

## 1.1 Introduction :

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 . 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.

## Les descripteurs de couleurs :

La couleur est un attribut indispensable dans la grande majorité des systèmes d'indexation et de recherche par le contenu Commencée par les travaux de Swain et Ballard [1] 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, 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. Une indexation couleur repose sur deux principaux choix : **l'espace colorimétrique** et **le mode de représentation de la couleur dans cet espace**.

### 1.2.1 Les espaces de couleurs :

Avant de sélectionner un type de description du contenu couleur, il convient de choisir un espace de couleurs. Une couleur est généralement représentée par trois composantes qui définissent un espace de couleurs. Il existe plusieurs espaces colorimétriques qui ont chacun certaines caractéristiques intéressantes :

1. L'espace RVB ou RGB : est très simple à utiliser, car c'est celui employé par de nombreux appareils de capture d'images qui effectuent leurs échanges d'informations uniquement en utilisant les triplets(R, V,B). On parle d'espace colorimétrique orienté matériel. 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 (figure (1.1)).

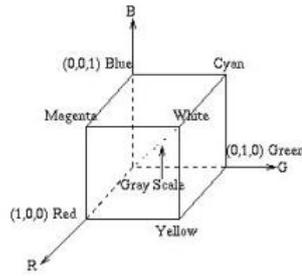


FIGURE 1.1: Cube de l'espace RGB.

2. L'espace HSV (Hue-Saturation-Value) : aussi connu sous le nom de système de cône hexagonal. 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 [2]. Le principe de cet espace est de caractériser les couleurs de façon plus intuitive et de se rapprocher de la perception humaine.

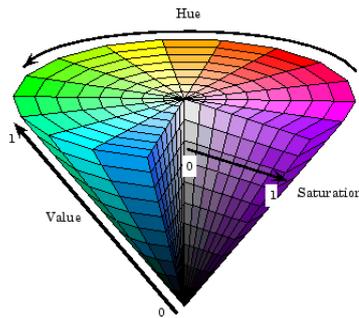


FIGURE 1.2: L'espace couleur HSV.

3. L'espace CIE XYZ : cet espace 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 [3]. La figure (1.3) 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.

### Le spectre visible dans le diagramme de chromaticité de la CIE (1931)

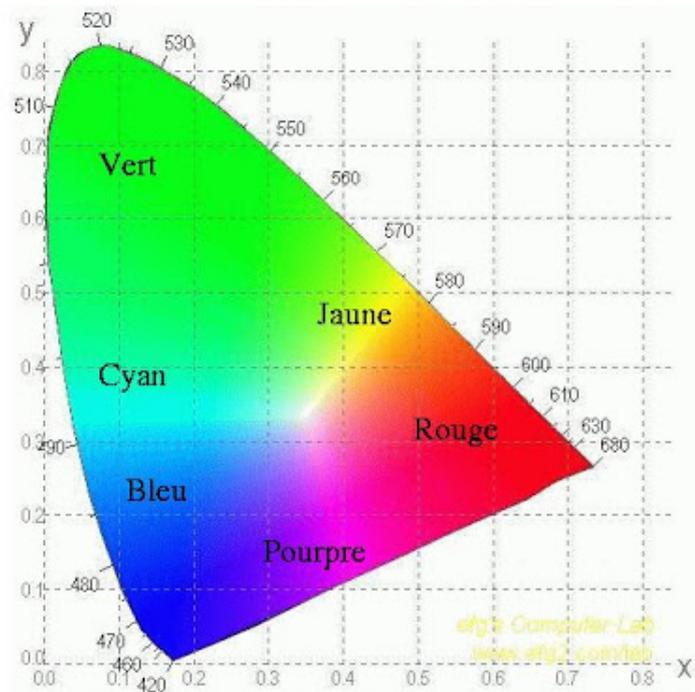


FIGURE 1.3: L'espace couleur XYZ.

4. L'espace CIE Luv : Il s'agit d'un système de perception standardisé qui peut être obtenu à partir du système XYZ . Il est permis d'utiliser une distance Euclidienne pour mesurer la différence entre deux couleurs . La composante L représente la réponse oculaire au niveau de luminance et la chrominance u et v.
5. L'espace CIE Lab : C'est un espace coloré avec une propriété de bon respecter Les distances entre les couleurs se rapprochent visuellement. Il défini à partir de l'espace XYZ par des relations non linéaires. Il est largement utilisé dans la recherche d'images par contenu.

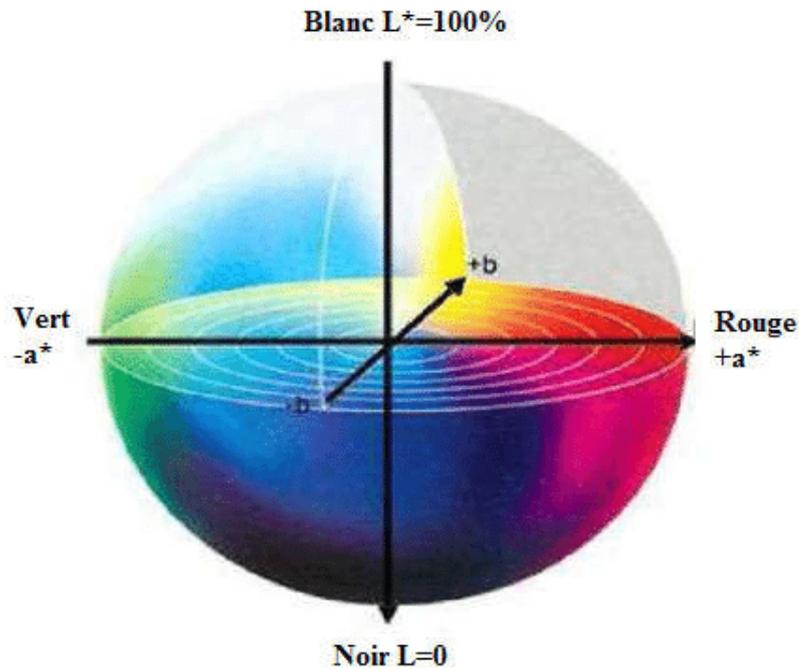


FIGURE 1.4: L'espace couleur CIE Lab.

6. L'espace couleur YCbCr : Une façon de représenter l'espace colorimétrique sur une vidéo. Nous convertissons donc les composants d'image RVB en composant de luminance Y et deux éléments de chrominance (c'est-à-dire les informations de couleur) désignent Cb et Cr.

### 1.2.2 L'histogramme :

L'histogramme de couleurs est le descripteur de couleurs le plus rencontré dans la littérature.

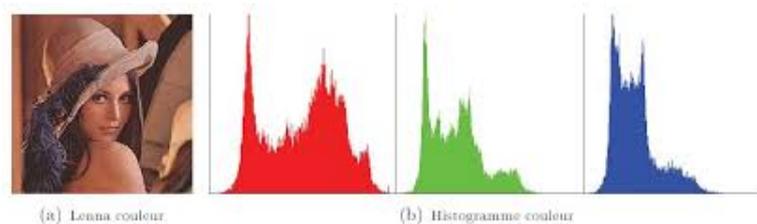


FIGURE 1.5: L'histogramme de couleur.

### 1.2.3 Corrélogramme de couleurs :

L'idée principale de cette méthode est inspirée des matrices de cooccurrence, développés au départ pour les images en niveaux de gris [4] .

### 1.2.4 Vecteur de cohérence de couleurs :

Le vecteur de cohérence de couleurs CCV, représente une autre variante, plus détaillée, de l’histogramme de couleurs. Il a été proposé par Pass [5]. Dans cette technique, chaque rang de l’histogramme peut être partitionné en deux catégories :

1. Cohérent, s’il appartient à une région de couleur uniforme.
2. Incohérent, sinon.

On note  $c_i$  le nombre de pixels cohérents dans le  $i$ ème rang de couleur, et  $nc_i$  le nombre de pixels incohérents. Le CCV d’une image est défini par le vecteur  $[(c_1, nc_1), (c_2, nc_2), \dots, (c_N, nc_N)]$ , tel que la somme  $(c_1 + nc_1, c_2 + nc_2, \dots, c_N + nc_N)$  donnera l’histogramme de couleurs de l’image.

#### 1.2.4.1 Descripteur par disposition de couleurs :

Ce descripteur vise à capturer la disposition spatiale (color layout) des couleurs dans une image. L’image est tout d’abord divisée en  $64(8 \times 8)$  blocs rectangulaires. Chaque rectangle est représenté par sa couleur dominante, qui est par définition la moyenne des couleurs des pixels constituant le rectangle considéré. On construit ainsi trois matrices  $(8 \times 8)$ , une par composante de couleur. Les coefficients quantifiés des transformées en cosinus discrète 2D (DCT - Discrete Cosine Transform) de chacune de ces matrices sont ensuite calculés. Ces coefficients sont parcourus en zigzag, à partir des fréquences les plus basses [6] .

#### 1.2.4.2 Descripteur par 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 [7] , 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’. Les différents nuages de points, sont déterminée par plusieurs méthodes dont la plus facile est celle basée sur l’histogramme.

#### 1.2.4.3 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.

– La Moyenne :

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

– La variance :

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

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

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

$P_{ij}$  : La valeur du pixel  $j$  pour le canal  $i$  ( RGB).  $N$  : nombre de pixel de l'image. Dans [8] 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.

### 1.3 Les descripteurs de texture :

La texture est le second attribut visuel largement utilisé dans la recherche d'images par le contenu. Elle permet de combler un vide que la couleur est incapable de faire, notamment lorsque les distributions de couleurs sont très proches. Fondamentalement la texture est définie comme la répétition d'un motif créant une image visuellement homogène. Plus précisément, la texture peut être vue comme un ensemble de pixels (niveaux de gris) spatialement agencés selon un certain nombre de relations spatiales, ainsi créant une région homogène. De ces définitions, les recherches sur la modélisation des textures se sont portées sur la caractérisation de ces relations spatiales [9] . De nombreuses approches et modèles [10] ont été proposées pour la caractérisation de la texture. Parmi les plus connues, on peut citer : **les méthodes statistiques, les méthodes fréquentielles et les méthodes géométriques**. Nous introduisons dans ce qui suit quelques représentations de la texture qui sont utilisées dans le domaine de la recherche d'images par le contenu. On peut distinguer deux types extrêmes de textures :

- les textures régulières.
- les textures aléatoires.

De nombreuses approches et modèles ont été proposées pour la caractérisation de la texture, parmi les plus connues, on peut citer : **Les méthodes statistiques et Les méthodes fréquentielles** .

#### 1.3.1 Les méthodes statistiques :

Ce sont les méthodes basées sur des évaluations quantitatives de la distribution de niveaux de gris. Parmi ces méthodes on peut citer :

##### 1.3.1.1 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  $I_g$  une image en niveau de gris correspondante à l'image  $I$  de  $N$  pixels. L'histogramme des niveaux de gris est calculé comme suit :

$$\text{hist}(i) = \frac{1}{N} \sum_{j=1}^N I_g(p) = i; (i \in \{1, 2, \dots, ng\}) \dots \dots \dots (4)$$

Les statistiques du premier ordre [11] , 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 suivants :

– La Moyenne :

$$\text{Moy} = \frac{1}{ng} \sum_{j=1}^{ng} i * \text{hist}(i) \dots \dots \dots (5)$$

– Ecart type :

$$E_{type} = \sqrt{\frac{1}{ng} \sum_{j=1}^{ng} (i - moy)^2 \text{hist}(i)} \dots \dots \dots (6)$$

– La variance :

$$\text{Var} = \frac{1}{N} \sum_{p=1}^N (I_g(p) - moy)^2 \dots \dots \dots (7)$$

### 1.3.1.2 Les matrices de longueur de plages (MLDP) :

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 :

$$\text{SLP} = \sum_{ng=0}^{Ng-1} \sum_{l=1}^L P_0(ng, l) \dots \dots \dots (8)$$

– La proportion de petites et grandes plages :

$$\text{RF1} = \frac{1}{\text{SLP}} \sum_{ng=0}^{Ng-1} \sum_{l=1}^L \frac{P_\theta(ng, l)}{l^2} \dots \dots \dots (9)$$

$$\text{RF2} = \frac{1}{\text{SLP}} \sum_{ng=0}^{Ng-1} \sum_{l=1}^L l^2 P_\theta(ng, l) \dots \dots \dots (10)$$

– L'hétérogénéité des niveaux de gris :

$$\text{RF3} = \frac{1}{\text{SLP}} \sum_{ng=0}^{Ng-1} (\sum_{l=1}^L P_\theta(ng, l))^2 \dots \dots \dots (11)$$

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

– L’hétérogénéité des longueurs de plages :

$$RF4 = \frac{1}{SLP} \sum_{l=1}^L (\sum_{ng=0}^{Ng-1} P_{\theta}(ng, l))^2 \dots\dots\dots(12)$$

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

– Le pourcentage des plages :

$$RF5 = SLP / R \dots\dots\dots(13)$$

$P_{\theta}(ng, l)$  : indique le nombre de plages de pixels ayant le niveau de gris de longueur l. Le paramètre  $\theta$  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.

### 1.3.1.3 Matrice de cooccurrence :

Les matrices de cooccurrence des niveaux de gris (GLCM (Grey Level Co-occurrence Matrices)) [12] , 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 [13] 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 . 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 apparaît à une distance relative d d’un pixel de niveau de gris j suivant une orientation  $\theta$  donnée.

$$P_{i,j}(d, \theta) = prob_{P_1 \in NG_i} [ P_2 \in NG_j / |p_1 - p_2| = d, (p_1, p_2) = \theta ] \dots\dots\dots(14)$$

Les matrices de cooccurrence sont exploitable par extraction des attributs numériques calculés appelés paramètres de texture. Haralick [14] 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  $N \times M$ .

### 1.3.1.4 La méthode de différence de niveaux de gris (GLDM(Grey Level Difference Method)) :

Cette méthode permet de calculer le nombre d'apparitions d'une différence de niveaux de gris donnée. Elle donne un aspect de la texture au sens de la différence de niveaux de gris entre les pixels.

### 1.3.1.5 Caractéristique de Tamura :

L'approche Tamura et al[15]. est intéressante pour la recherche d'images car elle décrit les textures possibles selon des concepts qui correspondent à la perception visuelle humaine. Les auteurs proposent 6 propriétés visuelles des textures : la grossièreté, le contraste, la direction, présence de lignes (linéarité), régularité et rugosité.

## 1.3.2 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 [16] . Ainsi les filtres de Gabor, ont eux aussi montrés leurs capacités de représenter une texture [17] . 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.

### 1.3.2.1 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 Fourier sera discrète. Elle est donnée par la formule suivante :

$$TFD_i(u, v) = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} I(i, j) \exp^{-f2\pi(\frac{ui}{N} + \frac{vj}{M})} \dots\dots\dots(15)$$

$$0 \leq u \leq N - 1 \dots\dots\dots(16)$$

$$0 \leq v \leq M - 1 \dots\dots\dots(17)$$

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.

### 1.3.2.2 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 suivante (18 et 19). Cette méthode est très utilisée en indexation d'images pour la caractérisation de la texture et notamment pour la segmentation des images texturées [18], ainsi dans le domaine de la reconnaissance faciale [19]. L'utilisation des filtres de Gabor consiste à analyser indépendamment différentes parties de l'espace de Fourier à l'aide de plusieurs filtres. Les principales étapes d'extraction des caractéristiques des textures de l'image en utilisant le filtre de Gabor sont :

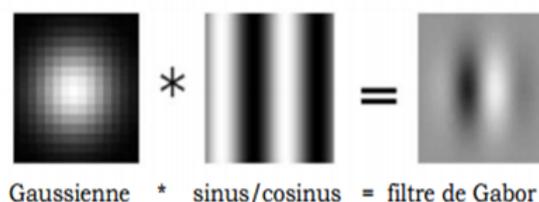


FIGURE 1.6: Représentation d'un filtre de Gabor 2D.

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

$$f_{Gb} = g(x', y') \exp(2\pi j[\mu_0(x - x_0)^2 + v_0(y - y_0)^2]) \dots \dots \dots (18)$$

$$g(x', y') = \exp\left(-\frac{x'^2}{2\sigma_x^2} - \frac{y'^2}{2\sigma_y^2}\right) \dots \dots \dots (19)$$

$x$  et  $y$  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).

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

#### 1. Calcul d'énergie :

L'énergie correspond à la somme de l'ensemble de valeurs obtenue après l'application du filtre de Gabor sur l'image, avec différentes orientations et échelles.

#### 2. Calcul de la moyenne :

La moyenne  $m, n$  : caractérise l'intensité lumineuse de l'image correspondant au niveau de gris moyen de l'ensemble des pixels de l'image.

#### 3. Calcul de l'écart-type :

L'écart-type  $m,n$  : caractérise la variation de l'intensité moyenne de l'ensemble des pixels. Elle correspond au changement du contraste de l'image.

### 1.3.2.3 Filtres spatiaux (Spatial domain filters) :

Les filtres spatiaux sont la façon la plus directe utilisée pour capturer les propriétés d'une texture d'image. On peut trouver : **Roberts** et **Laplacien**.

### 1.3.2.4 Les ondelettes :

Les ondelettes ont été introduites pour la première fois par Grossman et Morlet [20] 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 [21] et la compression d'image.



FIGURE 1.7: décomposition d' une image avec deux niveaux.

## 1.4 Les descripteurs de formes

La forme est l'un des attributs de bas niveau les plus utilisés pour décrire la structure géométrique du contenu visuel. Contrairement aux descripteurs de texture et de couleur, qui peuvent être utilisés pour décrire globalement l'image, les descripteurs de forme décrivent des propriétés locales aux régions composant l'image [22] . La représentation de la forme et ses techniques de description peuvent être classés en deux classes de méthodes : les méthodes de contour et les méthodes basées sur la région. Sous chaque catégorie, les différentes méthodes sont divisées en deux approches structurelles et globales.

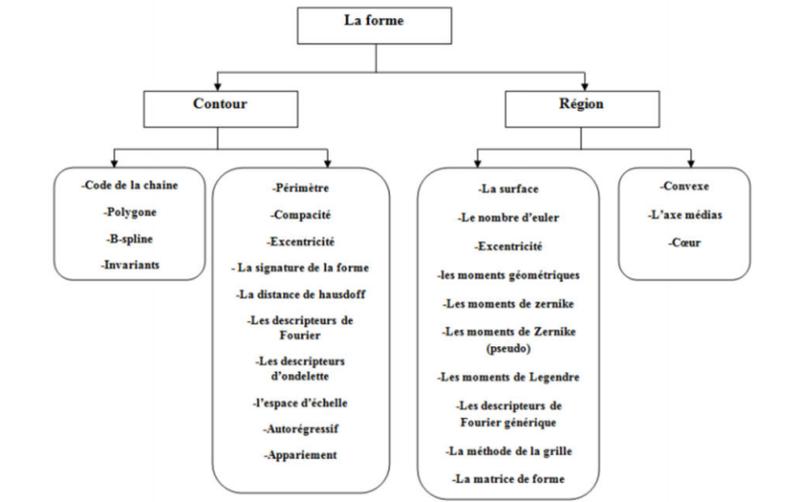


FIGURE 1.8: Classification des techniques de la représentation de forme. [22]

## 1.4.1 Représentation de la forme basée sur le contour

### 1.4.1.1 Les méthodes globales

#### La distance de Hausdorff ( Hausdorff distance)

Est une méthode classique utilisée pour localiser les objets dans une image et pour mesurer la similitude entre les formes.

#### Une signature de forme (Shape signature)

Il s'agit de représenter la forme par une fonction unidimensionnelle de type  $y = f(x)$ , extraite des points de contour de la forme. Notez que de nombreuses signatures officielles existent. Ils incluent le profil du centroïde, les coordonnées complexes et la distance du centroïde.

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

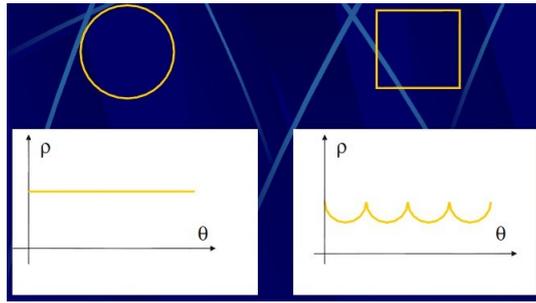


FIGURE 1.9: signatures de forme.

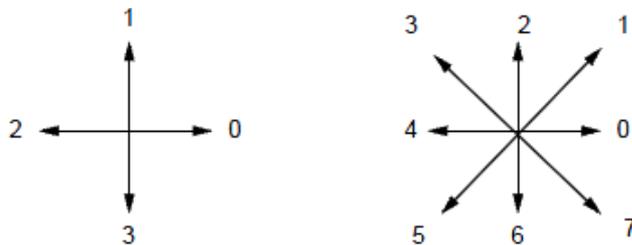
## Descripteurs de Fourier

Les descripteurs de Fourier DFs font partie des descripteurs les plus populaires pour les applications de reconnaissance de formes et de recherche d'images. Ils ont souvent été utilisés par leur simplicité et leurs bonnes performances en terme de reconnaissance [23]. De plus, ils permettent de décrire la forme de l'objet à différents niveaux de détails et ils sont peu sensibles au bruit. Les descripteurs de Fourier sont calculés à partir du contour des objets. Leur principe est de représenter le contour de l'objet par un signal 1D, puis de le décomposer en séries de Fourier.

### 1.4.1.2 Les méthodes structurales

#### La représentation par un code de la chaîne (Chaines de Freeman)

C'est la méthode la plus ancienne de description des contours dans les images et aussi la plus utilisée encore aujourd'hui. C'est une technique pour représenter les tendances de contour. Les tendances peuvent être représentées dans une connexion 4 bits (codage 2 bits) ou dans une connexion 8 bits (3 bits).



- Les codes de Freeman en 4-connectivité (à gauche) et en 8-connectivité (à droite).

#### La décomposition polygonale

La bordure de la forme est décomposée en segments de ligne par approximation polygonale (approximation d'un nuage de points par une droite unique). Les sommets du polygone sont utilisés en tant que primitives. La caractéristique de chaque primitive est exprimée en une chaîne à 4 éléments qui se compose des angles internes, de la distance à partir du sommet suivant, et ses coordonnées x et y. La similitude entre les 2 formes est mesurée par la distance entre les 2 chaînes de caractéristique.

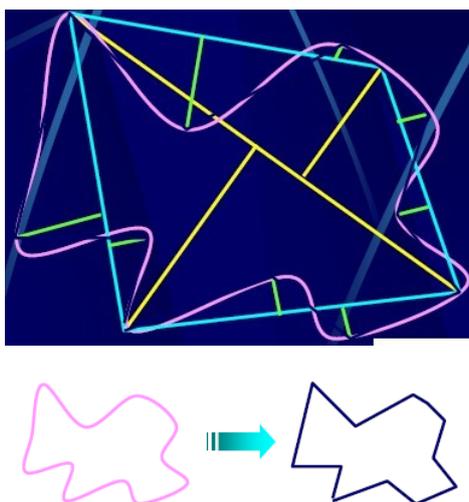


FIGURE 1.10: décomposée polygone

Pour plus d'efficacité et pour des raisons robustesse, seul un nombre fixe de sommets sont choisis dans chaque forme. C'est une méthode peu sensible au bruit.

### Approximation d'un nuage de points par une droite unique

Soit  $M_i = (x_i, y_i)$  les points en nombre  $N$ , que l'on cherche à approcher par une droite. Il y a 2 façons pour la faire :

-Approximation par régression linéaire.

C'est une approche aux moindres carrés : on recherche la droite :  $y = a_0 + a_1x$  qui minimise la distance suivante :

$$d_1^2 = \sum_i^N [y_i - (a_0 + a_1x_i)]^2 \dots\dots\dots(20)$$

La solution est donnée par :

$$A = X\#Y = (X^t X)^{-1} X^t Y \dots\dots\dots(21)$$

où  $X\#$  dénote la matrice pseudo-inverse de la matrice  $X$ , où  $X$ ,  $Y$  et  $A$  sont donnés par :

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_N \end{pmatrix} \dots\dots\dots(22)$$

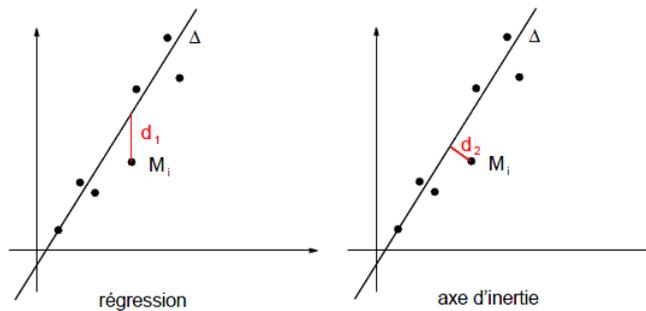
$$Y = [y_1, y_2, \dots, y_N]^t \dots\dots\dots(23)$$

$$A = [a_0, a_1]^t \dots\dots\dots(24)$$

-Approximation par axe principal d'inertie

C'est aussi une approche aux moindres carrés, mais on minimise dans ce cas la somme des distances de tous les points à la droite :

$$d_2^2 = \sum_{i=1}^N \frac{[(a_0 + a_1 x_i - y_i)]^2}{a_1^2 + 1} \dots\dots\dots(25)$$



### 1.4.2 Représentation de la forme basée sur la région

Dans cette approche, tous les pixels de la forme sont pris en considération pour obtenir la représentation de la forme, pas seulement les informations de contour

#### 1.4.2.1 Les méthodes globales

##### -Les moments géométriques

Les moments géométriques permettent de décrire une forme à l'aide de propriétés statistiques [25]. 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 :

$$m_{p,q} = \sum_{x=0}^{M-1} \sum_{y=0}^{n-1} x^p y^q f(x, y) \dots\dots\dots(26)$$

Où p+q est l'ordre du moment. le moment d'ordre 0  $m_{0,0}$  représente l'aire de la forme de l'objet. Les deux moments d'ordre 1  $m_{0,1}$  et  $m_{1,0}$ , associés au moment d'ordre 0  $m_{0,0}$  permettent de calculer le centre de gravité de la forme. Les coordonnées de ce centre sont :

$$x_g = \frac{m_{10}}{m_{00}}; y_g = \frac{m_{01}}{m_{00}} \dots \dots (27)$$

### -Les moments invariants de Hu

A partir des moments géométriques, HU [26] a proposé un ensemble de 7 moments invariants aux translations, rotation et changement d'échelle. Ces moments sont très utilisés dans la littérature pour la description des formes en vue d'une classification ou d'une indexation, mais sont assez sensibles aux bruits. Pour caractériser la forme des images, nous utilisons la famille de moments HU. Pour les calculer, on calcule d'abord les moments centrés ( $U_{p,q}$ ) et les moments centrés normalisés .

( $U_{p,q}$ ) par les équations suivantes :

$$m_{p,q} = \sum_{x=0}^{M-1} \sum_{y=0}^{n-1} (x - x_g)^p (y - y_g)^q f(x, y) \dots \dots \dots (28)$$

$$n_{p,q} = \frac{u_{p,q}}{u_{0,0}} \dots \dots \dots (29)$$

Avec :  $y = (2+p+q)/2 \dots \dots \dots (30)$

### -invariants du second

$$\varphi_1 = n_{20} + n_{02} \dots \dots \dots (31)$$

$$\varphi_2 = (n_{20} - n_{02})^2 + 4n_{11}^2 \dots \dots \dots (32)$$

### -invariants du troisieme ordre

$$\varphi_3 = (n_{30} - 3n_{12})^2 + (3n_{21} - n_{03})^2 \dots \dots \dots (33)$$

$$\varphi_4 = (n_{30} + n_{12})^2 + (n_{21} + n_{03})^2 \dots \dots \dots (34)$$

$$\varphi_5 = (n_{30} - 3n_{12})(n_{30} + n_{12}) [(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2 + (3n_{21} - n_{03})(n_{21} + n_{03}) [3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2]] \dots \dots \dots (35)$$

$$\varphi_6 = (n_{20} - n_{02}) [(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] + 4n_{11}(n_{30} + n_{12})(n_{12} + n_{03}) \dots \dots \dots (36)$$

$$\varphi_7 = (3n_{12} - n_{03})(n_{30} + n_{12}) [(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2 + (3n_{12} - n_{03})(n_{21} + n_{03}) [3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2]] \dots \dots \dots (37)$$

Le vecteur  $V = [\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6]$  constitue un vecteur d'index final. Alors, le calcul de la distance entre la forme de 2 images différences est donné par :

$$Dist(V_{req}, V_{cour}) = \left( \sum_{i=1}^6 |V(i)_{req} - V(i)_{cour}|^2 \right)^{1/2} \dots \dots \dots (38)$$

## -Les moments orthogonaux

Par opposition aux moments géométriques qui sont définis par rapport à une base quelconque  $(x^p y^q)$ , les moments orthogonaux, comme leur nom l'indique, sont définis dans une base orthogonale, ce qui évite la redondance des informations portées par chacun des moments. Les deux types de moments orthogonaux les plus utilisés sont : les moments de Legendre et les moments de Zernike, dont nous donnons les définitions ci-dessous.

## -Les moments de Legendre

Les moments de légende sont définis par des polynômes du même nom. Il est défini dans le carré unitaire  $[-1,1]$  et  $[-1,1]$ , qui se lie à la normalisation de l'objet dont nous voulons calculer les moments

Le polynôme de Legendre d'ordre  $n$  est donné par :

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

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

$$\forall (x, y) \in [-1, 1], \forall (p, q) \in N^2, \int \int_{-1}^1 p_m(x) p_n(y) dx dy = \frac{2}{2m+1} \delta_{mn} \dots\dots\dots(40)$$

$\delta_m$  représente la fonction de Kronecker. Les moments de Legendre d'ordre  $N$  sont donc donnés par :

$$\forall (x, y) \in [-1, 1]^2, \forall (p, q) \in N^2, N = p + q \dots\dots\dots(41)$$

$$L_{pq} = \frac{(2p+1) \times (2q+1)}{4} \int \int_{-1}^1 p_p(x) p_q(y) f(x, y) dx dy \dots\dots\dots(42)$$

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

## -Les moments de Zernike

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 [27]. 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 [28] , 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 orthogonaux de Zernike d'ordre p sont définis de la manière suivante :

$$A_{mn} = \frac{m+1}{\pi} \int_x \int_y I(x,y)[V_{m,n}(x,y)] dx dy \dots (43)$$

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 Où  $x^2 + y^2 \leq 1$ .....(44)

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

$$V_{mn}(r, \theta) = R_{mn}(r)e^{-jn\theta} \dots (45)$$

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

$$R_{mn}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s \frac{(m-s)!}{s! \left(\frac{m-|n|}{2} - 2\right)! \left(\frac{m-|n|}{2} - s\right)!} r^{m-2s} \dots (46)$$

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

### -La méthode basé sur la grille (Grid based method)

Le principe de cette méthode est qu'une grille de cellules est superposée à la forme, la grille est ensuite scannée de gauche à droite et de haut en bas. Le résultat est une image bitmap. Les cellules couvertes par la forme sont affectées à 1 et celles qui ne sont pas couvertes par la forme sont affectées à 0. La forme peut alors être représentée comme un vecteur caractéristique binaire. La distance de Hamming est utilisée pour mesurer la similitude entre les deux formes. Les avantages du descripteur de grille sont sa simplicité. Le principal problème de cette méthode est la normalisation de l'axe de rotation principal

#### 1.4.2.2 Les méthodes structurelles

##### Convexe (Convex hull)

Une région R est convexe si et seulement si pour deux points quelconques  $x_1, x_2 \in R$ , la ligne entière du segment  $x_1x_2$  est à l'intérieur de la région. [29] « Convex hull » d'une région est la région convexe H la plus petite qui satisfait la condition  $R \subset H$ . La différence H-R est appelé « carence convexe » D d'une région R. Une meilleure représentation de la forme peut être obtenue par un procédé récursif, ce qui conduit à un arbre de concavités.

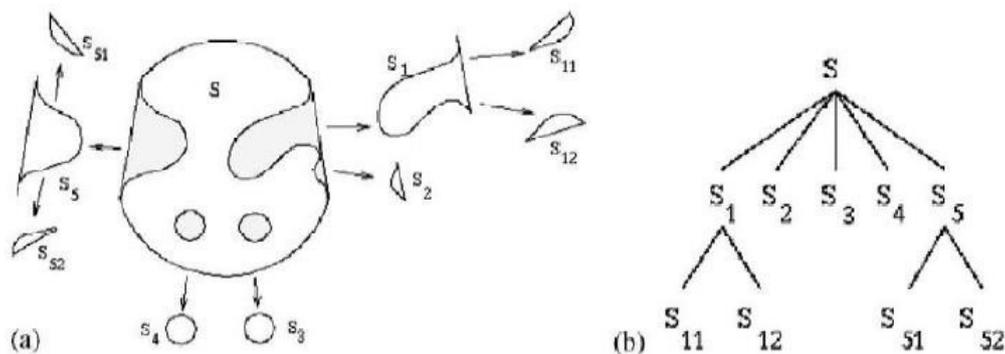


FIGURE 1.11: Convex et ses concavités et arbre de concavités [13]

### -L'axe médian (Medial axis)

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

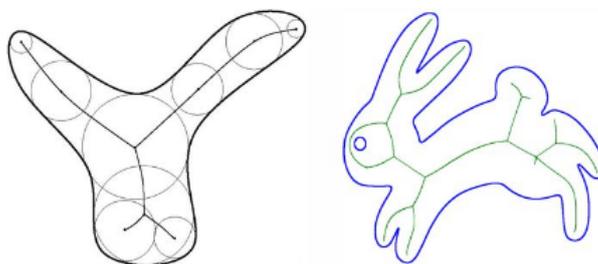


FIGURE 1.12: construction de l'axe median d'une forme

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

## 1.5 CONCLUSION :

Ce chapitre nous a permis de voir 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 [30] , et celles basées sur les régions [31].