

République Algérienne Démocratique et Populaire  
Ministère de L'Enseignement Supérieur et de la  
Recherche Scientifique

Université Mouloud Mammeri De Tizi-Ouzou



Faculté De Génie Electrique Et D'informatique  
DEPARTEMENT D'AUTOMATIQUE

**Mémoire de Fin d'Etude  
de MASTER ACADEMIQUE**  
Spécialité : **automatique et système**

*Présenté par*

Berradj Karima  
et  
Diaf Dihia

Mémoire dirigé par le **Pr. Moussa DIAF**

Thème

**RECHERCHE D'IMAGE PAR LE  
CONTENU**

*Mémoire soutenu publiquement le 27 Septembre 2018 devant le jury composé de :*

**M<sup>r</sup> Moussa Diaf**

Professeure, à UMMTO, Rapporteur

**M<sup>r</sup> Takfarinase Chelli**

Maitre assistant classe A, à UMMTO, Examineur

**M<sup>r</sup> Samir Yesli**

Doctorant en automatique, à UMMTO, Examineur

**M<sup>m</sup> Boukendour Ouiza**

Maitre assistant classe A, à UMMTO, Examineur

---

# Sommaire

Introduction générale	
Chapitre 1 : Indexation d'image	
1.1 Introduction	4
1.2 Attributs d'une image	5
1.3 La couleur	6
1.4 La texture	6
1.5 La forme	7
1.6 Distance de similarité	8
1.6.1 Distance de Minkowski	8
1.6.2 Intersection d'histogrammes	8
1.6.3 Distance quadratique	9
1.6.4 Distances entre distributions	9
1.6.5 Distance euclidienne	9
1.6.6 Distance d'Earth Mover (EMD)	9
1.7 Domaines d'application	10
1.8. Conclusion	10
Chapitre 2 Indexation d'images	
2.1 Introduction	11
2.2 Principe de l'indexation d'images	11
2.2.1 Phases d'indexation	11
) Indexation logique	12
) Indexation physique	12
2.2.2 Système d'indexation par le contenu	12
2.4 Les requêtes	13
2.4.1 Requête par mots clés	13
2.4.2 Requête par esquisse	14
2.4.3 Requête par le contenu	14
2.4.4 Requête par l'exemple	14
2.5 Descripteurs de couleur	15
2.5.1 Espace RVB ou RGB en anglais	15
2.5.2 Espace TSV	15
2.5.3 Espace Lab ou Luv	16
2.5.4 Espace couleur La*b*	16
2.5.5 Espace XYZ	16
2.5.6 Histogramme couleur	17
2.5.7 Le Corrélogramme	17
2.6 Descripteurs de forme	17
Les moments géométriques	18
) Moments de Hu	18
) Les moments de Legendre	19
2.6.2 Les moments orthogonaux	19
) Les descripteurs de Fourier	19
) Les moments de Zernike	19
2.7 Descripteur de Texture	20
2.7.1 Filters de Gabor	20
2.7.2 Les matrices de cooccurrence	21
8. Conclusion	21
Chapitre 3 Application et resultants	
3.1 Introduction	23
3.2 Mesure de similarité	24
3.2.1 Mesure de similarités entre les vecteurs de l'histogramme	24
3.2.2 Mesure de similarités entre les vecteurs de texture	24
3.2.3 Mesure de similarités entre les vecteurs de moments de Zernike	24
3.2.4 Mesure de similarités entre les vecteurs de la Combinaison	24
3.3 Evaluation d'un system	25
3.3.1 Le rappel	25
3.3.2 La précision	25

3.4	Les bases d'images utilisées.....	26
3.4.1	La base d'image COIL .....	26
3.4.2	La base d'image Cérébral Standard MR .....	26
3.4.3	La base d'image texture .....	27
3.4.4	La base d'image ORL.....	27
3.5	Testes et résultats .....	28
3.5.1	Evaluation sur la base COIL .....	28
3.5.2	Evaluation sur la base Visage .....	32
3.5.3	Evaluation sur la base Texture .....	34
3.5.4	Evaluation sur la base Texture .....	38
3.6	Interprétation des résultats .....	41
3.7	Conclusion	
	Conclusion générale .....	42

## Avant-propos

Ce mémoire a été effectué au Laboratoire Vision Artificielle et Automatique des Systèmes LVAAS, UMMTO.

Mes vifs remerciements vont tout d'abord à Monsieur **DIAF Moussa**, Professeur à l'UMMTO pour m'avoir proposé le thème de ce mémoire et m'avoir dirigée, aidée et conseillée tout le long de notre travail.

Nous exprimons notre reconnaissance aux membres du jury pour avoir bien voulu examiner ce travail.

Nous tenons à remercier aussi vivement Monsieur Merzouk Younsi pour nous avoir beaucoup aidées dans ce travail..

## Introduction générale

Dans le domaine d'imagerie, différentes techniques d'indexation et de recherche d'informations ont été proposées ces dernières années, avec le développement du multimédia et de l'informatique qui est devenu le cœur de tous les secteurs d'activités, il n'est pas rare d'avoir des quantités énormes et croissantes d'images dans des bases d'images numériques qui possèdent une problématique nouvelle.

L'image sous toutes ses formes (photographique, vidéo, graphique ou infographique) est aujourd'hui mise à la disposition de nombreuses personnes. Dans plusieurs domaines par exemple les progrès réalisés dans le domaine des réseaux de télécommunications permettent la transmission et le partage de l'information numérique localement et mondialement, avec des temps de transfert acceptables.

Le problème posé possède une certaine similarité. Il est nécessaire de retrouver les images semblables à celles que l'utilisateur recherche, avec un autre terme l'idée est de retrouver des images qui partagent un contenu qui peut être associé à différentes caractéristiques, nous avons choisi de travailler dans la famille des approches basées sur l'extraction de caractéristiques bas niveau. *, le haut niveau*

Les travaux de recherche dans le domaine de l'indexation de la base d'images par le contenu visent souvent la forme, la couleur et la texture des objets contenus dans une image, et pour gérer efficacement ces grandes bases de données, nous avons besoin d'un système de recherche et d'indexation d'images.

Les systèmes de recherche par le contenu peuvent chercher à reconnaître un objet tel un visage, un animal ou tout autre objet. Pour ce faire, ces systèmes nécessitent l'emploi de techniques et de descriptions spécifiques à l'objet ou à la scène

Notre travail se base sur une catégorie de systèmes qui cherchent à reconnaître les images similaires à une image requête. *ent*  
*à un autre jour*

Dans le premier chapitre, on commence notre introduction quelques rappels sur le traitement d'image. Nous présenterons ensuite sur les descripteurs utilisés dans notre travail qu'on va détailler dans le 3ème chapitre.

Dans le deuxième chapitre, nous présenterons quelques méthodes d'indexations et de recherche d'images par le contenu extraire dans la littérature

Le troisième et dernier chapitre sera consacrée aux tests et résultats.

Ce chapitre sera terminé par une conclusion générale.

# Chapitre 1

## *Notions sur l'indexation d'images*

### **1.1. Introduction**

Le traitement d'images constitue l'ensemble des méthodes permettant d'apporter des modifications dans une image en vue de l'améliorer ou d'en extraire des interprétations. Le traitement d'images fait appel à plusieurs disciplines comme les mathématiques, le traitement du signal, la technologie des composants électroniques etc. Sur le plan matériel, le traitement d'images nécessite une unité informatique rapide et de capacité mémoire élevée, un système d'acquisition des images comme les caméras et d'un équipement pour l'archivage et l'impression. Les applications en traitement d'images sont nombreuses et très diversifiées comme l'imagerie médicale, la biométrie, l'asservissement visuel, la robotique, la reconnaissance et le tri d'objets, la télédétection etc. Dans l'ensemble de ces applications, les opérations traitement d'images les plus rencontrées sont le prétraitement, la restauration, l'amélioration, la compression, la reconstruction, la segmentation etc.

Le prétraitement permet de compenser certains artefacts dus au système d'acquisition, de filtrer ou de lisser le bruit et d'améliorer et restaurer l'image. Ainsi, le prétraitement s'impose si l'image est bruitée, le contraste insuffisant et l'éclairement de l'image non uniforme. Si le filtrage consiste à éliminer ou réduire le bruit, le lissage s'intéresse aux petites variations autour de chaque pixel. En effet, le bruit d'images est un phénomène de brusque variation d'un pixel par rapport à ses voisins. Le filtrage de l'image s'effectue souvent dans le domaine fréquentiel et il peut être linéaire comme le filtre passe-bas ou non linéaire comme le filtre médian.

Après cette phase, il s'agit de procéder aux traitements proprement dits pour lesquels une panoplie de méthodes a été proposée selon la nature du problème. On peut, en effet, procéder à l'amélioration, la restauration, la segmentation etc.

L'amélioration permet une bonne exploitation visuelle par modification de l'information comme le rehaussement de contrastes, de certaines intensités pour mettre en valeur une région etc. La modification d'histogrammes est souvent utilisée dans ce cas. En effet, l'histogramme est une répartition des niveaux de gris ou des couleurs à partir duquel on peut corriger le contraste et l'échelle des couleurs pour des images surexposées ou sous-exposées sans altérer les informations contenues dans l'image.

La restauration permet de retrouver l'image d'origine à partir de sa version dégradée par ajout d'information connaissant le modèle de dégradation. Si l'on connaît le modèle du processus de dégradation, on applique le modèle inverse pour retrouver l'image d'origine.

La segmentation d'images, quant à elle, permet de partitionner l'image en régions homogènes de sorte que les régions soient distinctes les unes de autres. Elle permet, par exemple, de séparer des objets du fond. Il existe plusieurs techniques de segmentation d'images. Il s'agit de regrouper dans la même classe ou région les pixels ayant des propriétés communes. La binarisation est aussi opération de segmentation qui sépare les pixels en deux classes, le fond et l'objet. Aussi, la segmentation peut être effectuée de plusieurs manières. Elle est effectuée par approche basée régions, approche basées contours ou par classification des pixels.

Dans ce domaine nous s'intéressons à l'analyse d'image afin de l'indexer. La recherche d'images par le contenu où l'image constitue un support essentiel d'informations.

## 1.2. Attributs d'une image

L'image numérique est constituée de pixels, chacun représenté par un niveau de gris ou de couleur. Ces pixels sont présentés sous forme d'une matrice  $I(x,y)$ , où  $x$  et  $y$ , représentent les coordonnées cartésiennes d'un point de l'image et  $I(x,y)$  correspond au niveau de gris ou au niveau de couleur. Dans le cas d'une image monochrome, le pixel est codé sur un certain nombre de bits qui définissent le nombre de niveaux de gris de l'image. Le pixel codé sur 8 bits, l'image sera en 256 ( $2^8$ ) niveaux de gris mais codé sur 16 bits, l'image aura 65536 ( $2^{16}$ ) niveaux de gris. Cependant, avec les progrès technologiques, la grande majorité des images sont actuellement en couleur. La représentation des couleurs s'effectue selon le modèle de représentation. On peut représenter les couleurs à l'aide de leurs composantes primaires R (Rouge), V (Vert) ou B (Blanc). Les systèmes émettant de la lumière (écrans d'ordinateurs,...) sont basés sur le principe de la synthèse additive : les couleurs sont composées

d'un mélange de rouge, vert et bleu (modèle R.V.B.), et en anglais RGB (pour Red Green Blue). Ainsi dans le cas de ces images, chaque pixel est généralement codé sur trois octets (24 bits), un octet pour chacune de ces couleurs primaires. Un mélange quelconque de ces couleurs pures engendre une autre couleur, qu'on peut qualifier en termes de nuances selon différents critères plus ou moins intuitifs.

Un des problèmes rencontrés lors de la manipulation de grandes quantités d'images dans des grandes bases de données est la saturation, le stockage et la recherche de l'information. Ce problème peut être résolu par la technique d'indexation et de recherche d'image par le contenu dont les méthodes nombreuses. L'objectif est de pouvoir est d'interroger ces bases d'une manière ergonomique et intuitive pour l'utilisateur. En effet, Les systèmes d'indexation et de recherche d'images par le contenu permettant de recherchée des images dans une base des images et de déterminer les images qui ressemblent à l'image requête en fonction de leurs caractéristiques visuelles. De ce fait, il est nécessaire d'extraire les trois descripteurs que sont le niveau de gris, la couleur, la texture et la forme de chaque image de la base (hors ligne) et l'image requête (en ligne). Pour ce faire, on utilise les différents attributs ou descripteurs de l'image. Dans ce travail, nous nous intéressons aux descripteurs de couleurs, texture et de forme.

### **1.3 La couleur**

La couleur est un des attributs primaires le plus utilisés en indexation d'images car c'est le plus perceptuel. Rappelons qu'un pixel d'une image couleur est caractérisé par un vecteur aux trois composantes R, V et B pour Rouge, Vert et Bleu respectivement si l'image est représentée dans le système RVB. Les couleurs présentes dans une image sont décomptées dans l'histogramme des couleurs. La description des couleurs est essentielle pour la majorité du système de recherche d'images par le contenu. Dans ce cadre, on s'intéresse souvent à l'identification de l'espace couleur le plus discriminant [1], à l'étude des problèmes d'invariance aux conditions d'illumination et de prise de vue [2].

Pour le choix de l'espace colorimétrique, il en existe plusieurs parmi les quels l'espace HSV (Hue, Saturation et Value) ou ITS en français (Intensité, Teinte et saturation [3], Ainsi, La majorité des systèmes de recherche d'images par le contenu se base sur la description des couleurs de l'image.

### **1.4 La texture**

La texture est une notion qui existe dans presque tous les objets. Elle peut être définie de plusieurs manières. Dans le dictionnaire, la texture est la répétition spatiale d'un motif de base dans différentes directions de l'espace.

Mais la définition la plus adoptée est celle de Haralick [4] qui la décrit comme un phénomène à deux dimensions. La première dimension concerne la description d'éléments de base ou primitives (le motif) à partir desquels est formée la texture. La deuxième est relative à la description de l'organisation spatiale de ces primitives.

Les domaines d'application de l'analyse de la texture sont très variés. Les opérations principalement demandées sont celles relatives à la segmentation d'images et en reconnaissance en formes. C'est ainsi qu'on la retrouve appliquée en imagerie médicale, en télédétection, en industrie pour le contrôle de qualité ou le placement d'objets en utilisant un système de vision artificielle etc.

Pour l'analyse de la texture, plusieurs méthodes ont été proposées. Il s'agit de formaliser les descriptifs de la texture par des paramètres mathématiques qui serviraient à l'identifier. Dans ce sens, les critères visuels qui ont été retenus pour la texture sont le contraste, la granularité, l'orientation, la forme, la finesse, la régularité et la rugosité. Une multitude de méthodes, de variantes et de combinaisons de méthodes ont déjà été proposées dans la littérature et éprouvées en pratique [5]. On parle ainsi de méthodes structurelles qui tiennent compte de l'information structurelle et contextuelle d'une forme et sont particulièrement bien adaptées aux textures macroscopiques, de méthodes statistiques où la texture est décrite par des paramètres statistiques sont estimés pour chaque pixel de l'image, les méthodes spatio-fréquentielles et les méthodes fractales [6].

Parmi les méthodes statistiques permettant l'analyse la texture, les plus utilisées sont celles basées sur la matrice de cooccurrences, le filtre de Gabor et les ondelettes [7].

## 1.5 La forme

La forme est l'une des principales caractéristiques visuelles importantes. Elle est aussi l'une des caractéristiques de base pour décrire le contenu d'une image. Il existe aussi plusieurs descripteurs de forme que l'on peut classer en deux catégories à savoir, les descripteurs basés frontières et les descripteurs basés régions.

Les descripteurs de la première catégorie sont relatifs aux différents contours extraits de l'image. Ils ne peuvent pas détecter la structure interne de la forme puisqu'ils sont basés sur les contours uniquement. Par conséquent, elles sont limitées à un certains types d'applications. Les descripteurs de la deuxième catégorie différents et nombreux. Les plus utilisés en indexation d'images sont les moments invariants de HU [8] ou de Zernike [9] qui sont robustes aux transformations géométriques comme la translation, la rotation et le changement d'échelle.

## 1.6. Distance de similarité

Dans une base de données, les images sont classées en utilisant une mesure de distance entre les vecteurs de descripteurs associés à ces images. Pour la recherche et l'indexation d'image, on doit utiliser une distance de similarité entre les images. Ainsi, dans la pratique, il existe un grand nombre de mesures de similarité dans le cadre de la recherche d'images similaires. Ces mesures sont également utilisées pour la recherche de régions similaires tout comme elles sont aussi utilisées en classification automatique.

Les différentes distances les plus utilisées sont les suivantes:

### 1.6.1 Distance de Minkowski

Les métriques de Minkowski [10] est une généralisation des autres distances géométriques. Son expression est la suivante :

$$L_r(I1, I2) = \left[ \sum_{i=1}^n |I1(i) - I2(i)|^r \right]^{\frac{1}{r}} \quad (1)$$

où  $r > = 1$  est le facteur de Minkowski et  $n$ , la dimension de l'espace caractéristique.

Plus le paramètre  $r$  augmente, plus la distance  $L_r$  aura tendance à favoriser les grandes différences entre coordonnées.

Deux types de métriques de Minkowski se retrouvent dans la littérature :

- Norme  $L1$ , pour  $r=1$  :

$$L1(I1, I2) = \sum_{i=1}^n (|I1(i) - I2(i)|) \quad (2)$$

- Norme  $L2$  (distance Euclidienne). Pour  $r=2$  on a la distance euclidienne :

$$L2(I1, I2) = \sqrt{\sum_{i=1}^n (I1(i) - I2(i))^2} \quad (3)$$

### 1.6.2 Intersection d'histogrammes

L'intersection d'histogramme est une des plus anciennes techniques développées pour la comparaison d'histogramme et elle est la première méthode utilisée dans les systèmes CBIR (Content Base Image Retrieval). Elle a été proposée par Swain et Ballard [11] et son expression est la suivante :

$$d = (H^1, H^2) = 1 - \frac{\sum_{j=1}^n m (H_j^1, H_j^2)}{\sum_{j=1}^n H_j^2} \quad (4)$$

où  $n$  est le nombre de valeurs de chaque histogramme.

$H^1, H^2$  sont deux histogrammes normalisées.

Cette fonction est une mesure de similarité. Deux images sont similaires lorsqu'elles représentent une intersection de leur histogramme proche de 1.

### 1.6.3 Distance quadratique

L'inconvénient des mesures telles que la distance euclidienne ou l'intersection d'histogrammes est qu'elles comparent les composantes des vecteurs une par une, sans prendre en compte les autres composantes. Pour remédier à ce problème, on peut utiliser la distance quadratique [12] qui généralisée la distance Euclidienne. Elle est définie par:

$$D \quad q(I1, I2) = \sqrt{(I1 - I2)^T A (I1 - I2)} = \sqrt{\sum_{i=1}^n \sum_{j=1}^n (I1(i) - I2(i)(I1(j) - I2(j))a_i} \quad (5)$$

La matrice  $A = [a_i]$  permet de pondérer le poids des composantes voisines en fonction de leurs distances à la composante considérée. Cependant, cette mesure a une complexité quadratique.

### 1.6.4 Distance euclidienne

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

$$\text{Dist}(\text{Imreq}, \text{Imcour}) = (\sum_{k=1}^K |V(k)\text{imreq} - V(k)\text{imcour}|^2)^{\frac{1}{2}} \quad (6)$$

où  $i_r$  est l'image requête,  $i_c$ , l'image courante et  $V(k)i_r$ , le vecteur descripteur de l'image requête.

$V(k)i_c$  Le vecteur descripteur de l'image courante.

### 1.6.5 Distance d'Earth Mover (EMD)

La distance d'Earth Mover (EMD) [13] est basée sur la minimisation du coût nécessaire pour transformer une distribution en une autre distribution. Elle peut être appliquée pour calculer la similarité entre deux distributions ou entre deux ensembles de distributions. Elle est définie par l'expression suivante :

$$D \quad E \quad (I1, I2) = \frac{\sum_{i=1}^n \sum_{j=1}^n g_i d_{ij}}{\sum_{i=1}^n \sum_{j=1}^n g_i} \quad (7)$$

L'EMD peut travailler sur des histogrammes et n'a donc pas besoin de quantification de vecteurs. Cette mesure est intéressante pour calculer la distance entre deux images segmentées où chaque bloc serait représenté par une distribution. Cependant, l'EMD est très coûteuse au niveau du temps de calcul.

## 1.7 Domaines d'application

Les domaines d'applications de la recherche d'images par le contenu sont nombreux. On peut notamment citer :

- ) Le domaine de journalisme et publicité: cet univers contient de grosses bases d'images afin d'illustrer leurs articles ou supports publicitaires.
- ) Le domaine militaire qui utilise les applications les plus développées en reconnaissances d'engins via des images radars, des systèmes de guidage et l'identification de cibles par images satellitaires.
- ) Le secteur juridique où les services de polices et de gendarmeries disposent de grandes bases d'image et d'indices.
- ) Le monde médical pour des diagnostics médicaux.
- ) L'Architecture pour retrouver des aménagements intérieurs et des logements.

## 1.8. Conclusion

Dans ce chapitre nous avons présenté les trois types attributs que nous aurons à utiliser dans notre travail d'indexation d'images ainsi que les mesures de similarité les plus courantes. Cependant, lorsqu'il s'agit d'espaces de couleurs, certaines de ces mesures rencontrent certaines difficultés. Notons que certaines distances présentent des inconvénients. Ainsi, pour la distance euclidienne, la plus courante, si certains points sont visuellement proches, ils peuvent être aussi distants que des points visuellement éloignés. L'inconvénient de la distance quadratique, quant à elle, est que des vecteurs visuellement différents peuvent être similaires.

Dans le chapitre suivant nous avons détaillé plus profondément le sujet de l'indexation d'image et abrégé les méthodes les plus connus et utilisé dans ce domaine pour les trois descripteurs (couleur, texture et forme).

# Chapitre 2

## *Méthodes d'Indexation d'images*

### **2.1. Introduction**

Les rapides avancées des technologies de l'informatique ont révolutionné les moyens de communications permettant d'échanger de nombreux documents. Dans notre chapitre, nous présenterons le domaine d'indexation d'images et du système qui gère et utilise efficacement les bases d'images. Nous citons aussi quelques méthodes utilisées dans ce domaine. Notons que le thème de recherche d'images par le contenu devient un sujet très actif dans la communauté internationale depuis plus d'une dizaine d'années. Il s'agit surtout de trouver des caractéristiques permettant d'indexer l'image et de retrouver les similaires lors d'une requête en fonction du contenu des images.

### **2.2 Principe de l'indexation d'images**

Depuis quelques années, l'indexation d'images a émergé comme nouveau thème de recherche dans le domaine de l'analyse des images numériques. L'indexation a pour but de substituer à une image un représentant ou un descripteur moins encombrant qui la caractérise le mieux possible. Cela permettra une meilleure organisation des données, de limiter la quantité de données à examiner durant une recherche, d'y accéder rapidement et de confiner la recherche au maximum.

#### **2.2.1 Phases d'indexation**

Un système d'indexation comprend généralement deux phases de traitement à savoir, l'indexation logique et l'indexation physique.

) *Indexation logique*

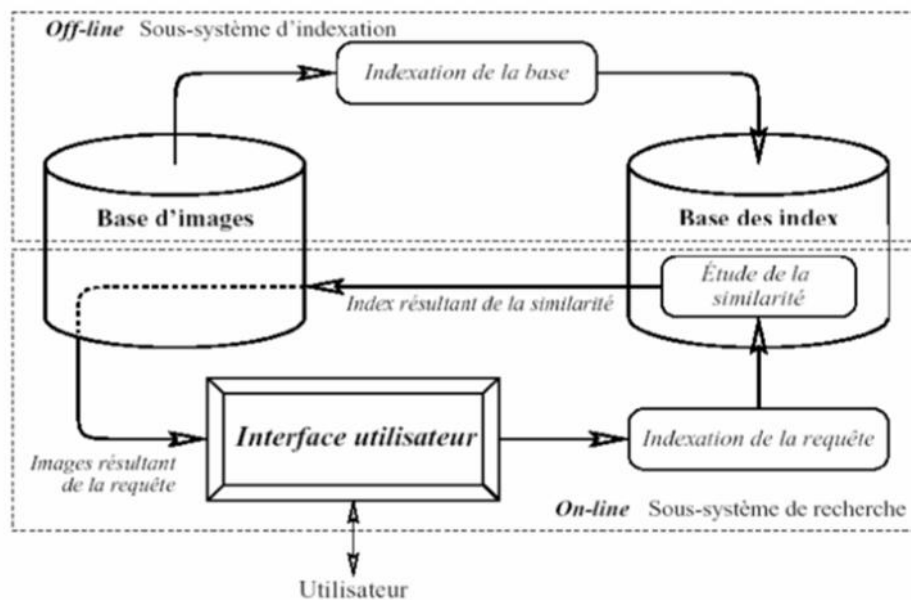
L'indexation logique consiste à extraire et à modéliser les caractéristiques de l'image qui sont principalement la forme, la couleur et la texture. Chacune de ces caractéristiques pouvant être considérée pour une image entière ou pour une région de l'image. Cette indexation logique doit être pertinente pour permettre un accès rapide aux documents recherchés et améliorer ainsi les performances système de recherche d'images.

) *Indexation physique*

L'indexation physique consiste à déterminer une structure efficace d'accès aux données pour trouver rapidement une information. De nombreuses techniques basées sur des graphes sous formes d'arbres ont été proposées.

## 2.3 Système d'indexation par le contenu

Un système d'indexation se compose de deux étapes principales. La première est en mode autonome (Off-ligne) concerne les prétraitements, l'organisation et la gestion de la base d'images de référence (sous système d'indexation de la base). La deuxième est la phase des consultations en ligne (On-line) pour les traitements propres à chaque nouvelle requête de l'utilisateur (sous-système de recherche) (Fig.2.1).



**Fig.2.1** Architecture du system d'indexation et de recherche par le contenu

Ce système s'exécute en deux étapes : l'étape d'indexation et l'étape de recherche.

Dans l'étape d'indexation, des caractéristiques sont automatiquement extraites à partir de l'image et stockées dans un vecteur numérique appelé descripteur.

Grâce aux techniques de gestion de la base de données, on peut stocker ces caractéristiques et les récupérer rapidement et efficacement. Dans l'étape de recherche, le système prend une ou des requêtes à l'utilisateur et lui donne le résultat correspond à une liste d'images ordonnées en fonction de la similarité entre leur descripteur visuel et celui de l'image requête en utilisant une mesure de distance.

## 2.4 Les requêtes

Les requêtes permettent au système de retrouver les images désirées par l'utilisateur. Suivant les besoins de l'utilisateur et le type de base de données images.

### 2.4.1 Requête par mots clés

Une des attentes des utilisateurs dans le domaine de recherche d'images se situe au niveau de sa sémantique c'est pour cela que la plupart des systèmes de recherche d'images développés utilisent des mots clés ou des descripteurs textuelles pour caractériser chaque image de la base (ex : recherche d'images sur Internet). Rappelons que Le modèle sémantique correspond à une description des notions stéréotypées qui apparaissent dans notre fond iconographique. Ce type de caractérisation comporte un certain nombre d'inconvénients. En effet, la description textuelle est une opération longue, coûteuse et difficile à élaborer car l'information externe est manuellement introduite par l'utilisateur, ce qui conditionne la qualité de recherche future, et puis elle ne décrit pas fidèlement le contenu de l'image car elle se fait de manière automatique à partir du nom, de la légende ou du texte qui l'entoure (Fig.2.2) et (Fig.2.3).



**Fig.2.2** Exemple de recherche par mot clés 'avocat'



Fig.2.3.Exemple de recherche d'images par mots clés dans Google

#### 2.4.2 Requête par esquisse

Dans ce cas, soit le système choisit quelques images au hasard dans la base d'images et les montre à l'utilisateur, ce dernier parcourt la base d'images pour choisit une requête et donne son image requête.

#### 2.4.3 Requête par le contenu

Comme son nom l'indique, le principe de cette méthode est d'identifier des images à partir de leur contenu, c'est à dire à partir des données de l'image elles même et non à partir du texte associé aux images.

L'indexation des images, qui se fait automatiquement, nécessite l'extraction des paramètres de celles-ci au préalable. Ces paramètres "quantifient" la couleur, la texture, l'intensité ou, encore, les formes contenues dans l'image et fournissent une "signature".

#### 2.4.4 Requête par l'exemple

Ici, l'image peut, par exemple, soit être fournie par l'utilisateur, soit être choisie par ce dernier dans la base d'images utilisée. Cette technique est simple et ne nécessite pas de connaissances approfondies pour manipuler le système. L'utilisateur fournit un exemple, le système en extrait des descripteurs visuels, puis recherche les images ayant des descripteurs similaires.

## 2.5 Descripteurs de couleur

De nombreuses méthodes de descriptions d'images proposent de caractériser la couleur dans certains espaces colorimétriques pour bénéficier des propriétés de ces derniers. Il existe plusieurs espaces de représentation de la couleur. En effet, l'espace couleur est basé sur le mélange de trois composantes qui donne une couleur. On parle de l'information couleur, ou intensité lumineuse et chromaticité. Le système de représentation de base est l'espace couleur RGB. Parmi les autres plus utilisés, on peut citer le HSV, CEI Lab, CEI Luv, XYZ. Ces systèmes sont décrits ci-après. Cependant, on peut citer le CMJ ou CMY en anglais, le TSL ou HSL en anglais, le YUV, le YIQ, le etc. [14]

### 2.5.1 Espace RVB ou RGB en anglais

C'est l'espace de base et le plus utilisé car la plupart des images d'origine sont codées dans cet espace couleur, ce qui ne nécessite pas de transformation inter-espace couleur, donc facilement applicable. Elles sont donc représentées par les trois paramètres primaires que sont les trois couleurs, le rouge, le vert et le bleu. L'espace RVB est utilisé couramment dans tous les systèmes de vision automatiques, mais il n'est pas forcément le mieux adapté, car si on aura un simple changement d'éclairage dans la scène modifié les 3 composantes et ses couleur ne sont pas toujours les mêmes. La figure 2.4 suivante montre une image originale et les trois images couleur dans l'espace RGB :



Fig.2.4 Images couleurs dans l'espace RGB

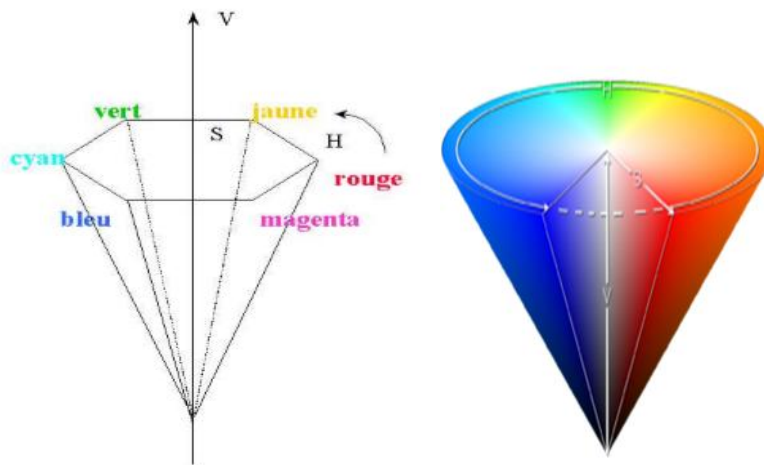
### 2.5.2 Espace TSV

Cet espace est basé sur les trois composantes définies par une approche psychologique et perceptuelle de la couleur la teinte, la saturation et la luminance ou valeur ou Hue, Saturation Value pour HSV en anglais.

Le modèle TSV a été créé en 1978 par Alvy Ray Smith [15]. C'est une transformation non linéaire de l'espace de couleur RVB. Ce système fait partie d'une famille des systèmes basés sur la perception des couleurs comme soit, la teinte, la saturation et la valeur.

Le modèle TSV est particulièrement utilisé dans les applications graphiques. Dans beaucoup de contextes applicatifs, un utilisateur est amené à choisir une

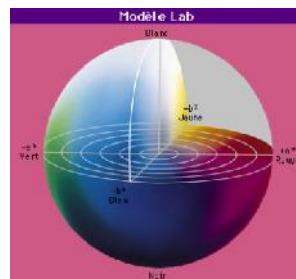
couleur qui sera appliquée à un élément utilisée de cette manière. La roue TSV est souvent employée. La représentation conique est bien adaptée pour représenter tout l'espace TSV en un seul objet (Fig.2.5)



**Fig.2.5** Représentation conique de tout l'espace TSV en un seul objet.

### 2.5.3 Espace Lab ou Luv:

Cet espace. Sont des espaces couleurs possèdent un grande avantages d'être uniformes. Ces espaces sont grandement utilisés dans les systèmes de comparaison d'images. Ils sont très utiles dans le cas de mélanges de pigments, par exemple, pour l'industrie graphique ou du textile.



**Fig.2.6** Espace de couleur Lab

### 2.5.4 Espace couleur La\*b\*

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

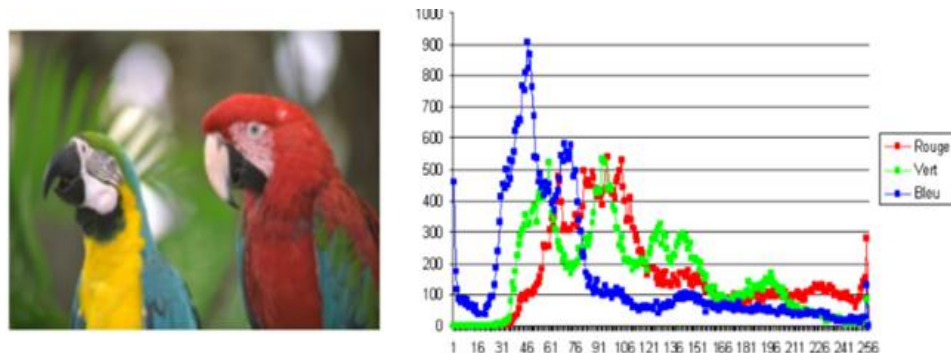
Le mode Lab permet de décrire l'ensemble des couleurs visibles indépendamment de toute technologie graphique, de cette façon il comprend la totalité des couleurs RGB et CMYK.

### 2.5.5 Espace XYZ

L'espace XYZ est un espace de couleur défini par la commission internationale de l'éclairage (CIE) en 1931. L'espace  $C X$  introduit la notion de *luminance*, intensité lumineuse subjective indépendante de la couleur, donnée par la composante  $Y$ . Il utilise deux autres composantes  $X$  et  $Z$ , choisies de telle manière qu'elles prennent toujours des valeurs positives pour décrire les couleurs visibles. Ceci a ouvert la voie à l'espace  $C x$  qui sépare parfaitement les notions de luminance  $Y$  et de *chrominance*  $x$ , sensation colorée indépendante de l'intensité, représentée sur le diagramme de chromaticité.

### 2.5.6 Histogramme couleur

L'histogramme est un outil fondamental du traitement d'images avec de très nombreuses applications. C'est la distribution des intensités de niveaux de gris ou de couleur. L'histogramme est définie comme une fonction discrète qui associé à chaque valeur d'intensité le nombre de pixels prenant cette valeur, c'est-à-dire comptant le nombre de pixel pour chaque intensité (couleur) de l'image et pour mieux visualiser le descripteur de ces intensités de l'image (Fig.2.7).



**Fig.2.7** Image couleur et son histogramme des trois composantes RVB

### 2.5.6 Le Corrélogramme

Les histogrammes de couleurs souffrent d'un manque d'information locale parce que ces derniers n'enregistrent pas la structure locale des couleurs mais seulement la composition globale. Le corrélogramme de couleurs enregistre la probabilité de présence d'une couleur dans un voisinage prédéfini d'une autre couleur. Une nouvelle représentation des informations locales et globales conjointement est nommée « histogrammes de blobs ». Un corrélogramme présente les corrélations spatiales entre les niveaux de gris en fonction de la distance inter pixel.

## 2.6 Descripteurs de forme

La forme est une des principales caractéristiques visuelles importantes. Elle est une des caractéristiques de base pour décrire le contenu d'une image.

Nous distinguons deux catégories de descripteurs de formes : les descripteurs basés régions et les descripteurs basés frontières. La forme est l'un des attributs de bas niveau les plus utilisés pour décrire la structure géométrique du contenu visuel. Même si la caractérisation du contenu géométrique s'est avérée complexe plusieurs primitives géométriques sont montrées leurs intérêts dans les systèmes de recherche.

La forme est intéressante pour retrouver certains concepts qui ne peuvent l'être autrement. Prenons le mot ballon, par exemple. Il n'y a pas de couleurs qui puissent caractériser un ballon puisque cet objet peut être de toutes les couleurs, par contre, il y a une forme très caractéristique. La forme peut avoir certains avantages également sur la texture. Nous présentons dans ce qui suit quelques méthodes de description de la forme.

### 2.6.1 Les moments géométriques

Les moments invariants sont un ensemble de fonctions non linéaires qui sont invariants à la translation aux changements d'échelle et à la rotation.

Les moments géométriques permettent de décrire une forme à l'aide des propriétés statistiques. Ils représentent les propriétés spatiales de la distribution des pixels dans l'image, parmi ces moments on peut citer ceux de Hu et ceux de Legendre.

#### ) Moments de Hu

Ce sont des moments issus des moments d'inertie, nommés également moments de Hilbert Ils se construisent à partir du moment normé

$$N_n = \frac{J^{m_k n}}{M_0^\alpha} \overline{M}_m \quad (2.1)$$

où  $\alpha = \frac{1}{2}(m + n) + 1$

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

$$\begin{aligned} h_1 &= N_{20} + N_{02} \\ h_2 &= (N_{20} - N_{02})^2 + 4N_{11}^2 \\ h_3 &= (N_{30} - 3N_{12})^2 + (N_{03} - 3N_{21})^2 \\ h_4 &= (N_{30} + N_{12})^2 + (N_{03} + N_{21})^2 \\ h_5 &= (N_{30} - 3N_{12})(N_{30} + N_{12})[(N_{30} + N_{12})^2 - 3(N_{03} + N_{21})^2] \\ &\quad + (3N_{21} - N_{03})(N_{03} + N_{21})[3(N_{30} + N_{12})^2 - (N_{03} + N_{21})^2] \\ h_6 &= (N_{20} - N_{02})[(N_{30} + N_{12})^2 - (N_{03} + N_{21})^2] \\ &\quad + 4N_{11}(N_{30} + N_{12})(N_{03} + N_{21}) \\ h_7 &= (3N_{12} - N_{30})(N_{03} + N_{21})[3(N_{30} + N_{12})^2 - (N_{03} + N_{21})^2] \\ &\quad + (3N_{21} - N_{03})(N_{30} + N_{12})[(N_{30} + N_{12})^2 - 3(N_{03} + N_{21})^2] \end{aligned} \quad (2.2)$$

) *Les moments de Legendre*

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

$$\forall x \in [-1,1], \forall n \in \mathbb{N}, P_n(x) = \frac{1}{2^n n!} \frac{d^n (x^2-1)^n}{d^n} \quad (2.3)$$

### 2.6.2 Les moments orthogonaux

Par opposition aux moments géométriques qui sont définis par rapport à une base quelconque  $(x, y)$ , les moments orthogonaux, comme leur nom l'indique, sont définis dans une base orthogonale, ce qui évite la redondance des informations portées par chacun des moments. On peut notamment citer ceux de Fourier et ceux de Zernike.

) *Les descripteurs de Fourier*

La transformée de Fourier permettent de représenter les pixels de l'image selon des modèles périodiques qui sont caractérisés par leurs fréquences et leurs directions. Ces éléments sont souvent exploités en indexation pour représenter l'objet dans des structures de données sous forme d'arbres à titre d'exemple. Le descripteur de Fourier générique est plus simple à calculer. En effet, les caractéristiques sont dans le domaine spectral et ont des meilleures performances d'indexation en raison de son analyse multi-résolution dans les deux directions radiales et circulaire de la forme.

) *Les moments de Zernike*

Les moments de Zernike ont fréquemment utilisés pour capturer les caractéristiques globales d'une image en reconnaissance et en analyse d'images. Un moment de Zernike est un nombre complexe qui contient deux valeurs distinctes : le module et la phase. Cependant, la manière habituelle de comparer deux descripteurs de Zernike ne prend en compte que le module des moments car cela apporte l'invariance en rotation. Dans le cadre de l'indexation et de la reconnaissance 2D et 2D-3D, cette perte d'information n'est pas anodine lorsque l'on compare des images et peut induire des résultats erronés et imprécis, Utiliser les informations de phase des moments de Zernike (avec le module) dans le processus de comparais on semble être une façon naturelle d'améliorer la mesure de similarité. Ce type des moments a été initialement introduit par Teague [16] et sont construits à partir des polynômes complexes et forment un ensemble orthogonal complet définie sur le disque unité. Ils sont invariants par rapport à la rotation et aux changements d'échelles et présentent des capacités intéressantes en termes de résistance aux bruits, efficacité informative et possibilité de reconstruction.

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

$$A_m = \frac{m+1}{\pi} \iint I(x, y) [V_{m,n}(x, y)] d d \quad (2.4)$$

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

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

$$V_{m,n}(r, \theta) = R_{m,n}(r) e^{-j\theta} \quad (2.5)$$

où  $R_{m,n}(r)$  est le polynôme radial orthogonal avec

$$R_{m,n}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s \frac{(m-s)!}{s! \left(\frac{m+|n|}{2}-s\right)! \left(\frac{m-|n|}{2}-s\right)!} r^{m-2s} \quad (2.6)$$

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

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

## 2.7 Descripteurs de Texture

Comme nous l'avons précisé précédemment, il n'existe pas de définition pertinente de la texture. Cependant, on peut rappeler que la texture est la répétition d'éléments de base construits à partir de pixels qui respectent un certain ordre. Elle se traduit aussi comme étant un arrangement spatial des pixels que l'intensité ou la couleur seules ne suffisent pas à décrire. Elle traduit l'aspect homogène d'une zone et peut être décrite selon ses propriétés spatiales et fréquentielles. Il existe plusieurs méthodes de description de la texture les plus utilisées concerne les propriétés fréquentiels s'appuie sur la transformée de Fourier, Filtre de Gabor, les Ondelettes, la matrice de cooccurrence, etc. Elle repose sur l'analyse d'une fonction de densité spectrale dans un domaine fréquentiel. La texture définie comme un mélange de signaux de fréquences, d'amplitudes et de directions différentes. Ces méthodes consistent à extraire l'énergie portée par le signal dans diverses bandes de fréquences.

### 2.7.1 Filtres de Gabor

Il s'agit de la somme de deux fonctions sinusoïdales, l'une paire et réelle, l'autre impaire et imaginaire modulées par une courbe gaussienne, la fonction  $g(x)$  est donc complexe. Dans le cas bidimensionnel, on va définir tout un ensemble de filtre gaussien séparable polairement. On subdivise ainsi le demi-plan fréquentiel  $(f - f)$  en  $N$  orientation et  $M$  bandes passantes. La fonction de transfert d'un filtre de Gabor bidimensionnel est donc :

$$G(x, y) = \cos(W_X X + W_Y Y) e^{-\frac{x^2+y^2}{\sigma^2}} \quad (2.7)$$

$$G(x, y) = \sin(W_X X + W_Y Y) e^{-\frac{x^2+y^2}{\sigma^2}} \quad (2.8)$$

### 2.7.2 Les matrices de cooccurrence

La matrice de cooccurrence mesure la probabilité pour qu'un couple de niveaux de gris, vérifiant une loi spatiale donnée, apparaisse dans l'image. Le niveau de gris d'un pixel d'une image dépend fortement des niveaux de gris des pixels voisins.

Nous avons implémenté la méthode de la matrice de cooccurrence à niveau de gris pour extraire les indices de textures, car elle est très utilisée en recherche d'images, et en général donne de bons résultats.

Cette approche est basée sur la probabilité jointe de la distribution des pixels dans l'image.

L'élément  $p_{d,\theta}(i, j)$  de la matrice de cooccurrence définit la fréquence d'apparition des couples de niveaux de gris  $i$  et  $j$  pour les couples de pixels séparés par une distance  $d$  selon la direction  $\theta$ . Cette matrice décrit les régularités observables dans les niveaux de gris des pixels d'une région. Le calcul de la matrice de cooccurrence nécessite le choix d'une distance et d'un angle de déplacement. Il a été noté par plusieurs chercheurs que la distance d'un pixel combinée avec des angles respectifs de  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  et  $135^\circ$  donne de bons résultats. C'est la solution que nous avons adoptée, ce qui nous donne à la fin quatre matrices de cooccurrence pour chaque image. Les matrices sont ensuite normalisées par la formule suivante :

$$p_{d,\theta}(i, j) = \frac{p_{d,\theta}(i, j)}{M \times N} \quad (1)$$

Pour une image de taille  $M \times N$

Afin d'estimer la similarité entre les matrices de cooccurrences, Haralick a proposé 14 caractéristiques statistiques extraites à partir de cette matrice. Actuellement, seulement les quatre caractéristiques les plus appropriées sont largement utilisées, et que nous avons calculés pour notre application sont : l'énergie, l'entropie, le contraste et le moment inverse de différence.

· **L'énergie**

$$\sum_i \sum_j (p_{d,\theta}(i,j))^2 \quad (2)$$

Ce paramètre mesure l'uniformité de la texture. Il atteint de fortes valeurs lorsque la distribution des niveaux de gris est constante ou de forme périodique. Dans ce dernier cas, les valeurs élevées d'énergie sont obtenues pour les matrices  $p_{d,\theta}(i,j)$  lorsque  $d, \theta$  correspond à la période.

· **L'entropie**

$$-\sum_i \sum_j (\log p_{d,\theta}(i,j) p_{d,\theta}(i,j)) \quad (3)$$

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). Elle est fortement corrélée (par l'inverse) à l'énergie.

· **Le contraste**

$$\sum_i \sum_j ((i-j)^2 p_{d,\theta}(i,j)) \quad (4)$$

Il mesure les variations locales des niveaux de gris. Il permet de caractériser la dispersion des valeurs de la matrice par rapport à sa diagonale principale. Ce paramètre est fortement non corrélé à l'énergie.

· **Le moment inverse de différence**

$$\sum_i \sum_j \frac{p_{d,\theta}(i,j)}{1+(i-j)^2} \quad (.)$$

Ce paramètre mesure l'homogénéité de l'image. Il est corrélé à une combinaison linéaire des variables énergie et contraste.

**8. Conclusion**

Dans ce chapitre nous avons présenté un aperçu sur le système actuel d'indexation d'images, puis nous avons abordé les méthodes d'indexations rencontrées dans la littérature, le système d'indexation est basé sur des descripteurs globaux calculés à partir des attributs couleur, texture et forme. Plusieurs moteurs de recherches d'images déployés sur le web sont basés sur les techniques d'indexation d'images précitées. Cette exploitation des méthodes d'indexation d'images sur le Web, a été rendue nécessaire du fait du contenu important des images et de leur abondance.

# Chapitre 3

## *Application et résultats*

### **3.1 Introduction**

L'indexation a pour but d'extraire et de représenter le sens d'un document de manière à ce qu'il puisse être retrouvé par l'utilisateur. Rappelons que notre objectif est d'extraire de l'image les descripteurs visuels qui permettent de retrouver efficacement des images similaires, de trouver une représentation pertinente de ces descripteurs et une mesure de similarité efficace et accéder rapidement à l'information. Puisque l'on souhaite construire des systèmes de recherche d'images qui soient utilisables par l'être humain, et que c'est l'être humain qui donne un sens à ce qu'il voit, il peut être intéressant de s'inspirer du système de perception humaine pour choisir les espaces visuels afin de s'approcher le plus possible de la compréhension qu'à l'être humain de l'image, et ainsi réduire le fossé sémantique.

Ainsi, dans notre recherche et indexation d'images par le contenu nous avons utilisé les trois descripteurs de couleur, de forme et de texture. Pour la couleur, nous avons utilisé les histogrammes des couleurs RVB qui sont des descripteurs très efficaces pour trouver les images similaires. Pour capturer les caractéristique de forme nous avons choisis le moment de Zernike qui a prouvé sa supériorité sur les autres fonctions de moments de par sa capacité de description et sa robustesse au bruit et aux déformations, les polynômes de Zernike sont définis sur une surface circulaire, Ils sont donc bien adaptés pour décrire la forme. Pour la texture, nous avons utilisé méthode de matrices de cooccurrences de Haralick.

## 3.2 Mesure similarité

Nous l'avons présenté dans le chapitre précédent sur les différents descripteurs essentiels dans l'image, nous avons utilisé trois distances de similarités. En effet, dans l'indexation d'image, on cherche à comparer les images pour déterminer les images qui ressemblent à une image requête. Pour ce faire, il faut est nécessaire d'extraire d'abord les trois descripteurs de couleur, de texture et de forme de chaque image de la base (hors ligne) et de l'image requête (en ligne).

Rappelons que pour la couleur, il s'agit de l'histogramme de l'espace RGB, pour la texture, il s'agit de la matrice de cooccurrence et pour la forme, nous calculons le moment de Zernike.

### 3.2.1 Mesure de similarités entre les vecteurs de l'histogramme

Pour mesurer la similarité entre un vecteur couleur de l'image requête  $H_{I_r}$  avec celui de l'image courante  $H_{I_b}$ , nous avons utilisé la distance euclidienne qui est donnée par la formule suivante :

$$D_n(H_{I_r}, H_{I_b}) = \sqrt{\sum_{i=1}^n (H(i)_{I_r} - H(i)_{I_b})^2} \quad (3.1)$$

### 3.2.2 Mesures de similarités entre les vecteurs de texture

Pour mesurer la similarité entre un vecteur texture de l'image requête  $V_{t_r}$  avec celui de l'image courante  $V_{t_b}$ , nous avons utilisé aussi la distance euclidienne. Elle est donnée par la formule suivante :

$$D_c(V_{t_r}, V_{t_b}) = \sqrt{\sum_{n=1}^N |V_{t_r}(n) - V_{t_b}(n)|^2} \quad (3.2)$$

$N$  est le nombre de composantes de notre vecteur caractéristique de texture.

### 3.2.3 Mesures de similarités entre descripteurs de moments de Zernike

La manière habituelle de comparer deux descripteurs de Zernike est une simple distance euclidienne entre les modules des moments :

$$D_z(H_{I_r}, H_{I_b}) = \sqrt{\sum_{i=1}^i (Z(i)_{I_r} - Z(i)_{I_b})^2} \quad (3.3)$$

### 3.2.4 Mesures de distance de similarités de la combinaison

La distance des similarités de la combinaison des trois types de descripteurs calculée par l'expression suivante:

$$\text{Dist}_g(V_r, V_b) = D_z + \beta D_n + \gamma D_c \quad (3.4)$$

Où  $\alpha$  est le coefficient des moments de Zernike,  $\beta$  le coefficient d'histogramme et  $\gamma$  le coefficient de la matrice de cooccurrence, avec  $\alpha + \beta + \gamma = 1$

### 3.3 Evaluation d'un système

À partir de nos résultats. On mesure l'efficacité d'une technique de recherche et d'indexation d'image en utilisant deux critères de mesure des performances sont le rappel et la précision.

En pratique, on cherche un bon compromis entre le rappel et la précision, afin d'évaluer un système on fait un graphique du rappel par rapport à la précision. Un exemple pour bien comprendre ces deux critères de mesures, j'ai trois documents pertinents sur un total de 10 dans ma base de données. Un algorithme sélectionne 5 documents, dont deux sont pertinents, quel est le rappel et la précision ?

La précision est de  $2/5$  alors que le rappel est de  $2/3$ .

Si le nombre d'image présentée est importante alors le taux de rappel est élevé, à l'inverse si le système possède de nombreux images similaires mais que ceux-ci n'apparaissent pas dans la liste des réponses, le rappel est faible.

Si la précision est élevée, cela signifie que peu des images inutiles sont proposées par le système et que ce dernier peut être considéré comme précis.

#### 3.3.1 Le rappel

Le rappel est défini par le nombre des images pertinentes retrouvés au regard du nombre de images pertinentes que possède la base de données. Cela signifie que lorsque l'utilisateur interroge la base il souhaite voir apparaître tous les images qui pourraient répondre à son besoin d'information.

On calcule le rappel avec la formule suivante :

$$R = \frac{|R_a|}{|R|} \quad (3.5)$$

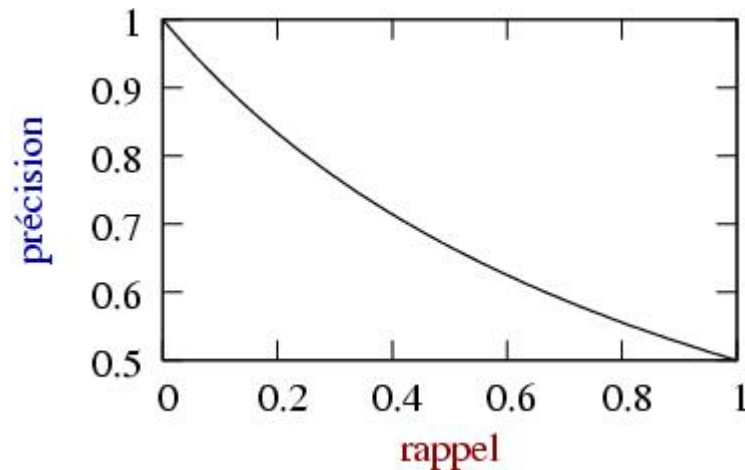
#### 3.3.2 La précision

La précision est le nombre des images pertinentes retrouvés rapporté au nombre des images total proposé par le moteur de recherche pour une requête donnée.

La précision est calculée par:

$$P = \frac{|R_a|}{|A|} \quad (3.6)$$

Une des manières de tenir compte à la fois du rappel et de la précision d'un système est d'exprimer les valeurs de précision en fonction des différents niveaux de rappel selon la courbe rappel/précision (Fig.3.1).



**Fig.3.1** La courbe rappel/précision

### 3.4 Les bases d'images utilisées

Une base d'image est une base de données qui regroupe de plusieurs centaines de milieu d'images, nous travaillons principalement sur quatre bases d'images généralistes dans le but d'évaluer et de valider notre système d'indexation et de recherche d'images par le contenu.

#### 3.4.1 La base d'image COIL (Columbia Object Image Library),

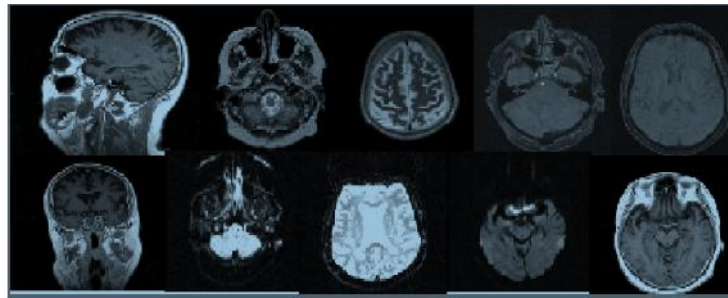
Dans cette base, il existe deux bases **COIL**. La COIL-100 qui comprend 1440 images en niveau de gris. Dans ce mémoire, c'est La COIL-100 que nous utiliserons. Elle comprend 7200 images en couleur (100 objet \* 72 image /objet) où chaque image a une taille de 128\*128 pixels (Fig.3.2).



**Fig.3.2** Objets utilisés dans COIL-100

#### 3.4.2 La base d'images Cerebral Standard MR

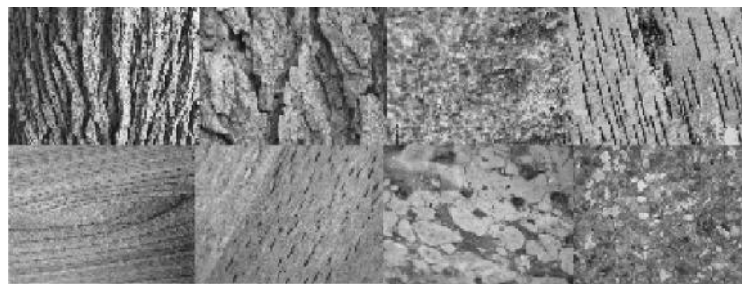
Cette base que nous avons utilisée contient 329 images en niveau de gris classée en plusieurs classes, chaque image à une taille de 512\*512 pixels La figure suivante montre un extrait de cette base



**Fig.3.3** Quelques exemples de la base Cérébral standard MR.

### 3.4.3 La base d'images texture

Elle contient 400 images en niveau de gris. Chaque image a une dimension de 640\*480.



**Fig.3.4** Quelques exemples de la base d'images de texture.

### 3.4.4 La base d'images ORL (Olivetti Research Laboratory)

La base de données ORL des visages est utilisée dans la reconnaissance faciale. Cette base comprend 40 images qui ont été prises à des moments différents et en fonction de l'éclairage, des expressions faciales et les détails du visage. La taille de chaque image est de 92\*112 pixels, avec 256 niveaux de gris par pixel. La figure 14 montre les 40 personnes utilise dans cette base:



**Fig.3.5** Les 40 personnes de la base ORL

### 3.5 Tests et résultats

En premier lieu, notre système va ordonner les distances calculées pour toutes les images de la base de la plus petite à la plus grande. Nous avons imposé l'affichage de dix images qui sont similaires à l'image requête que l'on a choisie de la plus proche à la plus loin.

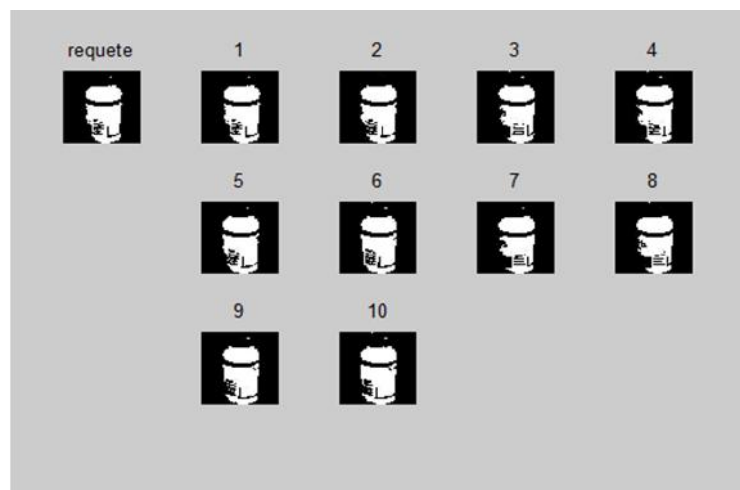
Nous présentons dans cette partie quelques courbes de rappel/précision, afin d'avoir une représentation visuelle global de la qualité de descripteur utilisé.

#### 3.5.1 Evaluation sur la base COIL

L'évaluation visuelle est la suivante :

- *Les résultats des moments de Zernike*

Les figures suivantes présentent les résultats de la recherche d'images par la forme en utilisant les moments de Zernike dans la base d'images COIL binaire. Nous avons choisi la requête 48.png.

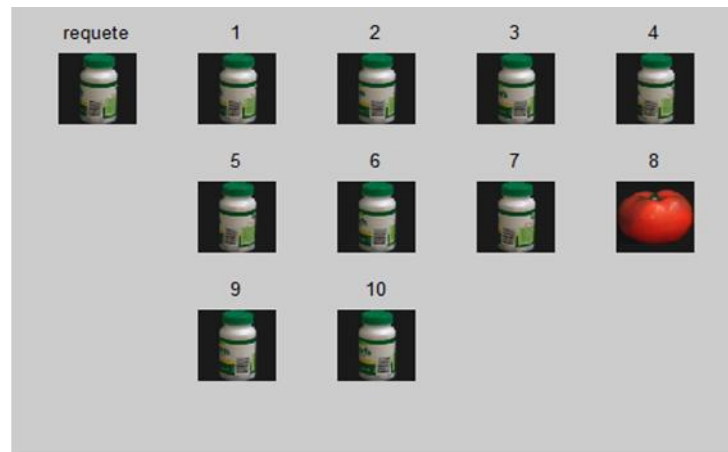


**Fig.3.6** Résultats de la recherche dans la base COIL binaire utilisant la Méthode des moments de Zernike.

A partir des résultats précédents, on peut remarquer que les moments de Zernike donnent des résultats très satisfaisants, ce qui permet de dire qu'ils sont de bons descripteurs de forme.

- ) *Les résultats d'histogramme RVB*

Les figures suivantes représentent quelques exemples de l'histogramme RVB que nous avons testé dans les bases COIL couleur avec les requêtes 48.png.



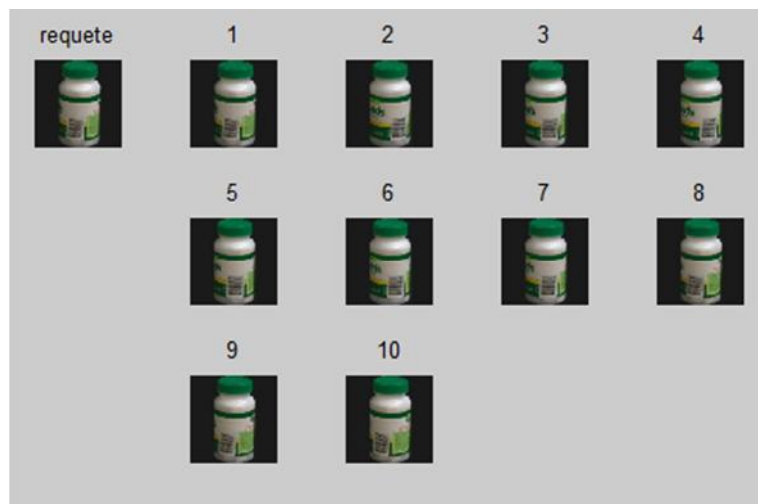
**Fig.3.7** Résultat de la recherche dans la base COIL binaire utilisant la méthode d’histogramme RVB

- *Les résultats pour la matrice de cooccurrence*



**Fig.3.8** Résultat de la recherche dans la base COIL binaire utilisant la matrice de cooccurrence.

- *Les résultats de la combinaison*



**Fig.3.9** Résultats de la recherche dans COIL binaire utilisant la combinaison.

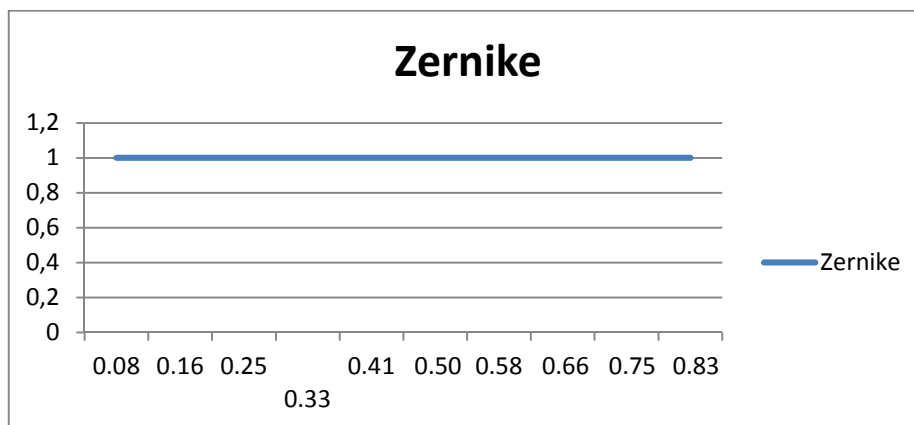
D'après les résultats nous concluons que les moments de Zernike et la combinaison sont les meilleures caractéristiques qui satisfont notre besoin.

L'évaluation par la mesure de la qualité des réponses montre que, à partir des résultats présentés dans la section précédente, nous pouvons calculer la table de rappel et précision. Par exemple, considérons la requête 48.png pour laquelle 10 images sont pertinentes dans la base **COIL**. Soit la liste des réponses du système  $\{1, 2, \dots, 10\}$ . Les images pertinentes sont marquées par la lettre P comme cela est montré dans le tableau suivant (Tab.1) :

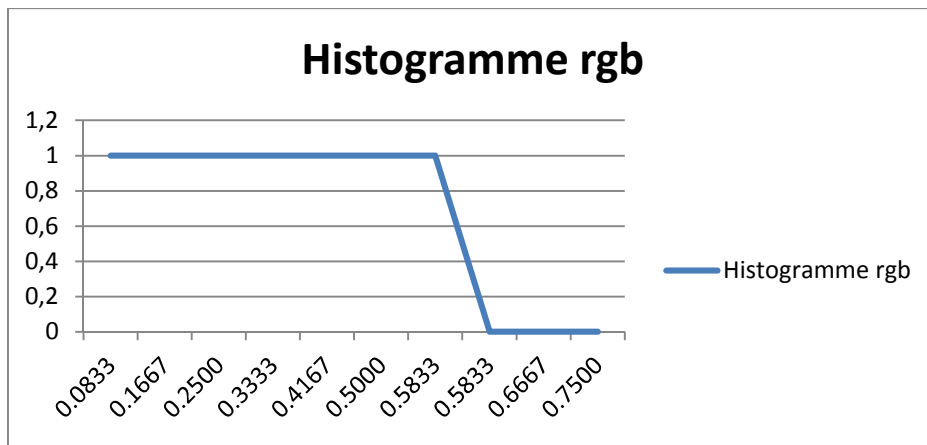
**Tab.1** Valeurs de rappel et de précision d'un exemple dans la base **COIL**

Image	Pertinente	Précision	rappel
1	P	1	0,08
2	P	1	0,16
3	P	1	0,25
4	P	1	0,33
5	P	1	0,41
6	P	1	0,5
7	P	1	0,58
8	P	1	0,66
9	P	1	0,75
10	p	1	0,83

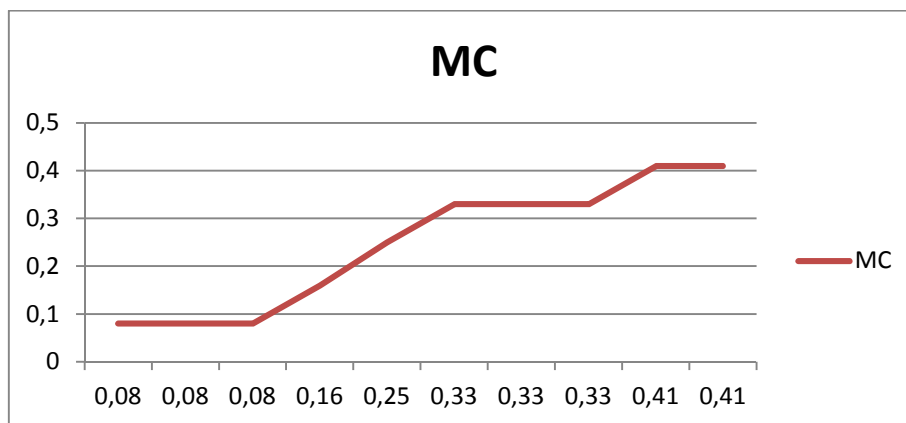
A partir du tableau Tab.1, nous obtenons la courbe de rappel/précision pour l'image 48.png avec le moment de Zernike (Fig.3.10), l'histogramme RGB (Fig.3.11), la matrice de cooccurrence (Fig.3.12) et combinaison des ces trois caractéristiques(Fig.3.13).



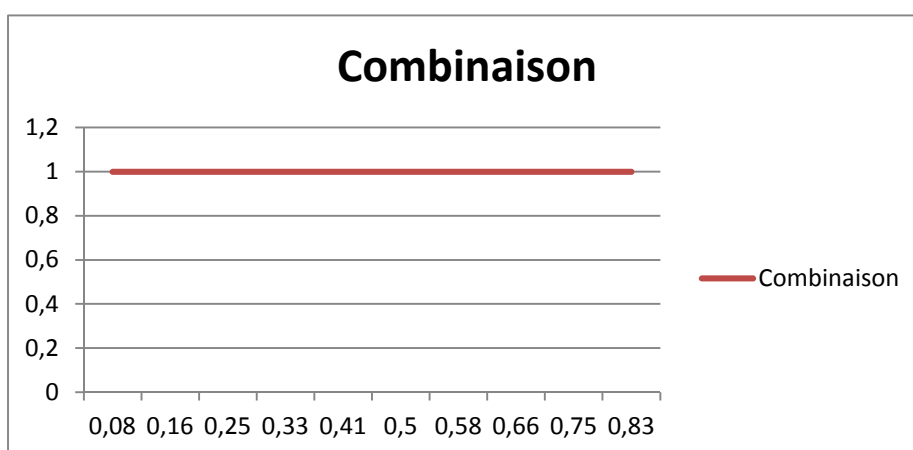
**Fig.3.10** Courbe de rappel/précision pour l'image 48.png en appliquant Zernike



**Fig.3.11** Courbe de rappel/précision pour l'image 48.png en appliquant l'histogramme RGB



**Fig.3.12** Courbe de rappel/précision pour l'image 48.png en appliquant la matrice de cooccurrence



**Fig.3.13** Courbe de rappel/précision pour l'image 48.png avec la combinaison

Le tableau ci-dessus montre les résultats de la base COIL on constate que les moments de Zernike et la combinaison sont les meilleurs, ainsi

l'histogramme NG qui donne des résultats bonne et la matrice de cooccurrence moins satisfaisantes

**Tab.1** Résultats de la base COIL

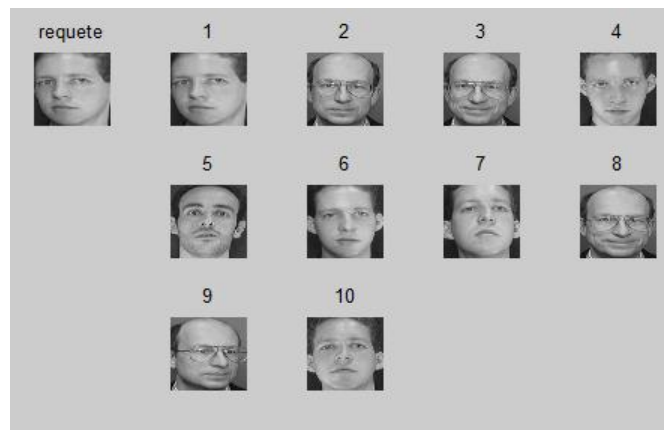
descripteur	Les moments de Zernike	L'histogramme NG	Matrice cooccurrence	Combinaison
Nombre images pertinentes dans 10	10	9	5	10
Précision	100%	90%	50%	100%
Rappel	83%	75%	42%	83%

### 3.5.2 Evaluation sur la base Visages

#### Evaluation visuelle

Les figures suivantes donnent un exemple avec la base ORL. L'image requête est l'image 4.JPG.

) *Les résultats de moment de Zernike*



**Fig.3.25** Résultat de la recherche dans la base ORL binaire utilisant les moments de Zernike.

) *Les résultats d'histogramme NG*



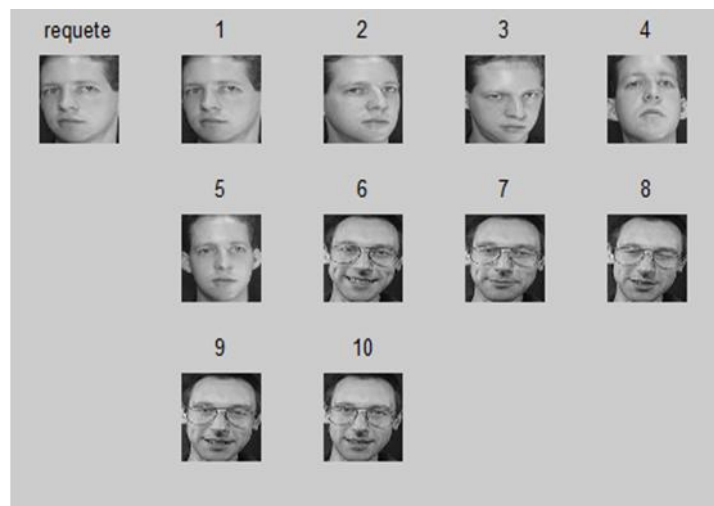
**Fig.3.26** Résultat de la recherche dans la base ORL binaire utilisant l'histogramme en niveau de gris

) Les résultats de matrice de cooccurrence



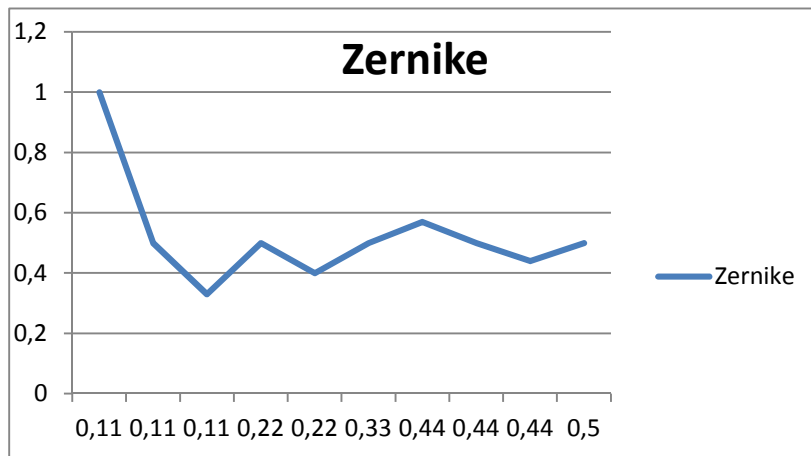
**Fig.3.26** Résultat de la recherche dans la base ORL binaire utilisant la Matrice de cooccurrence

) Les résultats de la combinaison

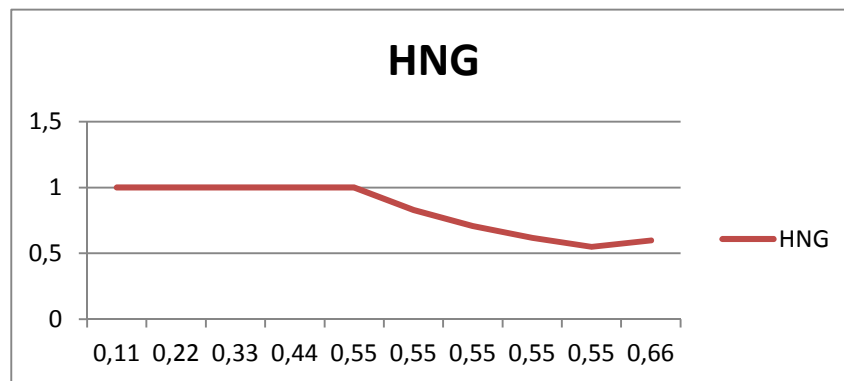


**Fig.3.27** Résultat de la recherche dans la base ORL binaire utilisant la Combinaison

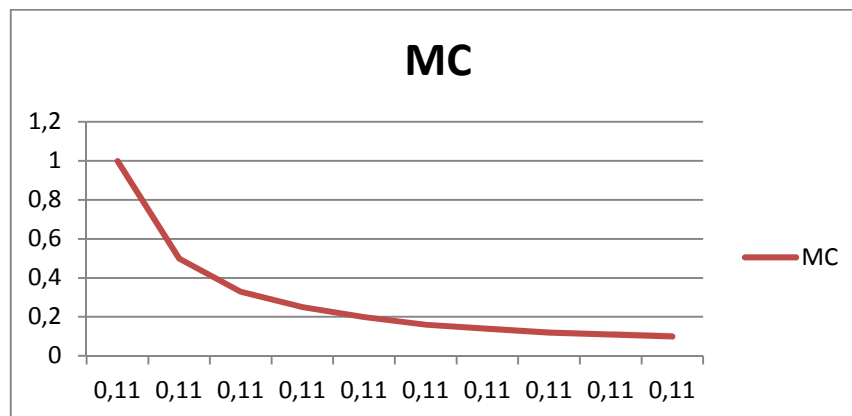
Pour la base Visages, nous obtenons les courbes de rappel/précision suivantes:



**Fig.3.14** Courbe de rappel/précision pour l'image 4.jpg en appliquant Zernike



**Fig.3.15** courbe de rappel/précision pour l'image 4.jpg en appliquant histogramme niveau de gris



**Fig.3.16** Courbe de rappel/précision pour l'image 4.jpg en appliquant la Matrice de cooccurrence

Le tableau ci-dessus montre les résultats de la base ORL. Nous constatons que les moments de Zernike, l'histogramme en niveau de gris et la combinaison sont les meilleurs. La matrice de cooccurrence est la moins satisfaisante.

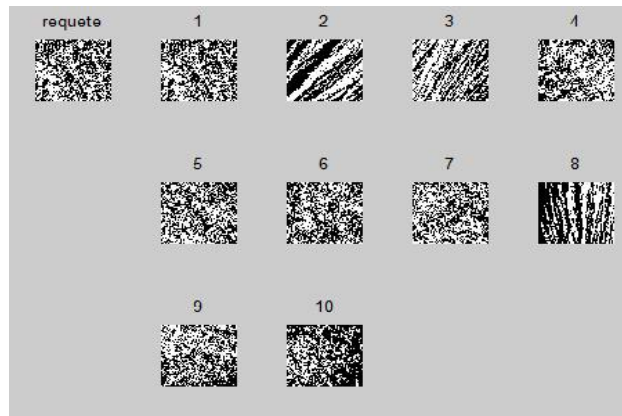
**Tab.1** Résultats de la base ORL

descripteur	Les moments de Zernike	L'histogramme NG	Matrice de cooccurrence	Combinaison
Nombre images pertinentes dans 10	5	5	1	5
Précision	50%	50%	10%	50%
Rappel	45%	45%	11	45%

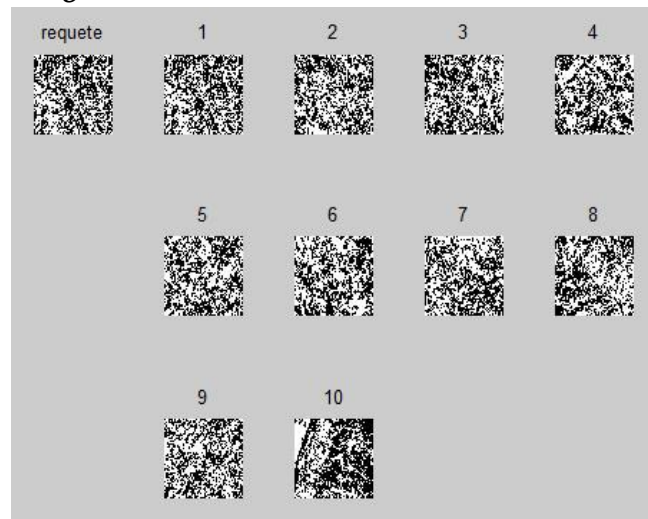
### 3.5.3 Evaluation sur la base Texture

L'évaluation visuelle est la suivante :

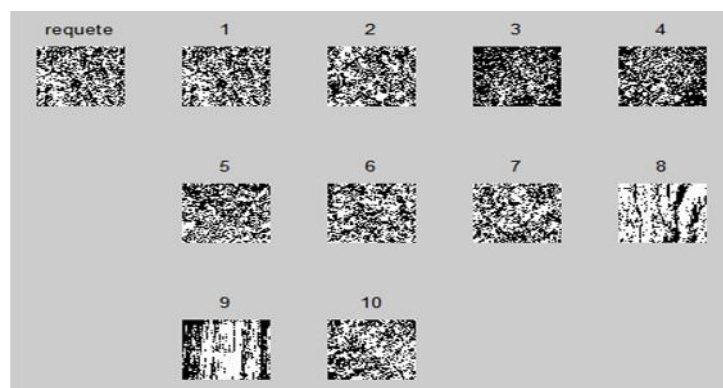
Les figures suivantes, parmi les dix images résultats, donne des exemples avec la base texture. L'image requête est la 280.jpg.

) *Les résultats de moment de Zernike*

**Fig.3.17** Résultats de la recherche dans la base texture utilisant les moments de Zernike

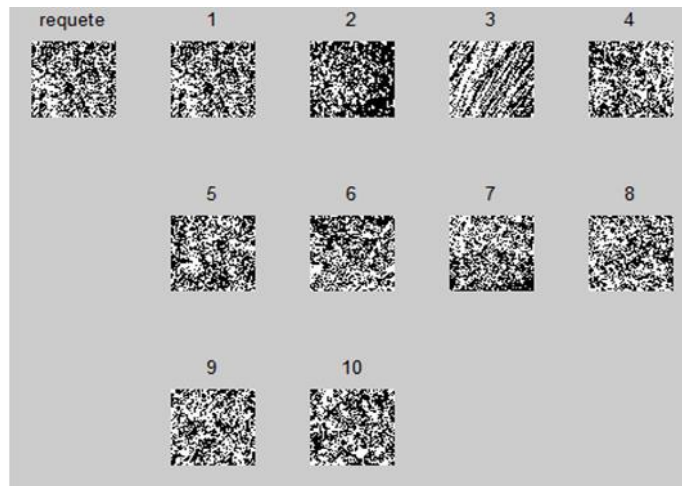
) *Les résultats d'histogramme NG*

**Fig.3.18** Résultats de la recherche dans la base texture utilisant l'histogramme en niveau de gris

) *Les résultats de matrice de cooccurrence*

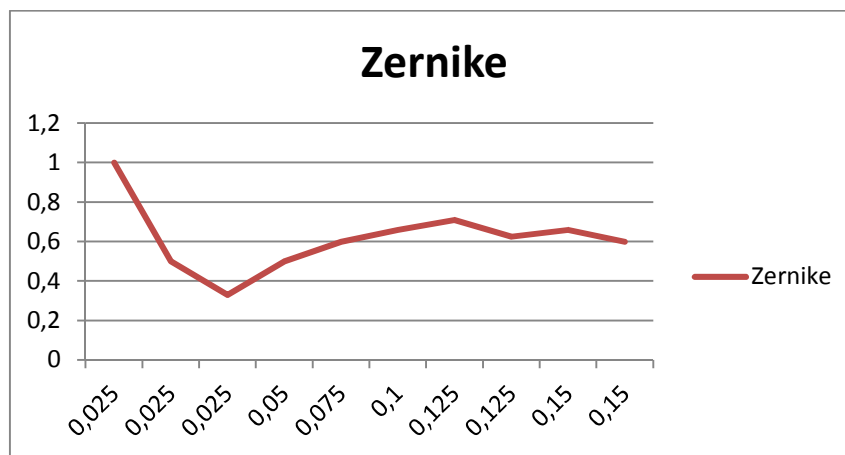
**Fig.3.19** Résultat de la recherche dans la base texture utilisant la matrice de cooccurrence

) Les résultats de la combinaison

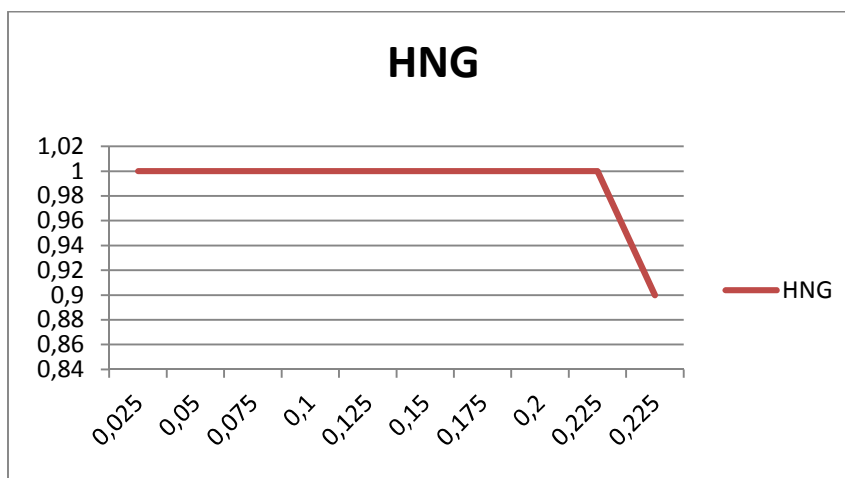


**Fig.3.20** Résultat de la recherche dans la base texture utilisant la combinaison

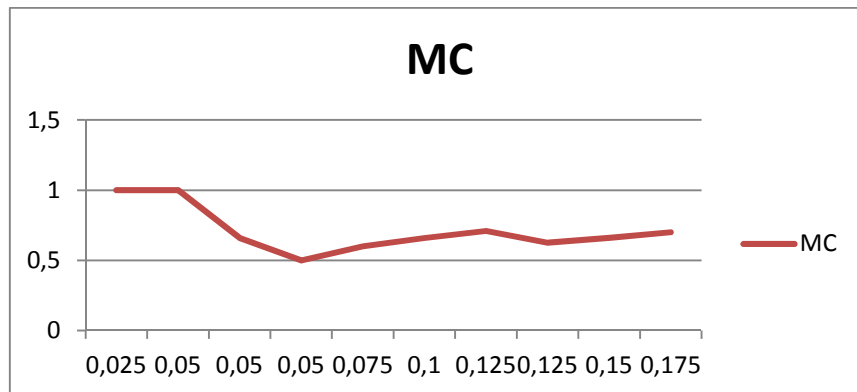
Les courbes de rappel/précision de la base Texture sont illustrées comme suit :



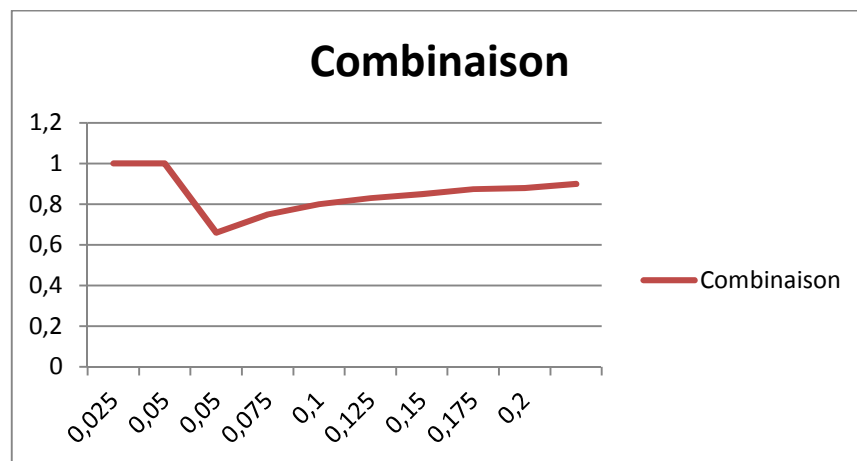
**Fig.3.21** Courbe de rappel/précision pour l'image 280.jpg en appliquant les moments de Zernike



**Fig.3.22** Courbe de rappel/précision pour l'image 280.jpg en appliquant l'histogramme en niveau de gris



**Fig.3.23** Courbe de rappel/précision pour l'image 280.jpg en appliquant la Cooccurrence



**Fig.3.24** Courbe rappel/précision pour l'image 280.jpg par la Combinaison

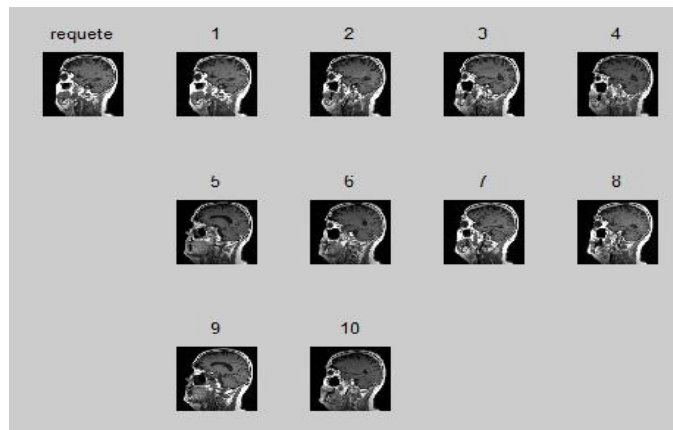
A partir de ces résultats, la combinaison il apparait que l'histogramme NG donne les meilleures dans cette base (Tab.4)

**Tab.4** Résultats de la base texture

descripteur	Moments Zernike	Histogramme niveau de gris	Matrice cooccurrence	Combinaison
Nombre images pertinentes dans 10	6	9	7	9
Précision	60%	90%	70%	90%
Rappel	15%	23%	18%	23%

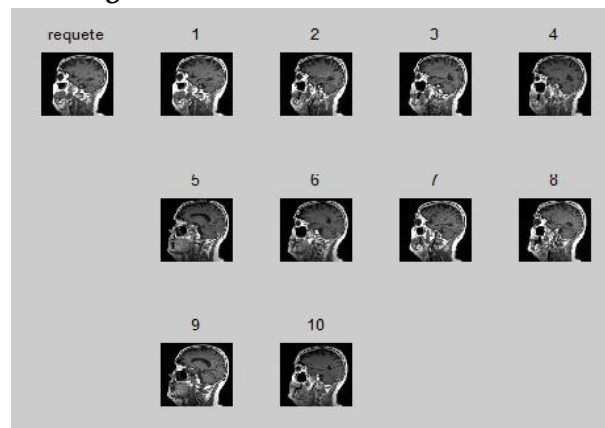
### 3.5.2 Evaluation sur la base Médicale

) *Les résultats de moment de Zernike*



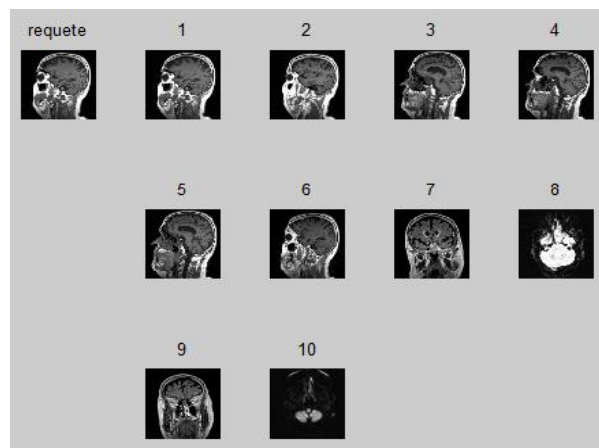
**Fig.3.25** Résultats de la recherche dans la base médicale utilisant les moments de Zernike

) *Les résultats d'histogramme NG*



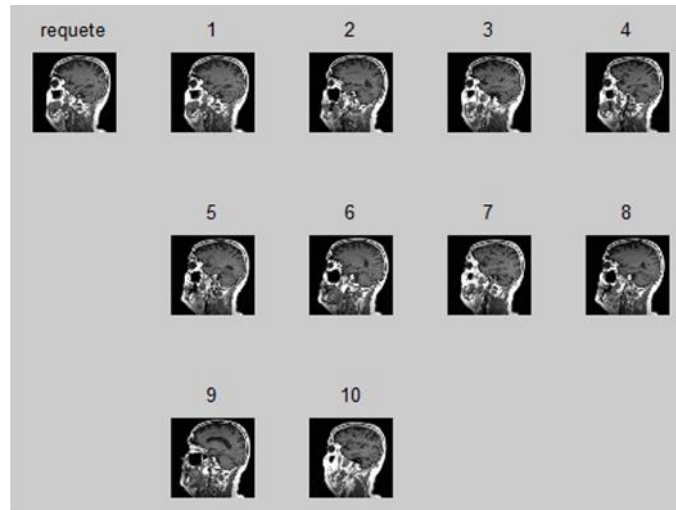
**Fig.3.25** Résultats de la recherche dans la base médicale utilisant l'histogramme en niveau de gris

) *Les résultats de matrice de cooccurrence*



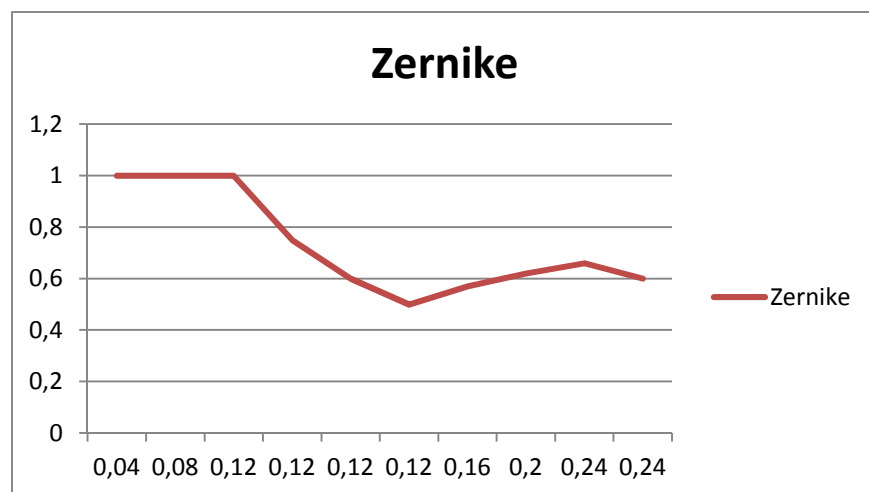
**Fig.3.26** Résultat de la recherche dans la base médicale utilisant la matrice de cooccurrence

) Les résultats de la combinaison



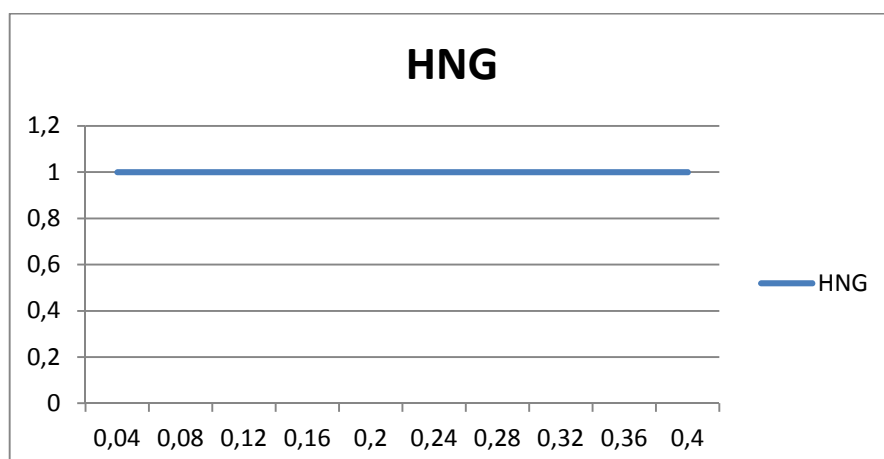
**Fig.3.27** Résultat de la recherche dans la base médicale utilisant la combinaison.

Pour la base Médicale, les courbes de

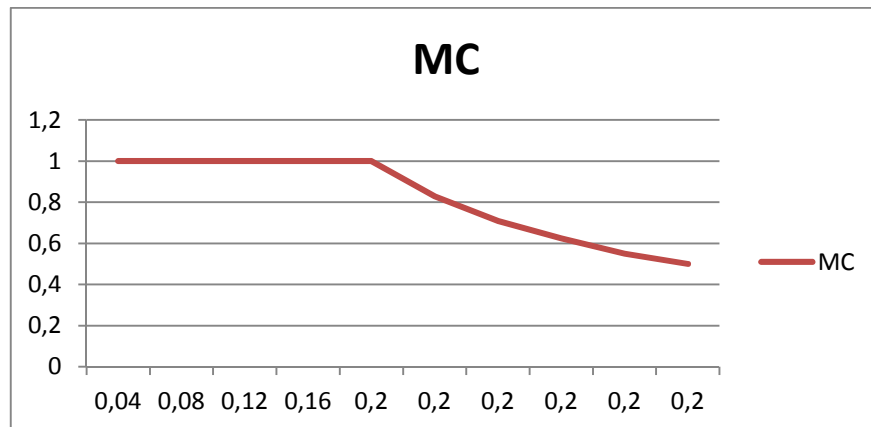


rappel/précision sont les suivantes :

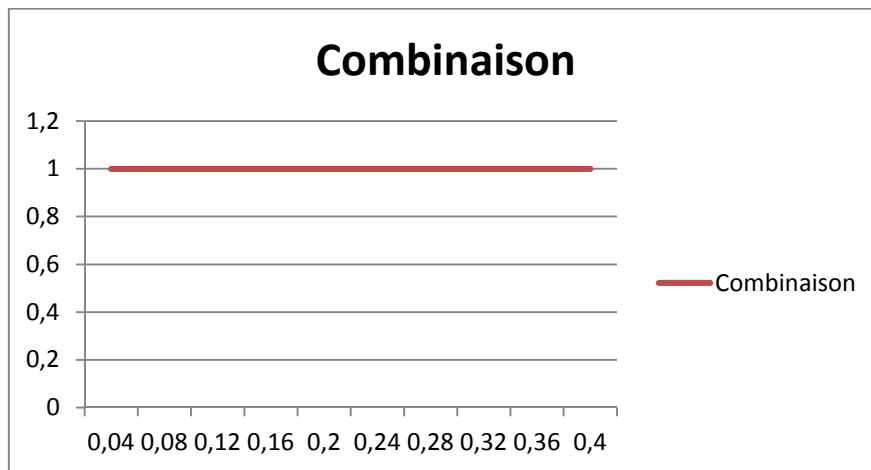
**Fig. 3.18** Courbe de rappel/précision pour l'image 6.jpg en appliquant Zernike



**Fig.3.19** Courbe de rappel/précision pour l'image 6.jpg en appliquant l'histogramme en niveau de gris



**Fig.3.20** Courbe de rappel/précision pour l'image 6.jpg en appliquant la Matrice de cooccurrence



**Fig.3.21** Courbe de rappel/précision pour l'image 6.jpg en appliquant la Combinaison

Le tableau ci-dessus montre les résultats de la base médicale on constate que l'Histogramme en niveau de gris et la combinaison sont les meilleurs, puis vient les moments de Zernike, et la matrice de cooccurrence moins satisfaisantes.

**Tab.5** Résultats de la base Medical

descripteur	Les moments de Zernike	L'histogramme niveau de gris	Matrice de cooccurrence	Combinaison
Nombre des images pertinentes dans 10	6	10	7	10
Précision	60	100%	50%	100%
Rappel	24%	66%	20%	66%

### 3.5 Interprétation des résultats

En comparant avec les résultats de toutes les bases, nous trouvons que les résultats avec la base COIL sont meilleurs que ceux avec la base médicale alors que pour la base texture, ils sont aussi bons. Les résultats avec la base visage sont très faibles. Nous pouvons expliquer ces observations par le fait que les images dans la base COIL sont plus simples que celles dans les autres bases. Dans la base COIL, les images contiennent un seul objet en avant avec le fond derrière. Ces images ne contiennent pas de scènes compliquées.

A partir de ces résultats avec quelques bases d'images différentes, nous donnons les quelques remarques suivantes. Pour la plupart des bases d'images, la combinaison de tous les descripteurs est la meilleure méthode. L'histogramme donne légèrement les meilleurs résultats que les moments de Zernike. Mais avec les moments de Zernike, on peut calculer et comparer des images plus rapidement. La texture avec la matrice de cooccurrence ne donne pas des résultats compétents sur les bases testées à cause de deux raisons. D'abord, la comparaison entre deux images par les mesures de texture est difficile. Ensuite, quatre chiffres des matrices de cooccurrences ne sont pas suffisants pour identifier la texture. La combinaison des descripteurs utilisés donne de meilleurs résultats dans toutes les bases testées.

### 3.6 Conclusion

Dans ce chapitre nous avons donné les trois attributs utilisés pour retirer les caractéristiques des images et les techniques essentielles pour mesurer la similarité entre ces images à partir des vecteurs résultants de notre système. Ensuite pour l'évaluation nous avons les deux critères rappel et précision. Nous constatons que les moments de Zernike et la combinaison des tous les descripteurs donnent de meilleurs résultats. Ces résultats diffèrent selon la base d'images utilisées. Les résultats les moins performants sont ceux obtenus de la base ORL.

## Conclusion générale

Dans ce mémoire, nous avons abordé le problème de la recherche et l'indexation d'images, plus précisément nous nous sommes intéressées sur la recherche et l'indexation d'images basée sur le contenu visuel. Notre choix a été motivé par la quantité d'images disponibles aujourd'hui et qui augmentent de jour en jour. Il s'agit d'extraire, à partir des images, une certaine information pertinente de couleur, de texture et de forme. Ces informations permettant une recherche efficace, pour atteindre cet objectif. Un système d'indexation et de recherche d'images par le contenu a été proposé, ce système fonctionne avec des descripteurs visuels qui sont extraits d'une image telle que les caractéristiques de l'histogramme RVB et niveau de gris, les caractéristiques de la matrice de cooccurrence et les caractéristiques des moments de Zernike pour permettre à l'utilisateur de proposer facilement une requête et de visualiser les résultats.

Nos résultats sont obtenus sur quatre bases de données images à savoir la base de données COIL, la base ORL, la base médicale, la base texture.

Pour l'évaluation de notre système, nous avons construit les courbes de rappel-précision pour chaque base.

Les résultats obtenus sont différents d'une méthode à une autre et d'une base à une autre, par exemple l'histogramme est la méthode utilisée par la couleur nous donne des résultats satisfaisants, ainsi que ceux fournis par les moments de Zernike qui sont très satisfaisants dans la base COIL et aussi dans les autres bases. Ceux de la matrice de cooccurrence sont moins bons dans la base médicale, texture et faible dans les autres.

Ce mémoire fait suite à deux mémoires de Magister soutenus dans ce même laboratoire. Il s'agit de ceux des références [17] et [18]. Notre objectif était de pousser ce travail jusqu'à son application concrète. Compte tenu que nous devions reprendre le travail, à la base, le temps ne nous a malheureusement pas permis d'arriver jusqu'à cette phase.

---

## Références bibliographiques

- [1] Gros (P). Meleam (G.), Drlon(R.), Schmid(C.) et Mistler(G.),-Utilisation de la couleur pour l'appariement et l'indexation d'images.-Technical report n RR\*3269.
- [2] Striker M.A. et Orenge M. – Similarity of color image. SPIE, Storage and Retrieval for Image Video Databases , pp. 381-392.-
- [3] [https://fr.wikipedia.org/wiki/Teinte\\_saturation\\_lumière](https://fr.wikipedia.org/wiki/Teinte_saturation_lumière)
- [4] R.M.Haralick, "Statistical and structural approaches to texture", Proceedings of the IEEE, mai 1979, number 5, vol.67
- [5] R.Bajcsy and Lieberman, "Texture gradient as a depth cue", Computer graphics and image processing 1976, vol.5.
- [6] A.Calway, "Image analysis using a generalised wavelet transform", Colloque institution of electrical engineers 1993, vol.DIG009
- [7] C.H.Chen, "A study of texture classification using spectral features", Proceedings of the 6th international conference of pattern recognition, Munich 1982.
- [8] Hu (M.K.) –Visual pattern recognition by moment invariants. IRE Transactions on Information Theory, vol. 8 n 2.1962, pp.179-187.
- [9] C. Cerjan, "The Zernike-Bessel representation and its application to Hankel transforms" [[archive](#)], J. Opt. Soc. Am. A 24 (2007) 1609.
- [10] [https://en.wikipedia.org/wiki/Minkowski\\_distance](https://en.wikipedia.org/wiki/Minkowski_distance)
- [11] M. Swain.D.H.Ballard.Colorindexing. International Journal of computer vision.32(11):11-32-1991
- [12] F. L. Hitchcock.  
The distribution of a product from several sources to numerous localities.  
*J. Math. Phys.*, 20:224-230, 1941.
- [13] ] [https://fr.wikipedia.org/wiki/CIE\\_XYZ](https://fr.wikipedia.org/wiki/CIE_XYZ), 27/06/2018
- [14] <http://www.profil-couleur.com/ec/106b-espace-rvb.php>
- [15] [https://fr.wikipedia.org/wiki/Alvy\\_Ray\\_Smith](https://fr.wikipedia.org/wiki/Alvy_Ray_Smith)
- [16] M.R. Teague. Image analysis via the general theory of moments. J. Optical Soc. Am. , 70(8) :920–930, August 1980

---

[17] AMROUNI Samia, Mémoire de Magister, recherche et classement d'image dans une base de données de grande taille à partir de requêtes relatives à leurs descripteurs, UMMTO 2015

[18] BEDOUHENE Saida, Mémoire de Magister, Recherche d'images par le contenu, UMMTO 2011