Chapitre 4

Segmentation d'images en régions

L'objectif de la segmentation d'images en régions est de partitionner une image en zones d'intérêt correspondant à des objets de la scène d'où elle est issue. Elle peut être située dans le cadre plus général de la segmentation de données.

Le problème de base de la segmentation de données est le suivant :

On connaît :

- un ensemble d'**entités**
- un ensemble d'**attributs** caractérisant ces entités
- des **relations topologiques** entre ces entités
- des **attributs relationnels** entre ces entités

On cherche :

- une (ou des) **partition(s) de ces données** ayant des propriétés intéressantes par rapport aux attributs et aux relations topologiques.

Les principaux problèmes qui se posent sont :

- définir les **propriétés des partitions** que l'on cherche
- concevoir des **algorithmes** permettant l'obtention de **partitions optimisant ces propriétés**.

Pour le problème spécifique de la segmentation d'images :

- les **entités** sont des **points d'images** par exemple bidimensionnelles ou tridimensionnelles.
- les **attributs** sont la **position dans l'espace** et les **luminances**.
- les **relations topologiques** sont la 4-connexité ou la 8-connexité.

• les **attributs relationnels** sont les caractéristiques des **frontières** entre deux ensembles d'entités.

On peut donc caractériser un problème de **segmentation d'images** par :

• **un ensemble de critères d'homogénéité** déterminant les propriétés des partitions de l'image que l'on cherche : ***modélisation analytique***.

• un algorithme utilisant ces critères de manière à segmenter l'image : *modélisation algorithmique*.

On est donc confronté à deux problèmes de base distincts :

• la définition de **critères de segmentation** spécifiques à chaque problème particulier de segmentation.

• la définition d'une **stratégie d'utilisation des critères** de segmentation qui peut être définie de manière générale.

On peut noter que dans le cas où les propriétés d'homogénéité des régions s'expriment simplement, il est possible de dériver un opérateur d'extraction de contour qui détecte les frontières entre les régions. On peut ensuite déterminer les régions par fermeture des contours ainsi obtenus. Un exemple de ce type d'approche est présenté ultérieurement. Mais souvent la recherche de filtres aisément implantables permettant de détecter les ruptures d'une propriété d'homogénéité d'ordre supérieur à 0 reste complexe. Ce constat a motivé le nombre important de travaux consacrés à la segmentation d'images en régions.

Les méthodes existantes peuvent être rangées dans deux catégories : les méthodes de "classification", et les méthodes de "croissance de régions". La différence principale entre ces deux types d'approches réside dans la stratégie d'utilisation de l'espace des luminances et des relations spatiales existant entre les pixels. Les méthodes de classification déterminent d'abord une partition de l'espace des luminances et se servent ensuite des relations de connexité pour déterminer les régions. Les méthodes de croissance de régions utilisent de manière simultanée ces deux types d'information.

4.1. Problématique : un problème d'optimisation NP-difficile

L'objectif de cette section est de montrer que la segmentation en régions dans des images de dimension quelconque se ramène à un problème d'optimisation. Nous verrons qu'il est le plus souvent NP-difficile, ce qui nécessite l'introduction d'heuristiques auxquelles est consacrée la section suivante.

La segmentation d'une image I utilisant un prédicat d'homogénéité P est communément définie [184, 98, 148] comme une partition $S = R_1, R_2, \dots R_n$ de I telle que :

1. $I = \cup\ R_i, \quad i \in [1 \dots n]$
2. R_i est connexe, $\forall i \in [1 \dots n]$
3. $P(R_i) =$ vrai, $\forall i \in [1 \dots n]$
4. $P(R_i \cup R_j) =$ faux, $\forall i \neq j$, pour tout couple (R_i, R_j) de régions connexes.

Il est important de remarquer que les conditions 1., 2., 3., 4. ne définissent pas, en général, une segmentation unique. Les résultats de segmentation dépendent par conséquent de l'ordre et de la manière avec lesquelles les données sont traitées et non pas seulement de l'information contenue dans l'image. Par exemple, dans beaucoup de méthodes une segmentation calculée sur la transposée d'une image n'est pas similaire à la transposée de la segmentation de l'image. Il est possible de réduire ce problème d'indétermination en ajoutant une contrainte d'optimisation d'une fonction C caractérisant la qualité d'une segmentation [127]. Soit Q une fonction caractérisant la qualité d'un élément de S (par exemple la variance des niveaux de gris des pixels d'une région). Soit C une fonction monotone et symétrique par rapport à la qualité de chaque élément de S : $C(S) = C(Q(R_1), \dots Q(R_n))$ (par exemple la moyenne arithmétique de $Q(R_1), \dots Q(R_n)$). Nous ajoutons aux axiomes 1. $\rightarrow$ 4. le suivant :

5. parmi toutes les segmentations possibles S vérifiant 1., 2., 3., nous recherchons la (ou une) segmentation S^* qui optimise la fonction de qualité c'est-à-dire telle que :

$$C(S^*) \leq C(S)\ \forall S \in S_P(I)$$

où $S_P(I)$ est l'ensemble des partitions de I, et C une fonction décroissante.

Très souvent C est additive, elle peut être par exemple dans le cas d'images 2-D la fonction C_1 définie comme suit :

$C_1(S) = \sum_{i=1,n} \sum_{(k,l) \in R_i} (I(k,l) - M_i)^2$; n étant le nombre des régions;

$M_i = \frac{\sum_{(k,l) \in R_i} I(k,l)}{Card(R_i)}$; $Card(R_i)$ étant le nombre de points d'une région R_i.

Il est important de remarquer que généralement les propriétés désirées pour les régions peuvent être décomposées hiérarchiquement de manière à simplifier leur optimisation [127]. Une suite $(P_1, C_1), \dots (P_n, C_n)$ sera ainsi utilisée au lieu d'un simple couple (P, C). Une telle décomposition permet de réduire, en le divisant le problème d'optimisation. Cette nécessité peut s'illustrer par un exemple simple. Supposons que nous voulions déterminer des angles droits à partir d'un ensemble de points. Dans une première étape, nous utiliserons un couple (P_1, C_1) permettant l'obtention de segments de droites à partir des points. Dans une deuxième étape, nous utiliserons un couple (P_2, C_2) permettant la formation d'angles droits avec les segments.

Il est clair que la condition 5. ne définit pas toujours une segmentation unique. En effet plusieurs segmentations ayant la même valeur minimale de C peuvent exister. Quoi qu'il en soit, cette condition permet d'obtenir une définition plus précise mais pose un problème algorithmique sous-tendant une complexité importante. C'est pourquoi la recherche d'algorithmes sous-optimaux est nécessaire.

4.2. Segmentation en régions par classification

Ces méthodes déterminent d'abord une partition de l'espace des luminances en utilisant les niveaux de gris présents dans l'image. On associe à chaque pixel la classe de niveaux de gris à laquelle il appartient. Les régions sont définies par les ensembles maximaux de pixels connexes appartenant à la même classe. Les prédicats d'homogénéité et les fonctions de qualité (voir paragraphe précédent) ne sont pas définis explicitement dans ce type d'algorithme. Le plus souvent la classification des luminances s'effectue à partir du calcul de l'histogramme de répartition dans l'image. On recherche les différents modes de l'histogramme et les "vallées" correspondantes. Les classes sont déterminées par les intervalles entre les vallées. Cette procédure fonctionne bien pour des images comprenant un nombre peu important d'objets ayant des niveaux d'intensité différents. Olhander [145] a amélioré cette idée en introduisant une classification récursive permettant d'opérer quand un nombre important d'objets est présent dans l'image. On définit un masque sélectionnant tous les pixels de l'image (un masque définit une zone de l'image). Pour chaque masque un histogramme de l'image masquée est calculé. Les modes de l'histogramme sont détectés permettant ainsi de segmenter l'espace des niveaux de gris. Les points de l'image sont étiquetés avec la classe correspondante. Les ensembles maximaux de pixels connexes appartenant à une même classe sont déterminés. Si l'histogramme comprend plus qu'un mode alors le masque est terminé. Sinon tous les ensembles connexes précédemment calculés sont utilisés pour générer des masques qui sont placés dans une pile de masques. Durant les itérations successives le masque courant détermine les pixels utilisés pour le calcul d'histogrammes. La classification est répétée pour chaque nouveau masque jusqu'à ce que la pile soit vide. Le nombre des modes de l'histogramme déterminés à chaque étape est un paramètre de cette procédure ; par exemple si ce nombre est réduit à l'unité on détermine à chaque étape le meilleur pic. Ohta, Kanade et Sakai [144] ont adapté cet algorithme pour la segmentation d'images couleur. Dans le cas où l'image contient un objet se détachant sur un fond, des techniques de classification par seuillage adaptatif ont été développées [34], [174]-[176]. Ces méthodes combinent l'information spatiale de l'image avec l'information de luminance pour aider la détermination des seuils. L'itération de ce type d'algorithme peut permettre la segmentation d'images comprenant plusieurs objets.

Chow et Kaneko [34] calculent en chaque point un seuil dépendant de l'histogramme de répartition des luminances sur son voisinage. Des voisinages carrés 33×33 ou 65×65 sont par exemple utilisés pour la détermination de ces "histogrammes locaux". De manière à éviter le calcul d'un histogramme pour chaque point Chow et Kaneko partitionnent l'image en blocs, calculent l'histogramme pour chaque bloc, déterminent un seuillage approprié pour chaque histogramme, et ensuite interpolent spatialement les valeurs des seuils de manière à obtenir un seuil adaptatif en chaque pixel.

Weszca, Nagel et Rosenfeld [176] suggèrent de déterminer un histogramme en prenant seulement en compte les pixels de laplacien faible. Seules les luminances à l'extérieur des zones de fort gradient donc a priori à l'intérieur de zones homogènes peu bruitées sont ainsi prises en compte.

Watanabe [174] propose de choisir une valeur de seuil qui maximise la somme des gradients calculée sur tous les pixels dont le niveau de gris est égal à la valeur du seuil.

Ces méthodes de "segmentation par classification" se révèlent efficaces si la classification des luminances permet de mettre en évidence les différentes régions homogènes de l'image. Dans le cas d'images comprenant des objets de luminances différentes se détachant sur un fond, cette approche donne de bons résultats. Mais lorsque les images sont bruitées et contiennent un nombre important d'objets la classification se révèle peu utilisable. Dans ce cas il est nécessaire d'utiliser les relations spatiales tout au long du processus de segmentation. Cette idée a été développée dans les méthodes dites de "croissance de régions".

4.3. Segmentation par croissance de régions

Ces techniques consistent à regrouper itérativement des ensembles de points connexes en régions plus importantes en utilisant des conditions dépendant de propriétés d'homogénéité. L'idée de base de ce type d'approche est de définir des critères de regroupement des pixels permettant l'obtention de régions homogènes. Les relations spatiales sont ainsi utilisées tout au long du processus de segmentation. Ces méthodes peuvent être subdivisées en deux classes : "agrégation de points" et "regroupement itératif d'ensembles de points".

4.3.1. Agrégation de points

Ces algorithmes associent à chaque pixel un vecteur de propriétés. Deux pixels sont regroupés si leurs vecteurs de propriétés sont "suffisamment similaires". Le résultat de la segmentation est constitué par les composantes connexes déterminées. Divers vecteurs de propriétés et diverses mesures de similarité ont été proposées [6, 81].

L'heuristique la plus simple consiste à réunir deux pixels si leur différence de niveau de gris est suffisament faible. Bryant [30] normalise cette différence par la moyenne des différences calculées sur tous les couples de pixels voisins de l'image. Asano et Yokoya [6] regroupent deux pixels si leur différence de niveaux de gris est faible par rapport à la plus grande différence existant entre chacun des pixels et ses voisins pris dans un petit voisinage carré.

Nagao et Matsuyama utilisent ce type d'algorithme pour la segmentation d'images couleur [139]. Un pixel est regroupé avec un noyau de région si la norme SUP entre son vecteur de composantes (dans R, V, B) et un vecteur de composantes quelconques d'un point du noyau est majorée par un seuil. Une expansion point à point permet d'obtenir un ensemble de régions dont l'amplitude de variation des intensités des points selon les composantes rouge (R), vert (V), bleu (B) est majorée par le seuil.

Des critères de regroupement plus sophistiqués attachant à chaque pixel un vecteur de propriétés dépendant d'un voisinage $k \times k$ autour du pixel ont été employés. Souvent le vecteur propriété est issu d'un opérateur de détection de contour [81, 80]. Les pixels non séparés par un contour sont associés. Les résultats obtenus dépendent évidemment du détecteur de contour utilisé.

Les limitations de ce type d'approche par agrégation de points sont liées au fait que les entités que l'on regroupe (les pixels) véhiculent peu d'informations. En effet ces méthodes consistent à déterminer des noyaux de régions et ensuite à regrouper itérativement des points à ces noyaux. Les seules informations dont on dispose pour décider d'un regroupement sont les caractéristiques du noyau de régions et la luminance du point. Les méthodes par regroupement itératif d'ensembles de points permettent d'utiliser des informations plus riches pour former les régions.

4.3.2. Regroupement itératif d'ensemble de points

L'idée de ces méthodes est de définir une succession de partitions de l'image par regroupement itératif de régions connexes. Cette stratégie permet de définir des heuristiques de regroupement dépendant de la similarité des propriétés des régions.

La première approche de ce type a été proposée par Muerle et Allen [138]. Ils suggèrent de regrouper deux régions adjacentes si leurs distributions de niveaux de gris sont suffisament similaires. Ils recommandent d'utiliser le test de Kolmogorov-Smirnov.

Brice et Fennema [29] effectuent la croissance de régions en partitionnant l'image en ensembles initiaux de points de même intensité. Ils regroupent ensuite séquentiellement les couples de régions adjacentes dont une fraction significative de la frontière a un contraste faible. Horowitz et Pavlidis utilisent un "quad tree" pour représenter l'image. Une stratégie de fusion et d'explosion ("split and merge") utilisant le quad-tree est ensuite mise en œuvre [98]. On obtient ainsi un ensemble de régions dont l'amplitude de variation des luminances est majorée par un seuil.

La méthode développée par Pong et al. [155] est basée sur l'utilisation séquentielle de plusieurs algorithmes : l'image est d'abord segmentée en utilisant un "sloped facet model" [82] ; ensuite la segmentation initiale sert d'entrée à un algorithme de croissance de régions. Deux régions adjacentes sont fusionnées si un critère de similarité est vérifié.

La plupart de ces méthodes regroupent itérativement tous les couples de régions adjacentes vérifiant certaines conditions. Le processus de croissance s'arrête quand plus aucun couple de régions adjacentes ne vérifie les conditions de fusion. Généralement le résultat obtenu dépend de l'ordre avec lequel les couples de régions sont regroupés. Ainsi le principe des méthodes les plus efficaces est de définir une hiérarchie de fusion.

4.3.2.1. Croissance hiérarchique de régions

On décrit un algorithme sous-optimal illustrant la définition de la segmentation proposée précédemment. L'idée de base de cet algorithme est d'optimiser la qualité globale de la segmentation (S) par une optimisation locale [127, 65]. Il consiste à regrouper parmi tous les couples de régions dont la réunion vérifie le prédicat (P) celui ayant la meilleure qualité locale (Q). Il s'écrit comme suit :

S = I

tant qu'un couple de régions connexes dont l'union vérifie P existe faire :

- regrouper parmi tous les couples (R_i, R_j) dont l'union vérifie P celui minimisant $Q(R_i \cup R_j)$
- mettre à jour S

Cet algorithme de regroupement optimise l'estimation globale de la qualité de la segmentation par optimisation locale de Q. Le fait que le coût de fusion (Q) soit optimisé sur l'image entière permet au processus de fusion d'être fortement guidé par les données. On ne définit pas de sens de parcours ou de traitement des données. L'ordre avec lequel l'image est traitée n'est pas déterminé a priori comme dans la plupart des algorithmes de segmentation. Cet algorithme peut être itéré pour utiliser une suite de critères de fusion : $(P_1, Q_1) \ldots (P_n, Q_n)$.

Au premier abord cet algorithme semble sous-tendre une complexité importante. En effet le choix du meilleur couple de régions connexes par rapport à (P_i, Q_i) amène en général des implantations coûteuses. Or si certaines conditions relatives aux (P_i, Q_i) sont vérifiées, une implantation de faible complexité algorithmique est possible [127]. On va maintenant décrire ces propriétés. On donnera tout d'abord quelques définitions, et ensuite les hypothèses nécessaires sur les (P_i, Q_i).

4.3.2.1.1. Définitions

Soit $E = e_1, e_2, \ldots e_n$ une partition initiale de l'image ; E peut être l'image ou une segmentation initiale.

Soit $P(E)$ l'ensemble des sous-ensembles de E.

Soit R l'ensemble des réels.

Soit k le nombre d'attributs associés à un sous-ensemble de E.

Soit A une fonction de $P(E)$ vers R^k qui associe un vecteur d'attributs à un sous-ensemble de E (par exemple pour une image de niveaux de gris : l'aire, la moyenne ou la variance de niveau de gris, ...). Soit m le nombre d'attributs associés à un couple de sous-ensembles de E.

Soit G une fonction de $P(E)^2$ vers R^m qui associe un vecteur d'attributs à un couple de sous-ensembles de E (par exemple pour une image de niveaux de gris : le gradient moyen à la frontière).

Soit S une partition de $E, S = \{E_1, E_2, \ldots E_p\}$.

On associe à S l'ensemble des attributs de ses éléments et l'ensemble des attributs de ses paires d'éléments respectivement $V(S)$ et $W(S)$:

$$\begin{aligned} V(S) &= \{A(E_1), A(E_2), \ldots A(E_p)\} \\ W(S) &= \{G(E_r, E_s); r, s \in [1 \ldots p]\} \end{aligned}$$

Chaque triplet $(S, V(S), W(S))$ définit et caractérise une segmentation de E.

On définit l'action consistant à fusionner deux éléments de S E_i et E_j. Cette opération consiste à définir un nouveau triplet :

$$(S, V(S), W(S))_{(i,j)} = (R, V(R), W(R))$$

$$\begin{aligned} R &= \{E_1, \ldots, E_{i-1}, E_i \cup E_j, E_{i+1}, \ldots, E_{j-1}, E_{j+1}, \ldots, E_{p-1}\} \\ V(R) &= \{A(_1), \ldots, A(E_{i-1}), A(E_i \cup E_j), \ldots, A(E_{p-1})\} \\ W(R) &= \{G(E_p, E_q), p \neq i, q \neq j, G(E_q, E_i \cup E_j) \ldots\} \end{aligned}$$

4.3.2.1.2. Conditions pour une implantation rapide

On suppose que les conditions (A_1) et (A_2) suivantes sont vérifiées :

(A_1) *: La suite de couples : (Prédicat de fusion, Coût de fusion)* $= (P, Q)$ *définissant les heuristiques de regroupement est telle que :*

$\forall E, F \in S$ *nous avons :*

$$\begin{aligned} P(E \cup F) &= P(A(E), A(F), G(E, F)) \\ Q(E \cup F) &= Q(A(E), A(F), G(E, F)). \end{aligned}$$

Cela signifie que le prédicat d'homogénéité et la fonction de qualité de la réunion de deux régions peuvent être calculés à partir de leurs attributs.
(A_2) *: Il existe deux fonctions T et U telles que :*

T est une fonction de $(R^k)^2$ vers R^k telle que :

$$\forall S_1, S_2 \in P(E); A(S_1 \cup S_2) = T(A(S_1), A(S_2))$$

Cela signifie que les attributs de la réunion de deux régions peuvent être calculés à partir des attributs des deux régions.

U est une fonction de $(R^m)^2$ dans R^m telle que :

$$\forall S_1, S_2, S_3 \in P(E); G(S_1, S_2 \cup S_3) = U(G(S_1, S_2), G(S_1, S_3))$$

Cela signifie que les attributs du couple $(S_1, S_2 \cup S_3)$ peuvent être calculés à partir de ceux de (S_1, S_2) et (S_1, S_3).

Les fonctions T et U permettent de mettre à jour aisément $(S, V(S), W(S))$ après une fusion. Ainsi, si un triplet initial $(S, V(S), W(S))$ et les fonctions T et U sont données, nous pouvons effectuer une suite de fusions en mettant à jour itérativement le triplet courant.

4.3.2.1.3. Remarques

Les conditions (A_1) et (A_2) permettent de réaliser une implantation de l' algorithme dont la complexité est faible. Dans beaucoup de problèmes de segmentation on peut trouver des critères de segmentation qui vérifient ces conditions. Le problème principal posé par la condition (A_1) est que le prédicat ainsi que le coût de fusion sont uniquement des fonctions du couple de régions. En effet, aucune information contextuelle (concernant, par exemple, les attributs des régions voisines) ne peut être insérée dans un couple (P_i, Q_i) vérifiant (A_1). Pour le cas de l'optimisation d'une fonction globale de qualité Ci monotone et symétrique par rapport aux qualités locales Q_i, cette restriction n'est pas gênante. En effet, une telle estimation de la valeur d'une segmentation ne nécessite pas, pour être optimisée, d'information contextuelle.

La caractérisation des attributs induite par (A_2) implique une restriction sur la nature des attributs utilisés. L'existence de la fonction T signifie que les attributs associés à une région peuvent être calculés à partir des attributs associés à un ensemble quelconque de régions formant une partition de la région. L'existence de la fonction U implique que les attributs associés à un couple de régions peuvent être déterminés à partir de ceux des couples du produit cartésien de deux partitions quelconques des deux régions. Pratiquement, quand un attribut intéressant pour un problème de segmentation ne vérifie pas la condition (A_2) nous cherchons à l'exprimer en fonction d'attributs la vérifiant. D'un point de vue théorique une telle décomposition existe toujours ; le pire des cas étant quand elle utilise tous les points de la région. Nous avons évidemment intérêt à employer des attributs qui peuvent être exprimés avec un nombre réduit d'attributs vérifiant (A_2). Nous allons illustrer ces remarques par le cas simple de la variance des niveaux de gris souvent utilisée pour la segmentation des images naturelles. La variance des niveaux de gris de deux régions R_1 et R_2 n'est pas uniquement fonction de la variance de R_1 et de la variance de R_2. Cependant elle peut être calculée si on connait l'aire, la somme des niveaux de gris, et la somme des carrés des niveaux de gris des deux régions R_1 et R_2. Ces trois attributs vérifient la propriété (A_2), en effet le vecteur attribut incluant respectivement ces trois attributs est associé à la fonction T suivante : $T((x_1, x_2, x_3), (y_1, y_2, y_3)) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$. Ainsi si nous voulons utiliser la variance des niveaux de gris, pour segmenter des images naturelles 2-D par exemple, nous inclurons dans les attributs associés aux régions : l'aire, la somme des niveaux de gris, et la somme des carrés des niveaux de gris. De même la matrice de covariance d'un ensemble de points 2-D ou 3-D peut être calculée aisément à partir des sommes des produits de coordonnées. Les mêmes remarques concernant la condition (A_2) peuvent être faites au sujet des attributs associés aux couples de régions tels que : nombre de points et gradient moyen le long de la frontière.

4.3.2.1.4. Implantation

On représente un triplet $(S, V(S), W(S))$ par un graphe valué. On associe à chaque nœud du graphe un élément de S et le vecteur attribut qui le caractérise. Nous associons à chaque arc du graphe un couple d'éléments de S et le vecteur attribut qui caractérise les relations existant entre les deux nœuds. Chaque nœud du graphe correspond à la réunion d'un ensemble de nœuds issus du graphe initial -le graphe initial étant obtenu à partir d'une segmentation initiale (dans le pire des cas l'image). Pratiquement des arcs sont créés seulement entre des régions vérifiant des propriétés intéressantes par exemple de connexité ou de recouvrement [127, 130].

Le processus de croissance de régions consiste donc à calculer le graphe initial et à itérer les deux actions suivantes :

1. détermination du couple de nœuds à regrouper.

2. mise à jour du graphe après le regroupement des deux nœuds correspondant.

L'étape 1 nécessite un accès rapide au couple de nœuds vérifiant le prédicat courant et optimisant la fonction de coût. Il est donc nécessaire de définir une structure de données permettant de représenter efficacement un ensemble d'objets muni d'une relation d'ordre et sur lequel doivent être appliqué les opérations suivantes :

- insérer

- ôter le minimum

La structure de **tas (heap)** permet d'effectuer ces opérations avec un coût de l'ordre du logarithme du nombre d'éléments pris en compte. On trouvera une description du tas et de son implantation dans [177].
L'étape 2 nécessite un **accès rapide** aux arcs incluant un nœud donné ainsi qu'**aux attributs** correspondant. On peut donc choisir une représentation du graphe permettant un accès direct à l'ensemble des arcs comprenant un nœud donné ainsi qu'aux attributs correspondants.
Le graphe est représenté par les **listes d'arcs associées à chaque nœud et des listes d'attributs associées aux nœuds et aux arcs**.

On spécifie ainsi un **problème de segmentation** par :

- un **graphe initial** issu d'une partition initiale et les attributs associés à ses nœuds et à ses arcs.

- **deux fonctions de mise à jour** attachées respectivement aux attributs-nœuds et aux attributs-arcs (T et U).

- **une suite de couples (Prédicat de fusion, Coût de fusion)** qui définit la stratégie de regroupement $((P_1, Q_1), \ldots (P_n, Q_n))$.

4.3.2.1.5. Complexité

Discutons maintenant de la complexité algorithmique de cet algorithme. Les calculs provoqués par la fusion de deux nœuds sont donc les suivants :

• calcul des attributs du nouveau nœud en utilisant la fonction de mise à jour T.

• calcul des attributs des arcs incluant le nouveau nœud en utilisant la fonction de mise à jour U.

Nous devons ajouter aux calculs précédents ceux dus à la détermination du couple à fusionner et à la mise à jour du tas c'est-à-dire :

• prendre la tête du tas jusqu'à ce qu'un arc valide soit extrait.

• calcul du coût attaché aux nouveaux arcs et rangement des pointeurs correspondants dans le tas.

Plus formellement, la complexité algorithmique peut être calculée comme suit :

Soient N et B respectivement le nombre de nœuds et le nombre d'arcs du graphe initial.

Soit F le nombre de fusions effectuées durant la segmentation.

Chaque nœud est connecté en moyenne avec $V = \frac{2B}{N}$ nœuds ; on suppose que V reste à peu près constant durant le processus.

Soit T_1 le nombre d'opérations requises pour le calcul de la fonction T.

Soit T_2 le nombre d'opérations requises pour le calcul de la fonction U.

Soit T_3 le nombre d'opérations requises pour le calcul du prédicat et de la fonction de coût.

Soit M le nombre moyen de nœuds adjacents vérifiant le prédicat de fusion pour un nœud donné du graphe initial ; on suppose que M reste constant durant le processus de regroupement.

Le calcul nécessité par la mise à jour de graphe lors des fusions est : $FT_1 + FT_2V$.

Pour mettre à jour **tas** lors d'une fusion nous insérons dans le tas en moyenne V nouveaux arcs et ainsi nous calculons V nouvelles valeurs de prédicats et de coût. De plus nous insérons dans le tas les nouveaux arcs qui vérifient le prédicat. Nous pouvons raisonnablement supposer que le nombre maximum d'éléments dans le tas est à peu près $V.M$.

Le nombre d'opérations requises pour la mise à jour du tas lors des fusions est donc approximativement :

$$FT_3V + FMV\ O(LOG(VM))$$

Le second terme de la somme ci-dessus correspond au coût de l'insertion des valeurs des nouveaux arcs créés par le regroupement.

La complexité totale de notre algorithme est donc :

$$O(FT_1 + FT_2V + FT_3V + FMVLOG(VM))$$

Ce résultat montre l'importance de V et M pour la complexité de notre algorithme. M caractérise l'état du graphe initial par rapport au prédicat de fusion. On a donc intérêt à utiliser une bonne méthode (!) pour déterminer une segmentation initiale. Nous remarquons que la complexité est une fonction de la "densité d'arc" du graphe. Nous aurons donc intérêt à limiter le nombre de relations entre les nœuds du graphe.

Cet algorithme de segmentation par croissance de régions s'applique pour des images de dimensions quelconques, monochromatiques et multichromatiques, ainsi que pour la segmentation de surfaces 3D [127, 130]. Il faut noter qu'il peut prendre en compte le résultat d'une détection de contours afin de contrôler la croissance des régions avec des propriétés de discontinuité [131]. Nous donnons un exemple d'application de cette méthode pour des images 2D monochromatiques [127].

4.3.2.1.6. Application aux images monochromatiques

De manière à appliquer cette méthode de segmentation aux images monochromatiques il faut établir une hiérarchie des propriétés désirées pour les régions. Ensuite, on attache à chaque propriété d'homogénéité un prédicat, une fonction globale de qualité, et une fonction locale de qualité. On détermine alors les attributs associés aux régions et aux couples de régions permettant de calculer le prédicat d'homogénéité et la fonction locale de qualité. On montre enfin que ces prédicats et ces fonctions vérifient les propriétés définies dans le paragraphe précédent.

La détermination de critères de segmentation s'appliquant à une classe importante d'images naturelles monochromatiques n'est pas aisée. En effet la multiplicité des facteurs physiques intervenant dans la formation d'une image rend difficile le calcul de fonctions d'homogénéité permettant de différencier les projections des divers objets de la scène. L'expérimentation montre que des critères simples donnent le plus souvent des résultats comparables à ceux obtenus à l'aide de critères compliqués et que l'on maîtrise mal. Il est donc intéressant d'utiliser une suite de critères simples permettant de caractériser une région homogène au sens des niveaux de gris.

Le premier critère basé sur l'amplitude des variations des niveaux de gris à l'intérieur d'une région permet de déterminer les régions de luminance quasiment uniforme (cette amplitude de variation est obtenue en effectuant la différence entre le niveau de gris le plus haut et le niveau de gris le plus bas).

Le deuxième critère, basé sur la variance des niveaux de gris, engendre la formation de régions homogènes mais bruitées. Les régions de luminance à peu près constante sont ainsi regroupées afin de former des régions de variance faible.

Le troisième critère permet de fusionner les couples de régions voisines de luminances moyennes sensiblement différentes, mais présentant une frontière avec un faible gradient.

On remarquera que ces trois critères ont une complexité calculatoire croissante. De manière générale, plus on avance dans le processus de regroupement, plus les informations nécessaires à la croissance des régions deviennent complexes. Mais en contrepartie, le nombre de régions diminuant, le coût du processus de fusion n'est pas affecté. L'ordre d'application des critères est important car il correspond à une description hiérarchique des régions recherchées. Il serait par exemple illogique d'utiliser d'abord la variance des niveaux de gris et ensuite leur amplitude de variation. En effet, lorsque le processus de regroupement débute avec tous les pixels de l'image la variance des niveaux de gris d'une région est moins significative et plus coûteuse à calculer que l'amplitude de leurs variations. Lorsque les régions de luminance uniforme se sont formées, alors seulement la variance devient un critère intéressant. De même, lorsque les régions homogènes bruitées sont déterminées alors le gradient sur la frontière devient significatif. Les informations utilisées par le processus de croissance de régions évoluent donc au cours de celui-ci, correspondant à des niveaux de description différents de l'image.

Plus concrètement, on suppose qu'une région peut être hiérarchiquement décrite comme un ensemble connexe de points tel que :

1. la somme des gradients calculés sur ses points intérieurs est faible, i.e. des contours ne la traversent pas,

2. la région peut être partitionnée en sous-régions dont la variance des niveaux de gris est faible,

3. chaque sous-région de faible variance de niveau de gris peut être elle-même décomposée en sous-régions dont les pixels ont à peu près la même luminance.

La troisième propriété peut être caractérisée par le prédicat P_1, la fonction de qualité globale C_1 et la fonction de qualité locale Q_1 :

Soit $I(i,j)$ l'image initiale Soit $S = \{R_1, R_2, \dots R_n\}$ une segmentation de I

$P_1(R_i) = [(MAX_i - MIN_i) < s_1]$; s_1 est un seuil ; MIN_i et MAX_i sont respectivement la plus petite et la plus grande valeur prises par I dans la région R_i.

$$C_1(S) = \sum_{i=1,n} \sum_{(k,l)\in R_i} ((I(k,l) - MAX_i)^2 + (I(k,l) - MIN_i)^2)$$

$$Q_1(R_i) = MAX_i - MIN_i$$

La seconde propriété peut être caractérisée par le prédicat P_2, la fonction de qualité globale C_2 et la fonction de qualité locale Q_2 :

$P_2(R_i) = [V(R_i) < s_2]$; s_2 est un seuil ; $V(R_i)$ est la variance des niveaux de gris des pixels de R_i.

s_2 est pris égal à s_1^2 de manière à ce que les régions issues de la première étape vérifient P_2 :

Notons $E(R_i)$ et $C(R_i)$ les moyennes respectives des niveaux de gris et du carré des niveaux de gris de la région R_i. Nous avons :

$MAX_i - MIN_i < s_1 \Rightarrow \mid I(k,l) - I(r,s) \mid < s_1 ; \forall (k,l)$ et (r,s) points de R_i (I est la fonction des niveaux de gris) $\Rightarrow (I(k,l) - I(r,s))^2 < s_1^2 \Rightarrow C(R_i) - E(R_i)^2 < s_1^2 \Rightarrow V(R_i) < s_1^2$

$C_2(S) = \sum_{(i,j)\in I} (I(i,j) - M_{(i,j)})^2$; $M_{(i,j)}$ est la moyenne des niveaux de gris de la région où (i,j) est inclus.

$$Q_2(R_i) = V(R_i)$$

La première propriété peut être exprimée par le prédicat P_3, la fonction de qualité globale C_3 et la fonction de qualité locale Q_3 :

Soient R_1 et R_2 deux ensembles de points connexes et S une segmentation.

1													
1	1												
1	1	1	1	1	1	2	2	2	2	2			
1	1	1	1	1	2	2	2	2	2	2	2	2	
1	1	1	1	1	2	2	2	2	2	2	2	2	
1	1	1	1	1	2	2	2	2	2	2			
1	1												
1													

Figure 4.1. *Les couples de points déterminés par les extrémités des segments définissent la frontière entre la région "1" et la région "2" : F("1","2").*

Soit $F(R_1, R_2)$ l'ensemble des couples de points définissant la frontière entre R_1 et R_2 (voir figure 4.1.).

Soit $N_F(R_1, R_2)$ le nombre de couples de points de $F(R_1, R_2)$.

Soit $D(R_1, R_2)$ le gradient moyen à la frontière entre R_1 et R_2 :

$$D(R_1, R_2) = (\sum ((i,j),(k,l)) \in F(R_1, R_2) \frac{\mid I(i,j) - I(k,l) \mid)}{(N_F(R_1, R_2))}$$

$C_3(S) = \sum_{((i,j),(k,l)) \in X_s(I)} \mid I(i,j) - I(k,l) \mid$; $X_s(I)$ est l'ensemble des couples de points connexes de l'image I inclus dans une même région S.

Nous supposons que pendant le processus de fusion :

$P_3(R1 \cup R_2) = [D(R_1, R_2) < s_3]$; s_3 est un seuil.

s_3 est pris supérieur à la valeur maximum de $D(R_1, R_2)$ sur les couples de régions adjacentes issues de l'étape précédente.

$$Q_3(R_1 \cup R_2) = D(R_1, R_2)$$

Le choix des différents seuils s_1, s_2, s_3 implique $P_1 < P_2 < P_3$.

$<$ est le symbole de la relation d'ordre partiel sur les prédicats :

Soient P et Q deux prédicats définis sur un ensemble E $P < Q$ signifie que si P est vrai alors Q est vrai.

On peut montrer que les conditions (A_1) et (A_2) définies précédemment sont vérifiées [127].

On spécifie ainsi ce problème de segmentation par les couples :

$$(P_1, Q_1), (P_2, Q_2), (P_3, Q_3)$$

Les attributs attachés à chaque région sont donc : l'aire, la somme des niveaux de gris, la somme des carrés des niveaux de gris, le maximum et le minimum des niveaux de gris ; et les attributs associés à chaque couple de régions : le nombre de couples de points définissant la frontière et la somme des différences de niveaux de gris entre ces points.

Lors d'une première étape on détermine une segmentation initiale afin de diminuer le coût du processus de regroupement. Nous utilisons la procédure MERGE de Pavlidis avec le prédicat d'homogénéité P_1.

Ensuite on construit le graphe d'adjacence correspondant aux régions obtenues de manière à appliquer la stratégie de regroupement définie par :

$$(P_1, Q_1), (P_2, Q_2), (P_3, Q_3)$$

La première étape de fusion permet d'obtenir des régions de luminance à peu près uniformes. Lors de la seconde étape des régions de niveaux de gris homogènes mais bruités se forment, grâce à la troisième les régions de luminance sensiblement différentes mais dont le gradient à la frontière est faible sont fusionnées. On ajoute une dernière étape qui élimine les petites régions.

4.3.2.1.7. D'autres critères de segmentation

On peut aussi utiliser des critères de segmentation basés sur une approximation de la distribution des niveaux de gris dans une région par des polynômes d'ordre 2.

Si on s'intéresse à la discrimination de régions ayant les mêmes statistiques d'ordre 1 les matrices de co-occurrence peuvent aussi être prises en compte par le processus de regroupement.

La même stratégie de regroupement peut être mise en œuvre pour approximer une surface avec des quadriques par morceaux [52].

4.4. Détection de régions par fermeture de contours

Une idée classique, par exemple développée par R. Deriche et J.-P. Cocquerez [46], est de détecter les régions en identifiant les contours fermés constituant leurs frontières. On détermine alors les régions non pas en recherchant des zones homogènes mais en détectant les points frontières entre deux zones homogènes de caractéristiques différentes.

Ces points frontières correspondent aux points de contour et peuvent être calculés avec un algorithme de détection de contours.

Il faut noter que cette approche prend le contrepied de la méthode précédente fondée sur la détermination d'ensembles connexes de points possédant des propriétés d'homogénéité. Les contours issus d'un algorithme de détection de contours sont rarement fermés, aussi les difficultés des approches "détection des régions par fermeture des contours" résident justement dans la fermeture de ces contours. Le succès de ce genre de méthodes dépend donc généralement de la qualité du détecteur de contours utilisé.

R. Deriche et J-P. Cocquerez utilisent les résultats obtenus par le détecteur de contours de R. Deriche décrit précédemment [43]. On remarquera que ce choix conduit à l'obtention de régions dont la distribution des niveaux gris correspond à une fonction de Heavyside (marche d'escalier) additionnée d'un bruit blanc.

L'originalité de cet algorithme est de proposer une méthode de fermeture basée sur un opérateur simple qui permet de fermer les contours en suivant les crêtes de gradient. Les régions sont ensuite obtenues par détection des composantes connexes maximales n'incluant pas de points de contours. Cette méthode a un coût algorithmique très faible (linéaire selon le nombre de points de l'image).

4.4.1. Description d'un algorithme de fermeture de contours

Le principe de cet algorithme [46] est de détecter, par balayage de l'image ligne par ligne avec un opérateur 3×3, les extrémités des contours, puis de fermer ces contours en les prolongeant par suivi des crêtes du gradient.

L'extraction des maxima locaux de la norme du gradient dans la direction du gradient permet l'obtention de contours d'épaisseur 1, le plus souvent.

Ainsi, une extrémité de contour peut être identifiée par un simple examen de son voisinage 3×3.

La topologie de chaque extrémité permet de définir une direction d'exploration pour la fermeture et à chaque configuration d'extrémités est associée la liste des voisins à examiner. Pour élaborer le chemin de crête issu d'un point extrémité il suffit d'examiner 3 voisins en fonction de la topologie de l'extrémité.

L'image après détection de contours et fermeture comprend les pixels marqués à 1 (les contours) et d'autres pixels marqués à 0 qui représentent des

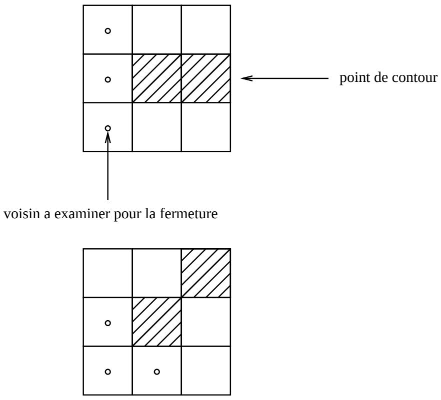

Figure 4.2. *Exemple de voisins à examiner pour la fermeture des contours, dans deux cas de configurations différentes.*

régions.

Le problème est donc d'affecter une même étiquette aux points marqués 0 appartenant à la même composante connexe de l'image. Cette opération est effectuée par simple balayage de l'image par un opérateur en L. Les classes d'équivalence sont mises à jour à chaque nouvelle application de l'opérateur.

4.5. Résultats expérimentaux

Nous présentons à titre indicatif des résultats obtenus avec l' algorithme de croissance optimale de régions. On remarque que les résultats obtenus avec ces deux types d'algorithmes (croissance de régions, fermeture de contours) sont comparables du moins *pour des images de scène d'intérieur*. On notera aussi que les contours des régions correspondent le plus souvent aux contours extraits par détection de contours. Les temps de calcul sur une station de travail SUN-3 sont pour une image 256×256 de l'ordre de 30 secondes CPU pour la croissance de régions.

Figure 4.3. *Image originale.*

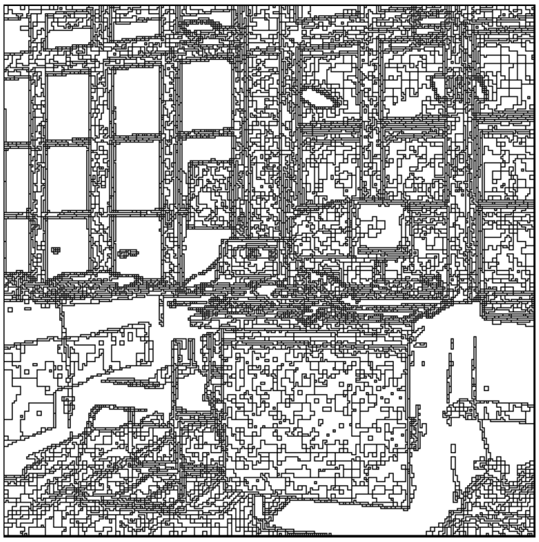

Figure 4.4. *Utilisation du critère max-min (croissance optimale de régions).*

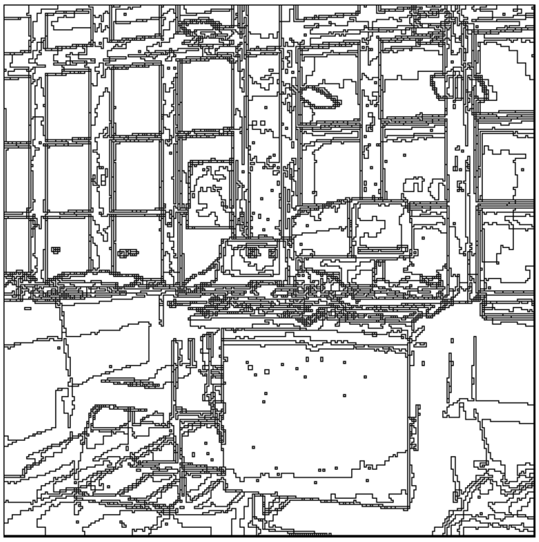

Figure 4.5. *Utilisation du critère moyenne (croissance optimale de régions).*

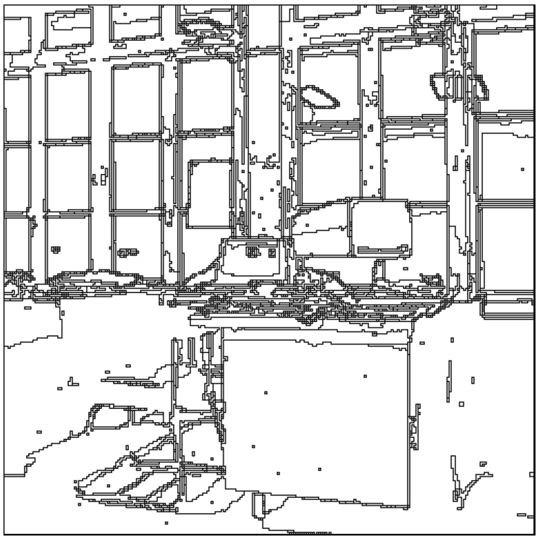

Figure 4.6. *Utilisation du critère gradient (croissance optimale de régions).*

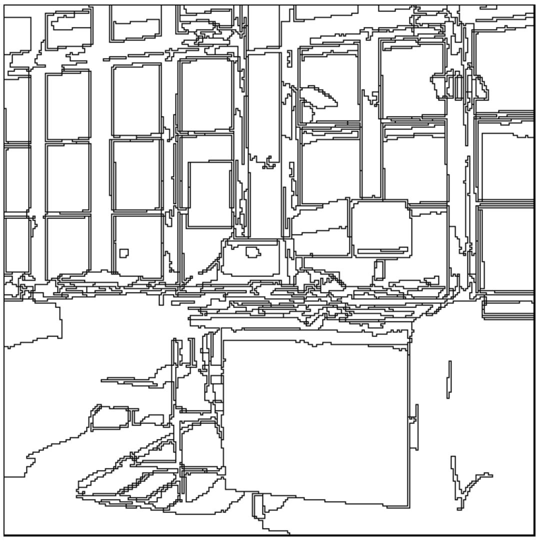

Figure 4.7. *Segmentation finale après élimination des petites régions (croissance optimale de régions).*

Figure 4.8. *Segmentation finale où les régions sont coloriées avec leurs moyennes de niveau de gris.*

Figure 4.9. *Image originale (coupe sagittale de tête obtenue par IRM).*

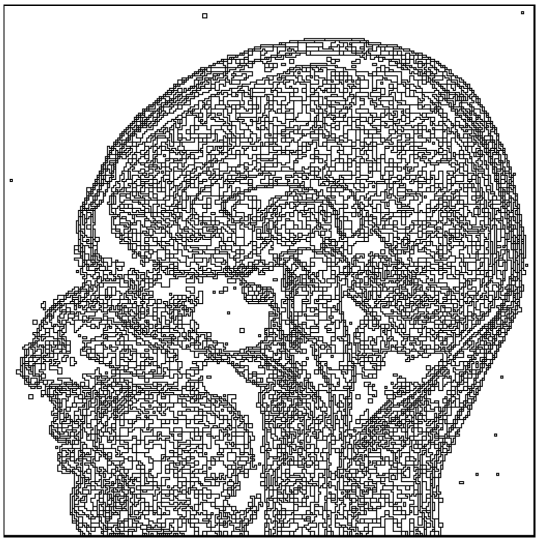

Figure 4.10. *Utilisation du critère max-min (croissance optimale de régions).*

Figure 4.11. *Utilisation du critère moyenne (croissance optimale de régions).*

Figure 4.12. *Utilisation du critère gradient (croissance optimale de régions).*

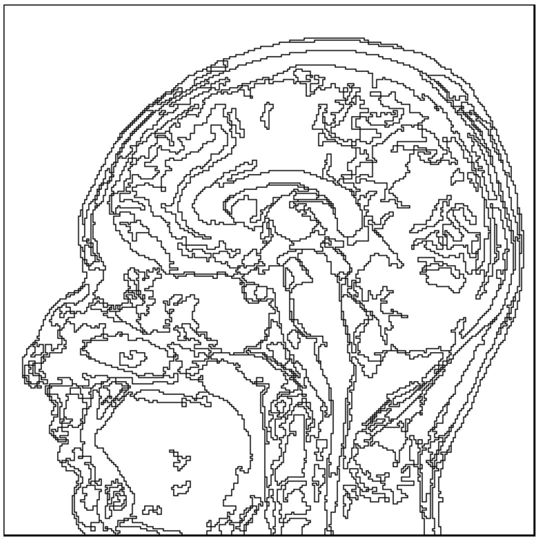

Figure 4.13. *Segmentation finale après élimination des petites régions (croissance optimale de régions).*

Figure 4.14. *Segmentation finale où les régions sont coloriées avec leurs moyennes de niveau de gris.*