilisateur.
La méthode traditionnelle de la recherche d'images se résumait en une problématique de recherche de mots en se basant sur les attributs textuels des images tels que le nom du fichier [2]. Cette approche nécessite de l'utilisateur une entrée manuelle des mots définissants l'image (légende) et qui la représentent fidèlement lors de la recherche. Cependant, cette méthode n'est pas toujours efficace suite aux problèmes de langue comme la fusion de sens qui donne des résultats différents. Deuxièmement, une requête par mots clés nécessite une longue et couteuse phase manuelle d'indexation des images. Cette méthode est devenue impraticable avec la taille toujours grandissante des bases de données multimédia actuelles et pour dépasser ces problèmes, une autre procédure a été proposée. C'est la recherche d'images par le contenu, en anglais : *Content Based Image Retrieval* (CBIR). Elle a été proposée par Kato au début des années 90 [3] et repose sur l'idée de chercher directement dans le contenu visuel des images telle que : texture, couleur, forme, les points d'intérêts… La recherche est exécutée à l'aide d'une image exemple proposée par l'utilisateur appelée la requête.
A partir de là, plusieurs auteurs ont proposés leurs méthodes servant à améliorer le domaine de la recherche d'images par le contenu. Le premier travail et le plus référencé sur l'indexation utilisant la couleur est la méthode d'intersection d'histogrammes proposée par Swain et Ballard [4].

Un histogramme 3D est construit pour chaque image de la base d'images. Les images de cette base semblables à une image requête sont trouvées en calculant l'intersection de l'histogramme de cette image avec ceux présents dans la base.

Les techniques d'indexation de données multidimensionnelles que sont les vecteurs descripteurs d'images sont décrites de façon exhaustive par Berrani [5] dans sa thèse, dont il utilise une recherche approximative des plus proches voisins afin de déterminer les images les plus proches de l'image requête de l'utilisateur.

Dans [6], l'auteur propose une approche de distribution couleur et de reconnaissance de contours d'objets en utilisant la distance EMD (*Earth Mover Distance*) pour la recherche d'images par le contenu. Le principe de son travail est basé sur les requêtes partielles dans lesquelles on cherche des images par comparaison d'attributs de régions similaires.

Les méthodes basées sur les points d'intérêt ont elles-aussi, fait l'objet de nombreuses recherches. Dans sa thèse, Loupias [7] propose une version multirésolution des points d'intérêts.

Une méthode utilisant la logique floue a été introduite dans [8]. La segmentation de l'image a plus de succès dans l'extraction des caractéristiques bas niveau de l'image. Julien Fauqueur [9] a étudié la recherche d'images par composition de catégories de régions issues de la segmentation couleur des images.

Dans [10] les auteurs ont proposés un système d'indexation de base de données images par le contenu, où chaque image est dans un premier temps décrite par un ou plusieurs vecteurs descripteurs invariants par rapport à certaines transformations. Ils ont proposé une classification bayesienne des vecteurs d'attributs des images. Ces vecteurs sont stockés sous forme arborescente.

Dans sa thèse, Jérôme Landré [11] a proposé une technique basée sur une décomposition multirésolution des images de la base paléontologique Trans'Tyfipal en utilisant l'analyse en ondelettes à l'aide de l'algorithme *lifting scheme*. L'auteur utilise une classification automatique des images afin de construire un arbre visuel de recherche.

Un autre système de recherche d'images par le contenu a été proposé par Houria ABED et al [12]. Ce système permet de rechercher les images de la base en fonction de leurs caractéristiques du bas niveau : la couleur et la texture. Dans cette approche, chaque image de la base est représentée par un descripteur qui stocke les descripteurs visuels des quadrants de l'image, obtenus par une décomposition récursive de l'image en arbre quaternaire selon un critère particulier (homogénéité de la couleur et la texture). La similarité des images est calculée à partir d'une distance notée Δ qui tient compte de la structure et des valeurs des nœuds des arbres quaternaires.

Rostom, dans sa thèse [13], a proposé une méthode de sélection adaptative des caractéristiques pertinentes et une méthode de classification hiérarchique multi-modèle pour la recherche d'images dans une base d'images hétérogènes.

Dans [14] les auteurs ont utilisé la transformée en ondelettes de Daubechies pour extraire le vecteur des caractéristiques de l'image. Dans [15], les auteurs ont proposé une technique de recherche d'images couleur basée sur la texture. Cette méthode utilise la matrice de cooccurrence de couleur pour extraire la caractéristique texture et mesure la similarité entre les deux images couleurs.

1.3 La recherche d'images par le contenu

La RIC est une technique de recherche qui consiste à identifier des images à partir de leur contenu sans aucune autre information. Dans les systèmes de recherche d'images par le contenu, le principe général consiste à calculer pour chaque image un vecteur descripteur d'une dimension *n* stocké dans une base de données, qui représentera l'image lors de la recherche. Notons que le calcul des descripteurs se fait hors ligne afin de ne pas pénaliser l'utilisateur avec des calculs longs.

De cette forte utilisation et des problèmes rencontrés lors de la manipulation de grandes quantités d'images (la structuration et le stockage) en résulte un dynamisme de recherche dans le domaine de l'indexation multimédia durant cette dernière décennie, donnant

naissance à de nombreuses méthodes d'indexation par le contenu, de recherche interactive et de navigation dans des bases.

1.4 Systèmes de recherche d'images par le contenu

Un système de recherche d'images par le contenu est défini par l'ensemble des différents éléments qui participent à la réalisation de la recherche d'images (méthodes, logiciel, les bases d'images…). La principale tâche d'un système de recherche d'images par le contenu est de rechercher une ou plusieurs images se trouvant dans la base d'images les plus similaires à une (ou plusieurs) requête(s) proposée(s) par l'utilisateur. Le résultat sera affiché suivant une distance de similarité entre le vecteur descripteur de l'image requête résumant son contenu et chaque vecteur de la base de données images.

La recherche d'images par le contenu passe par deux phases essentielles. La première est la phase d'indexation. C'est une étape durant laquelle le système exploite toutes les images de la base à partir desquelles il extrait les informations les plus intéressantes de leurs contenus visuels à savoir la couleur, la texture et la forme. Chaque image est donc représentée par un vecteur de caractéristiques qui soit le moins encombrant et qui la caractérise le mieux. Cette phase est dite hors ligne (*off-line*).

La deuxième phase contient, d'une part, l'extraction du vecteur des caractéristiques de l'image requête et, d'autre part, l'étape de la recherche. Ceci se fait par une fouille dans la base d'images indexée. Le système cherche, dans la base, les images qui vont être les plus ressemblantes à l'image requête proposée par l'utilisateur. Cette phase est la phase de la recherche et qui se fait en ligne (*on-line*). Le schéma de principe d'un système de recherche d'images par le contenu est donné sur la figure (1).

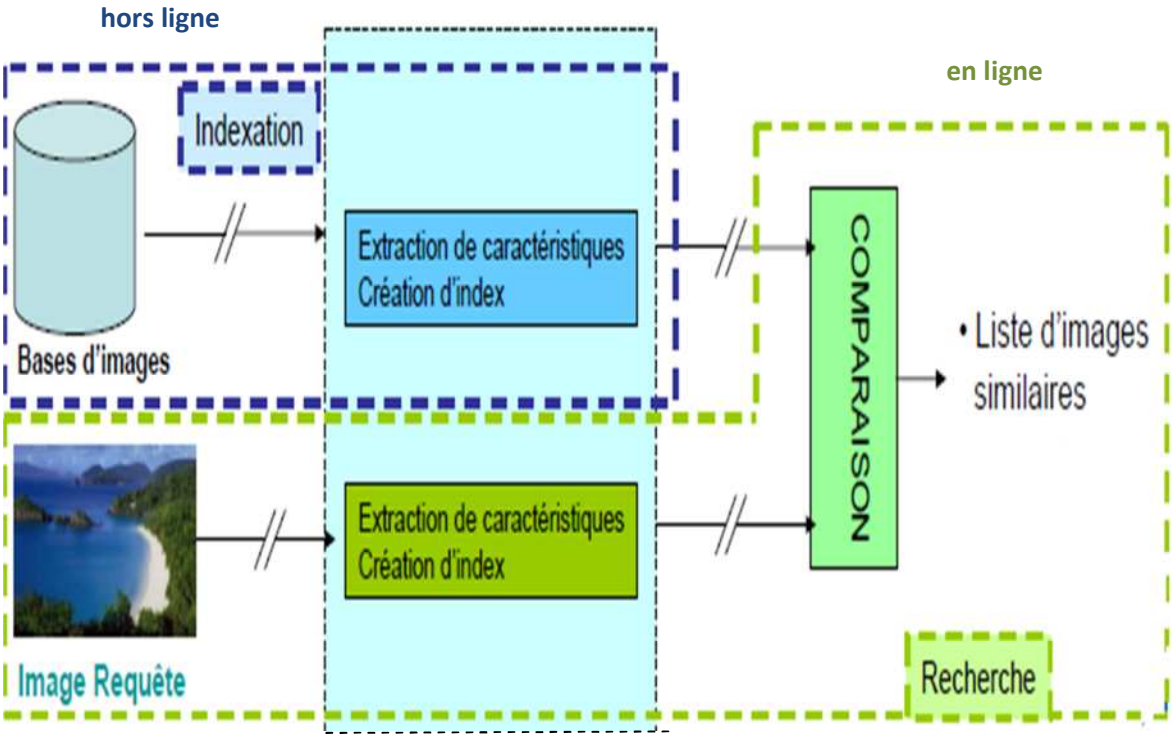


Fig.1. Architecture générale d'un système d'indexation et de recherche d'images par le contenu.

1.4.1 Principales composantes d'un système de RIC

La requête : Dans tout système de recherche, l'utilisateur doit exprimer ses souhaits, sous forme d'une requête au moteur de recherche par l'intermédiaire de l'interface utilisateur. Pour ce faire, celui-ci a le choix entre plusieurs types de requêtes dites par l'exemple, l'esquisse, par caractéristiques ou combinant exemple et texte.

Dans ce cas de la requête par l'exemple, l'utilisateur propose une image requête au système dont la recherche se fait par la comparaison du vecteur descripteur de cette requête avec les vecteurs descripteurs des images de la base. Le résultat sera une ou plusieurs images les plus similaires à l'image requête suivant une certaine distance de similarité. Dans la requête par esquisse (*sketch*), l'utilisateur spécifie sa requête par un dessin très simple à base de contours et de couleurs. Comme la recherche par image, ce type de requête demande l'indexation visuelle et le modèle de recherche par le contenu. Pour la requête par caractéristiques, l'utilisateur indique la ou les caractéristiques qu'il veut

utiliser pour trouver les images similaires comme, par exemple, la recherche d'images contenant 30% de rouge et 40% de vert. La requête combinant exemple et texte est une façon de renforcer l'image requête en lui associant un texte afin d'accéder à un niveau d'abstraction supérieur.

Les bases d'images: Une base d'images ou une banque d'images est définie par la donnée principale dans un système de recherche d'images par le contenu. Les bases d'images varient par leurs tailles qui peuvent atteindre un milliard d'images. Cette taille impose des contraintes sur la complexité des traitements effectués sur chaque image. Les bases d'images sont regroupées en deux grandes catégories pour la recherche et l'indexation. La première catégorie est dite *spécialisée*. Comme son nom l'indique, cette base d'images spécialisée contient des images d'un domaine particulier comme la base d'images médicales, tel que le dépistage du cancer du sein dans les bases mammographiques [16], et l'authentification des individus dans les bases d'images d'iris [17] etc. La deuxième catégorie est dite *généraliste* ou hétérogène. Les bases d'images généralistes contiennent des catégories d'images de sujet très différents comme par exemple le soleil, la montagne, la mer, les animaux etc.
Lors de la manipulation, la différence principale qui réside entre ces deux catégories de bases *spécialisées* ou *généralistes* est que les bases spécialisées ont le plus souvent un contenu sémantique homogène et la recherche d'images dans de telles bases est un problème spécifique à un domaine bien particulier.
Dans la catégorie de bases *généralistes* où les images ont des contenus assez différents, la recherche nécessite une classification en familles d'images spécialisées afin de passer aux étapes de l'indexation et de la recherche [11].

1.4.2 L'indexation

Dans les bases de données d'images, représentants des volumes d'informations de plus en plus considérables, une indexation automatique des images basée sur le contenu est devenue la méthode la plus adaptée

pour gérer ces masses de données. L'image étant représentée sous forme de matrice de pixels, son contenu brut ne peut être directement exploité pour la représentation d'images car la manipulation de tous les pixels de l'image est une opération pénible. Il est donc nécessaire de passer par une étape d'extraction des caractéristiques visuelles de l'image. Ce qui permet d'avoir une représentation plus facilement manipulable. Cette indexation consiste en l'extraction, de chaque image, un descripteur qui va la représenter fidèlement lors de la recherche et qui va être moins encombrant, ce qui permet une meilleure organisation des données images de la base, de limiter la quantité de données examinées durant une recherche, d'y accéder rapidement et de confiner la recherche au maximum. Cette étape est hors ligne car, durant cette phase l'utilisateur n'est pas connecté à la base d'images.

1.4.3 La recherche

Une fois la base d'images indexée, le système passe à l'étape de la recherche où l'utilisateur intervient avec sa requête. Cette dernière sera indexée de la même manière que celles de la base et représentée par un vecteur de caractéristiques. La recherche implique de fouiller dans la base, trouver les plus proches voisins de la signature associée à l'image requête. La performance d'un système RIC dépend essentiellement du choix des descripteurs employés et des techniques associées à leur extraction. La description des images peut être de bas niveau ou de haut niveau.

La description bas niveau se focalise seulement sur le contenu de l'image pour l'extraction des attributs visuels qui sont principalement la couleur, la texture, la forme, les points d'intérêt etc. L'extraction se fait grâce à des calculs mathématiques sur les pixels de l'image. C'est ce cas qui fera l'objet de notre étude.

Par contre, la description de haut niveau s'intéresse non seulement à la description visuelle du contenu de l'image mais aussi à l'interprétation du contenu de l'image, d'où l'appellation de description sémantique.

1.5 Attributs visuels

Les attributs visuels sont extraits à partir de l'image selon trois niveaux d'abstraction. Le premier niveau est basé sur une description globale des images, le deuxième est le niveau régions et le troisième représente la description par les points d'intérêt. Ces deux derniers permettent une description locale des images.
On distingue donc deux types d'attributs : les attributs globaux, et les attributs locaux.

- **Les attributs globaux**

Dans la recherche d'images par le contenu, la description globale d'images permet une recherche d'images se basant sur des similarités visuelles mesurées sur les images entières [18]. Une image est décrite globalement est représentée par un seul vecteur d'attributs. Les histogrammes de couleur ou de niveaux de gris sont des exemples de descripteurs globaux [4].

- **Les attributs locaux**

La description locale par région est construite par un ensemble de vecteurs d'attributs. Les attributs utilisées sont, généralement, extrait pour chaque sous région dans l'image [19]. La similarité entre deux images est calculée comme une combinaison des similarités entre les régions constituantes de chaque image.
Une image peut aussi être décrite par des points d'intérêt qui est une méthode qui effectue des recherches précises mais des temps de calcul trop élevé. Plusieurs approches ont été proposées pour la détection de ces points [20].

Dans notre travail, la recherche est effectuée par une description globale de caractéristiques visuelles qui sont la couleur, la forme et la texture.

1.5.1 La couleur

La couleur est un attribut largement utilisé dans les systèmes de recherche d'image par le contenu. La couleur permet de faciliter l'extraction et l'identification d'un objet dans une scène. Les travaux de Swain et Ballard [4] relatifs à l'indexation de la couleur constituent la référence et marquent les débuts de l'indexation de l'image. De nombreux travaux et méthodes existent pour l'indexation de la couleur. Cela fera l'objet d'une présentation détaillée dans le prochain chapitre.

1.5.2 La texture

La texture est une caractéristique facile à reconnaitre mais, en même temps, très difficile à définir. Cependant, une texture est un attribut représentant la distribution spatiale des niveaux de gris des pixels d'une région. Elle se traduit aussi par un arrangement spatial des pixels que la couleur, seule, ne suffit pas à décrire [21].

1.5.3 La forme

L'utilisation de descripteurs de forme n'a de sens que sur une image segmentée. Pour extraire les descripteurs d'une forme, la première opération consiste à définir sa fonction caractéristique. En général, elle est représentée sous la forme d'un masque dans lequel chaque pixel est représenté par le numéro de la région à laquelle il appartient [22]. C'est à partir de cette fonction, que sont calculés la plupart des descripteurs de forme, soit à partir de la région entière, soit à partir des contours seulement.

1.6 Stratégies de la recherche

La recherche d'images par similarité visuelle, à partir d'une image exemple, passe par la recherche des vecteurs voisins d'un vecteur q représentant la requête. Cette opération peut se décliner en trois stratégies à savoir, la recherche par les k plus proches voisins, La recherche par intervalle et la recherche à ε près. La première stratégie est

de chercher les k vecteurs les plus proches de q, au sens de la mesure de similarité associées aux vecteurs. Dans la recherche par intervalles, il s'agit de chercher les vecteurs r dont les composantes r_i sont tels que $|q_i - r_i| < s_i$ où $\{s_i\}$ représentent les seuils associés à chaque dimension. Dans la recherche à ε près, il s'agit de chercher les vecteurs se trouvant à une distance inférieure ou égale à un seuil ε, au sens de la mesure de similarité. En effet, si la signature d'une image est représentée par un ou plusieurs vecteurs, rechercher des images similaires revient à déterminer les vecteurs les plus proches au sens de la similarité qui leurs est associée.

La deuxième méthode concernant la recherche par intervalles n'a que très peu d'applications en RIC. Par contre, celle des k plus proches voisins et de la recherche à ε près sont les plus adaptées dans ce domaine. Cependant, la première est la plus claire mais le résultat, peut parfois contenir des vecteurs très éloignés, ce qui peut correspondre à des images non pertinentes.

Dans ce présent travail, la recherche d'images désirées suit la première stratégie. On cherche les douze vecteurs les plus proches du vecteur caractéristique de la requête.

1.7 Mesure de similarité entre images

L'idée principale de la RIC consiste à associer, à chaque image, un vecteur multidimensionnel représentant les caractéristiques de l'image en utilisant des fonctions de distance entre les vecteurs. Cette fonction quantifie la proximité des images dans l'espace des caractéristiques. Deux images sont considérées ressemblantes si la distance entre leurs descripteurs est faible. Cependant, pour mesurer cette similarité, plusieurs types de distances métriques peuvent être utilisés. Parmi ces distances, on peut citer celle de Minkowski, de Bhattacharya, de Mahalanobis et la distance quadratique. Avant de donner les expressions de ces distances, notons d'abord qu'un espace métrique K se définit

comme un ensemble non vide doté d'une application d, appelée distance, de K^*K dans R^+ et qui vérifie les trois axiomes suivants :

$\forall x, y, z \in K$

$$\left.\begin{array}{l} 1.\quad d(x,y) = 0 \Leftrightarrow x = y\ (identité) \\ 2.\quad d(x,y) = d(y,x)\quad (symétrie) \\ 3.\quad d(x,y) + d(y,z) \geq d(x,z)\quad (inégalité\ triangulaire) \end{array}\right\} \quad (1)$$

Parmi les distances les plus connues, on cite :

1.7.1 Distance de Minkowski

Cette distance est une famille de distances vectorielles. Soient I_1 , I_2 deux vecteurs de caractéristiques (vecteurs représentants de deux images). La formule générale de Minkowski s'écrit :

$$L_p(I_1, I_2) = (\sum_{i=1}^{n} |I_1(i) - I_2(i)|^p)^{\frac{1}{p}} \qquad p \geq 1 \qquad (2)$$

Où $I_1(i)$, $I_2(i)$ représentent les i^{émes} éléments des vecteurs I_1 et I_2 respectivement, p définie le facteur de Minkowski ($p \geq 1$), n représente la dimension de l'espace caractéristique.

- ✓ Pour p=1, on aura la distance de Manhattan (appelée city block) :

$$L_1\ (I_1, I_2) = \sum_{i=1}^{n} |I_1(i) - I_2(i)| \qquad (3)$$

- ✓ Pour p=2, on définit la distance Euclidienne :

$$L_2(I_1, I_2) = (\sum_{i=1}^{n} |I_1(i) - I_2(i)|^2)^{\frac{1}{2}} \qquad (4)$$

- ✓ Pour $p \to +\infty$, On définit la distance du maximum

$$L_\infty(I_1, I_2) = lim_{p\to\infty} \sqrt[p]{(\sum_{i=1}^{n} |I_1(i) - I_2(i)|^p)} = max_{i=1}^{n} |I_1(i) - I_2(i)| \qquad (5)$$

Les métriques de Minkowski sont simples d'utilisation. Par contre, leur calcul est réalisé en considérant que chaque composante du vecteur apporte la même contribution à la distance.

Afin de rendre compte de l'importance relative des composantes du vecteur les unes par rapport aux autres, les distances de Minkowski pondérées sont préférables.

$$L_p^w(I_1, I_2) = (\sum_{i=1}^{n} w_i |I_1(i) - I_2(i)|^p)^{\frac{1}{p}} \qquad p \geq 1 \qquad (6)$$

où w : Le vecteur de pondération à n composantes.

1.7.2 Distance quadratique

Contrairement à la distance de Minkowski qui traite les éléments de manière équitable, la distance quadratique favorise les éléments les plus ressemblants. Hafner et *al.* [23] propose une formule qui permet de déterminer une matrice de similarité A. Cette distance est donnée par l'expression suivante :

$$D_Q = \sqrt{(I_1 - I_2)^T A (I_1 - I_2)} \qquad (7)$$

où $A = [a_{ij}]$ est la matrice de similarité. a_{ij}, la distance entre deux éléments des vecteurs I_1 *et* I_2.

$$a_{ij} = 1 - \frac{d_{ij}}{\max(d_{ij})} \qquad (8)$$

d_{ij} est la distance dans l'espace considéré et max (d_{ij}), le maximum global de cette distance.

1.7.3 Distance de Mahalanobis

La distance de Mahalanobis permet de prendre en compte la distribution des classes. En effet, l'amplitude des différents attributs peut varier fortement, favorisant les valeurs des attributs élevés dans le calcul de la distance. Elle est définie par l'expression suivante :

$$D_{Mah} = \sqrt{(I_1 - I_2)^T C^{-1} (I_1 - I_2)} \qquad (9)$$

où C est la matrice de covariance entre l'ensemble des descripteurs d'images. Dans le cas où les dimensions des caractéristiques sont

indépendantes, C ne comporte que la variance. Et la formule (9) se simplifie comme suit:

$$D_{Mah}(I_1, I_2) = \sqrt{\left(\frac{\sum_{i=1}^{n}\left(I_1(i)-I_2(i)\right)^2}{C_i}\right)} \quad (10)$$

Si C est une matrice identité alors cette formule devient la distance Euclidienne.

1.7.4 Intersection d'histogrammes

La méthode d'intersection d'histogrammes a été proposée par Swain et Ballard [4]. Elle consiste en la mesure de la partie commune entre deux histogrammes H_1 et H_2.

Elle est calculée par l'expression suivante :

$$D_{Int}(H_1, H_2) = \frac{\sum_{i=1}^{n} \min(H_1(i), H_2(i))}{\sum_{i=1}^{n} H_2(i)} \quad (11)$$

où n est le nombre de valeurs de chaque histogramme. Deux images présentant une intersection d'histogrammes proche de 1 sont considérées similaires. Cependant cette mesure n'est pas une métrique parce qu'elle est non symétrique.

1.7.5 Distance Earth Mover Distance

La distance Earth Mover Distance (EMD) [6], permet d'évaluer la dissemblance entre deux distributions multidimensionnelles dans un espace de fonctions. Elle consiste aussi à minimiser le coût de transformation d'une signature en une autre sous certaines contraintes de déplacement des classes de descripteurs. Elle est alors définie comme suit :

$$D_{EMD}(I_1, I_2) = \frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} g_{ij} d_{ij}}{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} g_{ij}} \quad (12)$$

où d_{ij} est la distance entre deux pixels d'indices i et j, g_{ij} représente le flot optimal entre deux distributions, dont le coût total est

$$D_{EMD}(I_1, I_2) = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} g_{ij} d_{ij} \quad (13)$$

Ce coût est minimal sous les contraintes suivantes :

$$g_{ij} \geq 0 \quad \forall\, i\, et\, j \tag{14}$$

Cette contrainte n'autorise que des mouvements des composantes de I_1 vers I_2.

$$\sum_{i=1}^{n_1} g_{ij} \leq I_2\,(j) \qquad \forall j \tag{15}$$

$$\sum_{j=1}^{n_2} g_{ij} \leq I_1(i) \qquad \forall i \tag{16}$$

Ces deux contraintes limitent la quantité de composantes déplacées de I_1 et la quantité de composantes reçues I_2.

$$\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} g_{ij} = min(\, I_1(i), I_2(i)). \tag{17}$$

Cette formule exprime le maximum de déplacement de composantes possibles.

1.7.6 La distance de Kolmogorov Smirnov

Cette distance est appliquée aux distributions cumulées I^c(i) , elle est définit par la formule suivante :

$$D_{KS}(I_1^c, I_2^c) = \max_{\mathrm{i}}(|I_1^c(\mathrm{i}) - I_2^c(\mathrm{i})|) \tag{18}$$

1.7.7 La distance de Cramer Von Mises

Est une distance qui s'applique également sur les distributions cumulées. Sa formule s'écrit :

$$D_{CVM}(I_1^c, I_2^c) = \sum_i (I_1^c(i) - I_2^c(i))^2 \tag{19}$$

1.7.8 Distance de Kullback-Leibler

Cette mesure exprime l'entropie relative de deux distributions et elle s'exprime par la formule suivante :

$$D_{KL}(I_1, I_2) = \sum_i I_1(i) log \frac{I_1(i)}{I_2(i)} \tag{20}$$

1.7.9 La Divergence de Jeffrey (JD)

Cette distance est une extension de la mesure de Kullback-Leiber, la différence est que la divergence de Jeffrey est symétrique et stable.

$$D_{JD}(I_1, I_2) = \sum_i (I_1(i) log \frac{I_1(i)}{\widehat{I_i}} + I_2(i) log \frac{I_2(i)}{\widehat{I_i}}) \quad (21)$$

où $\widehat{I_i}$ représente la moyenne de $I_1(i)$ *et* $I_2(i)$, $\widehat{I_i} = \frac{(I_1(i) + I_2(i)}{2}$

1.8 Quelques systèmes RIC

Depuis son début, le domaine de la recherche d'images par le contenu a été l'objet d'études des groupes de recherche, ce qui a permis le développement de ce domaine et des systèmes RIC. Ces systèmes se différencient par leur mode de fonctionnement. Certains se basent sur les descripteurs visuels tels que la couleur, la texture et la forme et d'autres servent à classifier des images par catégories et même ceux capables d'intégrer la sémantique dans leurs recherche. Dans cette partie nous allons lister quelques systèmes les plus utilisés.

- ✓ **QBIC** (Query By Image Content) [24] est le premier système de recherche d'images par le contenu visuel développé par IBM. Il combine les techniques du domaine de la vision par ordinateur avec celles du domaine des bases de données. La recherche est basée sur l'indexation des textures des régions des images dans l'espace couleur de Munsell amélioré. Elle est réalisée par le biais des procédures de reconnaissance de formes, d'extraction des contours des objets de l'image, des textures et de l'histogramme de couleurs.

- ✓ **Virage** [25] Développé par Virage Inc. Virage tel que QBIC est un système de recherche d'images et propose des requêtes portant sur la couleur, la localisation des couleurs, la texture et la structure de l'image.

✓ **Netra** [26] : Développé au sein de l'université de Santa Barbara. Il emploie une description par région segmentée dont chacune est caractérisée par la couleur, la texture et la forme. La caractérisation de la texture est assurée par les filtres de Gabor et les ondelettes, pour caractériser la forme, ce système utilise les courbures de forme. Pour la mesure de similarité, ce système emploie la distance Euclidienne.

✓ **Windsurf** [27] :C'est un système basé sur la décomposition en ondelettes des images, suivie par une segmentation des régions à l'aide des nuées dynamiques et par l'extraction d'attributs colorimétriques et de texture. La comparaison entre les régions de l'image requête et celles de la base d'images est réalisée par la distance de Mahalanobis.

✓ **Photobook** [28] : Développé par MIT Media Laboratory, ce système propose une recherche sur trois attributs essentiels la couleur, la texture et la forme. Photobook est très adapté pour les bases d'images spécialisées. Pour le calcul de la similarité il utilise un grand nombre de méthodes telles que la distance Euclidienne et celle de Mahalanobis.

✓ **Cortina** [29] : La description des images est issue de la norme MPEG-7 et des mots issus du texte autour des images dans les pages Web pour construire son index d'images et la requête est effectuée soit par mot clés ou par exemple. Le système Cortina est un système destiné aussi pour la classification d'images en catégories dont le regroupement d'images est réalisé avec l'algorithme des k plus proches voisins.

✓ **Blobworld** [30]: Est un système de recherche d'images par le contenu qui repose sur les régions caractérisées par la couleur, la texture, la forme et la localisation spatiale, en procédant à une recherche par région exemple.

✓ **VisualSeek** [31]: Ce système repose sur des images segmentées. Les caractéristiques sont un ensemble de couleurs définies dans l'espace HSV, les ondelettes pour caractériser la texture ainsi que les relations spatiales entre régions. Le système permet à l'utilisateur de crayonner les régions d'intérêt et sélectionner la couleur appropriée puis paramétrer les attributs de type spatial. Pour mesurer la similarité le système utilise la distance quadratique et la distance Euclidienne.

✓ **Ikona** [32] : Le système Ikona est un système de recherche d'images par le contenu développé par l'INRIA. La recherche effectuée par ce système est basée sur une requête exemple utilisant les attributs couleur, texture et forme, ainsi qu'un système de bouclage de pertinence qui adapte la recherche suivant les modifications apportées aux résultats par l'utilisateur au cours de la recherche.

✓ **SIMPLIcity** [19] : Les images de la base sont classifiées automatiquement en catégories. La recherche s'effectue dans l'une de ces catégories. Les images sont segmentées en régions, dont les ondelettes ont pour rôle de caractériser chaque région.

La liste est encore ouverte et tous ces systèmes sont malheureusement difficilement comparables parce qu'ils travaillent sur des bases d'images différentes. Ces bases ne sont pas connues dans leur intégralité par l'utilisateur ce qui rend impossible de calculer une quelconque efficacité ou précision.

1.9 Domaines d'application

Les systèmes RIC deviennent plus en plus indispensables dans certains domaines tels que :

- Le domaine judiciaire : Les services de police possèdent de grandes collections d'indices visuels (visages, empreintes, …) qui vont être

exploitées par des systèmes de recherche d'images par le contenu afin d'identifier la personne recherchée.

- Les agences de journalisme et de publicité : Les agences de journalisme et de publicité est une autre application de la recherche d'images. Elles maintiennent en effet de grosses bases d'images afin d'illustrer leurs articles ou leurs supports publicitaires.
- Le domaine militaire : les systèmes de recherche d'images par le contenu sont très utilisés dans les applications militaires, tel que la reconnaissance d'engins ennemis via images radar, système de guidage et identification de cibles via images satellite.
- D'autres applications telles que le diagnostic médical afin de détecter les organes malades chez les patients, la gestion d'œuvres d'art pour explorer et chercher les peintures similaires, les moteurs de recherche d'images sur internet et la gestion de photos personnelles etc.
-

1.10 Conclusion

La recherche d'image est une branche en plein développement. Dans ce chapitre, nous avons donné un état de l'art sur la recherche d'images par le contenu et son importance dans le domaine de traitement d'images. Ensuite, nous avons décrit le principe de fonctionnement des systèmes de recherche d'images qui se compose essentiellement d'une donnée principale qui est la base d'images dont l'utilisateur fouille pour extraire une ou plusieurs images dépendamment de son choix, ainsi une requête exposée par cet utilisateur qui soit sous forme d'une image, esquisse ou une requête par caractéristiques. Nous avons présenté les différentes distances employées pour la mesure de similarité entre images. Et enfin, nous avons donné une liste de quelques systèmes existants.

Chapitre II

Méthodes d'indexation d'images

2.1 Introduction

Rappelons que la recherche d'images par le contenu (RIC ou *CBIR*) est une technique de traitement d'images qui vise à rechercher des images à l'aide d'une requête portant sur les caractéristiques visuelles d'une image telle que la couleur, la texture et la forme. Le résultat de cette recherche sera une image ou un ensemble d'images ordonnées dans l'ordre croissant de la valeur de la distance. La construction d'une base des index commence par l'extraction des caractéristiques appropriées des images. La couleur, la forme et la texture sont des caractéristiques importantes pour la description du contenu des images. Par conséquent, beaucoup de recherches se font sur la couleur, la forme et la texture afin de pouvoir les utiliser dans la reconnaissance et l'interprétation des images.

2.2 La description de la couleur

La couleur est un attribut indispensable dans la grande majorité des systèmes d'indexation et de recherche par le contenu [33]. Commencée par les travaux de Swain et Ballard [4] dont les auteurs ont

pu caractériser une couleur par son histogramme. De nombreux descripteurs sont proposés dans la littérature et nous pouvons considérer qu'ils forment deux grandes catégories. La première catégorie concerne les descripteurs relatifs à l'espace couleur [34], où il s'agit de représenter les principales couleurs d'une image, tout en fournissant des informations sur leur importance, leur distribution colorimétrique. Dans la deuxième, les descripteurs incluent des informations spatiales relatives à la distribution dans le plan image de la couleur, à la connexité entre couleurs. L'approche la plus courante et la plus rencontrée dans la littérature est l'histogramme couleur. De très nombreux auteurs ont proposé diverses manières d'utiliser l'histogramme comme descripteur. Deb et. al [35] ont proposé un histogramme dans l'espace HSV pour décrire la couleur. Ainsi que diverses distances associées qui permettent de mesurer la similarité entre deux histogrammes. Dans [4], Swain et Ballard ont décrit la couleur d'une image à l'aide de son histogramme couleur et ont défini l'intersection d'histogrammes couleur comme mesure de similarité entre deux images.

2.2.1 L'espace couleur

Une couleur est généralement représentée par trois composantes qui définissent ce qu'on appelle un espace couleur [34]. Les espaces les plus utilisés dans la recherche d'images par le contenu sont : l'espace RGB, l'espace CIE L*a*b*, l'espace CIE L*u*v*, l'espace HSV (ou HSL, HSB).

Espace RVB : Défini pour la première fois en 1931 par la compagnie internationale de l'éclairage CIE. Il est le système couramment le plus utilisé qui se base essentiellement sur trois couleurs fondamentales le rouge, le vert et le bleu et dont toutes les autres couleurs dérivent.
Il se représente sous forme d'un cube de la figure (1).

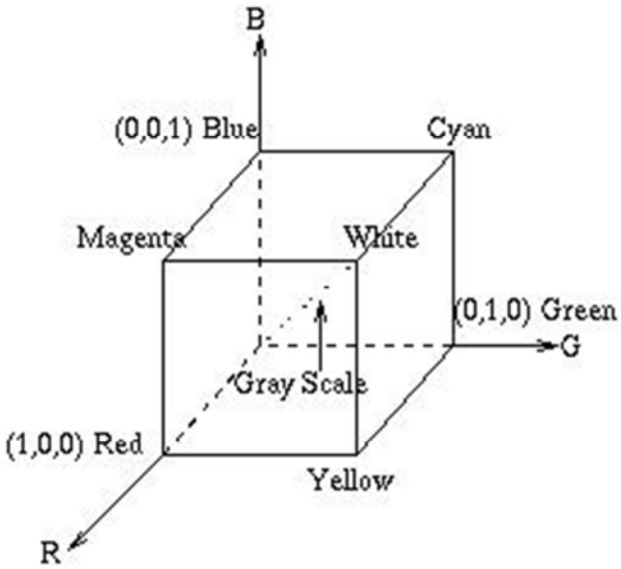


Fig.1. Cube de l'espace RVB.

Cet espace couleur est sensible aux changements d'illumination, et ne correspond pas au processus de perception humaine.

Espace HSV (HSI) (Hue-Saturation-value) : La projection d'une image dans cet espace permet de séparer les trois informations relatives à la teinte (H), qui représente la couleur pure (rouge, jaune, vert, bleu…etc.), la saturation (S) qui donne des informations sur la quantité de blanc dans la couleur pure et la valeur (V) correspondant à l'intensité lumineuse qui indique si la couleur est claire ou sombre[11].

On représente cet espace par un cône, dont la teinte est représentée par les angles du cône, la saturation est donnée par les rayons du cône et la valeur est définie par la hauteur du cône, illustré par la figure (2).

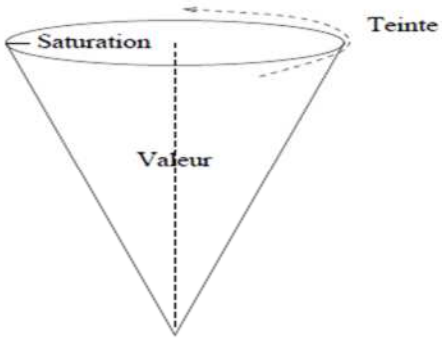


Fig.2. L'espace couleur HSV.

Cet espace présente l'avantage de simuler le comportement visuel humain. Mais il a l'inconvénient d'être non uniforme ce qui se traduit par le fait que la distance entre deux couleurs visuellement proches calculée dans cet espace peut être très grande.

Espace XYZ : La C.I.E. a défini un autre espace de représentation d'une couleur qui prend en compte la sensibilité de l'œil. Il s'agit de l'espace XYZ. Il s'obtient simplement à partir du système RGB et à l'aide d'une matrice de passage [36]. La figure (3) donne une représentation de cet espace, Y contient la luminance, X et Z l'information de chrominance. Cet espace lui aussi n'est pas perceptuellement uniforme ce qui le rend rarement utilisé dans les systèmes de recherches d'images. C'est-à-dire que les différentes nuances d'une même teinte (par exemple vert) peuvent être très éloignées dans l'espace.

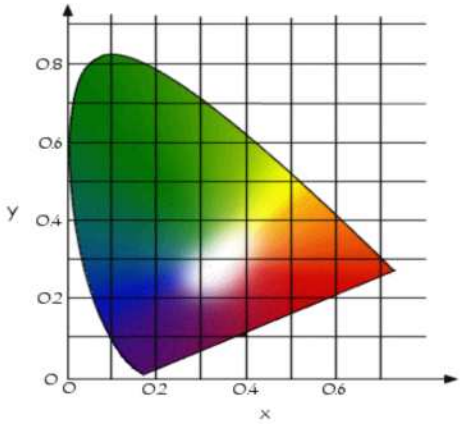


Fig.3. L'espace couleur XYZ.

Espace CIE Luv : C'est un système perceptuellement uniforme que l'on peut obtenir à partir du système XYZ. Il permet l'utilisation d'une distance Euclidienne pour mesurer l'écart entre deux couleurs. La composante L représente la réponse de l'œil à un niveau de luminance, u et v représentent la chrominance.

Espace CIE Lab : C'est un espace couleur qui possède la bonne propriété de respecter les distances entre les couleurs visuellement proches. Donc il est perceptuellement uniforme. De même que le système précédent, il est

défini à partir de l'espace XYZ par des relations non-linéaires. Il est très utilisé dans le domaine de la recherche d'image par le contenu. Dans [11] Landré a fait sa représentation de couleur dans deux espaces couleur l'espace HSV et l'espace CIE Lab pour avoir une meilleure description de la couleur.

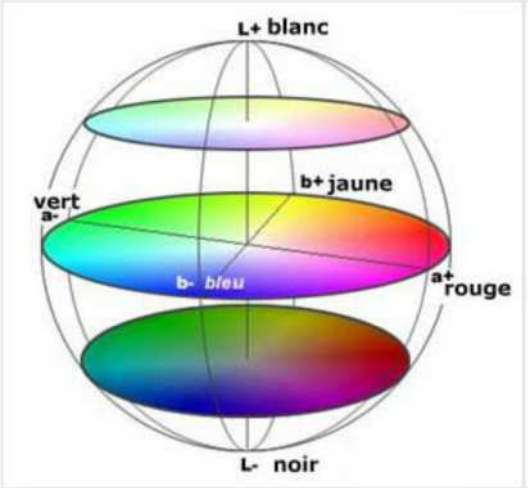


Fig.4. Espace CIE Lab.

Espaces YUV, YIQ et UCbCr : Ces espaces sont semblables. YUV est utilisé pour le codage des couleurs dans le système de télévision PAL, YIQ pour le système NTSC, dont Y représente l'intensité, UV et IQ sont les deux composantes de chrominance. Pour Cr et Cb correspondent aux complémentaires du rouge et du bleu.

2.2.2 Histogramme couleur

Une représentation pratique et la plus connue de la couleur est l'histogramme. Cette méthode a été utilisée pour la première fois par Swain et Ballard [4]. Hafner et al [37] ont défini un histogramme par le nombre d'apparitions d'un élément dans un ensemble. Malgré que ce descripteur donne une information sur la présence de chaque couleur dans l'image, il ne tient pas en considération sa localisation spatiale.
Cette méthode sera utilisée dans notre travail pour la caractérisation de la couleur et sera détaillée au chapitre suivant.

2.2.3 Les moments statistiques

Contrairement à la méthode d'histogramme qui utilise une distribution complète de la couleur, cette approche permet de calculer seulement des caractéristiques dominantes de couleur telles que la moyenne, la variance et le moment du troisième ordre et qui sont donnés par les formules suivantes.

1. La Moyenne :

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} P_{ij} \quad (1)$$

2. Le contraste (contraste dans une image)

$$\sigma_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(P_{ij} - \mu_i)^2} \quad (2)$$

3. Le moment d'ordre 3 désigne la quantité de la lumière dans une image.

$$S_i = \left(\frac{1}{N}\sum_{j=1}^{N}(P_{ij} - \mu_i)^3\right)^{\frac{1}{3}} \quad (3)$$

P_{ij} : La valeur du pixel j pour le canal i (RGB), N : nombre de pixel de l'image.

Dans [34] les auteurs ont même prouvé que les méthodes utilisant des moments statistiques marchent plus vite et donnent des résultats meilleurs que les méthodes d'histogrammes.

2.2.4 Cohérence spatiale

Contrairement à l'histogramme, ce descripteur sert à combler l'absence de l'information spatio-colorimétrique de l'image. La cohérence spatiale est calculée pour chaque classe de couleur identifiée. On définit d'abord un histogramme de connexité donné par la formule suivante :

$$H_s(c) = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} \delta(\mathrm{I_s}\ (\mathrm{i,j}),\ \mathrm{c})\ \alpha(\mathrm{i,j}) \quad (4)$$

I_s est une image segmentée de taille M*N, c est la couleur du pixel (i,j), δ est le symbole de Kronecker et α(i, j) est définie par :

$$\alpha(\mathrm{i,j}) = \begin{cases} 1 & \text{si } \forall \mathrm{k,k'} \in (-\mathrm{W,W}) \quad \mathrm{I_s(i+k\ ,j+k')} = \mathrm{I_s(i,j)}. \\ 0 & \text{si non} \end{cases} \quad (5)$$

La fenêtre (2W+1)*(2W+1) représente le degré de compacité souhaitée.

La cohérence spatiale est donnée par le rapport suivant :

$$\mathrm{SCR(c)} = \frac{H_s(\mathrm{c})}{H(c)} \quad (H : \text{Histogramme des couleurs}) \tag{6}$$

2.2.5 Les couleurs moyennes

Cette méthode sert à déterminer les différentes distributions moyennes de couleur dans une image. Pour une image I représentée dans l'espace RGB et comportant M*N pixels, la caractéristique couleur moyenne est donnée par la formule suivante :

$$\bar{x} = (\bar{R}_{moy}, \bar{G}_{moy}, \bar{B}_{moy})^t \tag{7}$$

$$\bar{R}_{moy} = \frac{1}{N*M} \sum_{p=1}^{N*M} R(p) \tag{8.a}$$

$$\bar{G}_{moy} = \frac{1}{N*M} \sum_{p=1}^{N*M} G(p) \tag{8.b}$$

$$\bar{B}_{moy} = \frac{1}{N*M} \sum_{p=1}^{N*M} B(p) \tag{8.c}$$

R(p), G(p), B(p) sont les niveaux de couleur du pixel P dans chaque canal le rouge, le vert et le bleu respectivement.

2.2.6 Le corrélogramme

L'idée principale de cette méthode est inspirée des matrices de cooccurrence, développés au départ pour les images en niveaux de gris [38]. Le Corrélogramme a pour rôle de représenter les corrélations spatiales entre les niveaux de gris en fonction de la distance inter pixels utilisant la norme L_∞.(voir la formule (9)). Les corrélogrammes ont été utilisées dans [39] pour l'indexation et la recherche d'image par le contenu.

$$|P_1 - P_2| = max\{|x_1 - x_2|, |y_1 - y_2|\} \tag{9}$$

$|P_1 - P_2|$:La distance entre P_1 *et* P_2 de coordonnées (x_1, y_1) *et* (x_2, y_2).
Pour une distance $k \in \{1,2, \dots d\}$ fixée à priori, le corrélogramme de la couleur c_i par rapport à la couleur c_j est défini par :

$$Cor_{i,j}^{k} = prob_{p1 \in I_{ci}}[P2 \in I_{cj} / |P_1 - P_2| = k] \quad (10)$$

I_{ci}, I_{cj} : Les ensembles de pixels dont la couleur est c_i et c_j respectivement.

Pour tout pixel de couleur c_i dans une image I de dimension N*N, $Cor_{i,j}^{k}$ exprime la probabilité de trouver un pixel de couleur c_j à une distance k.

Pour comparer deux images I et I′, la mesure de similarité proposée par [40] est dérivée de la distance de Manhattan L_1.

2.2.7 Le vecteur de cohérence de couleur

Ce descripteur décrit par Pass et al. [41] représente une autre variante, plus détaillée, de l'histogramme. Chaque rang de l'histogramme peut être partitionné en cohérent et non cohérent. La formule de la cohérence de couleurs est donnée par le vecteur suivant : $[(\alpha_1, \beta_1), (\alpha_2, \beta_2), \ldots, (\alpha_N, \beta_N)]$, tel que la somme $(\alpha_1 + \beta_1, \alpha_2 + \beta_2 \ldots, \alpha_N + \beta_N)$ donnera l'histogramme de couleurs de l'image. L'avantage qu'apporte cette approche est l'ajout de l'information spatiale à l'histogramme et cela à partir de leur raffinement. Mais, cette méthode présente l'inconvénient d'amplifier la sensibilité aux conditions d'illumination.

2.2.8 Les couleurs dominantes

Les couleurs dominantes caractérisées par la signature s telle que s $=\{s_i = (m_i, w_i)\}$ est un ensemble de nuages de points. Chaque nuage est représenté par son mode m_i définissant le maximum local de sa densité de probabilité, et le nombre w_i de pixels qui appartient au nuage.

Les signatures par couleurs dominantes, proposées dans [42], ne stockent que les couleurs qui appartiennent à l'image, elles ne stockent pas les cellules vides. La mesure de similarité la plus utilisée pour comparer ces signatures est l'Earth Mover Distance 'EMD'. La distance entre deux distributions s_1 et s_2 est donnée par le cout minimum de travail nécessaire pour transformer s_1 en s_2. Les différents nuages de points, sont déterminée par plusieurs méthodes dont la plus facile est celle basée sur l'histogramme.

2.3 La description de la texture

La texture d'une image peut être interprétée comme la régularité d'apparition de couples de niveaux de gris selon une distance donnée dans l'image. La texture est un attribut très riche en information sur la structuration des pixels dans l'image, ce qui a permis le développement de beaucoup de méthodes pour extraire le maximum de ces informations. En pratique, on distingue deux grandes classes de textures, qui correspondent à deux niveaux de perception à savoir les macrotextures et les microtextures.

Le but de l'analyse d'une texture est de formaliser les descriptifs de cette dernière par des paramètres mathématiques qui serviraient à l'identifier. Pour cet objectif, une multitude de méthodes, de variantes et de combinaisons de méthodes sont proposées dans la littérature et éprouvées en pratique. Elles sont classifiées en trois principales catégories qui sont principalement : les méthodes structurelles, les méthodes statistiques et les méthodes fréquentielles.

Le premier type tient compte de l'information structurelle et contextuelle et est particulièrement bien adaptée pour le cas des textures macroscopiques qui sont des textures construites sur une primitive bien identifiable qui se répète sur une maille régulière. Les étapes d'analyse sont d'abord l'identification des éléments constitutifs, puis la définition des règles de placement.

2.3.1 Les Méthodes statistiques

Consistant à évaluer des propriétés statistiques d'une région ou d'un certain voisinage autour d'un pixel (telle que la variance, la moyenne, la corrélation, l'entropie…) en se basant sur la distribution des niveaux de gris. On trouve les caractéristiques statistiques du premier ordre se basant sur l'histogramme des niveaux de gris de l'image et les statistiques du second ordre tenant compte des relations inter-pixels dans une image en niveau de gris.

2.3.1.1 Les statistiques du premier ordre

Les caractéristiques statistiques du premier ordre sont des méthodes ne tenant en compte qu'un pixel à la fois, utilisant les propriétés de l'histogramme des niveaux de gris de l'image. Soit *Ig* une image en niveau de gris correspondante à l'image I de N pixels. L'histogramme des niveaux de gris est calculé comme suit :

$$hist(i) = \frac{1}{N}\sum_{p=1}^{N} Ig(p) == i \qquad (i \in \{1,2,..,ng\}) \qquad (11)$$

Les statistiques du premier ordre [43], permettent de donner une certaine information sur les textures présentes dans cette image. Les attributs extraits sont : la moyenne, l'écart type et la variance de la distribution des niveaux de gris dans l'image. Donnés par les formules (12), (13) et (14).

La moyenne :

$$Moy = \frac{1}{ng}\sum_{i=1}^{ng} i * hist(i) \qquad (12)$$

Ecart type:

$$E_{type} = \sqrt{\frac{1}{ng}\sum_{i=1}^{ng}(i - moy)^2\, hist(i)} \qquad (13)$$

La variance :

$$Var = \frac{1}{N}\sum_{p=1}^{N}(Ig(p) - moy)^2 \qquad (14)$$

2.3.1.2 Les statistiques du second ordre

Ce sont celles qui exploitent les paires de pixels en incluant une certaine dépendance spatiale. Elles sont très exploitables pour la caractérisation de la texture. On se limite de citer les méthodes les plus connues.

1) Les matrices de cooccurrence

Les matrices de co-occurrence des niveaux de gris [44] (Gray Level Co-occurrence Matrix GLCM) est une méthode qui représente la référence en analyse statistique de texture d'une image. Les matrices de cooccurrence sont très riches en information de texture et servent souvent de méthode comparative pour les nouvelles approches. Les matrices de cooccurrence font l'objet de plusieurs recherches. Dans [45] les auteurs ont donné une nouvelle technique pour un système de recherche d'image par le contenu, utilisant les matrices de cooccurrence des motifs. Elle est, en effet, simples à mettre en œuvre et offrent de bonnes performances.

Le principe de cette méthode est de mesurer la probabilité d'apparition des paires de valeurs de pixels situés à une certaine distance dans l'image (la formule (15)). Elle est basée sur le calcul de la probabilité $P_{i,j}(d,\theta)$ qui représente le nombre de fois où un pixel de niveau de gris i apparait à une distance relative d d'un pixel de niveau de gris j suivant une orientation θ donnée.

$$P_{i,j}(d,\theta) = prob_{p_1 \in NGi}[p_2 \in NGj | p_1 - p_2| = d, \llcorner(p_1,p_2) = \theta]. \quad (15)$$

Il est généralement admis que toute l'information texturale dérivable des matrices de cooccurrence est obtenue pour les angles 0°, 45°, 90° et 135°, avec une distance inter-pixels égale à 1. D'où la présence de quatre matrices de cooccurrence.

Les quatre matrices sont alors calculées suivant les équations suivantes où *(k,l)* sont les coordonnées d'un pixel de niveau de gris i∈$[0, ng_{max} - 1]$et (m,n) celles du pixel de niveau de gris j∈$[0, ng_{max} - 1]$. ng_{max} représente le niveau de gris maximum dans l'image.

$$P_{i,j}(d,0) = \left|\left\{\left((k,l),(m,n)\right) \in (N*M)^2 tq\left(k-m=0, |l-n|=d, I_{k,l}=i, I_{m,n}=j\right)\right\}\right| \quad (16.a)$$

$$P_{i,j}(d,45) = \left|\begin{matrix}\left\{\left((k,l),(m,n)\right) \in (N*M)^2 tq \\ (k-m=d, l-n=-d) \vee (k-m=-d, l-n=d), I_{k,l}=i, I_{m,n}=j\right\}\end{matrix}\right| \quad (16.b)$$

$$P_{i,j}(d,90) = \left|\{\big((k,l),(m,n)\big) \in (N*M)^2 tq(|k-m| = d, l-n = 0), I_{k,l} = i, I_{m,n} = j\}\right| \quad (16.c)$$

$$P_{i,j}(d,135) = \left| \begin{matrix} \{\big((k,l),(m,n)\big) \in (N*M)^2 tq \\ (k-m = d, l-n = d) \vee (k-m = -d, l-n = -d), I_{k,l} = i, I_{m,n} = j\} \end{matrix} \right| \quad (16.d)$$

Les matrices de cooccurrence sont exploitable par extraction des attributs numériques calculés appelés paramètres de texture.

Haralick [38] a proposé 14 attributs pouvant être extraits de ces matrices de cooccurrence. Généralement, les quatre attributs utilisés dans la recherche d'image par le contenu sont : l'énergie, le contraste, l'entropie et le moment inverse de différence, obtenus après normalisation des matrices $P_{i,j}(d,\theta)$ par N*M.

✓ Energie

L'énergie est un paramètre qui mesure l'uniformité de la texture. Il atteint de fortes valeurs lorsque la distribution des niveaux de gris est constante ou de forme périodique. Elle est donnée par la formule suivante.

$$E = \sum_i \sum_j (P_{ij}(d,\theta))^2. \quad (17)$$

✓ Contraste

$$Cont = \sum_i \sum_j \Big((i-j)^2 P_{ij}(d,\theta)\Big) \quad (18)$$

Cette valeur est élevée lorsque la texture présente un fort contraste.

✓ Entropie

Ce paramètre mesure le désordre dans l'image. Contrairement à l'énergie, l'entropie atteint de fortes valeurs lorsque la texture est complètement aléatoire (sans structure apparente).

$$Ent = -\sum_i \sum_j (P_{ij}(d,\theta) \log(P_{ij}(d,\theta))). \quad (19)$$

✓ Moment inverse de différence (MID)

Il mesure l'homogénéité de l'image. Ce paramètre reflète l'existence de plages uniformes de texture. Plus cette valeur est élevée, plus la texture est grossière.

$$MID = \sum_i \sum_j \frac{P_{ij(d,\theta)}}{1+(i-j)^2} \quad (20)$$

2) La matrice de longueur de plages

Les longueurs de plages sont une manière de représenter une texture par une estimation de la probabilité d'observer des plages des niveaux de gris constants dans une certaine direction. Une plage est définie par l'ensemble de pixels consécutifs et dans une direction donnée ayant le même niveau de gris. Une longueur de plage définit le nombre de pixel que contient cette dernière.

De même que la méthode des matrices de cooccurrence, la matrice de longueur de plages est exploitable par extraction de plusieurs attributs.

- Le nombre de longueurs de plages :

$$SLP = \sum_{ng=0}^{Ng-1} \sum_{l=1}^{L} P_\theta\,(ng, l) \quad (21)$$

- La proportion de petites et grandes plages (resp. $RF1$ et $RF2$):

$$RF1 = \frac{1}{SLP} \sum_{ng=0}^{Ng-1} \sum_{l=1}^{L} \frac{P_\theta(ng,l)}{l^2} \quad (22)$$

$$RF2 = \frac{1}{SLP} \sum_{ng=0}^{Ng-1} \sum_{l=1}^{L} l^2 P_\theta\,(ng, l) \quad (23)$$

- L'hétérogénéité des niveaux de gris:

$$RF3 = \frac{1}{SLP} \sum_{ng=0}^{Ng-1} (\sum_{l=1}^{L} P_\theta(ng, l))^2 \quad (24)$$

RF3 mesure la dispersion des plages entre les niveaux de gris.

- L'hétérogénéité des longueurs de plages :

$$RF4 = \frac{1}{SLP} \sum_{l=1}^{L} (\sum_{ng=0}^{Ng-1} P_\theta\,(ng, l))^2 \quad (25)$$

Ce paramètre donne une information sur la dispersion des plages entre les longueurs.

- Le pourcentage des plages :

$$\mathrm{RF5} = SLP/K. \quad (26)$$

$P_\theta(ng, l)$: indique le nombre de plages de pixels ayant le niveau de gris ng, de longueur l. Le paramètre θ indique la direction de la plage.

Ng correspond au nombre de niveaux de gris dans l'image et L définit la longueur de la plage maximale.

K : Le nombre total des pixels de l'image.

2.3.2. Les méthodes fréquentielles

Analyser une image par son spectre fréquentiel nous donne beaucoup d'informations sur celle-ci. L'analyse fréquentielle d'images est une analyse qui est la plupart du temps déterminante de la texture étudiée. On cite la transformée en ondelette qui est très utilisée dans la description et la classification de la texture [46]. Ainsi les filtres de Gabor, ont eux aussi montrés leurs capacités de représenter une texture [47].

Cependant une analyse fréquentielle est résultat d'une transformation de Fourier dans le domaine fréquentiel, ce qui nous amène à passer par la transformation de Fourier pour les signaux images.

2.3.2.1 La transformée de Fourier

La transformée de Fourier (TF) permet de passer d'une représentation de l'image dans le domaine spatial à sa représentation dans le domaine fréquentiel. Les résultats de la TF nous informent sur le contenu fréquentiel global à partir duquel certains paramètres caractéristiques texturaux peuvent être extraits.

Pour une image I définie contenant N lignes et M colonnes, la représentation de la transformée de Fourrier sera discrète. Elle est donnée par la formule suivante :

$$TFD_I(u,v) = \frac{1}{N*M}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} I(i,j)e^{-j2\pi(\frac{ui}{N}+\frac{vj}{M})} \quad (27)$$

$0 \leq u \leq N-1$

$0 \leq v \leq M-1.$

Cette TFD bidimensionnelle n'est pas toujours facile à interpréter. On s'intéresse principalement à son module, appelé spectre de Fourier.

Pour une texture directionnelle, le spectre de puissance contient des pics sur une ligne perpendiculaire à la direction principale de la texture. Si la texture est périodique, on retrouvera cette périodicité dans le spectre.

Pour le cas d'une texture fine, le spectre sera maximum dans les hautes fréquences, et inversement, si elle est plutôt grossière, le maximum sera proche des basses fréquences. En effet, le spectre de Fourier rend compte de la rugosité d'une texture.

2.3.2.2 Les filtres de Gabor

La méthode de Gabor se base sur une formule dite fonction de Gabor, qui est définie par association d'une courbe de Gauss avec une sinusoïde orientée. Voir la formule (28). Cette méthode est très utilisée en indexation d'images pour la caractérisation de la texture et notamment pour la segmentation des images texturées [47], ainsi dans le domaine de la reconnaissance faciale [48].

L'utilisation des filtres de Gabor consiste à analyser indépendamment différentes parties de l'espace de Fourier à l'aide de plusieurs filtres.

Gaussienne * sinus/cosinus = filtre de Gabor

Fig.7. Représentation d'un filtre de Gabor 2D

L'expression du filtre de Gabor en deux dimensions est donnée par la formule suivante :

$$f_{Gb} = g(\mathrm{x}', \mathrm{y}')\exp(2\pi j[\mu_0(x - x_0)^2 + \upsilon_0(y - y_0)^2] \quad (28)$$

$$g(\mathrm{x}', \mathrm{y}') = \exp(-\frac{\mathrm{x}'^2}{2\sigma_{\mathrm{x}}^2} - \frac{\mathrm{y}'^2}{2\sigma_{\mathrm{y}}^2}) \quad (29)$$

σ_x et σ_y : sont des constantes d'espace de l'enveloppe gaussienne qui déterminent l'étendue de l'onde suivant les axes x et y respectivement.

(x_0, y_0) : représente le point d'origine où s'applique la fonction f_{Gb} (f_{Gb} est maximale en ce point).

$(x', y') = (xcos\gamma + ysin\gamma, -xsin\gamma + ycos\gamma)$ (γ: représente l'angle de rotation de (x', y') par rapport à (x, y) et définit l'orientation de l'enveloppe gaussienne dans le domaine spatial).

Un filtre de Gabor est un filtre de convolution obtenu en appliquant cette fonction à un masque de convolution.

Un filtre de Gabor est complètement défini par la connaissance de six paramètres à savoir l'orientation de l'enveloppe gaussienne γ, l'orientation de l'onde sinusoïdale θ , les deux paramètres de position (μ_0, υ_0) ou (F_0, γ) et les deux paramètres d'étalement $(\sigma_\mu^2, \sigma_v^2)$ ou bandes radiale et transversale B et Ω tels que :

$$\theta = arctg\frac{\mu_0}{\upsilon_0} \tag{30}$$

$$F_0 = \sqrt{{\mu_0}^2 + {\upsilon_0}^2} \ , \ B = \log_2 \frac{\pi\sigma_\mu F_0 + \alpha}{\pi\sigma_\mu F_0 - \alpha} \tag{31}$$

$\Omega = 2Arctg\frac{\alpha}{\pi\sigma_\upsilon F_0}$, $\quad avec\ \alpha = \sqrt{\frac{\log(\frac{1}{s})}{2}}$: désigne un seuil de troncature.

Les paramètres de texture sont donc déterminés en calculant la moyenne et l'écart type des images filtrées par Gabor. Le vecteur caractéristiques de la texture va contenir un ensemble de valeurs égal au nombre d'échelles multiplié par le nombre d'orientations.

2.3.2.3 La transformée en ondelettes

L'analyse par ondelettes est un outil mathématique capable de transformer un signal d'énergie finie dans le domaine spatial en un autre signal d'énergie finie dans le domaine spatio-fréquentiel. Les composantes de ce nouveau signal sont appelées les coefficients d'ondelettes. Ces coefficients renseignent sur la variation locale des niveaux de gris autour d'un pixel donné de l'image. Son utilisation est justifiée par le fait qu'elle offre une nouvelle représentation de l'image dans laquelle l'information texturale est facilement accessible.

Les ondelettes ont été introduites pour la première fois par Grossman et Morlet [49] comme un outil mathématique d'analyse des signaux sismiques, et depuis, elle a été largement utilisée dans des applications diverses en traitement d'images tels que le débruitage et la restauration [50] et la compression d'images.
Cette méthode fera l'objet de notre étude et sera détaillée au chapitre suivant.

2.4 La description de la forme

La forme est un descripteur très important dans l'indexation d'images. Les descripteurs (ou paramètres) de forme sont des nombres qui représentent chaque forme et permettent de les classer [51]. La description d'une forme dans une image se fait par une description basée région ou par une description basée contour.

2.4.1 Les descripteurs géométriques

Ce sont des descripteurs de formes qui nécessitent une segmentation en région préalable de l'image, Ils permettent de distinguer les différents types de formes que peuvent prendre les objets d'une scène. Ces moments sont calculés sur les différentes régions de l'image [52].

- La surface relative S_k d'une région R_k qui correspond au nombre de pixels contenus dans cette région par rapport au nombre total de pixels de l'image de N lignes et M colonne :

$$S_k = \frac{card(R_k)}{N*M} \tag{32}$$

- Le centre de masse des pixels de la région est défini comme suit :

$$P = (P_i, P_j) = \left(\frac{\sum_{i \in R_k} i/card(R_k)}{M}, \frac{\sum_{j \in R_k} j/card(R_k)}{N} \right) \tag{33}$$

- La longueur du contour de la région qui définit le nombre de pixels en bordure de la région :

$$l_k = card(contour(R_k)) \quad (34)$$

- La compacité qui traduit le regroupement des pixels de la région en zones homogènes non trouées

$$C_k = \frac{{l_k}^2}{S_k} \quad (35)$$

Ces attributs permettent d'obtenir des informations sur la géométrie des régions de l'image.

2.4.2 Les moments géométriques

Les moments géométriques permettent de décrire une forme à l'aide de propriétés statistiques [53]. Ils sont simples à manipuler mais leur temps de calcul est très long. La formule générale des moments géométriques pour décrire les formes présentes dans une image, est définie comme suit:

$$m_{p,q} = \sum_{p=0}^{m} \sum_{q=0}^{n} x^p y^q \ f(x,y) \quad (36)$$

L'ordre du moment est p+q, $f(x,y)$ représente un objet dans l'image.
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* $m_{1,0}$ permettent de calculer les coordonnées (x_c, y_c) du centre de gravite de l'objet.

$$x_c = \frac{m_{1,0}}{m_{0,0}} \qquad et \qquad y_c = \frac{m_{0,1}}{m_{0,0}}$$

Une caractéristique très importante de ces moments est qu'ils permettent d'avoir des informations sur l'ellipse équivalente à l'objet. Ce qui revient à la détermination des axes principaux de l'ellipse. Il faut ramener les moments d'ordre 2 au centre de gravite :

$$\mathrm{m_{2,0}}^{\mathrm{g}} = \mathrm{m_{2,0}} - \mathrm{m_{0,0}}\ \mathrm{x_c}^2 \quad (37.a)$$

$$\mathrm{m_{1,1}}^{\mathrm{g}} = \mathrm{m_{1,1}} - \mathrm{m_{0,0}}\ \mathrm{x_c}\ \mathrm{y_c} \quad (37.b)$$

$$m_{0,2}{}^{g} = m_{0,2} - m_{0,0}\, y_c{}^{2} \quad (37.c)$$

On peut déterminer l'angle d'inclinaison de l'ellipse α

$$\alpha = \frac{1}{2} arctan \frac{2m_{1,1}{}^{g}}{m_{2,0}{}^{g} - m_{0,2}{}^{g}} \quad (38)$$

A partir des moments géométriques, Hu [54] a introduit sept moments invariants aux transformations géométriques, dits moments de Hu.

- **Les moments de Hu**

Hu [54] a permit de décrire une forme à l'intérieur d'une image à l'aide de propriétés statistiques. Les moments de Hu sont simples à manipuler, robustes aux changements d'échelle, la translation et la rotation, mais leur temps de calcul est très long et sont sensibles aux bruits. Cette méthode a été le sujet de plusieurs recherches. Dans sa thèse, Landré [11] a caractérisé les formes des images de la base paléontologique par les moments de Hu pour des imagettes approximation résultantes d'une décomposition en ondelettes.
Ce descripteur fera l'objet de notre caractérisation de forme et il sera détaillé au chapitre suivant.

2.4.3 Les moments orthogonaux

Comme leur nom l'indique, les moments orthogonaux sont définis dans une base orthogonale, ce qui évite la redondance des informations portées par chacun des moments. Les moments orthogonaux les plus connus sont les moments de Legendre et les moments de Zernike.

2.4.3.1 Les moments de Legendre

Les moments de Legendre se basent sur les polynômes de Legendre. Ils sont définis sur un carré unité [-1,1] x [-1,1] ce qui nous mène à normaliser les objets dont on veut calculer ces moments.
Le polynôme de Legendre d'ordre n est défini comme suit :

$$\forall x \in [-1,1], \forall\, n \in \mathbb{N},\ P_n(x) = \frac{1}{2^n n!}\frac{d^n(x^2-1)^n}{dx^n} \quad (39)$$

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

$$\forall (x,y) \in [-1,1]^2, \forall\, (m,n) \in \mathbb{N}^2,\ \iint_{-1}^{1} P_m(x)\, P_n(y) dxdy = \frac{2}{2m+1}\delta_{mn} \quad (40)$$

δ_{mn} : représente le symbole de Kronecker.
Les moments de Legendre d'ordre n sont donnés par la formule suivante :

$$\forall (x,y) \in [-1,1]^2, \forall\, (p,q) \in \mathbb{N}^2, \qquad N = p + q,$$
$$L_{pq} = \frac{(2p+1)(2q+1)}{4}\iint_{-1}^{1} P_p(x)\, P_q(y) f(x,y) dxdy = \frac{2}{2m+1}\delta_{mn} \quad (41)$$

$f(x,y)$ *:* Le niveau de gris d'un pixel de l'image I sur lequel on calcule le moment.

2.4.3.2 Les moments de Zernike

Une autre méthode fréquemment utilisées pour extraire les caractéristiques globales d'une image en reconnaissance et en analyse d'images est la méthode des moments de Zernike. Introduite pour la première fois par Teague [55]. Ce descripteur de forme a prouvé sa supériorité sur les autres fonctions de moments ce qui revient à sa capacité de description et sa robustesse aux bruits et aux déformations. Dans [56], les auteurs ont tenté d'améliorer les moments de Zernike, en remplaçant le terme factoriel par la formule de Stirling afin de limiter le temps de calcul de ces moments.

Les moments de Zernike sont définis comme des polynômes complexes [43] qui forment un ensemble orthogonal complet du disque unité $x^2 + y^2 \leq 1$.

Mathématiquement, les moments de Zernike sont définis avec un ordre p et une répétition q sur un ensemble $\boldsymbol{D}$, tel que :

$$\boldsymbol{D} = \{(\boldsymbol{p},\boldsymbol{q}) \,/\, \boldsymbol{0} \leq \boldsymbol{p} < +\infty, \boldsymbol{0} \leq |\boldsymbol{q}| \leq \boldsymbol{p} \; \boldsymbol{et} \; |\boldsymbol{p} - \boldsymbol{q}| = \boldsymbol{pair}\}. \quad \boldsymbol{p} \in \mathbb{N}^* \; \boldsymbol{et} \; \boldsymbol{q} \in \mathbb{N}^*$$

$$A_{pq} = \frac{p+1}{\pi} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y)[V_{p,q}(x,y)]^* \tag{42}$$

(* définit le complexe conjugué)

En coordonnées polaires, les polynômes de Zernike sont connus sous la formule suivante :

$$V_{p,q}(r,\theta) = R_{p,q}(r) exp(-jn\theta) \tag{43}$$

Où $R_{p,q}(r)$ est le polynôme radial orthogonal.

$$R_{p,q}(r) = \sum_{s=0}^{(p-|q|)/2} \frac{(-1)^s (p-s)!}{s!\left(\frac{p+|q|}{2}-s\right)!\left(\frac{p-|q|}{2}-s\right)!} r^{p-2s} \tag{44}$$

$$r = \sqrt{x^2 + y^2}$$

r donne la longueur du vecteur de l'origine du pixel *(x,y) et θ=atang2(y/x).*

2.4.4 La matrice norme de gradient

La norme de gradient [56] est une autre méthode de caractérisation de forme qui présente le degré de différence d'intensité d'un pixel par rapport à ses voisins. Elle permet de distinguer les différentes régions et les différents objets dans une image. La norme de gradient d'une image peut être calculée à l'aide de différents filtres (filtre de Sobel, Perwit, le filtre Laplacien, etc…). Le principe de cette méthode est basé sur une convolution de l'image niveaux de gris avec les masques $h_x \; et \; h_y$, pour le calcul des gradients directionnels selon x et y. Deux image sont donc extraites par cette convolution sont $G_x \; et \; G_y$.

La matrice norme de gradient correspondante sert à combiner, en chaque pixel de l'image, les approximations des gradients horizontaux et verticaux

pour obtenir une approximation globale de la norme du gradient. Elle est donnée par la formule suivante.

$$G = \sqrt{G^2{}_x + G^2{}_y} \quad (45)$$

2.5 Conclusion

Dans ce présent chapitre nous avons dressé un état de l'art sur les méthodes existantes pour l'indexation d'images. Au départ, nous avons tenté de définir les différents attributs les plus pertinents pour la caractérisation d'une image qui sont la couleur, la texture et la forme. Nous avons défini la caractéristique couleur qui est généralement représentée par trois composantes qui définissent un espace couleur. On a aussi vu que parmi les méthodes de caractérisation de la couleur celles qui ne tiennent pas compte de l'information sur la localisation spatiale des couleurs dans l'image et celles qui tiennent compte de cette information. On a ensuite passé à définir la texture qui peut être caractérisée par des calculs statistiques sur les pixels de l'image (matrice de cooccurrence…) ou par une analyse fréquentielle (ondelette et filtre de Gabor…). La troisième caractéristique est la forme qui est aussi un attribut essentiel pour représenter une image, ce qui a conduit au développement d'une multitude de méthodes, celles basées sur la caractérisation contours [57], et celles basées sur les régions [58].

Dans ce chapitre nous avons choisi les histogrammes couleurs HSV pour la caractérisation de la couleur, les ondelettes pour la texture et enfin nous avons préféré les moments de Hu pour la caractérisation de la forme.

Chapitre III

Tests et résultats

3.1 Introduction

Une image numérique est représentée par une matrice de valeurs codées sur un certain nombre de bits dont le traitement ne permet pas souvent d'obtenir des temps de réponse escomptés. Dans certains cas, il est nécessaire d'utiliser une représentation de dimension réduite pour mieux la caractériser. Ainsi, des attributs caractéristiques de l'image sont extraits à l'aide de fonctions mathématiques arrangés sous la forme d'un vecteur dit descripteur de l'image.

L'objectif de ce chapitre est de tester les méthodes choisies pour la caractérisation globale d'une image dans le but de rechercher des images par leur contenu visuel en se basant sur les trois attributs, de couleur, de texture et de forme.

La caractérisation de la couleur est réalisée par la méthode des histogrammes. La base Coil 100 [59] représentant des images couleur est utilisée pour nos tests et ce, en utilisant l'histogramme de l'image HSV où H est la teinte (Hue), S, la saturation et V, la luminosité, la valeur ou la brillance. Les autres bases d'images sont en niveaux de gris. La couleur est donc indexée par les histogrammes niveaux de gris. Pour la texture, la

méthode choisie est la méthode des ondelettes. Concernant la forme, nous avons opté pour les moments de Hu pour leur facilité d'implémentation et leur invariance aux transformations géométriques. Nous avons aussi appliqué ces méthodes sur des images requêtes dégradées par un bruit gaussien pour, ensuite, conclure sur leurs sensibilités à ce type de bruit.

3.2. Indexation de la couleur

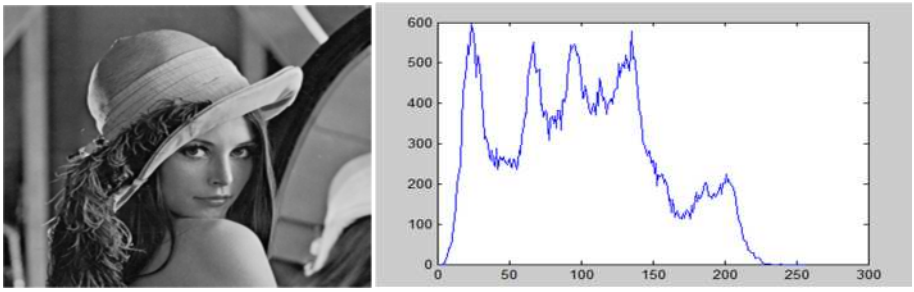
Comme nous l'avons déjà vu au chapitre précédent, la couleur est un attribut fondamental dans l'indexation d'une image. Pour des bases d'images en niveaux de gris, l'histogramme est en niveau de gris alors que pour les images en couleur, la caractérisation se fait par un histogramme couleur dans l'espace HSV.

3.2.1 Histogramme couleur

En indexation, une représentation pratique la plus connue de la couleur est sans doute l'histogramme. Cette méthode a été utilisée pour la première fois par Swain et Ballard [4]. Hafner et *al.* [23]. Le calcul de l'histogramme se fait en comptant, pour chaque couleur, le nombre de pixels de cette couleur contenus dans l'image ou par utilisation de trois histogrammes suivant l'espace couleur utilisé, un pour chaque canal. Si l'image I est de dimension M*N, contenant un nombre C de couleurs, son histogramme normalisé pour chaque couleur c est donc calculé par la formule suivante :

$$h(c) = \frac{1}{M*N} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \delta(\text{couleur}(i, j), c) \quad \forall c \in C \qquad (1)$$

où δ est le symbole de Kronecker tel que δ(x, y)= 1 si x=y et δ(x, y)=0, sinon.



(a) (b)

Fig.1 (a) Image de Lena. (b) son histogramme des niveaux de gris.

L'histogramme est très utilisé dans la recherche d'images par le contenu pour sa simplicité de calcul, son invariance aux changements d'échelle et aux transformations géométriques. Dans l'espace HSV, une image couleur I est représentée par trois images I_H, I_S et I_V, relatives aux trois composantes respectivement de cet espace : la teinte (H), la saturation (S) et la valeur (V).

3.2.2. Mesure de similarité entre histogrammes

Un histogramme détermine la distribution statistique de chaque couleur. La distance usuelle la plus utilisée pour comparer deux histogrammes est la distance euclidienne. Elle est donnée par la formule suivante :

$$Dist(H_{req}, H_{cour}) = \sqrt{\sum_{k=1}^{K} |h_{req}(k) - h_{cour}(k)|^2} \quad (2)$$

où H_{req} est l'histogramme de l'image requête et $h_{cour}(k)$ celui de l'image en cours. K représente la dimension du vecteur histogramme (K=3*256 pour une image couleur HSV).

3.3. Indexation de la texture

La caractérisation de la texture nous permet d'avoir une information sur l'organisation des niveaux de gris dans une image. Dans ce travail, nous avons choisi la méthode basée sur la transformation d'ondelettes de Daubechies.

3.3.1 La transformée en ondelettes continue

Dans le domaine du traitement du signal, l'analyse par ondelettes est un outil mathématique capable de transformer un signal d'énergie finie dans le domaine spatial en un autre signal d'énergie finie dans le domaine spatio-fréquentiel. Les composantes de ce nouveau signal sont appelées les coefficients d'ondelettes.

Rappelons que cette méthode a été introduite pour la première fois par Morlet [49] comme un outil mathématique d'analyse des signaux sismiques. Elle a été largement utilisée dans des applications diverses en traitement d'images tels que le débruitage et la restauration d'images. Elle a été prouvée efficace pour la caractérisation de la texture [46], dont l'idée est d'extraire l'énergie portée par le signal dans les bandes de fréquences diverses.

3.3.2. Principe de la décomposition en ondelettes discrètes

La transformée en ondelettes continue consiste à décomposer un signal d'entrée x(t) en une série de fonctions d'ondelettes $\psi_{a,b}(t)$ qui dérivent d'une fonction mère ψ(t) donnée par des opérateurs de dilatation 'a' et de translation 'b'.

$$C(a,b) = \int_{-\infty}^{+\infty} x(t)\ \ \psi^*_{a,b}(t)dt \quad (3)$$

$$\text{Avec } \psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), \quad (4)$$

ψ est à moyenne nulle et à énergie finie avec $a \neq 0$, * désigne le complexe conjugué.

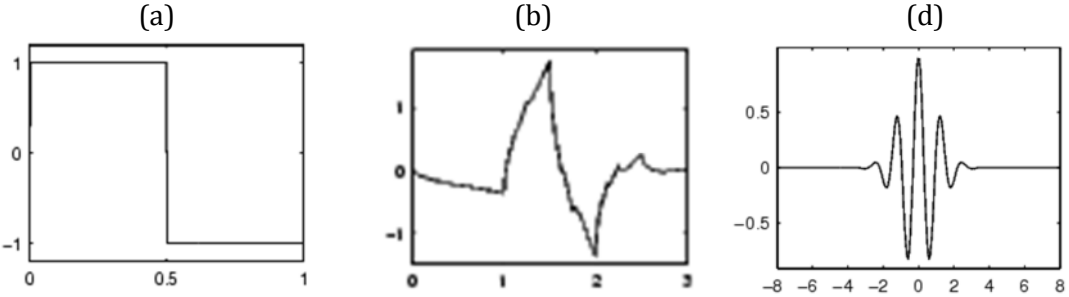


Fig.2 Quelques ondelettes les plus usuelles : (a) ondelette de Haar, (b)ondelette de Daubechies db2, (c)ondelette de Morlet.

Dans l'objectif d'appliquer efficacement la transformée en ondelettes aux signaux discrets telles que les images, il convient de discrétiser les coefficients de dilatation a et de translation b.

$a = {a_0}^m$ et $b = nb_0{a_0}^m$

a_0 , b_0 , n, $m \in \mathbb{Z}$.

Si on choisit $a_0 = 2$ et $b_0 = 1$ on se place alors dans le cas dyadique.

La fonction de la formule (3) s'écrit :

$$C(m,n) = 2^{\frac{-m}{2}} \int_{-\infty}^{+\infty} \mathrm{x(t)}\, \psi(2^{-m}\mathrm{t} - n)\mathrm{dt} \qquad (5)$$

3.3.4 Analyse Multirésolution

L'analyse multirésolution introduite pour la première fois par Mallat [60] est un outil de traitement du signal qui permet de décomposer un signal à plusieurs résolutions. Les éléments de cette décomposition permettent de reconstruire le signal original.

Elle est définie par une suite de sous-espaces $\{V_j\}_{j\in\mathbb{Z}}$ de $L^2(\mathbb{R})$, emboîtés les uns dans les autres. Le passage de l'un à l'autre est le résultat d'un changement d'échelle. Ces sous-espaces sont appelés des espaces d'approximation à l'échelle vérifiant les propriétés suivantes :

$$\left.\begin{array}{l} \forall\, j \in \mathbb{Z},\ V_{j+1} \subset V_j \\ \overline{\cup\, J \in \mathbb{Z},\ V_j} = L^2(\mathbb{R}) \\ \cap\, j \in \mathbb{Z},\ V_j = 0 \\ \forall\, j \in \mathbb{Z},\ f(x) \in V_j \Leftrightarrow f(2^{-1}x) \in V_{j+1} \\ \forall\, k \in \mathbb{Z},\ f(x) \in V_0 \Leftrightarrow f(x-k) \in V_0 \end{array}\right\} \qquad (6)$$

L'idée principale de l'analyse multirésolution est de projeter un signal $f(x) \in L^2(\mathbb{R})$ appartenant à un sous-espace V_j sur un sous-espace d'approximation V_{j+1} et un sous-espace de détail W_{j+1} pour réduire la résolution de moitié. La figure (3) donne le schéma de l'analyse multi-résolution.

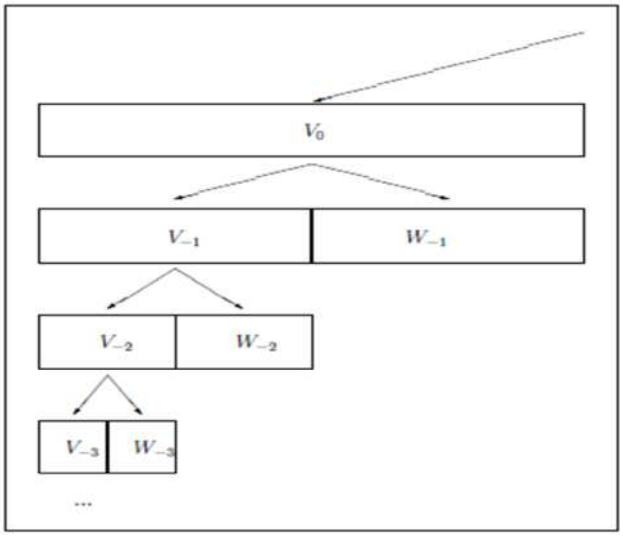


Fig. 3. Principe de l'analyse multirésolution.

On définie deux opérateurs de projection A_j et D_j qui projettent respectivement le signal sur V_{j+1} et W_{j+1}. L'analyse multirésolution se base sur une fonction d'échelle $\phi(t) \in L^2(\mathbb{R})$ qui engendre la base orthonormée V_{j+1} par dilatation et translation, et une fonction d'ondelette $\psi(t) \in L^2(\mathbb{R})$ qui engendre par dilatation et translation une base orthonormée W_{j+1}.

V_{j+1} et W_{j+1} sont complémentaires tel que

$$V_j = V_{j+1} \oplus W_{j+1}. \tag{7}$$

Les fonctions de base dilatées sont données par les relations suivantes :

$$\phi_{j,n}(t) = 2^{\frac{-j}{2}} \phi\left(2^{-j}t - n\right) \quad n \in \mathbb{Z} \tag{8}$$

$$\psi_{j,n}(t) = 2^{\frac{-j}{2}} \psi\left(2^{-j}t - n\right) \quad n \in \mathbb{Z} \tag{9}$$

L'approximation est caractérisée par la relation suivante :

$$A_j f = \sum_n < f, \phi_{j,n} > \phi_{j,n} \tag{10}$$

$< f, \phi_{j,n} > = \int_{-\infty}^{+\infty} f(t) \phi^*_{j,n}(t)\, dt$: Le produit scalaire de f par ϕ.

Les coefficients détails se calculent comme suit :

$$D_j f = \sum_n < f, \psi_{j,n} > \psi_{j,n} \tag{11}$$

$< f, \psi_{j,n} > = \int_{-\infty}^{+\infty} f(t) \psi^*_{j,n}(t)\, dt$: Le produit scalaire de f par ψ.

Stéphane Mallat a donné un algorithme d'analyse (ou décomposition) en ondelettes qui permet d'obtenir une analyse multirésolution du signal. Cet algorithme travaille par filtrage de l'image suivant les lignes puis les colonnes par deux filtres, $\tilde{g}$ passe-haut et $\tilde{h}$ passe-bas. $\tilde{h}$ va permettre de repérer les basses fréquences dans l'image (l'approximation) et $\tilde{g}$ les hautes fréquences (les détails). $\tilde{h}$ et $\tilde{g}$ sont construits à partir des fonctions ψ et ϕ. La figure (4) donne le schéma de décomposition de Mallat.

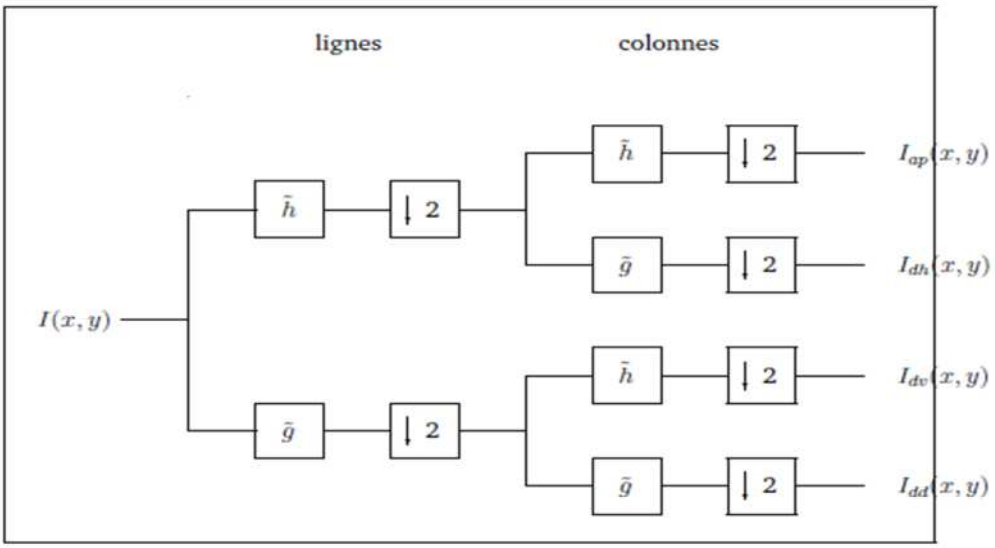


Fig.4. Algorithme d'analyse de Mallat.

Dans le cadre de notre travail, nous avons utilisé les ondelettes de Daubechies 4 qu'on appelle db2 qui sont des ondelettes à support compact [61] permettant d'utiliser des filtres de taille finie. Nous avons effectué trois niveaux de décomposition. Pour chaque niveau de décomposition, nous avons extrait l'énergie, la moyenne et l'écart type des trois imagettes correspondantes aux : détail horizontal, détail vertical et au détail diagonal. En effet, à partir de la transformée en ondelettes on peut extraire des attributs de différents types et à différents niveaux de résolution. L'image d'approximation donne des informations sur les régions qui composent l'image, d'une résolution fine à une résolution grossière. Les images de détails donnent des informations horizontales, verticales et diagonales sur l'image. L'un des indices le plus utilisé pour caractériser la texture dans le

plan spatio-fréquentiel est la mesure d'énergie sur les images de détails. L'indice énergie est une mesure locale de la distribution des coefficients d'ondelette $C(i,j)$ en fonction de la fréquence, de l'orientation et de l'échelle. L'expression de l'énergie exprimée sous forme normalisée est donnée par :

$$E = \frac{1}{N*M}\sum_{i=1}^{N}\sum_{j=1}^{M} C(i,j)^2 \quad (12)$$

Le deuxième indice, extrait lui aussi par les images de détails, que nous exploitant pour la caractérisation de la texture est la moyenne. Elle mesure la moyenne des valeurs prises par les échantillons du signal image. Sa formule est donnée comme suit :

$$M = \frac{1}{N*M}\sum_{i=1}^{N}\sum_{j=1}^{M} |C(i,j)| \quad (13)$$

Un autre indice extrait est l'écart type qui représente la racine carrée de la variance. L'écart type indique comment, en moyenne, les valeurs de la variable sont groupées autour de la tendance centrale (moyenne arithmétique).

$$\sigma^2 = \frac{1}{N*M}\sum_{i=1}^{N}\sum_{j=1}^{M} (C(i,j) - M)^2 \quad (14)$$

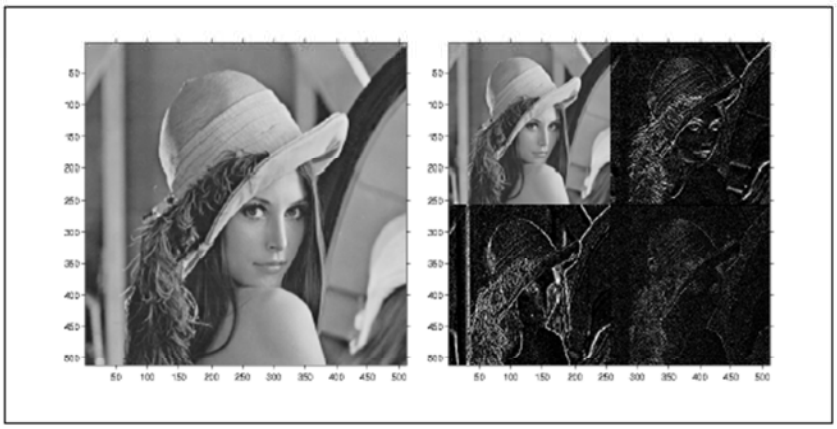


Fig.5. Un exemple de décomposition en ondelettes.

Le vecteur descripteur de texture extrait, contient les valeurs des énergies, des moyennes et des écarts type de toutes les images de détail résultantes des trois niveaux de décomposition. Il est constitué de 27 valeurs.

Soient :

- $Eh_1, Eh_2, Eh_3, Mh_1, Mh_2, Mh_3, Eth_1, Eth_2$ et Eth_3 les énergies, les moyennes et les écarts type des images de détails horizontaux suivant les trois décompositions de niveau 1, 2 et 3 respectivement.
- $Ev_1, Ev_2, Ev_3, Mv_1, Mv_2, Mv_3, Etv_1, Etv_2$ et Etv_3 les énergies, les moyennes et les écarts type des images de détails verticaux suivant les trois décompositions de niveaux 1, 2 et 3 respectivement.
- $Ed_1, Ed_2, Ed_3, Md_1, Md_2, Md_3\ Etd_1, Etd_2$ et Etd_3 les énergies, les moyennes et les écarts type des images de détails diagonaux suivant les trois décompositions de niveaux 1, 2 et 3 respectivement.

3.3.3. Mesure de similarité entre les vecteurs de texture

Pour mesurer la similarité entre un vecteur descripteur de texture de l'image requête $V_{texture\ req}$ avec celui de l'image courante $V_{texture\ cour}$, nous utilisons la distance Euclidienne comme le cas de l'équation (2) précédente. Elle est donnée par la formule suivante :

$$Dist\left(V_{texture\ req}, V_{texture\ cour}\right) = \sqrt{\sum_{n=1}^{N} \left|V_{texture\ req}(n) - V_{texture\ cour}(n)\right|^2} \quad (15)$$

Où N représente la dimension de notre vecteur caractéristique de texture.

3.4 Indexation de la forme

Contrairement aux attributs couleurs et textures qui s'intéressent à la description du contenu général de l'image, les attributs formes permettent de caractériser les différents objets contenus dans l'image. Donc l'indexation de la forme est aussi une étape très importante dans la recherche d'images par le contenu.

3.4.1. Les moments de Hu

Les moments de Hu [54] permettent de décrire la caractéristique forme à l'aide de propriétés statistiques. Ils sont simples à manipuler, robustes aux changements d'échelle, la translation et la rotation, mais leur temps de

calcul est très long et sont sensible aux bruits. Leur calcul se base sur les moments géométriques, ce qui est montré dans le chapitre précédent. Les moments centrés sont définis comme suit :

$$\mu_{pq} = \sum_{p=0}^{m} \sum_{q=0}^{n} (x - x_c)^p (y - y_c)^q \ f(x, y) \quad (16)$$

Centrés et normalisés, ces moments sont donnés par 'expression suivante :

$$\eta_{pq} = \frac{\mu_{pq}}{s^{\frac{p+q}{2}+1}} \quad (17)$$

Où $p+q \geq 2$ et s représente la surface de l'objet.

A partir des moments normalisés, Hu dans [54] a introduit les sept moments invariants aux transformations géométriques appelées les moments de Hu, donnés par les formules suivantes :

$$M_1 = \eta_{20} + \eta_{02} \quad (18)$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (19)$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (20)$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (21)$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - (3\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (22)$$

$$M_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (23)$$

$$M_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (24)$$

3.4.2 Mesure de similarité entre les vecteurs de forme

Le vecteur caractérisant la similarité est composée des sept moments de Hu.

$$V_{hu} = [M_1, M_2, M_3, M_4, M_5, M_6, M_7] \quad (25)$$

La mesure de la similarité sera un résultat d'une distance Euclidienne entre le vecteur des moments de Hu correspondant à la requête et celui correspondant à une image courante de la base d'image. Voir la formule (26).

Soient $V_{hu\,req}$ et $V_{hu\,cour}$ les deux vecteurs des moments de Hu de l'image requête et celui de l'image courante da la base respectivement.

$$Dist(V_{hu\ req}, V_{hu\ cour}) = \sqrt{\sum_{k=1}^{7} |V_{hu\,req}(k) - V_{hu\ cour}(k)|^2} \quad (26)$$

3.5 Protocole d'évaluation d'un système de recherche d'images par le contenu

Dans le cas général, lorsqu'un utilisateur interroge une base de données, que ce soit un logiciel documentaire ou un moteur de recherche, il attend un nombre de réponses (sous forme de documents). À partir de l'ensemble de réponses obtenues mis en regard de l'attente de l'utilisateur, on peut mesurer les performances de l'algorithme de recherche mis en œuvre pour retrouver un document.

Pour évaluer un système de recherche d'images par le contenu nous nous intéressons à deux critères qui concernent sa capacité de reconnaissance. Ils s'agissent principalement du rappel et de la précision.

Ainsi, soit A un ensemble des images pertinentes pour une requête donnée et B, l'ensemble des images retournées par le système.

3.5.1 Le rappel

Le rappel est défini par le nombre d'images retrouvées au regard du nombre d'images pertinentes que possède la base de données. Cela signifie que lorsque l'utilisateur interroge la base il souhaite voir apparaître toutes les images qui pourraient répondre à son besoin d'informations. Le rappel est donc un rapport défini par le nombre d'images pertinentes retrouvées sur le nombre d'images pertinentes dans la base d'image. Il est définit comme suit :

$$Rappel = \frac{|A \cap B|}{|A|} \in [0,1] \quad (27)$$

3.5.2 La précision

La précision est le nombre d'images pertinentes retrouvées rapporté au nombre d'images total proposé par le moteur de recherche pour une requête donnée. Donnée comme suit :

$$Precision = \frac{|A \cap B|}{|B|} \in [0,1] \quad (28)$$

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. Un rappel faible indique qu'une partie de l'information pertinente ne lui sera pas accessible. Une précision faible est expliquée par une forte concentration des informations non pertinentes fournies dans les résultats, ce qui cause une non satisfaction de l'utilisateur. Le rappel et la précision sont des mesures importantes dans un système de recherche d'images par le contenu, mais une seule paire de valeur rappel-précision ne peut pas indiquer la performance du système. Il est donc évident de dessiner une courbe rappel-précision.

La courbe rappel-précision peut être obtenue en interpolant les valeurs du rappel et de la précision par la méthode de TREC [64]. Elle donne la pertinence des réponses du système de recherche d'images aux requêtes proposées. La courbe rappel-précision est illustrée par la figure (6).

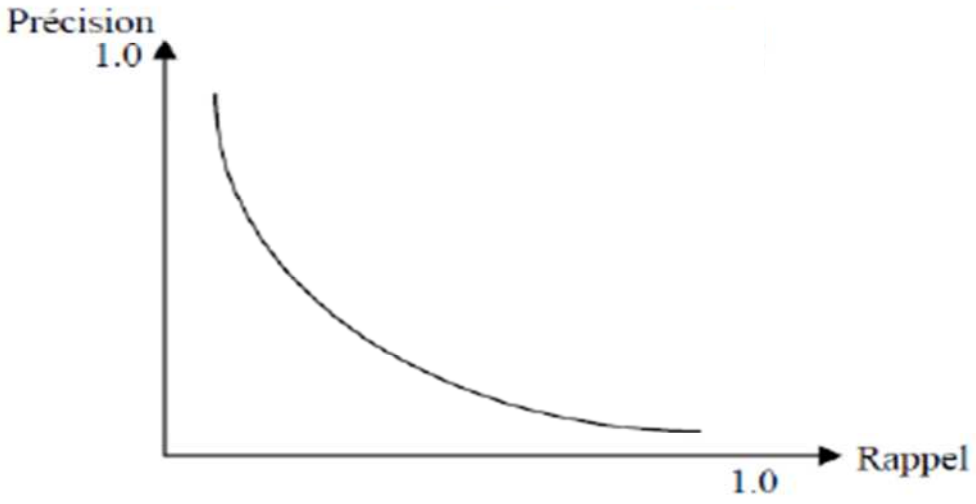


Fig.6. Allure d'une courbe de rappel-précision [18].

Une courbe rappel-précision est dite idéale si la précision est égale à un pour toutes les valeurs du rappel.

3.6 Les bases d'images utilisées

Cette section présente les quatre bases d'images utilisées, qui sont disponibles sur internet. Ces bases d'images sont très différentes par leur contenu et permettent de valider nos méthodes.

La base ORL [62]

Commençant par la base ORL qui est une base qui contient 400 images de visages de 40 personnes ayant chacune des prises de vues différentes. Les images sont en niveaux de gris de dimension 112x92 (Figure(7)).

La base COIL 100 [59]

La base Coil 100(Columbia Object Image Library) est une base d'images qui contient des images en couleur de 100 objets, photographiés sous 72 positions différentes chacun, donnant donc 7200 images. Cette base est utilisée surtout dans le domaine de reconnaissance d'objets. Ces images sont tous de même taille 128*128pixels et sous format png (Figure(8))

Fig.7. Les 40 personnes de la base ORL.

Fig.8. Les différents objets de la base COIL 100.

La base cerebral standard MR

Notre base contient 329 images médicales classées en plusieurs classes. Les images sont en niveau de gris et de dimension 512x512.

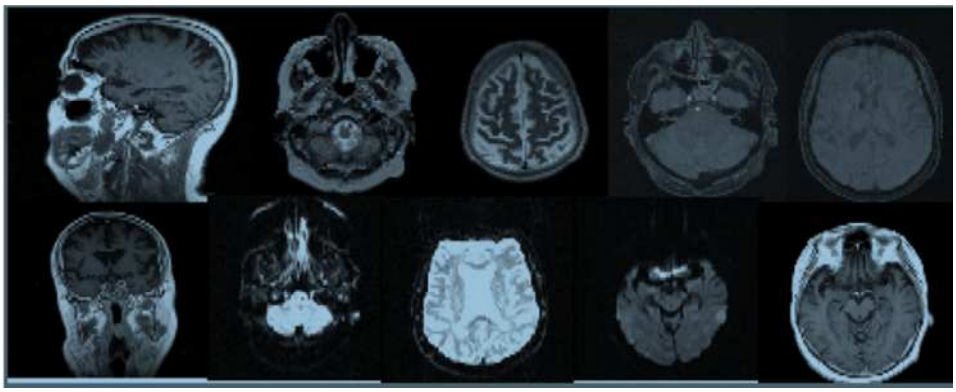


Fig.9. Quelques exemples de la base Cerebral standard MR.

La base d'images de texture

La quatrième base d'images est la base d'images texturées. Elle contient 400 images d'essai en niveau de gris de dimension 640*480.(Fig.10)

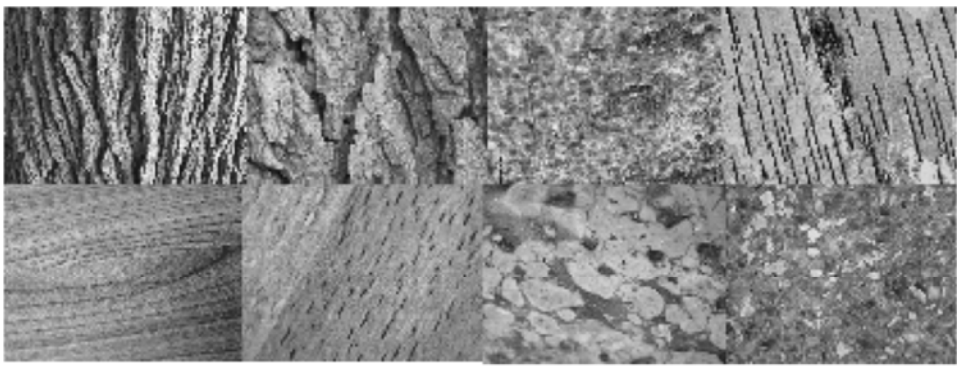


Fig.10. Quelques exemples de la base d'images de texture.

Avant de passer à l'étape d'extraction des caractéristiques, nous redimensionnons les images de toutes les bases à une dimension de 256x256 pixel.

3.7 Tests et résultats

Notre évaluation est basée sur les tableaux correspondants au douze images retrouvées pour quelques requêtes dans les différentes bases. Après une extraction des vecteurs caractéristiques, le système passe à l'étape où il parcourt toute la base pour faire une comparaison entre le vecteur requête et tous les vecteurs de la base. En utilisant la distance euclidienne, le système a ordonné les images de la base, de la distance la plus petite à la distance la plus grande. Nous avons procédé l'affichage de, seulement, douze images résultantes de la recherche.

Nos résultats correspondants aux douze images les plus similaires à la requête suivant la méthode utilisée sont affichées selon le score, de gauche à droite et de haut en bas tel que chaque image est accompagnée d'un paramètre *d* qui exprime la distance entre l'image requête et l'image correspondante. La requête est affichée en haut et à gauche des douze images résultantes.

3.7.1 La recherche d'images par histogramme

Les figures (11), (12) et (13) présentent les résultats de la recherche d'images par la caractéristique couleur dans les trois bases d'images qui contiennent des images en niveaux de gris utilisant la méthode de l'histogramme en niveau de gris. La figure (14) donne les résultats de la recherche dans la base Coil couleur contenant des images couleur à l'aide de l'histogramme HSV.

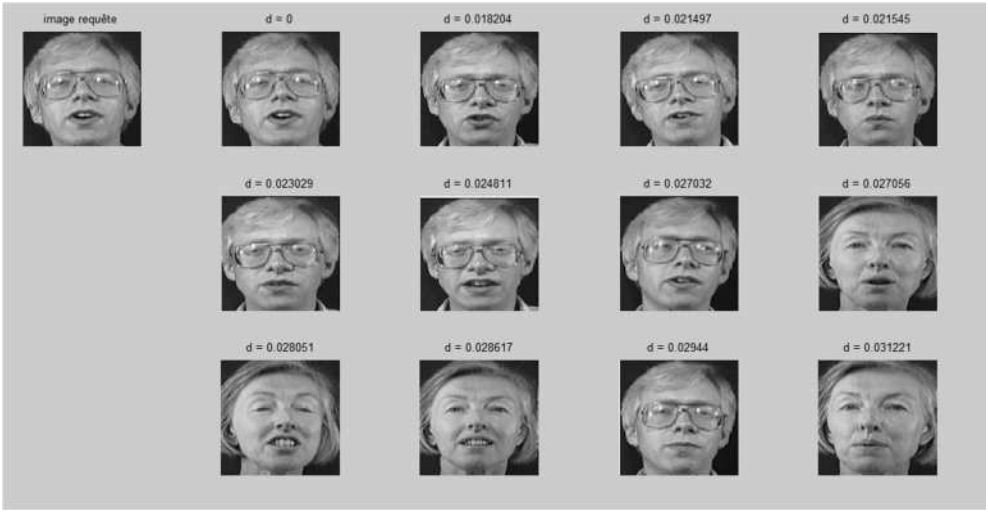


Fig.11. Résultat de la recherche dans la base ORL en utilisant la méthode l'histogramme en niveau de gris.

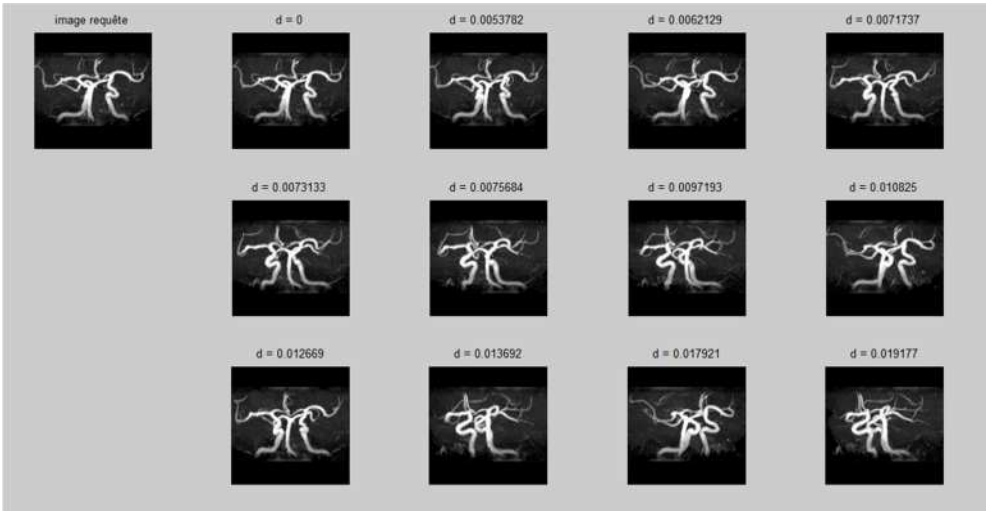


Fig.12. Résultat de la recherche dans la base d'images médicales en utilisant la méthode de l'histogramme en niveau de gris.

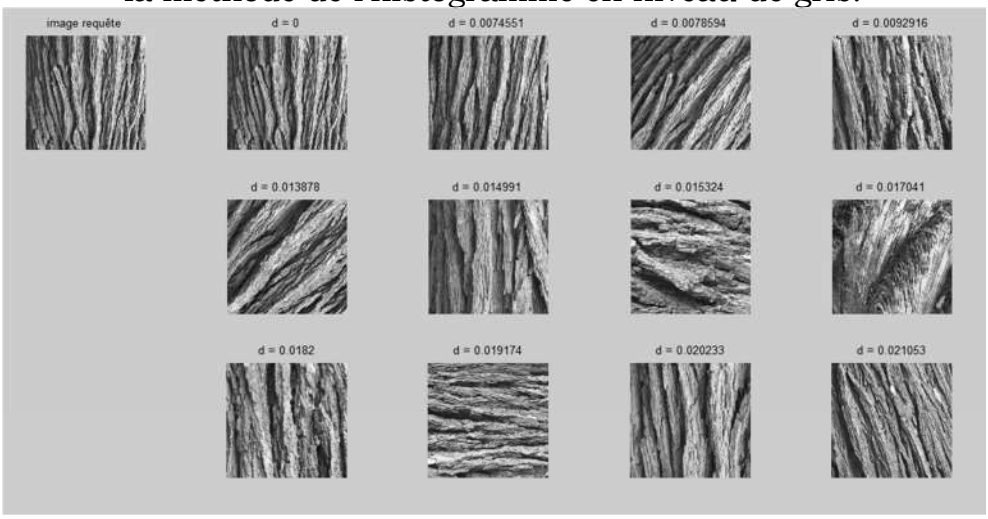


Fig.13. Résultat de la recherche dans la base des images texturées utilisant la méthode de l'histogramme en niveau de gris.

Fig.14. Résultat de la recherche dans la base Coil couleur utilisant un histogramme couleur HSV.

A partir des résultats précédents, nous pouvons remarquer qu'un histogramme de niveau de gris donne des résultats très satisfaisants ce qui permet de dire qu'il est un bon descripteur de la couleur, et de même pour l'histogramme HSV dans la base d'images en couleur.

3.7.2 La recherche d'images par la méthode des ondelettes

La texture est caractérisée par la méthode de décomposition en coefficients d'ondelettes de Daubechies, db2.

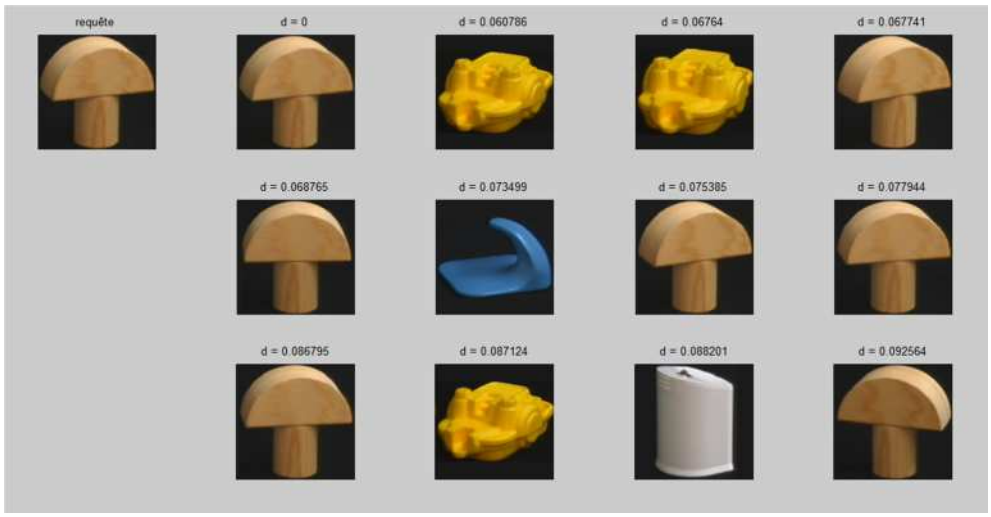


Fig.15. Résultat de la recherche dans la base d'images Coil100 en utilisant la méthode des ondelettes.

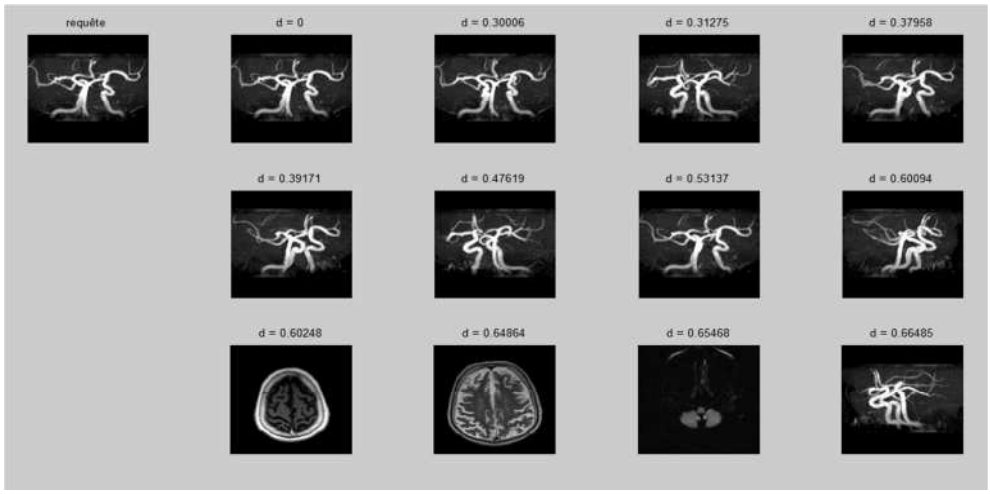


Fig.16. Résultat de la recherche dans la base d'images médicales en utilisant la méthode des ondelettes.

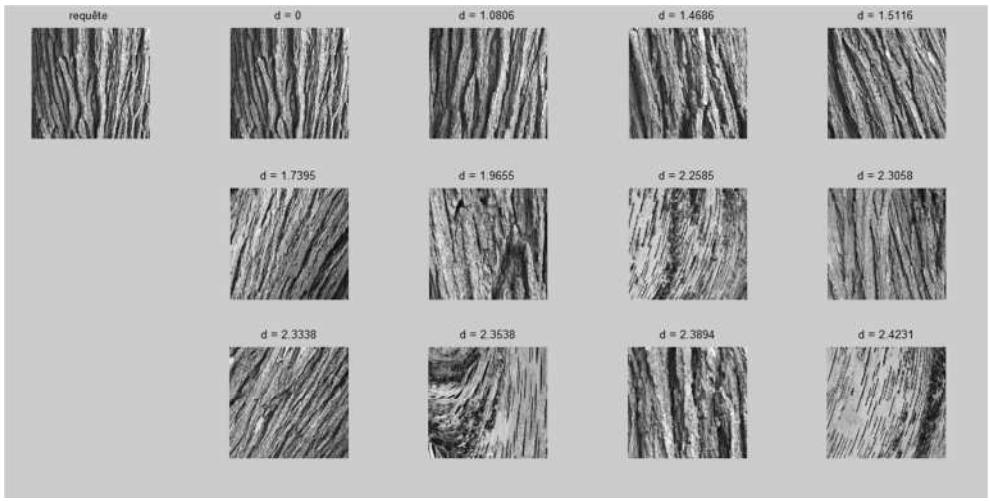


Fig.17. Résultat de la recherche dans la base d'images texturées en utilisant la méthode des ondelettes.

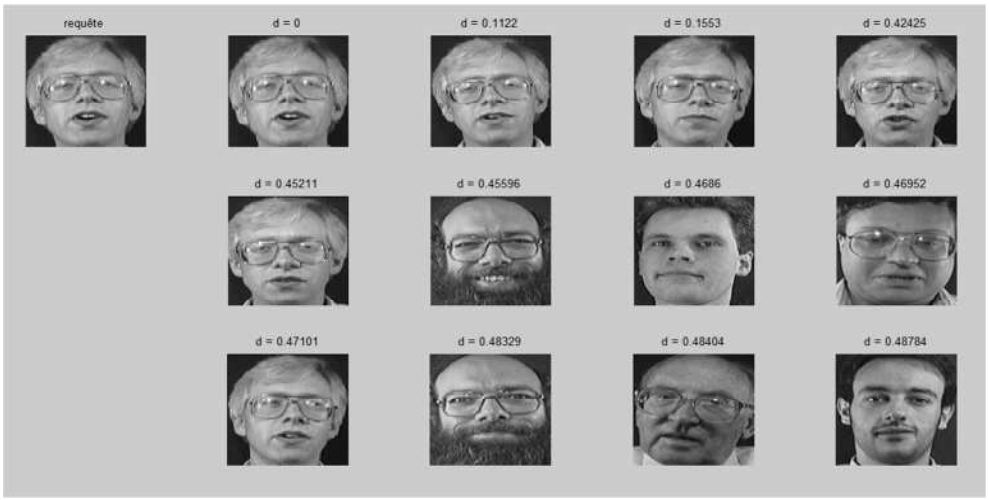


Fig.18. Résultat de la recherche dans la base d'images ORL en utilisant la méthode des ondelettes.

Après les tests effectués sur l'ensemble des bases d'images, cette méthode à donné de bon résultats pour retrouver les images similaires à une image requête ce qui prouve sa capacité dans le domaine de la recherche d'images par le contenu texture.

3.7.3 La recherche d'images par la méthode des moments de Hu

Pour une application des moments de Hu, nous avons binarisé les images par un seuillage d'Otsu [63]. Ensuite, nous avons testé l'invariance des moments de Hu par rapport aux transformations géométriques. Le tableau suivant donne les résultats d'un test sur une image de la base de visage ORL.

Tab.1. Résultat du test de l'invariance des moments de Hu par rapport aux transformations géométriques.

Moment de Hu pour l'image originale de M1 jusqu'à M7	**Moment de Hu Rotation (45^o) de M1 jusqu'à M7**	**Moment de Hu Translation de 25 pixels de M1 jusqu'à M7**	**Moment de Hu Changement d'échelle de M1 jusqu'à M7**
6.708170895014926E-4	6.714963578472591E-4	6.718660417915651E-4	6.725898992268611E-4
7.220987933770049E-9	7.341998508550092E-9	7.139489856793856E-9	3.137868359950515 3E-9
6.593641291533323E-13	6.384429216714055E-13	6.555332039783197E-13	2.726327010371615E-12
6.526494607308844E-15	7.523464135814487E-15	6.841803170344383E-15	2.040258120547275 3E-14
5.450107091688783E-29	-1.736526780735399 7E-28	1.5249595086304892E-31	2.607847113136761E-28
2.548311072441718E-19	2.289172297012794E-19	1.9548504333449945E-19	-6.914290041727305E-19
-3.204606694581758E-28	1.1625340203514935E-28	-3.698618467516282E-28	-2.733907507686838 2E-27

Les résultats donnés dans le tableau (1) sont presque identiques ce qui montrent que les moments de Hu sont invariant aux différentes transformations géométriques.

Les quatre figures suivantes donnent les résultats de l'application des moments de Hu pour les quartes bases utilisées. Pour une comparaison des résultats, nous avons effectué nos tests sur les mêmes images requêtes utilisées par l'application des deux méthodes précédentes.

Fig.19. Résultat de la recherche dans la base Coil 100 en utilisant les moments de Hu.

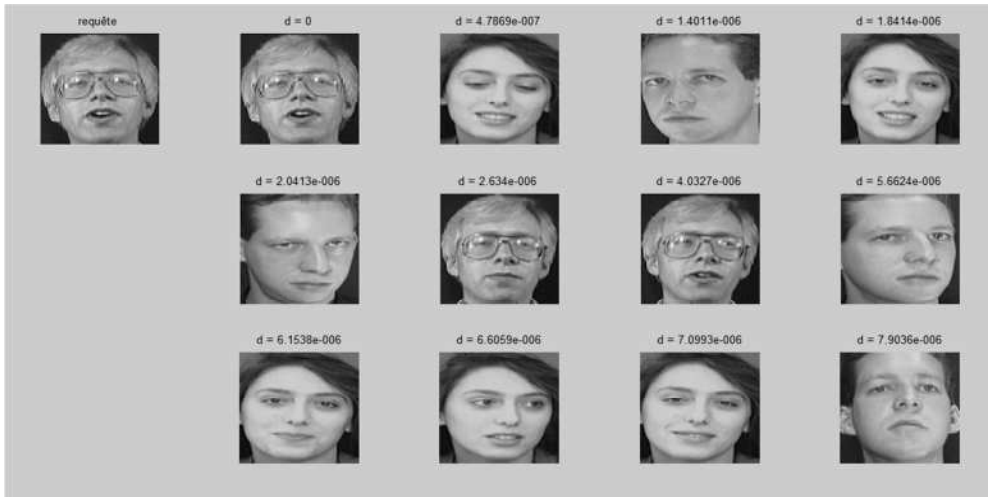


Fig.20. Résultat de la recherche dans la base ORL en utilisant les moments de Hu.

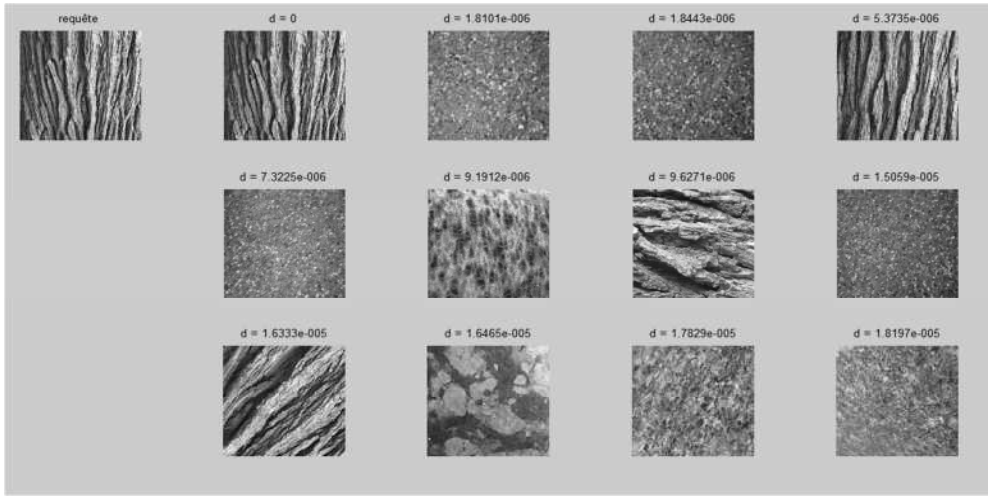


Fig.21. Résultat de la recherche dans la base d'images texturées en utilisant les moments de Hu.

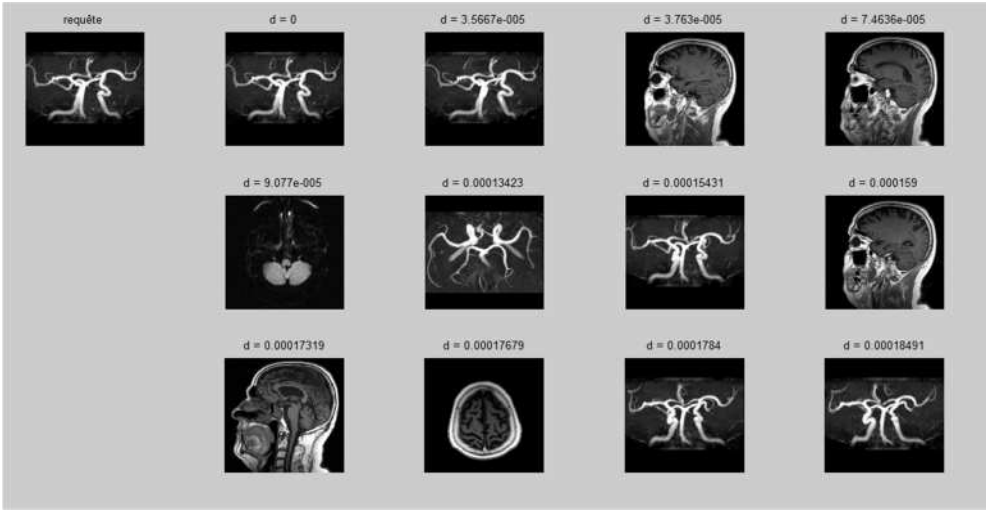


Fig.22. Résultat de la recherche dans la base d'images médicales en utilisant les moments de Hu.

On peut constater que les moments de Hu représentent une autre méthode qui peut rechercher des images similaires à une image requête proposée.

3.8 La qualité des réponses

A partir des images résultantes données par la figure(11), on remarque que la méthode d'histogramme niveau de gris permet une bonne extraction d'image similaire à l'image requête proposée.

La méthode extrait 8 images pertinentes parmi les 10 images pertinentes présentes dans la base de visage ORL. Pour remplir le tableau (2) de rappel-précision obtenu par la méthode de l'histogramme en niveaux de gris, il faut noter les états de chaque image résultante de image1 à image12 en terme de rappel et précision.

Tab.2. Le tableau contenant les valeurs de rappel-précision pour une requête donnée.

Image affichées	Précision	Rappel
image1	1/1	1/10
Image2	2/2	2/10
image3	3/3	3/10
image4	4/4	4/10
image5	5/5	5/10
image6	6/6	6/10
image7	7/7	7/10
image8	7/8	7/10
image9	7/9	7/10
image10	7/10	7/10
image11	8/11	8/10
image12	8/12	8/10

Mais pour dessiner la courbe Rappel-précision, nous devons passer par une étape d'interpolation des valeurs du rappels et celles de la précision. Dans notre travail, nous avons choisis la méthode de TREC [64].

3.8.1 Principe de la méthode de TREC [64]

Le principe de la méthode est de commencer par une interpolation des valeurs du rappel $r_{calculé}$ pour avoir les valeurs interpolées r_{int} . L'idée générale est de fixer les valeurs du rappel entre 0 et 1 avec un pas de 0.1 (onze nouvelles valeurs de rappel). Pour le calcul des valeurs de la précision interpolées P_{int}, on suit l'algorithme suivant :

$$P_{int}(r_{int}) = \max_{r_{calculé} \geq r_{int}} [P(r_{calculé})] \quad (29)$$

Les résultats de cette interpolation sont donnés dans le tableau (3).

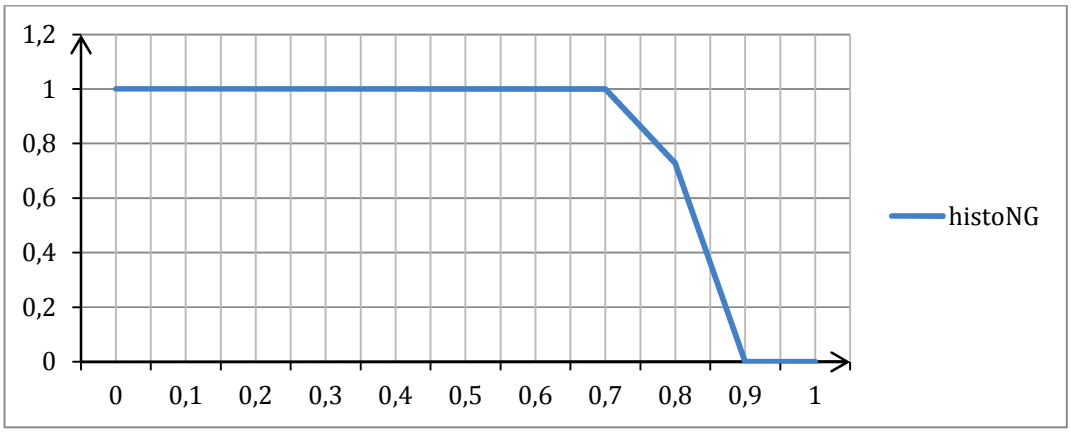
Tab. 3. Résultats de l'interpolation du rappel et la précision.

P_{int}	1	1	1	1	1	1	1	1	0.727	0	0
r_{int}	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1

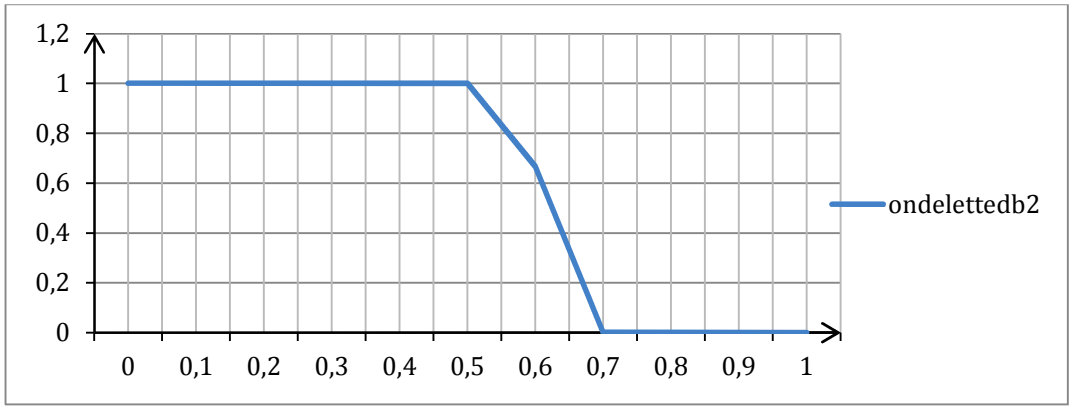
Les figures suivantes donnent les courbes du rappel et précision pour la base ORL suivant les résultats donnés par les figures (11), (18) et (20) avec les différentes méthodes et pour une même image requête.

Le résultat de l'application des différentes méthodes sur cette base est donné par la figure (24), dont on peut conclure à propos du changement de la précision en fonction du rappel et de comparer entre les différentes méthodes utilisées.

(a)



(b)



(c)

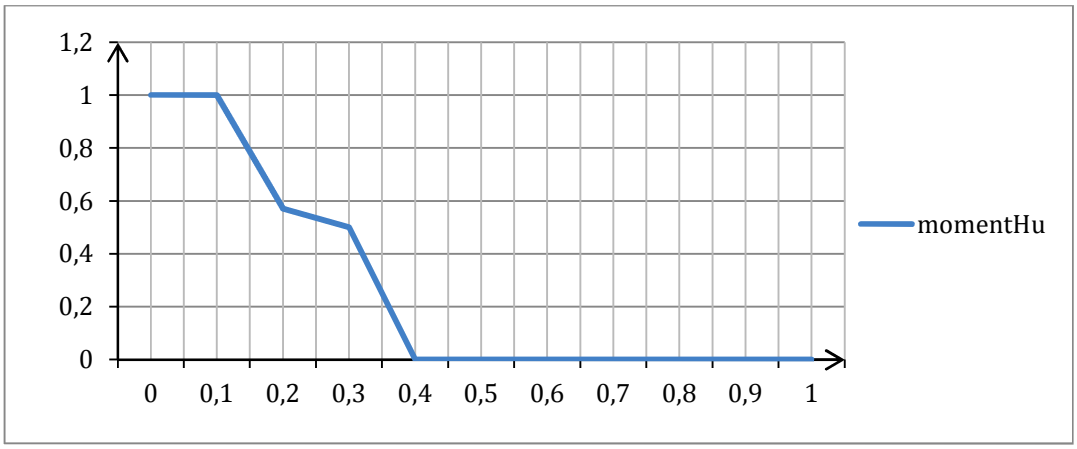


Fig.23. Les courbes rappel-précision pour une image requête dans base ORL. (a) : par la méthode de l'histogramme niveau de gris. (b) : par la méthode des ondelettes db2. (c) par la méthode des moments de Hu.

De la même procédure, on trace les courbes pour toutes les bases d'images.

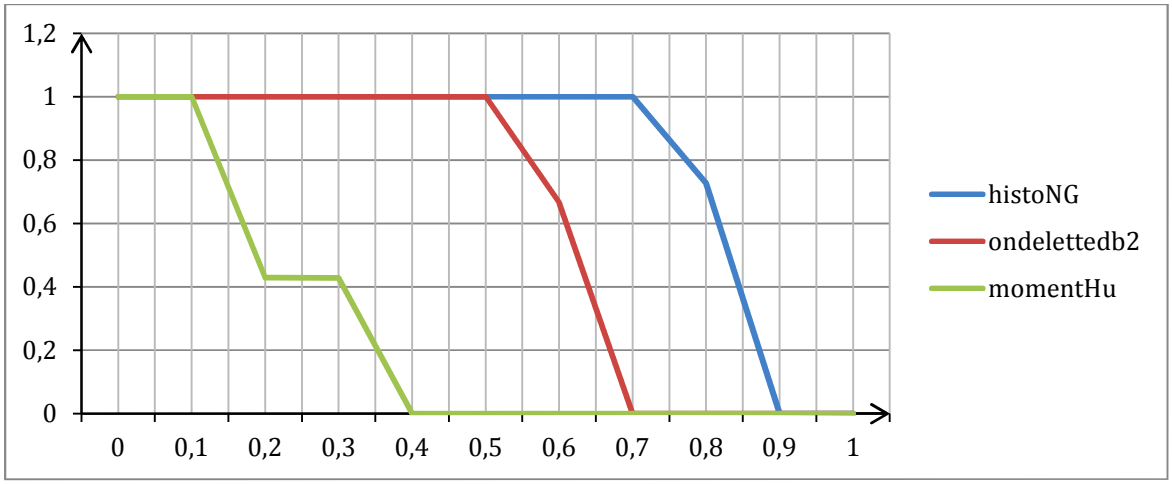


Fig.24. Les courbe rappel-précision pour la base ORL.

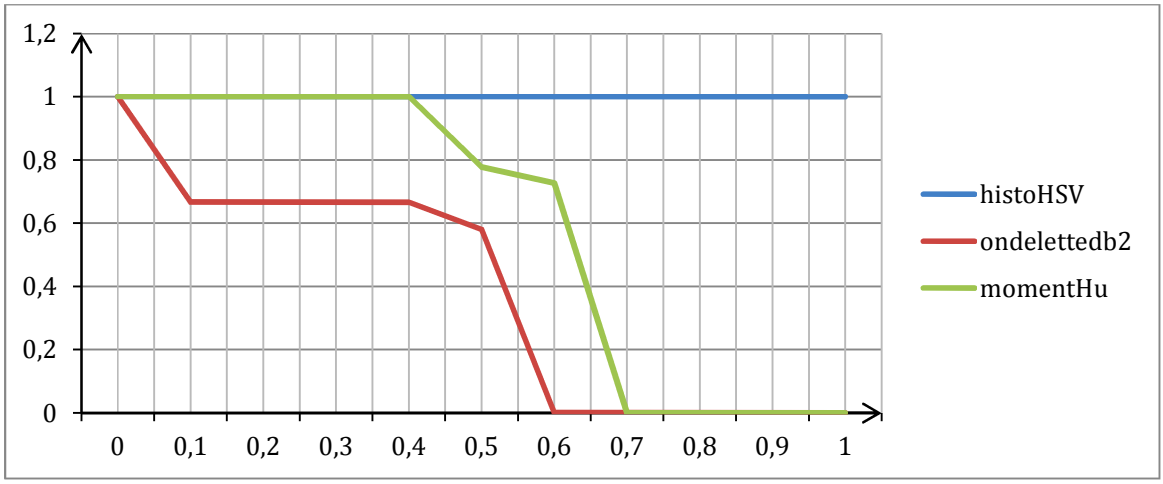


Fig.25. Les courbe rappel-précision. Base Coil100.

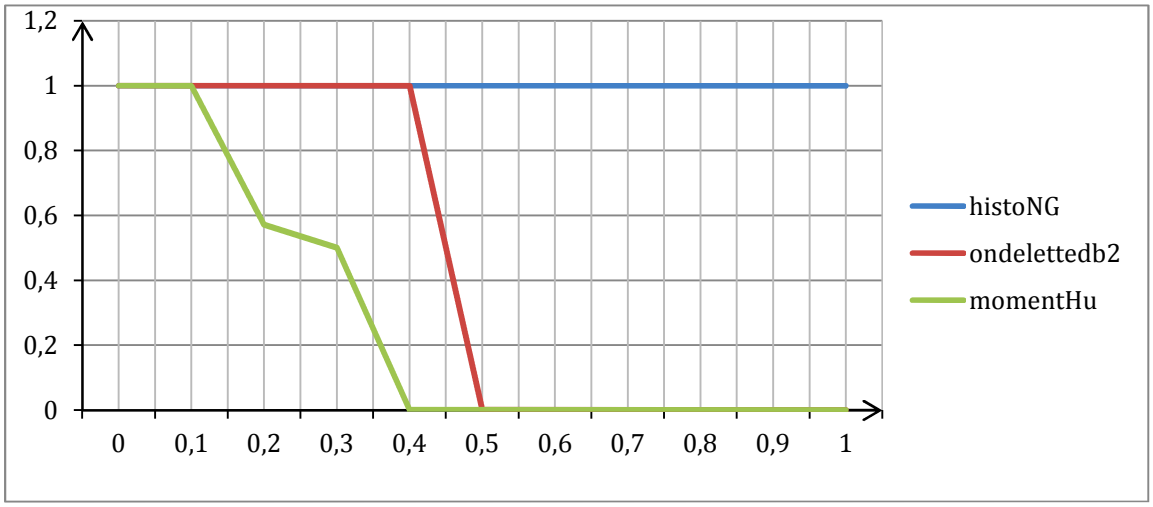


Fig.26. Les courbe rappel-précision. Base Cerebral Standard.

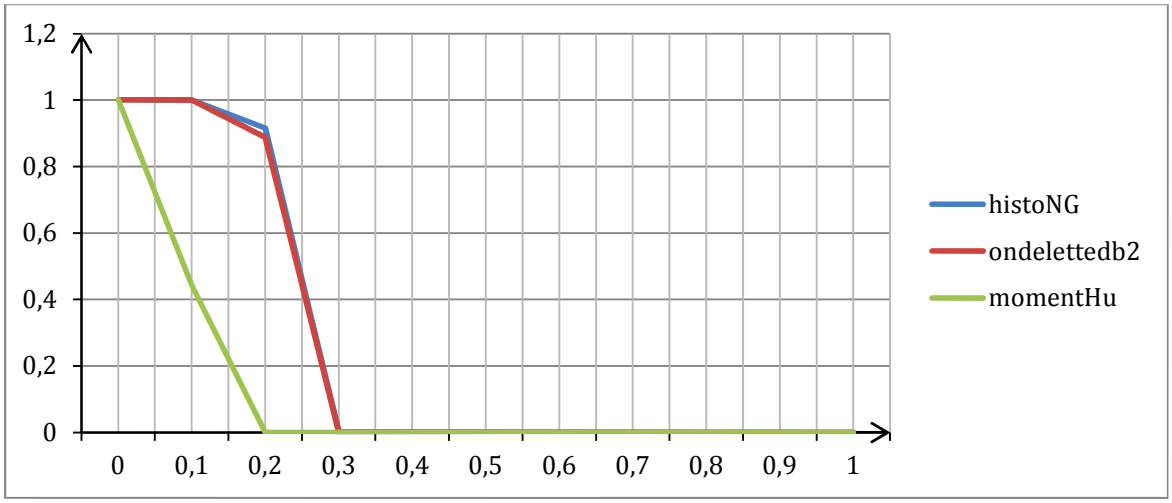


Fig.27. Les courbe rappel-précision. Base image texturées.

3.8.2 Interprétation des résultats

On peut facilement constater par interprétation des courbes rappel-précision, que les histogrammes sont de bons descripteurs de la couleur, ce qui nous permet de dire que la méthode basée sur l'histogramme, que ce soit en niveau de gris ou en couleur HSV, donnent des résultats encourageants pour les quatre bases d'images. Dans La base d'images ORL, la méthode d'histogramme a prouvé son efficacité, ce qui se traduit par la détection de huit images d'une même personne (figure (11)). Ce qui nous a permis de dire que cette méthode est très bonne la recherche d'images de personne, mais cela reste toujours limité aux cas d'images en niveaux de gris.

Les résultats obtenus par la méthode des ondelettes sont très encourageants pour la recherche d'images. Cette méthode a été un descripteur excellent pour l'extraction de l'information texture, ce qui peut s'expliqué par sa capacité d'extraire un nombre importants de résultats pertinents dans toutes les bases d'images.

De même, les moments de Hu qui sont des descripteurs de forme (région), ont pus répondre à nos besoins mais d'une manière insuffisante, ce qui est illustré par les figures (20) et (21). La cause principale revient au manque de formes bien précises dans les images après leur binarisation, ce qui est le cas dans la base d'images texturées, la base médicales et aussi les images le la base ORL, qui contiennent beaucoup de détails. Mais, ce qui est le contraire dans la base COIL100, dont la binarisation a donné des régions très homogènes et le résultat est très satisfaisant (figure(19)).

Les courbes rappel-précision dans les figures (24), (25), (26) et (27) montrent le résultat de chaque méthode dans les quatre bases. L'histogramme a prouvé sa supériorité par rapport aux autres méthodes dans toutes les bases d'images.

Cependant, la figure (27) présente des courbes sur une bande de rappel étroite. Ce qui revient au contenu de cette base (images texturées) qui contient 40 images pertinentes pour chaque image requête. La méthode de l'histogramme niveaux de gris et les ondelettes ont pu extraire un nombre intéressant d'images pertinentes parmi les 12 affichées. Voir les figures (13) et (17). On peut dire due les résultats pour cette base sont aussi très satisfaisants.

3.9 Le bruitage d'images

On passe à l'étape de bruitage de notre image requête par un bruit gaussien de variance $\delta^2 = 0.1$. Cette étape va être suivie par une étape de lissage. Le lissage est effectué par un filtre gaussien de taille 11 et de déviation standard σ=2.

L'image résultante sera une image moins bruitée mais qui présente un flou de lissage. L'objectif de cette étape est de tester la sensibilité au bruit des méthodes étudiées.

Nous testons nos méthodes sur une image requête dégradée, dont l'image originale est présente parmi les images de la base. Le résultat affiche douze images résultantes, organisées de la distance la plus petite à la distance la plus grande, de haut en bas et de gauche à droite. L'image requête est affichée en haut et à gauche des images et son image originale est affichée en bas de la requête.

Les figures suivantes donnent quelques résultats.

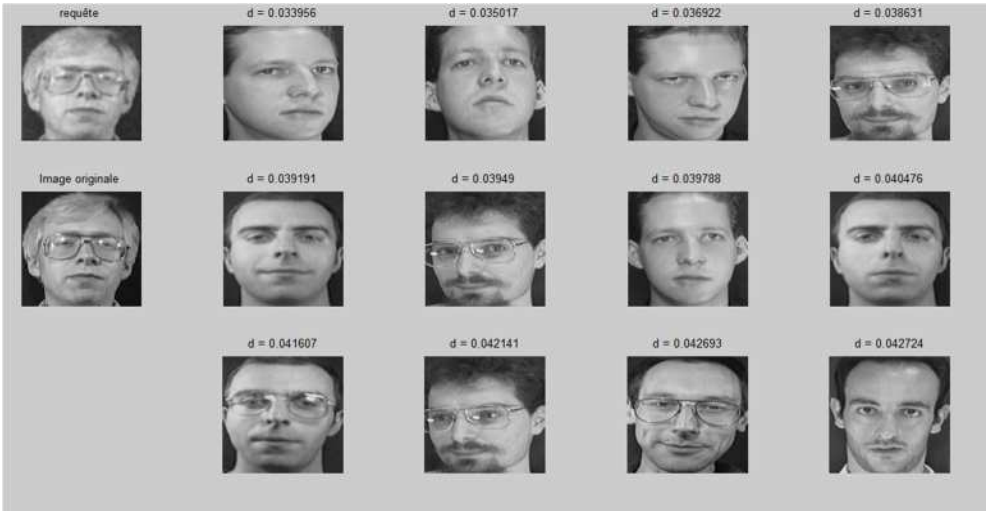


Fig. 28. Résultat de la recherche d'images dans la base ORL pour une image requête dégradée en utilisant la méthode d'histogramme NG.

Fig. 29. Résultat de la recherche pour une image requête dégradée utilisant l'histogramme HSV.

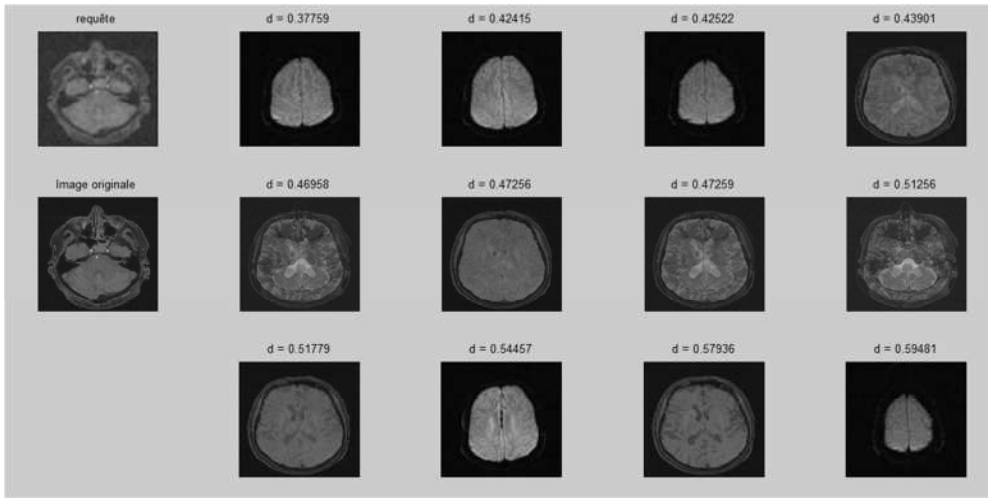


Fig.30. Résultat de la recherche pour une image requête dégradée en utilisant la méthode des ondelettes.

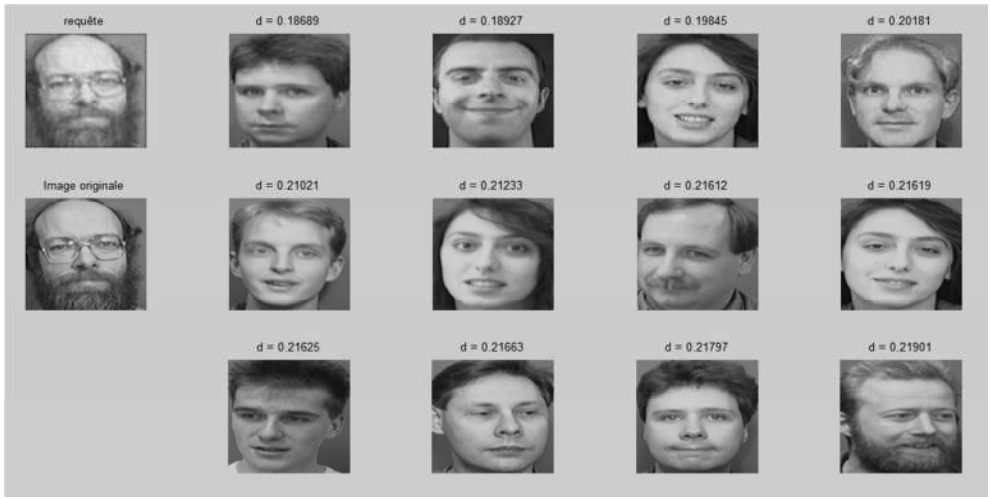


Fig.31. Résultat de la recherche dans la base ORL pour une image requête dégradée en utilisant la méthode des ondelettes.

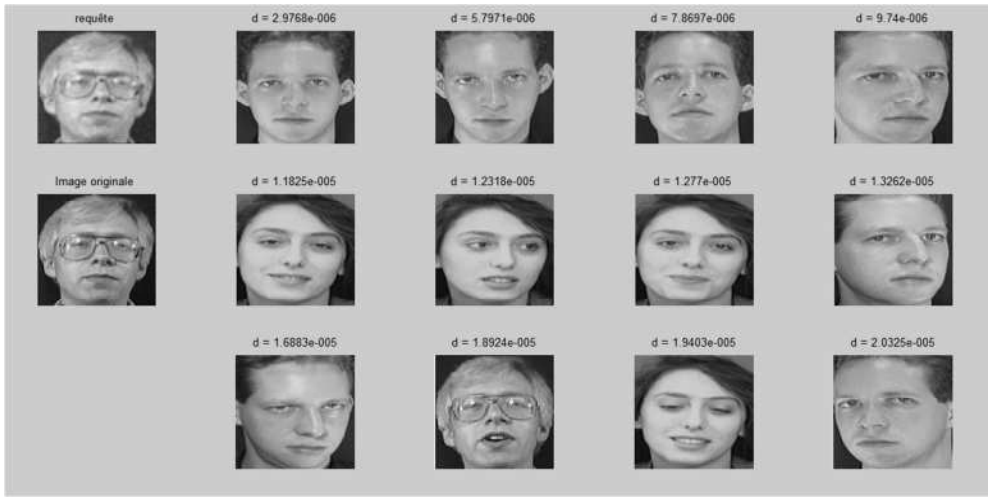


Fig.32. Résultat de la recherche dans la base ORL en utilisant les moments de Hu pour une image requête dégradée.

3.9.1 Interprétation des résultats

Les figures précédentes montrent les résultats de test de nos méthodes pour une image requête dégradée. Les résultats de la recherche d'images ne sont pas satisfaisant et parfois sont complètement hors sujet. Notamment, l'histogramme HSV et l'histogramme NG qui présentent des mauvais résultats, comme le montrent les figures (28) et (29). Et ce qui revient à la présence de nouvelles couleurs imposées par le bruit gaussien.
Les moments de Hu, ont aussi prouvé une grande sensibilité à la présence des perturbations.

L'application de la méthode des ondelettes a donné, elle aussi, des résultats non pertinents, tel qu'il est illustré par les figures (30) et (31). On remarque qu'elle est aussi très sensible au bruit.

3.10 Conclusion

Dans ce présent chapitre, nous avons commencé notre travail par la présentation des différentes méthodes que nous avons choisis pour la recherche d'images par le contenu. Les histogrammes niveau de gris ont été utilisés pour la recherche d'images dans les bases en niveau de gris. Les histogrammes HSV sont préférables pour les images en couleur. Nous avons aussi caractérisé la texture par la méthode des ondelettes db2. Ainsi, nous avons choisi les moments de Hu pour la recherche d'images par la forme.

Dans la plupart des résultats, les méthodes de la recherche par la couleur et la texture ont donné de bons résultats que les moments de Hu. On peut conclure que les histogrammes présentent de bonnes méthodes pour la recherche d'images par le contenu. Ainsi que les ondelettes sont elles aussi une méthode très efficace dans la recherche d'images par la caractéristique texture, ce qui revient aux résultats offerts par cette méthode.

Nous avons aussi étudié l'influence du bruit sur la sensibilité des méthodes étudiées. L'image requête a été dégradée par un bruit gaussien. Nous avons ensuite tenté de récupérer les images avec débruitage par un filtre gaussien. Le résultat est une image requête floue avec moins de bruit.

L'analyse des résultats révèle que la reconnaissance des images similaires par les différentes méthodes est très faible. Ce qui nous permet de conclure à la sensibilité des méthodes étudiées aux bruits.

Conclusion Générale

Dans ce présent mémoire nous avons abordé un domaine très intéressant du traitement d'image, c'est la recherche d'images par le contenu.

Au début de notre travail, nous avons donné un bref état de l'art sur la recherche d'images. Ensuite nous avons présenté la structure générale d'un système de recherche d'images par le contenu et ses composantes essentielles. La description des attributs visuels essentiels dans une image, qui sont la couleur, la texture et la forme, est prise en compte. Par la suite, nous avons étudié les différentes méthodes d'extraction des trois attributs visuels définis.

Notre système de recherche d'images par le contenu fonctionne avec des vecteurs descripteurs représentants le contenu visuel et global des images. L'étude théorique des différentes méthodes nous a permis de sélectionner les méthodes adéquates pour la description de l'image dans notre cas. La couleur a été extraite par la méthode des histogrammes niveaux de gris et les histogrammes HSV, la description de la texture est faite avec les ondelettes de Daubechies dont le vecteur signature comporte l'énergie, la moyenne et l'écart types de chaque imagette résultante, et enfin la méthode des moments de Hu a été choisie pour construire le vecteur descripteur de forme.

Après avoir testé les trois méthodes, nous avons construit les courbes de rappel-précision pour chaque base d'images pour une évaluation de notre système de recherche d'image par l'une des méthodes choisies.

Les résultats obtenus ont été différents d'une méthode à une autre et selon chaque base. Ces résultats sont en général très prometteurs et

répondent aux besoins de la recherche d'images par le contenu. Les histogrammes par exemple, sont performants pour la recherche d'images par la couleur, ainsi que les ondelettes qui ont elles aussi donné des résultats satisfaisants, ce qui revient à leur capacité de détection des détails présents dans une image. Pour les moments de Hu la recherche est moins bonne, sauf pour la base Coil 100 où les images présentent un seul objet à binariser. La forme de l'objet est donc bien précise ce qui a facilité la caractérisation de la forme pour les moments de Hu, contrairement aux autres bases d'images où les images présentent beaucoup de détails.

Nous avons ensuite testé la capacité de notre système pour reconnaitre les images similaires à une image requête dégradée. Le résultat a montré que les histogrammes et la méthode des moments de Hu ainsi la méthode des ondelettes sont très sensibles aux bruits, cela peut être expliqué par la présence de nouvelles couleurs ainsi qu'un changement de contours introduit par le bruit.

Les perspectives et les directions de recherche à suivre pour améliorer le présent travail portent essentiellement sur le choix d'autres méthodes pour la caractérisation des différents attributs et qui soient moins sensibles aux bruits. En particulier, nous pensons à un bouclage de pertinence ainsi qu'à combiner les approches textuelles avec les méthodes proposées pour contourner le problème du fossé sémantique. Nous envisageons aussi de faire appel à une classification des images de chaque base en familles d'images, ainsi la recherche s'effectue uniquement sur une famille.

Bibliographie

[1] J. Jeon, V. Lavrenco, R. Manmatha. Automatic image annotation and retrieval using crossmedia relevance models. In Preceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR'03. 2003.

[2] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain. Content-based image retrieval at the end of the early years, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (12) (2000) 1349–1380.

[3] T. Kato, K. Hirata. Query by visual exemple in content-based image retrieval, Proc.EDB192. Lecture Notes in computer Science, 1992, p. 56-71.

[4] M. J. Swain et D. H. Ballard. Color indexing. International journal of computer vision, 7(1):11–32, 1991.

[5] S. A. Berrani. Recherche approximative de plus proches voisins avec contrôle probabiliste de la précision : application à la recherche d'images par le contenu. PhD thesis,Université de Rennes 1, février 2004.

[6] S. Cohen. Finding colors and shape patterns in images. PhD thesis, Stanford University, 1999.

[7] E. Loupias. Indexation d'images : aide au télé-enseignement et similarités préattentives. PhD thesis, Institut National des Sciences Appliquées de Lyon, 2000

[8] P. Lambert and H. Grecu. Indexation par descripteurs flous : Application à la recherche d'images. 18ème colloque sur le traitement du signal et des images (GRETSI01), II :372–379, 2001.

[9] J. Fauqueur. Contributions pour la recherche d'images par composantes visuelles. PhD thesis, Université de Versailles - Saint-Quentin, 2003.

[10] L. Amsaleg, P. Gros, R. Mezhoud. Mise en base d'images indexées par des descripteurs locaux : problèmes et perspectives. Institut National de Recherche en Informatique et en Autoatique. INRIA 2000.

[11] J. Landré. Analyse multi-résolution pour la recherche et l'indexation d'images par le contenu dans les bases de données application à la base d'image paléontologique Trans'Tyfipal. Université de Bourgogne. Thèse de Doctorat, décembre 2005.

[12] H. ABED, L. ZAOUI, Z. GUEZZEN. Fusion Couleur Texture dans l'Indexation et la Recherche des Images. JIG'2007 - 3èmes Journées Internationales sur l'Informatique Graphique.

[13] R. KACHOURI. Classification multi-modèles des images dans les bases hétérogènes. Université d'Evry-Val d'Essonne. Thèse de doctorat, 2010.

[14] Ela Yildizer, Ali Metin Balci, Mohammad Hassan, Reda Alhajj, Efficient content-based image retrieval using Multiple Support Vector Machines Ensemble, *Expert Systems with Applications, Volume 39*, Issue 3, 15 February 2012, Pages 2385-2396.

[15] WangXing-yuan, ChenZhi-feng, YunJiao-jiao, An effective method for color image retrieval based on texture, *Computer Standards & Interfaces Volume 34* (2012) Pages 31–35.

[16] Cheikhrouhou, I., Djemal, K., Masmoudi, D., Maaref,H., and Derbel, N. 2009. Empirical descriptors evaluation for mass malignity recognition. In the First International Workshop on Medical Image Analysis and Description for Diagnosis Systems MIAD'09.

[17] H. Jlassi, and K. Hamrouni. Detection of blood vessels in retinal images. International Journal on Image Graphics, 10(1) : 57-72. 2010

[18] S. Bedouhene. Recherche d'images par le contenu. Mémoire de magister. Université Tizi Ouzou. 2011.

[19] [WAN 01] WANG J. Z., LI J., WIEDERHOLD G., « SIMPLIcity : Semantics-sensitive Integrated Matching for Picture Libraries », IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2001.

[20] Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., and Gool,L. A comparaison of affinene region detectors. International Journal of Computer Vision 2(65). 43-72. 2005

[21] J. Zhang, T.Tan Brief review of invariant texture analysis methods. Pattern Recognition 35 (2002) 735–747.elsvier

[22] C. L. Jordan, T. Ebrahimi, m. Kunt, Progressive content- basedshape compression for retrieval of of binary images, Computer vision and Image Understanding 71 (2) (1998), 198-212.

[23] J. Hafner et al, "Efficient color histogram indexing for quadratic form distance functions", IEEE trans. Pattern Analysis and Machine Intelligence, Vol. 17, pp 729-736, 1995.

[24] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q.Huang,B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D, Steele, and P. Yanker. Query by image and video content: The qbic system. IEEE Computer, 28(9) :23 32,1995.

[25] GUPTA A. *et al.*, « The Virage image search engine : an open framework for image management », *SPIE Storage and Retrieval for Image and Video Databases*, vol. 2670, 1996.

[26] W. Ma and B. S. Manjunath. Netra : A toolbox for navigating large image databases. *Multimedia Syst, 7(3) :184 198, 1999*

[27] S. Ardizzoni, I. Bartolini, and M. Patella. Windsurf : Region-based image retreival using wavelets. DEXA Workshop, pages 167-173, 1999.

[28] A. Pentland, R. W. Picard, and S, Sclaroff. Photobook : content- based manipulation of image databases. *Int, J, Comput, Vision,18(3) :233 254, 1996*

[29] T. Quach, U. Monich, B.S. Manjunath. A system of large scale, content based web image retrieval. Universite California Santa Barbara 2004.

[30] C. Carson, M. thomas, S. Belongie, J. M. Hellerstein, and J, Malik. Blobworld : A system for region-based image indexing and retrieval. In Visual ’99 : Proceedings of the Third International Conference on Visual Information and Information Systems, Pages 509 516, London, UK,1999. Springer-Verlag.

[31] J. R. Smith and S. F. Chang. Visualseek : A fully automated contentbased image query system. In ACM Multimedia Conference, Page 87 98, 1996.

[32] BOUJEMAA N., FAUQUEUR J., FERECATU M., FLEURET F., GOUET V., SAUX B. L., SAHBI H., « IKONA : Interactive Generic and Speci_c Image Retrieval », *International workshop on Multimedia Content-Based Indexing and Retrieval (MMCBIR'2001), Rocquencourt, France*, 2001.

[33] Stehling, R. O., Nascimento, M. A., and A. X . Falcao . On Shapes of Colors` for Content-based Image Retrieval. *In ACM International Workshop on Multimedia Information Retrieval,* 2000, 171-1

[34] M. A. Stricker and M. Orengo. Similarity of color images . In SPIE, Storage and Retrieval for image Video Databases, pages 381-392, 1995.

[35] S. Deb, Y. Jhang, An Overview of Content-based Image Retrieval Techniques, *Proceedings of the 18t International Conference on Advanced Information Networking and Application (AINA’04*), 2004 IEEE.

[36] K Houari. Recherche d’images par le contenu. Université MENTOURI. Constantine**.** Thèse de Doctorat science en informatique. 2010

[37] H. S. Sawhney and J. L. Hafner. EFFICIENT COLOR HISTOGRAM INDEXING. Machine Vision Group.IBM Almaden Research Center. 1994 IEEE.

[38] R.M. Haralick. Statistical and structural approachs to texture. Proceedings of the IEEE In Proceedings of the IEEE, Vol. 67, No. 5, pages 786–804, may 1979.

[39] J. Huang, R. Kumar, M. Mitra, W. Zhu, W. Zahib, Image indexing using color correlogram, in: IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, June (1997) 762–768.

[40] M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra, and T. S. Huang. Supporting similarity queries in MARS. In Proceedings of the 5th ACM International Multimedia Conference, Seattle, Washington, 8-14 Nov. '97, pages 403-413, 1997.

[41] G.Pass, R. Zabih and J. Miller, "Comparing images using color coherence vectors", *MULTIMEDIA '96: Proceedings of the fourth ACM international conference on Multimedia*, p. 65--73, 1996

[42] Y. Rubner. Perceptual metrics for image database navigation. Rapport Technique CS-TR- 99-1621, Stanford University, 1999

[43] W.Press, B. Flanney,S. Teukolsky,and W. Vetterling(1987). Numerical Recipes. The Art of Scientific Computing.

[44] R.M.Haralick,K. Shanmugam and Its’hak Dinstein. Textural features for image classification. Reprinted by permission from IEEE, Vol. SMC-3,No.6,.pages 610-621, november 1973.

[45] N. Jhanwar, S. Chaudhuri, G. Seetharaman, B. Zavidovique .Content based image retrieval using motif cooccurrence matrix Image and Vision Computing 22 (2004) 1211–1220.Elsevier.

[46] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, C. Roux Wavelet optimization for content-based image retrieval in medical databases, Medical Image Analysis. 227–241. 2009 Elsevier.

[47] M. Jian, L. Liu, F, Guo. Texture Image Classification Using Perceptual Texture Features and Gabor Wavelet Features. Asia-Pacific Conference on Information Processing. 2009 IEEE.

[48] L. L. Huang , A. Shimizu, H. Kobatake. Robust face detection using Gabor filter features. Pattern Recognition Letters 26 (2005) . 1641-1649.Elsevier.

[49] A. Grossmann and J. Morlet. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.*, 15(4) :723–736, 1984.

[50] A. Khare and U. S. Tiwary . A New Method for Deblurring and Denoising of Medical Images using Complex Wavelet Transform. Proceedings of the 2005 IEEE. Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005.

[51] S.X. Liao et M. Pawlak. On image-analysis by moments. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(3) :254–266, March 1996.

[52] S. Lafon, R. R. Coifman, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S., W. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. 7426–7431,PNAS,May 24,. Vol. 102, no. 21, 2005.

[53] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision*. PWS Publishing, seconde edition edition, 1999.

[54] M.K. Hu. Visual pattern recognition by moments invariants, computer methods in image analysis. *Transactions on Information Theory*, 8, 1962.

[55] M.R. Teague. Image analysis via the general theory of moments. J. Optical Soc. Am., 70(8) :920–930, August 1980.

[56] H. Delingette, J. Montagnat, Shape and topologiy constraints on parametric active contours. Computer Vision and Image Underst-anding, 83(2) : 140-171. 2001.

[57] H. Delingette, J. Montagnat, Shape and topologiy constraints on parametricactive contours. Computer Vision and Image Understanding, 83(2) : 140-171. 2001.

[58] Z. Huang, J. Leng. Analysis of Hu's Moment Invariants on Imag Scaling and Rotation. Proceedings 2nd International Conference on Computer Engineering and Technology (ICCET). (pp. 476-480). Chengdu, China. IEEE 2010.

[59] Base Columbia. http://www.cs.columbia.edu/CAVE/research/softlib/coil-100.html

[60] S. Mallat. A theory for multi-resolution signal decomposition : The wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, p. 674-693. 1989.

[61] C. Vonesch, T. Blu, M. Unser. Generalized Daubechies Wavelet Families. IEEE Transaction on Signal Processing, Vol.55, No. 9, p. 4415-4429. September 2007.

[62] http:/ /www.cl.cam.ac.uk/Research/DTG/attarchive :/data/att_faces.tar.Z

[63] N. OTSU, « A Threshold Selection Method from Gray-Level Histograms », IEEE transactions on Systems, Man and Cybernetics, 9(1), p. 62-66, 1979.

[64] C. D. Manning, P. Raghavan, H. Schütze. An Introduction to Information Retrieval. Cambridge University Press. Cambridge, England. 151-177. 2009.