
SEGMENTATION D'IMAGE

Sommaire

2.1	Introduction	18
2.2	Définition de la segmentation	19
2.3	Objectifs de la segmentation	20
2.4	Exemples d'applications de la segmentation	20
2.5	Les méthodes de segmentation	20
2.6	Présentation de l'algorithme FCM.	30
2.7	Conclusion	39

2. 1. Introduction

Une image est une représentation planaire d'une scène ou d'un objet situé en général dans un espace tridimensionnel. L'information élémentaire associée à chaque point de l'image est transcrite en niveau de gris ou en couleur. L'analyse pertinentes de l'image en regard de l'application concernée, les traiter et les interpréter.

La segmentation est un traitement de bas -niveau qui consiste à créer une partition de l'image en sous-ensembles appelés régions. Une région est un ensemble connexe de points images (pixel) ayant des propriétés communes (intensité, texture,. . .) qui les différencient des pixels des régions voisines. Les connaissances utilisées sont le plus souvent des domaines de l'image numérique et du traitement du signal, donc sémantiquement assez pauvres. Il n'y a pas de méthode unique de segmentation d'une image, le choix d'une technique est lié à plusieurs facteurs. Du fait de cette diversité, il est difficile de définir de manière absolu une « bonne » segmentation.

2.2 Définition de la segmentation

Classiquement, on peut définir la segmentation comme étant une partition de l'image I en un nombre N de sous-ensembles R_i appelés régions est un ensemble connexe de points image (pixels) ayant de même propriétés (texture, intensité,...) qui les différencient des pixels des régions voisines. La région i est notée R_i , homogènes dans le sens d'un ou plusieurs critères, comme le niveau de gris, tels que :

- La segmentation doit être complète (c'est-à-dire, chaque pixel doit être affecté à une classe).
- Les pixels appartenant à la même région doivent être connectés.
- Les régions doivent être disjointes.

$$\text{En termes mathématiques} \quad \left\{ \begin{array}{l} \bigcup_{i=1}^n R_i = I \\ R_i \neq \emptyset \quad \forall i=1 \dots n. \\ R_i \cap R_j = \emptyset \quad \forall i, j \text{ avec } i \neq j \end{array} \right.$$

Où R_i est l'ensemble des régions formant l'image I

Dans un contexte médical, les régions R_i correspondent aux différentes structures anatomiques constituant les différentes régions d'intérêt. La détermination automatique du nombre de régions N est une vraie problématique [5]. Souvent, la valeur N est supposée être connue comme étant une information à priori sur les structures anatomiques en investigation.

Il est difficile de définir d'une manière absolue, une bonne méthode de segmentation. Le choix d'une technique est lié à :

- La nature de l'image (éclairage, contours, texture, etc.)
- Aux opérations en aval de la segmentation (compression, reconnaissance des formes, mesures, etc.)
- Aux primitives à extraire (droites, régions, textures, etc.)
- Aux contraintes d'exploitation (temps réel, espace mémoire, etc.)

2.3 Objectifs de la segmentation [6]

- L'extraction d'éléments et d'informations qualitatives de l'image.
- Fournir des régions homogènes
- Localiser les contours de région.
- Fournir une description de haut niveau : représentation par un graphe dont les nœuds sont des régions, chaque région est connectée à ses voisines et chaque région porte une étiquette qui donne des informations comme la taille, la couleur, la forme,...
- Réduire le bruit.

2. 4. Exemples d'applications de la segmentation

La segmentation est utilisée dans de nombreuses applications dont certaines concernent:

- La reconnaissance de la parole: Pal et Majumder [7] ont été les premiers à introduire le concept de sous-ensemble flou dans ce domaine. Ils décrivent des techniques permettant la discrimination de voyelles et de différentes voix d'orateurs appliquées à la langue Telugu (une des nombreuses langues en usage en Inde).
- La reconnaissance des caractères écrits peut avoir de nombreuses applications, telles que, par exemple, le traitement automatique des documents.
- Le traitement des eaux usées.
- La segmentation d'images RMN (Résonance Magnétique Nucléaire).

2.5 Les méthodes de segmentation [8]

La segmentation est une étape primordiale en traitement d'image. À ce jour, il existe de nombreuses méthodes de segmentation, que l'on peut regrouper en quatre principales classes :

- Segmentation basée région
- Segmentation basée contour
- Segmentation basée classification ou le seuillage des pixels en fonction de leur intensité.
- Segmentation basée la coopération entre les trois premières segmentations.

2.5.1 Segmentation basée région

Ce sont des approches utilisant des techniques et de localisation d'ensembles connexes de pixels ayant des propriétés similaires. Parmi elles, on distingue.

2.5.1.1. Méthode par division (Top-Down)

Ce sont des Méthodes qui opèrent en découpant l'image principale en régions plus petites suivant un critère d'hétérogénéité. Le découpage s'arrête quand il n'y a plus de régions inhomogènes [9]. La méthode Quatre est souvent utilisée pour réaliser ce découpage (**Figure 2.1**). Il s'agit d'une décomposition récursive en quatre sous régions carrées des régions hétérogènes en commençant par l'image entière qui est de taille $N \times N$, où $N = n^2$. La figure 1.4 illustre un exemple de résultat de découpage obtenu avec cette méthode.

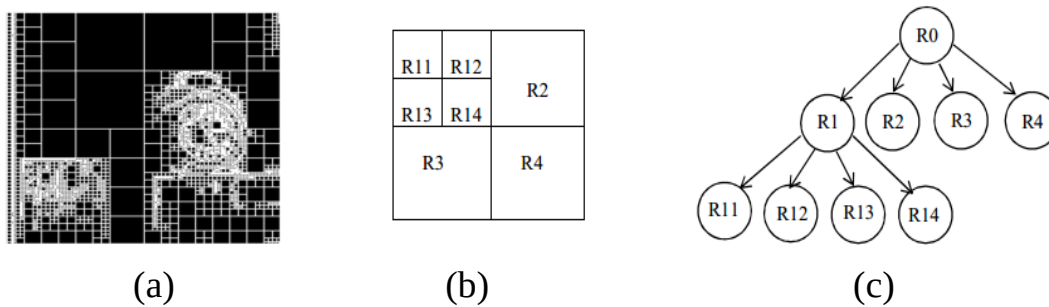


Figure 2.1 : Principe de division de région. (a) Blocs de division obtenus, (b) Décomposition de l'image en blocs, (c) Diagramme hiérarchique de la décomposition.

2.5.1.2. Méthode par fusion (Bottom-up)

Ces méthodes opèrent en regroupant des pixels ou des groupes de pixels selon un critère d'homogénéité pour obtenir un ensemble de régions homogènes [9]. Elles suivent une hiérarchie arborescente comme montrée dans la **Figure 2.2**. Leurs inconvénients majeurs c'est que le résultat final (image segmentée) dépend fortement du choix des régions de départ.

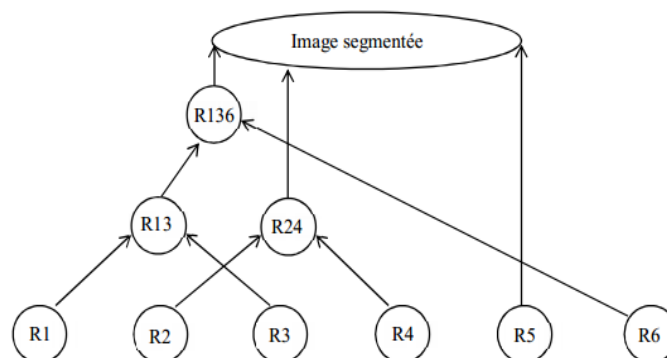


Figure 2.2 : Hiérarchie de la méthode de segmentation par fusion

1.5.1.3 Méthodes par division / fusion (split and merge)

Proposées par Horowitz [10], elles réunissent l'ensemble des algorithmes utilisés dans les techniques précédemment citées (méthodes par fusion et méthodes par division). Le processus de segmentation se déroule en deux étapes : la division en blocs (régions) suivant un critère d'hétérogénéité est effectuée en premier lieu ; puis, suit l'étape de fusion où toutes les régions adjacentes qui découlent de l'étape de division et vérifiant une heuristique d'homogénéité sont regroupées.

2.5.2 Segmentation basée contour

De façon général, un contour est défini comme étant la frontière entre deux régions, la détection du contour est équivalente à la détection de la discontinuité entre ces deux régions. Les approches contour opèrent à trouver les zones de variations significatives d'intensité lumineuse (niveaux de gris) ou de couleur dans l'image. Nous pouvons citer les approches se basant sur les différences finies comme l'opérateur de gradient, l'opérateur Laplacien et les différents filtres à savoir : le filtre de Sobel, Prewitt et Roberts ou bien des approches analytiques comme le filtre de Canny [11]. Mais ce genre de techniques est peu exploitable car elles donnent souvent des contours non fermés, bruités ou non détectés ; une utilisation des propriétés des régions comprise entre ces contours pourrait nettement améliorer la détection de ce derniers [12]. *Figure 2.3* illustre les résultats obtenus avec quelques filtres.

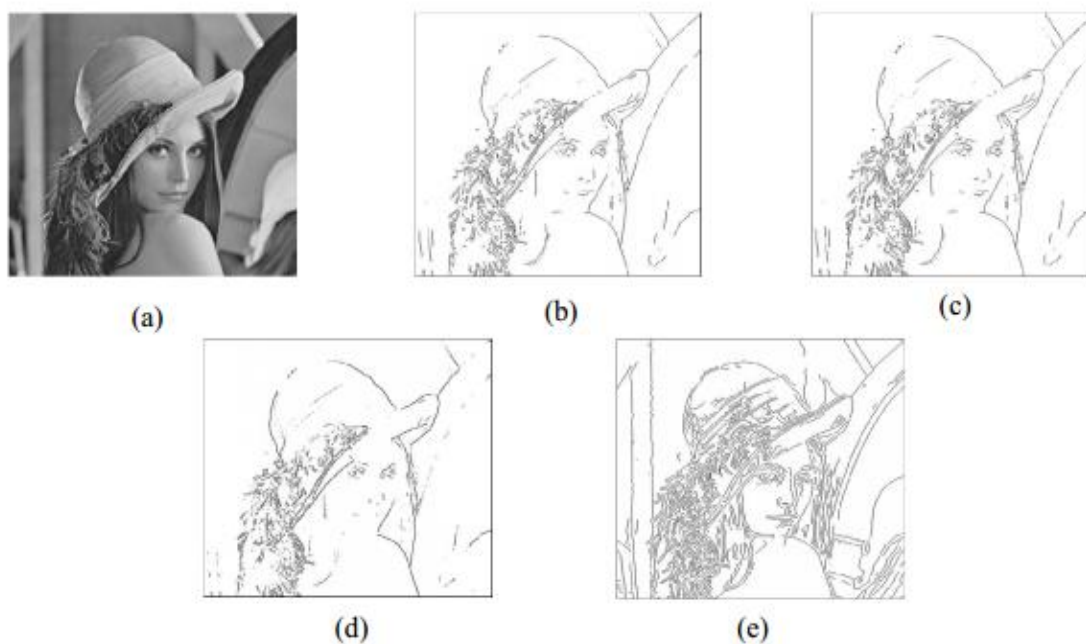


Figure 2.3 : Illustration de détection de contours, (a) Image Lenna originale ; (b) Détection de contours par filtre de Sobel ; (c) Détection de contours par filtre de Prewitt ; (d) Détection de contours par filtre de Roberts ; (e) Détection de contours par filtre de Canny

2.5.2.1 Segmentation par seuillage

Dans beaucoup d'applications de traitement d'images, on considère que les niveaux de gris des pixels appartenant à un objet sont relativement différents de ceux du fond [13] [14] [15]. Le seuillage est une technique simple et efficace pour séparer les objets du fond. Toutefois, la difficulté apparaît dans le choix des seuils optimaux pour une séparation optimale. Soit ' I ' l'image originale et ' I_s ' l'image segmentée. Segmenter I en N classes revient à chercher $N-1$ seuils. La classification de chaque pixel est définie comme suit :

$$I_s(x, y) = k, \text{ si } S_k \leq I(x, y) < S_{k+1} \quad \text{avec } k = 0, \dots, N$$

Où x, y sont les coordonnées du pixel ; $S_0, \dots, S_N$ sont les différents seuils de segmentation

A titre d'exemple, nous allons un seuillage manuel qui comporte quatre étapes :

1. Observation de l'histogramme d'image.
2. Choix des seuils dans les vallées.
3. Définition des classes de régions par intervalle de couleurs.
4. Classification des pixels.

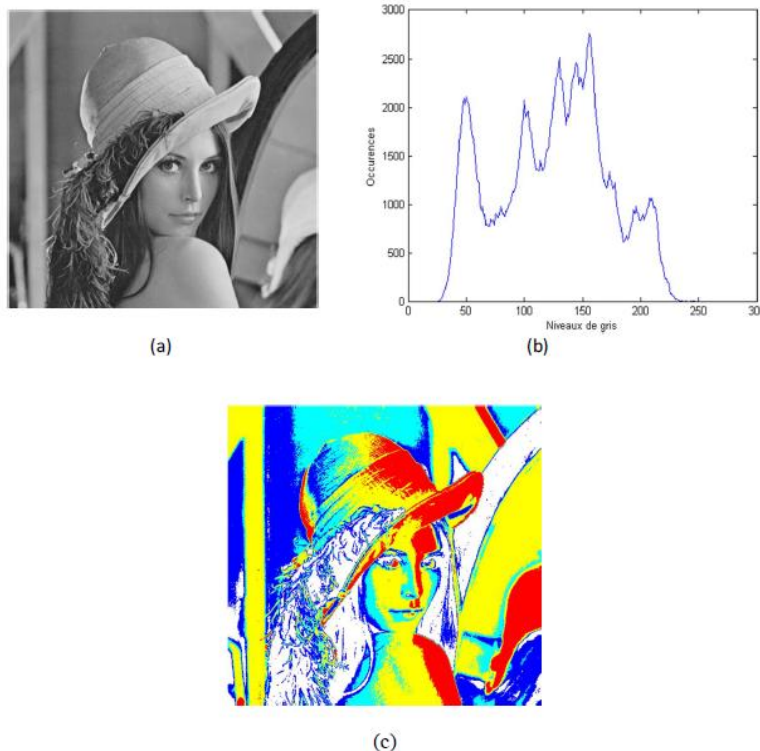


Figure 2.4 : Exemple de seuillage manuel de l'image Lina. (a) Image originale ; (b) Histogramme de l'image ; (c) Seuillage en 5 classes ($S_1, S_2 = 0, 70, 113, 136, 185, 255$)

Dans la littérature, on y trouve plusieurs classifications des méthodes de segmentation par seuillage [14] [16] [17]. Les auteurs les ont classées en deux catégories :

- **Les méthodes non paramétriques** : qui ont la réputation d'être robustes ; elles sont basées sur l'optimisation d'un ou plusieurs critères comme, par exemple, maximiser la variance interclasse. La plus connue d'entre elles est la méthode d'Otsu [16].
- **Les méthodes paramétriques** : qui sont basées sur l'hypothèse que les densités de probabilité des niveaux de gris des différentes classes sont gaussiennes, une approximation de l'histogramme de l'image à traiter par une combinaison linéaire de gaussiennes est toujours effectuée avant l'opération de seuillage, ou les seuils optimaux sont à l'intersection de ces dernières [14].

2.5.3 Segmentation basée classification [18]

Cette méthode consiste à regrouper et à classer les pixels d'une image en classes en fonction de leurs propriétés. A chaque point de l'image est associé un vecteur d'attributs. La classification est alors effectuée sur ces vecteurs d'attributs de façon à aboutir à un nombre restreint de régions homogènes au sein de l'image.

Donc la classification est définie comme une procédure dans laquelle les pixels similaire d'une image sont identifiés et regroupés dans une même classe. Il existe deux grandes tendances: ***La classification non supervisée*** et ***La classification supervisée***.

2.3.5.1 Méthodes basé sur la classification Supervisées

Les méthodes supervisées nécessitent un ensemble d'images ou de régions d'intérêt dans l'image déjà étiquetée. Soit, les données à analyser $X = \{X_j \in R^p, 1 \leq j \leq N\}$, où X_j est un vecteur forme représentant le pixel j . Cet ensemble de données est décomposé en deux sous-ensembles : un ensemble d'apprentissage et un ensemble de test. L'ensemble d'apprentissage est construit par étiquetage manuel des images ou des régions d'intérêt en C classes de tissus ($C_1 \dots C_n$) par un ou plusieurs experts (radiologues).

Chaque classe C_j Se voit donc affecter un ensemble d'apprentissage E_i , et les données de l'ensemble de test sont segmentées en fonction des E_j . Parmi ces

méthodes on peut citer: la segmentation Bayésienne, la segmentation par les champs de Markov, réseaux de neurones, segmentation bayésiennes, etc.

a. Méthodes bayésiennes

La segmentation bayésienne consiste à calculer, pour chaque vecteur forme, la probabilité conditionnelle $P(C_i / X_j)$ pour chacune des C classes ($C_1 \dots C_n$) à l'aide de la règle Bayes:

$$P(C_i / X_j) = \frac{P(C_i) \cdot P(X_j / C_i)}{\sum_{k=1}^C P(C_k) \cdot P(X_j / C_k)} \quad \forall i \in [1 \dots C] \quad (2,1)$$

La segmentation bayésienne admet:

- Soit une approche paramétrique, dans laquelle l'intensité d'un voxel (est un point à trois dimensions) est considérée comme la combinaison linéaire de probabilités d'appartenance attachées à chaque class (en effectuant l'hypothèse que les données sont conformes à des distributions paramétrique). La probabilité conditionnelle $P(C_i / X_j)$.
- Est modélisée par une fonction dépendant d'un vecteur de paramètre θ . Le problème est alors pour chaque classe C_i d'estimer le meilleur paramètre θ connaissant l'ensemble d'apprentissage E_i de cette classe.
- Soit une approche non paramétrique, dans ce type d'approche, les probabilités conditionnelles $P(C_i / X_j)$ sont supposées quelconques. La forme non paramétrique doit permettre de rendre compte le plus fidèlement possible de la réelle distribution statistique des niveaux de gris dans l'image.

b. Réseaux de Neurones [19]

Un *réseau de neurones* est constitué d'un certain nombre d'unités élémentaires (neurones) fortement interconnectés (neurones) fonctionnant en parallèle et organisée en couches avec un flux d'informations à anticipation.

L'architecture du perceptron multicouche est organisée de la manière suivante: les signaux flux successivement dans les différentes couches de l'entrée à la couche de sortie. Les couches intermédiaires sont appelées couches cachées. Pour chaque couche, chaque unité élémentaire calcule un produit

scalaire entre un vecteur de poids et le vecteur de sortie donné par la couche précédente. Une fonction de transfert est alors réalisée à la suite de faire une entrée pour la couche suivante. Les poids des connexions sont ajustés au cours du processus d'apprentissage pour obtenir la relation d'entrée/sortie désirée du réseau. De nombreux modèles de réseaux existent (réseaux de Hopfield, perceptrons multicouche, etc.),

2.3.5.2 Méthodes basé sur la classification non-Supervisées

L'intérêt des méthodes non supervisées est qu'elles ne nécessitent aucune base d'apprentissage et par là même aucune tâche préalable d'étiquetage manuel n'est requise. Les algorithmes non supervisés les plus répandus tendent à minimiser une fonction coût, prototype d'une classe étant un point connu dont l'appartenance à la classe est garantie et où chaque pixel est assigné à la classe qui lui est la plus proche. Selon la certitude de la classification que nous voulons obtenir, et la relation entre les classes, nous pourrions distinguer plusieurs méthodes de classification.

a. l'algorithme K-Means

L'algorithme k-means est l'algorithme de clustering le plus connu et le plus utilisé, du fait de sa simplicité de mise en oeuvre. Il partitionne les données d'une image en K clusters. Contrairement à d'autres méthodes dites hiérarchiques, qui créent une structure en « arbre de clusters » pour décrire les groupements, k-means ne crée qu'un seul niveau de clusters. L'algorithme renvoie une partition des données, dans laquelle les objets à l'intérieur de chaque cluster sont aussi proches que possible les uns des autres et aussi loin que possible des objets des autres clusters. Chaque cluster de la partition est défini par ses objets et son centroïde[20].

Soit $X = \{x_k/K=1, \dots, n\}$ un sous-ensemble d'un espace n-dimensionnel. Une partition de X en C classes peut être représentée par des sous-groupes mutuellement disjoints $C_1 \dots C_c$, tel que $C_1 \cup \dots \cup C_c = X$.

Chaque vecteur x_k est assigné à une et seulement une classe qui lui est la plus proche. La fonction objective correspondante s'exprime sous la forme:

$$J = \sum_{i=1}^c \sum_{j=1}^n U_{ik} d^2(x_k, v_i) \quad (2,2)$$

Où:

- v_i désigne le prototype (centroïde) de cette classe.
- d mesure la similarité entre le prototype et le vecteur de données.
- U_{ik} est un indicateur binaire $\in \{0,1\}$ désignant l'appartenance ou non du vecteur x_k à la classe C_i ($U_{ik}=1$ si $x_k \in C_i$ et $U_{ik}=0$ sinon, $\forall i=1 \dots C$ et $x_k \in X$), L'indicateur d'appartenance U_{ik} s'exprime comme suit:

$$U_{ik}(x_k, v_i) = \{1 \text{ si } d^2(x_k, v_i) = \min C(x_k, v_i) \text{ ou } 0 \text{ sinon}\}.$$

L'optimisation se fait de manière itérative. Les étapes de cette classification pouvant être résumé comme suit:

- Initialiser l'ensemble des centroides avec des éléments, soit pris au hasard de l'ensemble de données, soit d'après des connaissances a priori.
- Assigner chaque point au centroïde le plus proche.
- Calculer les nouveaux centroides des classes construites. Les centroïdes étant les centres des classes, ils sont calculés par la moyenne des éléments de la classe.
- Générer une nouvelle partition de l'ensemble de données en assignant chaque point au centroïde le plus proche.
- S'il y a changement dans la partition, réitérer depuis l'étape 3 [5].

b. L'algorithme Fuzzy C-Maens [5]

La modélisation de l'imprécision s'effectue en considérant des frontières graduelles au lieu de frontières nettes entre les classes. L'incertitude s'exprime par le fait qu'un pixel possède aussi bien des attributs qui l'assignent à une classe qu'à une autre. La classification floue assigne donc, non pas à un pixel une étiquette relative à une classe unique, mais son degré d'appartenance à chacune des classes. Ces valeurs expriment l'appartenance incertaine d'un pixel à une région et sont appelées degrés d'appartenance. Le degré d'appartenance se situe dans l'intervalle $[0, 1]$ et les classes obtenues ne sont pas forcément disjointes. Dans ce cas, les données X_j ne sont plus assignées à une classe unique, mais à plusieurs par l'intermédiaire de degrés d'appartenance U_{ij} du vecteur X_j à la classe i . Le but des algorithmes de classification est non seulement de calculer les centres de classe B mais aussi l'ensemble des degrés d'appartenance des vecteurs aux classes.

Si U_{ij} est le degré d'appartenance de X_j à la classe i , la matrice $U_{C \times N} [U_{ij}]$ est appelée matrice de C-partitions floues si et seulement si elle satisfait aux conditions:

$$\forall i \in \{1 \dots C\}, \forall j \in \{1 \dots N\} \begin{cases} U_{ij} \in [0,1] \\ 0 < \sum_{j=1}^N U_{ij} < 1 \end{cases} \quad (2,3)$$

$$\forall j \in \{1 \dots N\} \sum_{i=1}^c U_{ij} = 1 \quad (2.4)$$

La fonctionnelle à minimiser (2.5), et les solutions (2.6), (2.7), au problème du FCM sont décrites par les formules suivantes:

$$J(B, U, X) = \sum_{i=1}^c \sum_{j=1}^n (u_{ik})^m d^2(x_j, b_i) \quad (2,5)$$

$$b_i = \frac{\sum_{k=1}^n u_{ik}^m \cdot x_k}{\sum_{k=1}^n u_{ik}^m} \quad (2,6)$$

$$u_{ij} = \left[\sum_{k=1}^c \left(\frac{d^2(x_j, b_i)}{d^2(x_j, b_k)} \right)^{2/(m-1)} \right]^{-1} \quad (2,7)$$

c. L'algorithme Possibilités C-Maens [15][21]

Krishnapuram et Keller ont proposé une approche possibiliste des c-moyennes appelée Possibilistic C-Means, ou PCM. Leur approche est censée conduire à une meilleure performance en présence de bruit. Mais leur travail est motivé essentiellement par le désir de remédier au caractère relatif des degrés d'appartenance générés par les FCM, qui est dû à la contrainte probabiliste que la somme de tous les degrés d'appartenance d'un objet aux différentes classes soit égale à 1. Une condition nécessaire pour passer à des degrés d'appartenance "absolus" est donc de lever cette contrainte, ou, plutôt, de l'assouplir afin que les degrés d'appartenance ne soient plus dépendants les uns des autres. Cependant, bien qu'incarnant cette notion de typicalité, l'algorithme PCM présente des performances plutôt médiocres comparées à celles de l'algorithme FCM. Dans

le cas général, l'algorithme PCM a tendance à produire des classes concentriques, i. e les classes obtenues ont des centres très proches, voire confondus.

L'algorithme PCM

L'algorithme PCM permet de partitionner un ensemble $X = \{x_1, x_2, \dots, x_n\} \in R^P$ en C classes, en minimisant la fonction objective:

Principe d'algorithme PCM

$$J(B, U, X) = \sum_{i=1}^C \sum_{j=1}^N (u_{ij})^m d^2(x_j, b_i) + \sum_{i=1}^C \beta_i \sum_{j=1}^N (1 - u_{ik})^m \quad (2,8)$$

$$u_{ik} \in [0,1] \text{ et } 1 \leq i \leq c \text{ et } 1 \leq k \leq N \quad (2,9)$$

$$0 < \sum_{j=1}^N u_{ij} < N \text{ et } 1 \leq i \leq c \quad (2,10)$$

Algo 2.1 Principe d'algorithme PCM

- u_{ij} C'est la distance euclidienne entre l'élément k et le centre de la classe i .
- $V = \{v_1, v_2, \dots, v_n\}$ Est l'ensemble des prototypes des classes C_k dont la distance aux x_i est notée d_{ik} .
- $\beta = \{\beta_1, \beta_2, \dots, \beta_n\}$ est l'ensemble des termes de pénalité den PCM des données w_i Est à déterminer par l'utilisateur, et représente une évaluation de l'extension de la classe k . La minimisation de $J(U, V, \beta)$ consiste à minimiser pour toutes les classes: atypiques associés à chacune des classes :

$$J(U, V, \beta) = (u_{ij})^m d^2(x_i, b_i) + \beta_i (1 - u_{ik})^m \quad (2.14)$$

En dérivant la formule (2.14) par rapport à u_{ik} et en annulant cette dérivée on obtient:

$$u_{ij} = \left[1 + \left(\frac{d_{ik}^2}{\beta_i} \right)^{\frac{1}{m}} \right]^{-1} \quad (2,15)$$

Où le paramètre représente w_i le diamètre de la classe et peut être défini a priori par initialisation. Les points aberrants (les points très éloignés de toutes

les classes) se voient alors dotés de petits poids pour toutes les classes et dont l'influence dans la phase de classification est réduite. Le PCM souffre, cependant, du problème de coïncidence de classes. On peut avoir, dans certains cas, des classes confondues tandis que des sous-groupes de données normaux fonction objectif est obtenue quand toutes les classes sont coïncidentes.

Des résultats plus ou moins satisfaisants ont été obtenus avec le PCM parce que l'optimisation mène à un minimum local mais pas à un minimum global. Cette propriété n'est pas toujours satisfaite, d'un point de vue théorique.

2.6. Présentation de l'algorithme FCM

Cet algorithme est une extension de l'algorithme de K-Mean en introduisant une notion de floue dans la définition du degré d'appartenance. Son principe est de regrouper des individus dans k classes qui soient le plus homogènes et naturelles possibles. Les groupes obtenus doivent contenir des pixels les plus semblables, et entre groupes différents les individus doivent être le plus différents possible. En fait, la modélisation de l'imprécision s'effectue en considérant des frontières graduelles au lieu de frontières nettes entre les classes. L'incertitude s'exprime par le fait qu'un pixel possède bien des attributs qui l'assignent à une classe qu'à une autre. La classification floue assigne donc, non pas à un pixel une étiquette relative à une classe unique, mais son degré d'appartenance à chacune des classes. Ces valeurs expriment l'appartenance incertaine d'un pixel à une région et sont appelées degré d'appartenance. Le but des algorithmes de classification est non seulement de calculer les centres de classe B mais aussi l'ensemble des degrés d'appartenances des vecteurs aux classes. [22]

2.6.1 Principe et utilisation

Evolution du K-Means, il fait intervenir l'appartenance floue des points à un cluster. La complexité de l'algorithme reste plus faible mais il permet un clustering plus fin puisqu'on détermine pour chaque point un degré d'appartenance aux différents clusters. Cette information plus nuancée que l'appartenance définitive à une classe peut se révéler utile pour des traitements ultérieurs de l'information. Cet algorithme garde, par contre, une forte sensibilité à l'initialisation tout comme le K-Means.

2.6.2 Description de l'algorithme

La phase d'initialisation est identique, mais va s'accompagner de l'initialisation de valeurs d'appartenance u_{ij} du vecteur x_i au cluster de centre. Ces valeurs sont comprises entre 0 et 1.

Après cette phase d'initialisation, la résolution du problème de clustering se fait par la minimisation de la fonction objective définie par (2.16), m étant un paramètre constant supérieur à 1 (généralement pris égal à 2). [23]

$$J \leftarrow \sum_{i=1}^c \sum_{j=1}^n u_{ik}^m d^2(x_j, b_i) \quad (2.16)$$

2.6.3 Les principales étapes [24]

Les principales étapes de l'algorithme Fuzzy C-means sont :

- La fixation arbitraire d'une matrice d'appartenance.
- Le calcul des centroïdes des classes.
- Le réajustement de la matrice d'appartenance suivant la position des centroïdes.
- Calcul du critère de minimisation et retour à l'étape 2 s'il y a non convergence de critère.

2.6.4 Description de la méthode

On considère l'espace de n points de dimension p suivant :

$$x : \begin{bmatrix} x_1^1 & \dots & x_1^j & \dots & x_1^p \\ \dots & \dots & \dots & \dots & \dots \\ x_i^1 & \dots & x_i^j & \dots & x_i^p \\ \dots & \dots & \dots & \dots & \dots \\ x_n^1 & \dots & x_n^j & \dots & x_n^p \end{bmatrix} \quad (2.17)$$

On suppose que les n points peuvent être groupés en c clusters $c < n$

Les clusters sont décrits par leurs centres :

$$V = [v_i^1, v_i^2, v_i^3, \dots, v_i^j, \dots, v_i^p] \quad \text{et } 1 \leq i \leq C \quad (2.18)$$

On considère la matrice de proximité suivante :

$$u : \begin{bmatrix} u_{11} & \dots & u_{1k} & \dots & u_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ u_{i1} & \dots & u_{ik} & \dots & u_{in} \\ \dots & \dots & \dots & \dots & \dots \\ u_{c1} & \dots & u_{ck} & \dots & u_{cn} \end{bmatrix} \quad \forall \quad \begin{cases} k = 1, \dots, n \\ i = 1, \dots, c \end{cases} \quad (2,19)$$

u_{ik} représente le degré d'appartenance du point x_k au centre v_i

$$u_{ij} = \left[\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)} \right]^{-1} \quad (2,20)$$

d_{ik} représente la distance entre v_i et v_k

$$u_{ij} = \left[\sum_{j=1}^c \left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{2/(m-1)} \right]^{-1} \quad (2,21)$$

, si on choisit la distance euclidienne [25] .

L'algorithme utilise dans notre approche [19].

Etape 1 : fixer les paramètres.

- Les entrées : $\mathbf{X} = (x_j, j = 1 \dots N)$ l'ensemble des vecteurs forme. C : nombre de classes.
- ε : Seuil représentant l'erreur de convergence (par exemple $\varepsilon = 0.001$)
- m : degré de flou, $m \in (1.5, \dots, 3)$

Etape2 : initialiser la matrice degrés d'appartenance U par valeur aléatoire dans l'intervalle $[0,1]$.

Etape 3 : mettre à jour la matrice prototype B par la relation.

$$b_i \leftarrow \frac{\sum_{k=1}^n u_{ik}^m \cdot x_k}{\sum_{k=1}^n u_{ik}^m} \quad (2,22)$$

$$j^{ancien} \leftarrow \sum_{i=1}^c \sum_{j=1}^n (u_{ik})^m d^2(x_j, b_i) \quad (2,22)$$

Etape 4 : mettre à jour la matrice des degrés d'appartenance par la relation

$$u_{ij} \leftarrow \left[\sum_{k=1}^c \left(\frac{d^2(x_j, b_i)}{d^2(x_j, b_k)} \right)^{2/(m-1)} \right]^{-1} \quad (2,23)$$

$$j^{nouveau} \leftarrow \sum_{i=1}^c \sum_{j=1}^n (u_{ik})^m d^2(x_j, b_i) \quad (2,21)$$

Etape 5 : répéter les étapes 3 a 4 jusqu'à satisfaction du critère d'arrêt qui s'écrit :

$$\|j^{ancien} - j^{nouveau}\| \leq \varepsilon$$

Les sorties : la matrice d'appartenance U et les centres des classe B .

2.6.5 Les avantages du FCM [26]

- Facilité de construction et d'interprétation des règles : les règles sont formulées de manière naturelle par les experts, comme en symbolique.
- Interprétation numérique entre les règles : lorsque deux règles sont activées en même temps car leurs prémisses sont en partie vérifiées, alors la conclusion proposée du système peut prendre une valeur intermédiaire entre les conclusions proposées par celle-ci.
- Intégration de la connaissance à priori.
- Robustesse vis à vis des incertitudes.

2.6.5 Les inconvénients du FCM [27]

L'algorithme flou c-moyens possède plusieurs inconvénients qui affectent à la performance de cet algorithme

- la spécification a priori du nombre de clusters.
- Pour de petites valeurs de nous obtenons le meilleur résultat, mais au détriment de plus nombre d'itérations.
- les mesures de distance euclidienne peuvent inégalement pondérer les facteurs sous-jacents.
- Un temps de calcul énorme.
- Sensibilité à la partition initiale (vitesse, minima locaux)
- La sensibilité au bruit.
- On s'attend à faible degré d'appartenance (ou même pas) pour les valeurs aberrantes(les points bruyants).

2.7 Concoction

Dans ce chapitre nous nous intéressons aux techniques de la classification automatique, ou segmentation (clustering). Nous avons vu que la segmentation (clustering) permet de regrouper des objets (individus ou variables) en un nombre limité de groupe ou de classes (segmentes, ou cluster).

La segmentation d'images est l'un des problèmes phares du traitement d'images. Elle consiste à partitionner l'image en un ensemble de régions connexes. L'intérêt de ces régions est de pouvoir être manipulées ensuite via des traitements de haut niveau pour extraire des caractéristiques de forme, de position, de taille, etc.

Le problème est évidemment très mal posé, car on ne sait jamais dire quelle est la segmentation idéale. L'idée est bien sûr que la région se rapproche de la notion d'objet, au sens courant du terme. Néanmoins, on peut dégager des propriétés plus raisonnables qu'on cherche à obtenir dans un algorithme de segmentation, en particulier:

- Stabilité: la segmentation obtenue ne doit pas varier beaucoup lorsque les conditions d'acquisition varie légèrement (bruit, illumination, point de vue, . . .).
- Régularité: les régions obtenues doivent être simples à manipuler (taille suffisante, forme régulière, . . .).