

Segmentation et classification

6.1. Principe de segmentation

La segmentation d'images est l'opération la plus importante dans un système de traitement d'images, car elle est située à l'articulation entre le traitement et l'analyse des images. La segmentation consiste à créer une partition d'une image A en sous-ensembles R_i appelés régions tel que est de partitionner une image en plusieurs régions

$$\forall i \ R_i \neq \emptyset$$

$$\forall i, j \ R_i \cap R_j = \emptyset$$

$$A = \cup_i R_i$$

Une région est un ensemble convexe de points images ayant les propriétés communes (intensité, texture, ...).

6.2. Techniques de segmentation

On peut distinguer quatre grandes approches de segmentation : par seuillage, par régions, par contour (frontière) et approche de la classification.

6.2.1. Approche contour

Il s'agit dans cette approche de rechercher les contours qui délimitent les régions homogènes de l'image. Un contour est un ensemble de pixels formant une frontière entre deux ou plusieurs régions voisines. Cependant, cette méthode ne conduit pas directement à une segmentation de l'image car les contours obtenus sont rarement convexes, il faut donc procéder à une fermeture des contours si on souhaite une partition complète de l'image.

Il existe plusieurs méthodes de segmentation basées sur l'approche contour qu'on peut regrouper en trois catégories : les méthodes dérivatives, par filtrage optimal et les contours actifs.

6.2.2. Approche région

Contrairement à la segmentation d'images par l'approche contour, la segmentation par région consiste à décomposer l'image en régions selon des critères d'homogénéité.

On distingue quatre types de méthodes :

✓ Croissance des régions

Appelé aussi méthode d'agrégation du pixel. C'est une méthode ascendante, elle consiste à choisir initialement un ensemble de pixels comme des noyaux des régions à partir desquels

la croissance des régions s'effectuera. Chaque région doit respecter un critère d'homogénéité comme la variance $var(R)$ du niveau de gris de l'image associée au point de la région R est inférieure à un seuil.



Figure 6.1. Segmentation par (a) approches contours (b) approches régions

✓ Division des régions

Division des régions ce sont des méthodes descendantes consistent à diviser l'image ou une partie de l'image en région en utilisant des partitions élémentaires connues comme le maillage ou l'arbre quaternaire (quadtree). le maillage peut être régulier ou irrégulier de type rectangulaire, triangulaire ou polygone quelconque. La division d'une région R en sous régions se fait si R ne remplit pas la contrainte d'homogénéité fixée. La décision peut aboutir à un nombre important de régions, la solution consiste à fusionner les régions adjacentes similaires. La figure () montre le principe de la méthode.

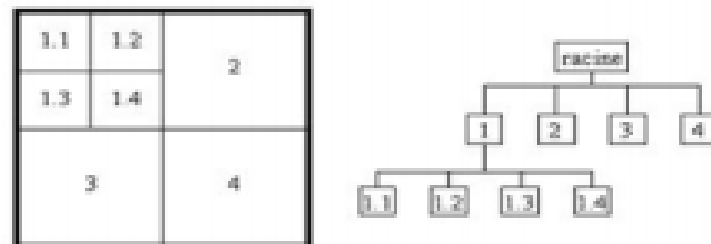


Figure 6.2. Principe de la segmentation par division des régions

6.2.3. Segmentation d'images par classification

Cette technique consiste à regrouper des pixels ayant des vecteurs caractéristiques aussi similaires que possible sans la considération de la disposition spatiale des pixels. De ce fait, les pixels appartenant à une classe peuvent former plusieurs régions non-adjacentes dans l'image mais ayant les mêmes propriétés. Plusieurs algorithmes sont utilisés dans la segmentation par classification, on cite entre autres l'algorithme K-means.

6.2.4. Segmentation par seuillage

La segmentation par seuillage permet de répartir les pixels en classes en fonction de leurs niveaux de gris. Les classes sont alors déterminées par des seuils.

Soit $\{0,1,2, \dots, L-1\}$ l'ensemble des niveaux de gris de l'image I , la segmentation par seuillage consiste à répartir les pixels en K classes $(C_1, C_2, \dots, C_K)$ à partir d'un ensemble de seuils $T = (t_1, t_2, \dots, t_{K-1})$. Un pixel de niveau de gris $I(x, y)$ est affecté à la classe C_k si $t_k \leq I(x, y) < t_{k+1}$ avec $k = 0, 1, 2, \dots, K-1$ et $t_0 = 0$ et $t_K = L-1$.

Si le nombre de seuils est supérieur à 1, on parle de multi-seuillage ou multi-level thresholding.

6.3. Seuillage d'images

Le seuillage (Thresholding en Anglais) permet d'extraire des objets de leurs fonds en fonction d'un seuil. Le problème majeur dans le seuillage réside dans la détermination du seuil. Plusieurs méthodes ont été proposées pour le calcul d'un seuil puis étendues au problème du multi-seuillage.

Parmi les méthodes de seuillage on cite :

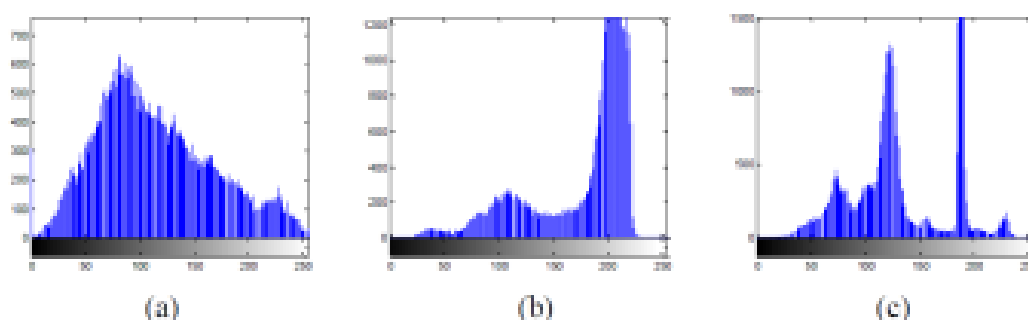
6.3.1. Seuillage globale

Pour les méthodes de seuillage globale on définit un même seuil pour tous les pixels de l'image sur l'exploitation de l'histogramme de toute l'image.

A partir de l'allure de l'histogramme, on peut déduire le nombre de classes ainsi que la position relative des seuils.

- Histogramme uni-modal → formé d'un seul pic → Objet ou fond
- Histogramme bimodal → caractérisé par deux modes séparés par une vallée → Objet sur un fond
- Histogramme multi-modal → comporte plus de deux modes séparés par des vallées → plusieurs classes

Les seuils doivent être localisés dans les vallées.



Histogramme unimodal

Histogramme bimodal

Histogramme multimodal

Figure 6.3. Différents types d'histogramme

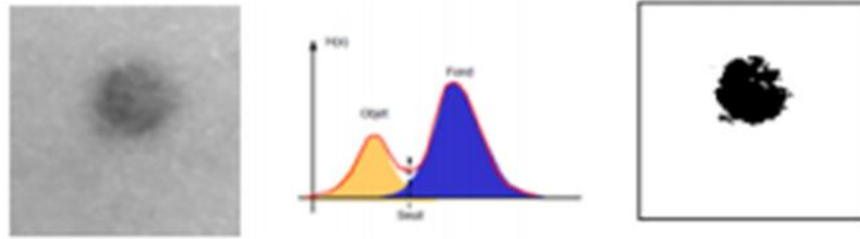


Figure 6.4. Exemple de seuillage en 2 classes par histogramme

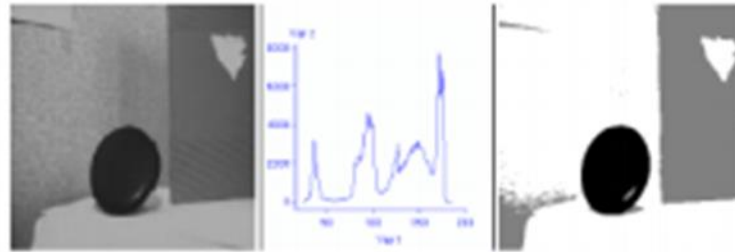


Figure 6.5. Exemple de seuillage en 4 classes par histogramme

6.3.2. Seuillage local

Dans le seuillage locale, la valeur des pixels voisins est considérée pour le calcul des seuils. La méthode consiste à diviser l'image en blocs non chevauchants, puis pour chaque bloc un seuil est déterminé selon l'histogramme local i.e. du bloc en question. Une autre approche consiste à déterminer un seuil pour chaque bloc en fonction des pixels voisins. Pour chaque pixel centré $I(x, y)$, le seuil $t(x, y)$ est évalué en fonction de la moyenne et l'écart-type des niveaux de gris des pixels de la fenêtre :

$$t(x, y) = \mu(x, y) + \alpha\sigma(x, y)$$

avec α un paramètre réel fixe.

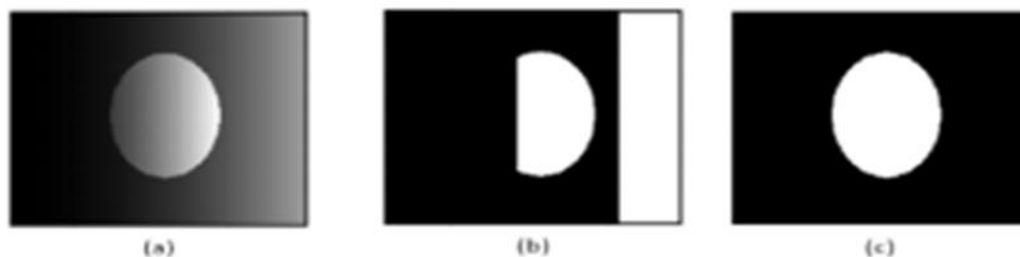


Figure 6.6. Illustration du seuillage local (a) image originale , (b) image segmentée avec un seuil global , (c) image segmentée avec un seuil local.

6.3.3. Seuillage par détection de vallées

On suppose que chaque classe correspond à une plage de niveaux de gris. La position des minima de l'histogramme H permet de fixer les $(m - 1)$ seuils nécessaires pour séparer les m classes. les seuils t_i sont obtenus par :

$$H(t_i) = \text{Min} [H(k)] \text{ pour } k \in]m_i, m_{i+1}[$$

où m_i et m_{i+1} sont les valeurs moyennes (ou les modes) de l'intensité lumineuse dans les classes C_i et C_{i+1} .

Malgré le développement de techniques robustes visant à faciliter la détection des vallées, cette méthode, bien que simple, est très peu appliquée car les histogrammes traités sont le plus souvent bruités et uni-modaux.

6.3.4. Seuillage par minimisation de variance

Méthode de Fisher : L'objectif de la méthode est de trouver le seuil t qui minimise la somme des moments d'ordre 2 de deux classes (meilleur seuillage dans cette approche correspond à une minimisation de la variance intra-classe).

Soit $H(x)$ l'histogramme de l'image, la variance d'une classe i est donné par :

$$var_i(s) = \sum_{x \in c_i} (x - G_i)^2 H(x)$$

où $G_i(s)$ est le centre de gravité de la classe i qui est donné par :

$$G_i(s) = \frac{\sum_{x \in c_i} x H(x)}{H(x)}$$

Le problème revient à trouver le seuil S qui minimise la somme des variances :

$$S_{opt} = \min_s (var_1(s) + var_2(s))$$

Plus récemment, Otsu a proposé de réaliser une maximisation de la variance inter-classe.

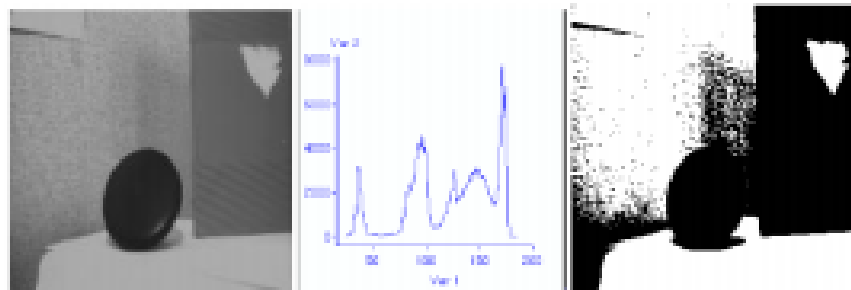


Figure 6.7. Exemple de seuillage par minimisation de variance

6.4. Morphologie mathématique

La morphologie mathématique est un outil qui permet d'explorer la structure géométrique des objets dans une image en utilisant la théorie des ensembles. De ce fait les opérateurs de la mise en œuvre sont des opérateurs ensemblistes.

6.4.1. Élément structurant

L'élément structurant est un ensemble ou surface de géométrie connue. Il est caractérisé par sa forme, sa taille et son origine. Pour la forme, on peut avoir un diamant, un disque, une ligne ...

Si A est l'objet à traiter et B l'élément structurant, la morphologie mathématique consiste à déplacer B sur A comme un opérateur de convolution. A chaque position de B, des opérations logiques sont effectuées entre A et B

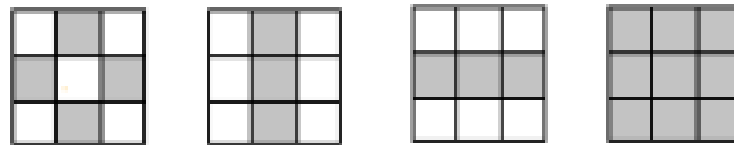


Figure 6.8. Exemples d'éléments structurants

6.4.2. Operations morphologiques de base

Dilatation

La dilatation sert à grandir les objets dans une image binaire. Donc elle permet d'ajouter à l'objet des pixels qui lui sont proche dans des directions qui dépendent de l'élément structurant. Ce dernier parcourt l'objet en gardant son centre à l'intérieur de l'objet.

Soient A et B deux ensembles, on définit la dilatation de A par B, notée $A \oplus B$, par :

$$c = A \oplus B = \{x / A \cap B_x \neq \emptyset\}$$

L'exemple suivant permet de montrer un exemple de dilatation.



Figure 6.9. Image originale en (a), Images dilatées en 4-connexité (b) et en 8-connexité (c)

Erosion

L'érosion sert à rétrécir les objets dans une image. Lorsque l'élément structurant parcourt l'objet, l'ensemble des pixels non recouvert par l'élément structurant forment l'ensemble érodé.

Soient A et B deux ensembles, on définit l'érosion de A par B, notée $A \ominus B$, par :

$$c = A \ominus B = \{x / B_x \subset A\}$$

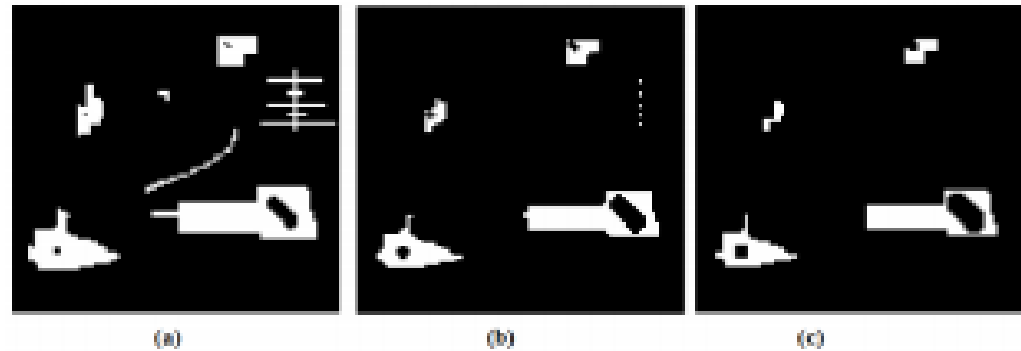


Figure 6.10. Image originale en (a), Images érodées en 4-connectivité (b) et en 8-connectivité (c)

Fermeture

C'est une dilatation suivie d'une érosion avec le même élément structurant.

La fermeture de l'ensemble A par l'élément structurant B, notée $A \bullet B$ est définie par :

$$c = A \bullet B = (A \oplus B) \ominus B$$

La fermeture permet le recouvrement des objets proches et la suppression des trous.

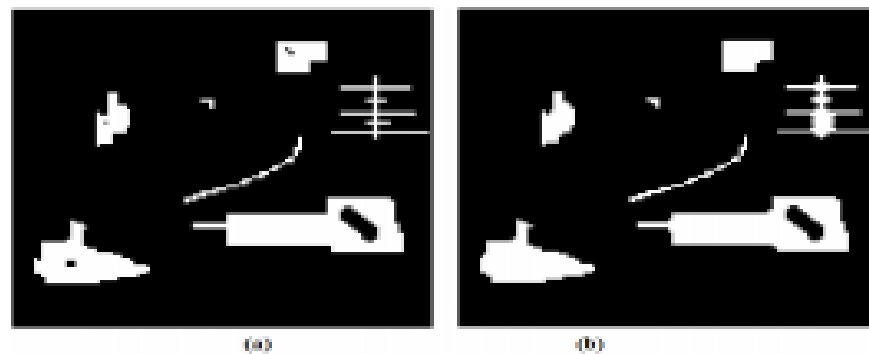


Figure 6.11. (a) Image originale ; (b) Résultat d'une fermeture

Ouverture

L'ouverture est une érosion suivie d'une dilatation avec le même élément structurant.

L'ouverture de l'ensemble A par l'élément structurant B, notée $A \circ B$ est définie par :

$$c = A \circ B = (A \ominus B) \oplus B$$

L'ouverture lisse les contours, supprime les petits îlots.

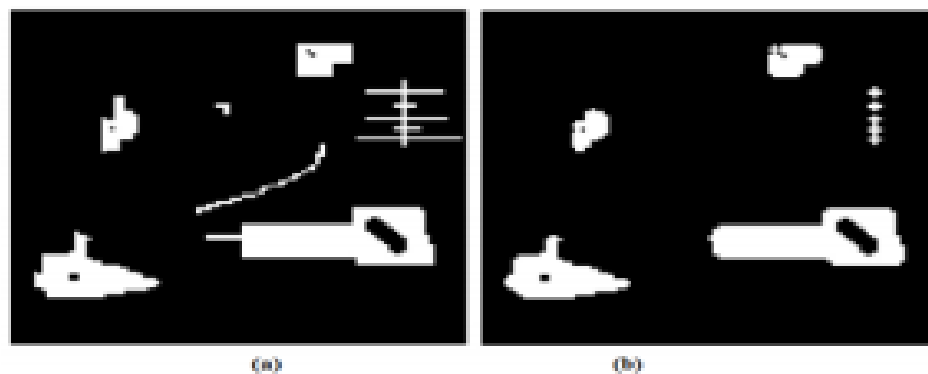


Figure 6.12. (a) Image originale ; (b) Résultat d'une ouverture

6.5. Extraction de paramètres et classification d'objets

6.4.1 Extraction des paramètres

L'extraction de paramètres caractéristiques -en anglais, « feature extraction »- est une étape de grande importance. Si elle est mal conçue, il sera difficile, voire impossible, d'effectuer une classification efficace.

L'extraction des caractéristiques consiste à l'obtention de l'information la plus pertinente d'une forme pour la présenter au système de reconnaissance. Les caractéristiques ou primitives d'une forme constituent une représentation réduite de la forme d'où une perte irréversible d'information s'impose.

Nous disposons d'une image binaire contenant des objets ou régions avec des contours ou frontières parfaitement définis (localisés, lissés, fermés), l'objectif est d'extraire des paramètres de forme afin de caractériser chacun des objets ou région. Parmi les paramètres caractéristiques, on cite les moments géométriques et les moments de Zernike.

Les moments géométriques

La méthode des moments géométriques, comme toutes les méthodes statistiques, permet d'extraire des paramètres propres à la forme. Hu a élaboré un théorème énonçant qu'un ensemble infini de moments calculés à partir d'une image permet de la décrire uniquement et complètement. Il a aussi créé une version de ces moments invariants en translation, en rotation et en homothétie.

Les moments d'ordre $p + q$ d'une fonction $f(x, y)$ sont définis par :

$$m_{pq} = \iint_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy$$

où p et q sont l'ordre du moment par rapport à x et y respectivement. Dans le cas d'une image binaire où :

$$f(x, y) = \begin{cases} 1 & (x, y) \in \text{objet} \\ 0 & \text{sinon} \end{cases}$$

L'équation du moment devient :

$$m_{pq} = \sum_x \sum_y x^p y^q$$

Pour rendre les moments invariants par rapport à la translation, l'image doit être centrée par rapport à son centre de gravité. Le moment central résultant aura la forme :

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q$$

où x_c et y_c sont les coordonnées du centre de gravité de l'image de l'objet données par :

$$x_c = \frac{m_{10}}{m_{00}}, \quad y_c = \frac{m_{01}}{m_{00}}$$

$$\mu_{00} = m_{00}$$

$$\mu_{10} = \mu_{01} = 0$$

$$\mu_{11} = m_{11} - (m_{00}x_c y_c)$$

$$\mu_{20} = m_{20} - (m_{00}x_c^2)$$

$$\mu_{02} = m_{02} - (m_{00}y_c^2)$$

$$\mu_{21} = m_{21} - (m_{20}y_c) - (2m_{11}x_c) - (2m_{00}x_c^2 y_c)$$

$$\mu_{12} = m_{12} - (m_{02}x_c) - (2m_{11}y_c) - (2m_{00}y_c^2 x_c)$$

$$\mu_{30} = m_{30} - (3m_{20}x_c) + (2m_{00}x_c^3)$$

$$\mu_{03} = m_{03} - (3m_{20}y_c) + (2m_{00}y_c^3)$$

Les moment centrés et normalisés η_{pq} sont obtenus en divisant chaque moment centré μ_{pq} par γ :

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}, \quad \gamma = \frac{p+q}{2} + 1$$

Ces moments sont caractérisés par une invraiance par rapport à la translation et le changement d'échelle.

$$\Phi_1 = \eta_{20} + \eta_{02}$$

$$\Phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\Phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\Phi_4 = (\eta_{20} - \eta_{02})^2 + (\eta_{21} - \eta_{03})^2$$

$$\Phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) / ((\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2) + (3\eta_{21} - \eta_{03}) / (3\eta_{30} - 3\eta_{12}) + (\eta_{21} - \eta_{03})(3(\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2)$$

$$\Phi_6 = (\eta_{20} - \eta_{02}) / ((\eta_{20} - \eta_{02})^2 - (3\eta_{21} - \eta_{03})^2) + \eta_{11}(\eta_{03} - \eta_{12})(\eta_{21} - \eta_{03})$$

$$\Phi_7 = (\eta_{21} - \eta_{03})(\eta_{30} - 3\eta_{12}) / ((\eta_{30} + \eta_{12})^2 - (3\eta_{21} - \eta_{03})^2) - (\eta_{30} + 3\eta_{12})(\eta_{21} - \eta_{03}) / (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2$$

Le calcul de ces momoents se fait comme suit, après la segmentation de l'image ,on calcule ces moments m_{pq} , puis les moments centraux μ_{pq} et enfin les moments invariants Φ_k .

Les moments de Zernike

Les moments de Zernike sont basés sur le principe des polynômes orthogonaux. De ce fait, la reconstruction (en général approchée) de la forme à partir de ces moments est possible. Un ensemble relativement petit de moments de Zernike peut caractériser la forme globale d'un objet, Les moments d'ordre bas représentent la forme globale d'un objet et celles d'ordre plus

supérieur le détail de cet objet. Plus on fait appel à des moments de Zernike d'ordre plus élevé, plus est la précision de reconstruction de l'image de l'objet. Les moments de Zernike sont qualifiés par leur invariance par translation, par changement d'échelle et par rotation.

Les polynômes de Zernike notés par $V_{nm}(x, y)$ sont donnés par :

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho)e^{jm\theta}$$

où n entier positif ou nul, m entier positif ou négatif selon n ; ρ longueur du vecteur de l'origine au pixel (x, y) et θ l'angle entre le vecteur ρ et l'axe des abscisses dans le sens des aiguilles d'une montre

R_{nm} est un polynôme radiale donné par :

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s}$$

Les moments de Zernike sont les projections de l'image sur la base des fonctions orthogonales $V_{nm}(\rho, \theta)$

Les moments de Zernike d'ordre n avec une répétition m pour une fonction image $f(x, y)$ qui s'annule en dehors du cercle unitaire ont pour expression :

$$A_{nm} = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y) V_{nm}^*(\rho, \theta) dx dy$$

Pour une image numérique :

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(\rho, \theta)$$

Pour calculer les moments de Zernike d'une image donnée, le centre de l'image est pris comme origine

Pour ramener l'image à l'intérieur du cercle unitaire, on réalise une normalisation des rayons de tous les pixels de l'image : Partant du centre de gravité, on balaye tous les pixels de l'image pour chercher le plus grand rayon, celui-ci est considéré comme unitaire.

Tous les autres rayons sont divisés par ce rayon maximum pour inclure les pixels correspondants dans le cercle unité.

Remarques

Les paramètres caractéristiques devront être :

Discriminatifs : les paramètres caractéristiques doivent être significativement différents pour des formes appartenant à des classes différentes.

Fiables : pour des formes d'une même classe, les paramètres caractéristiques doivent être similaires.

De nombre réduit : le coût, la complexité et les exigences en temps de calcul d'un système de reconnaissance de formes croissent avec le nombre de paramètres utilisés.

6.4.2 Méthodes de classification

L'objectif des méthodes de classification s'est de classer des individus en exploitant leur similarité mesurée par les paramètres caractéristiques. Les individus peuvent être :

Les pixels : les attributs sont ponctuels ou calculés dans un voisinage centré sur le pixel.

Les sous-images : l'image est découpée en fenêtres. Chaque fenêtre est classée en utilisant les attributs calculés sur la totalité de la fenêtre

Les régions : obtenues par une méthode de segmentation.

La classification sert à interpréter les formes issues de la segmentation (reconnaissance de formes).

A chaque individu, on associe un vecteur de paramètres caractéristiques. Ce vecteur représente l'individu dans l'espace des attributs.

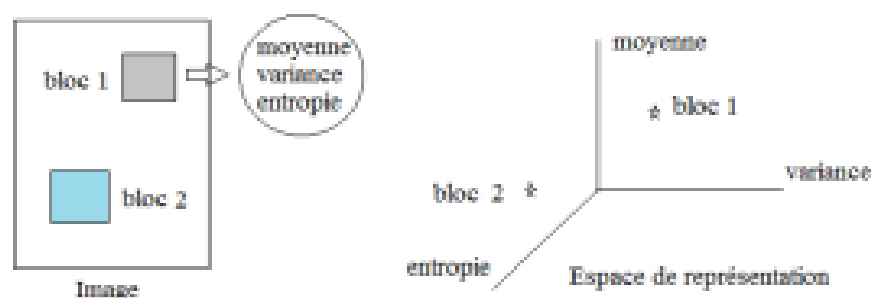


Figure 6.13. Espaces de représentation des individus

Pour la classification, il existe deux approches :

a- Classification non supervisée

C'est le partitionnement de l'ensemble d'individus en fonction des critères de proximité des vecteurs d'attributs dans l'espace de représentation. Cette méthode ne demande pas des connaissances a priori sur l'image.

b- Classification supervisée

Dans ce cas la classe d'un certain nombre d'individu est connu apriori, cette connaissance permet de construire un partitionnement de l'espace de représentation.

Les individus de cette classe inconnue sont classés en fonction de leur position dans l'espace de représentation grâce à ce découpage. Dans ce cas il est nécessaire d'avoir un ensemble d'individu dont la classe d'appartenance est connue, c'est l'ensemble d'apprentissage. Parmi les méthodes de la classification supervisée on cite la méthode des K plus proches voisins (kppv).

