

Chapitre 6

Segmentation en régions

6.1. Introduction

La segmentation en régions est un traitement intermédiaire qui, à partir de l'image originale ou de l'image filtrée, s'attache à reconstruire les régions qui correspondent aux objets présents dans l'image. La segmentation transforme l'image en une image où chaque pixel est caractérisé par un label correspondant à la région à laquelle il appartient. Une fois les régions reconstruites, l'étape d'analyse tente en général de les décrire à l'aide d'attributs qui permettent de reconnaître les objets contenus dans l'image.

Pour le cas spécifique des images couleur, on suppose généralement que les différentes couleurs présentes dans l'image correspondent principalement aux différentes orientations et propriétés de réflexion des surfaces des objets observés ainsi qu'à la présence d'ombres. Les méthodes de segmentation analysent ces couleurs afin de distinguer les différents objets qui composent une scène observée par une caméra couleur. Les différences entre les couleurs présentes dans l'image sont plus ou moins bien mises en évidence selon le système de représentation de la couleur utilisé [VAN 03].

Chapitre rédigé par Alain CLÉMENT (Université d'Angers) et Laurent BUSIN (Université du Littoral Côte d'Opale) et Olivier LEZORAY (Université de Caen Basse Normandie) et Ludovic MACAIRE (Université Lille 1).

Les méthodes de segmentation d'images couleur décrites dans ce chapitre supposent qu'une région est constituée de pixels connexes partageant des propriétés colorimétriques similaires ou formant une texture couleur spécifique. Les méthodes relatives à cette approche, dite approche régions, recherchent dans l'image des sous-ensembles de pixels connexes dont les couleurs sont homogènes ou forment une texture.

Nous n'allons pas présenter ici une liste exhaustive des méthodes de segmentation d'images couleur. Le lecteur trouvera une revue complète des méthodes de segmentation dans [TRÉ 04], [CHE 01], [PAL 07]. Nous allons plutôt tenter de mettre en évidence les concepts inhérents aux méthodes de segmentation les plus récentes.

L'information couleur étant multi-dimensionnelle, l'analyse d'images couleur requiert la manipulation de structures de données comme les histogrammes qui peuvent être coûteuses en espace mémoire. Le parcours de telles structures peut également nécessiter un nombre important d'opérations. Dans la première partie de ce chapitre, nous allons donc décrire une implémentation originale de l'histogramme couleur, appelée histogramme compact, qui permet de réduire l'espace mémoire occupé ainsi que la complexité algorithmique de son exploitation.

Les techniques de construction de régions sont basées sur l'analyse conjointe de la distribution des couleurs au sein de l'espace couleur et de la répartition des couleurs au sein du plan image. La deuxième partie de ce chapitre décrit des méthodes de classification spatio colorimétrique qui privilégient l'analyse dans l'espace couleur tandis que la troisième partie présente des techniques d'analyse de graphes qui modélisent l'interaction entre les pixels au sein du plan image.

Enfin, la question de l'évaluation de la qualité de la segmentation reste un problème ouvert. La dernière partie présente une des approches retenues par une grande partie de la communauté car basée sur la comparaison entre une segmentation automatique et plusieurs partitions effectuées par des observateurs humains.

6.2. Histogrammes compacts

L'histogramme d'une image est la représentation d'une fonction discrète qui à chaque valeur de l'image associe le nombre de pixels ayant cette valeur. Il peut être vu comme la densité de probabilité d'une variable aléatoire dont l'image constitue un ensemble de réalisations. Quand les valeurs de l'image sont des vecteurs tridimensionnels comme c'est le cas des images couleur, on parlera d'histogramme multidimensionnel.

6.2.1. Histogramme multidimensionnel classique

Dans sa représentation classique, l'histogramme H^c d'une image de N pixels à P composantes d'un espace couleur c codées chacune sur Q bits est constitué par un tableau P -dimensionnel comportant $2^{P \cdot Q}$ cellules. Afin de stocker les effectifs, chaque cellule doit pouvoir contenir un nombre entier positif au maximum égal à N , donc codé sur un nombre de bits E tel que

$$E > \log_2(N), \quad (6.1)$$

le codage informatique des nombres faisant en général appel à des multiples de 8 bits. L'histogramme multidimensionnel occupe alors $2^{P \cdot Q} E$ bits. Dans le cas d'une image couleur ($P = 3$) de taille 512×512 pixels et dont chaque composante est codée sur 256 niveaux ($Q = 8$), cette valeur représente un minimum de 36 Mo. Pour une image multi-composantes à 10 composantes de même résolution, l'histogramme multidimensionnel occupe plus de $2,5 \cdot 10^{18}$ Mo.

Devant la difficulté à manipuler un tel volume de données, il existe couramment deux stratégies :

- La première consiste à effectuer des projections de l'histogramme multidimensionnel sur un nombre d'axes p réduit ($p < P$). Dans une approche marginale ($p = 1$), les P histogrammes monodimensionnels n'occupent alors plus qu'un minimum de $2^Q \cdot P \cdot \log_2(N)$ bits. L'inconvénient de cette stratégie est de perdre tout ou partie de la corrélation existant entre les différentes composantes de l'image au risque de voir disparaître certains modes (zones de haute densité) significatifs de l'espace multidimensionnel d'origine.

- La seconde stratégie consiste à requantifier chaque composante de l'image sur q bits ($q < Q$) ce qui revient à effectuer une classification *a priori* de l'image. Les conséquences de cette pré-segmentation sont imprévisibles, et en contradiction avec les efforts déployés lors de l'utilisation de caméras à haute résolution colorimétrique ou spectrale.

6.2.2. Histogramme multidimensionnel compact

Soit C le nombre de cellules occupées dans l'histogramme multidimensionnel d'une image à P composantes codées chacune sur Q bits, c'est à dire le nombre de P -uplets (ou couleurs) de l'image. En pratique, le nombre C est toujours très inférieur à $2^{P \cdot Q}$ le nombre de cellules disponibles. À nombre de pixels constant, le nombre de cellules inoccupées augmente avec le nombre P de composantes de l'image. Le principe de l'histogramme multidimensionnel compact [CLé 01] consiste à ne coder que les C cellules effectivement occupées en triant ces cellules par ordre lexicographique sur les

P composantes de l'image. Ce codage utilise deux tableaux, le premier de dimension $C \times P$ pour stocker les P -uplets colorimétriques, et le second de dimension $C \times 1$ pour les effectifs correspondants. À titre d'illustration, le tableau 6.1 présente un extrait de l'histogramme compact d'une image couleur.

R	V	B	effectif
0	0	5	13
0	0	23	5
...
255	10	0	21
255	251	254	3

Tableau 6.1. Extrait de l'histogramme compact d'une image couleur RVB à 8 bits par composante.

Pour un nombre E de bits par cellule tel que défini dans l'équation (6.1), l'espace mémoire occupé par l'histogramme compact vaut $C.E + C.P.Q$ bits, soit une compression, par rapport au codage classique, d'un facteur

$$\frac{2^{P.Q}}{C.(\frac{P.Q}{E} + 1)}. \quad (6.2)$$

Le cas le plus défavorable pour l'histogramme compact serait que chaque pixel de l'image présente une valeur différente ($C = N$). Dans ce pire cas, pour une image couleur 24 bits de résolution 512×512 , le facteur de compression est supérieur à $27 : 1$.

Afin d'illustrer le gain d'espace mémoire apporté par l'histogramme compact, le tableau 6.2 présente quelques résultats comparatifs obtenus sous Matlab© en utilisant les types de données *Double* (64 bits) et *Uint8* (8 bits) pour le codage respectif des effectifs et des valeurs colorimétriques.

6.2.2.1. Algorithme

Le calcul de l'histogramme compact repose sur un tri lexicographique des niveaux des N pixels de l'image structurés en P sous-clés dans une relation d'ordre définie par l'ordre des P composantes de l'image. Nous préconisons l'algorithme de tri rapide par dichotomie qui est plus rapide que toutes les méthodes itératives pour $N > 20$ et plus rapide d'un facteur 1,5 à 2 que le tri par tas pour $N > 1000$ [PRE 92]. Il n'est pas dans notre objectif d'exposer ici le détail d'un algorithme de tri dont la mise en œuvre ne présente pas de difficulté particulière. Pour plus de précisions sur cet algorithme et sa complexité, le lecteur se référera à [CLé 01].


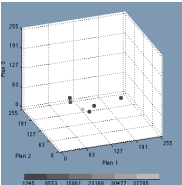
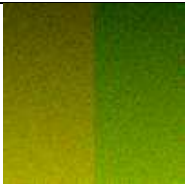
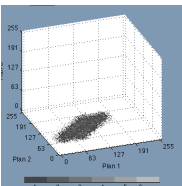

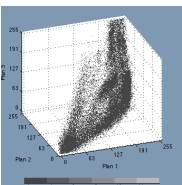
Image ($Q = 8$ bits)	Histogramme	Volume (Mo) histogramme classique	Volume (Mo) histogramme compact
 RVB 6 couleurs pures		128	$6,29.10^{-5}$
 RVB de synthèse		128	0,07
 RVB naturelle		128	0,32

Tableau 6.2. Comparaison des volumes mémoire des histogrammes multidimensionnels classique et compact.

6.2.2.2. Propriétés

Le principal atout de l'histogramme compact est de réduire, sans perte d'information, le volume mémoire occupé par un histogramme multidimensionnel. Le gain d'espace mémoire est d'autant plus élevé que le nombre de composantes colorimétriques est grand. Par construction, l'histogramme compact et l'histogramme classique présentent des différences structurelles qui engendrent des différences de coût algorithmique lors de l'accès aux données.

6.2.2.2.1. Parcours complet de l'histogramme

Le parcours complet de l'histogramme classique nécessite P boucles imbriquées de taille 2^Q et $2^{P \cdot Q}$ comparaisons pour identifier les C effectifs non nuls. Le parcours complet de l'histogramme compact nécessite seulement d'itérer C fois la lecture de $P + 1$ valeurs. L'histogramme compact permet donc de réduire de manière non négligeable la complexité algorithmique de toutes les opérations qui consistent à lister et compter les différentes valeurs d'une image, à savoir par exemple, la représentation

du graphe de l'histogramme, les calculs statistiques ou encore la construction d'une palette de couleurs.

6.2.2.2.2. Recherche de couleurs particulières

L'histogramme classique est un ensemble directement indexé par ses coordonnées colorimétriques. L'accès à un P -uplet particulier est immédiat et une simple comparaison suffit pour vérifier que l'effectif est non nul. L'histogramme compact est trié dans l'ordre lexicographique et la recherche d'un P -uplet particulier est semblable à celle d'un mot dans un dictionnaire. Les stratégies applicables sont celles établies pour rechercher des éléments dans une liste triée : recherche linéaire, recherche par dichotomie, recherche aléatoire, etc. L'histogramme compact sera donc nettement moins avantageux que l'histogramme classique pour ce type d'opération.

6.2.2.2.3. Recherche d'effectifs

Les effectifs ne sont directement adressables dans aucun des deux types d'histogrammes. La recherche d'effectifs s'apparente donc à un parcours complet favorisant l'histogramme compact.

6.2.2.3. Étiquetage en composantes connexes

L'étiquetage en composantes connexes d'un histogramme multidimensionnel consiste à effectuer un balayage de tous les P -uplets d'effectif non nul afin de réunir sous une même étiquette les P -uplets connexes dans l'espace colorimétrique. Cette étape est essentielle pour de nombreuses méthodes de classification de pixels par analyse d'histogrammes [TRÉ 04]. L'histogramme compact est une structure séquentielle triée dans laquelle un déplacement par itération n'est pas linéaire avec un déplacement dans l'espace géométrique engendré par les composantes colorimétriques. Il en résulte que les méthodes d'étiquetage classiques [COC 95] ne sont pas directement applicables. Nous présentons ici le principe de l'algorithme d'étiquetage compact développé par Ouattara et Clément [OUA 07].

Comme dans un cas classique, l'étiquetage de l'histogramme compact est effectué en un seul parcours grâce à une table d'équivalence des étiquettes. Dans ce cas, seuls les voisins prédécesseurs d'un P -uplet sont à considérer (voir tableau 6.3).

$(i-1, j-1)$	$(i-1, j)$	$(i-1, j+1)$
$(i, j-1)$	(i, j)	

Tableau 6.3. *Voisins en 2 dimensions du point de coordonnées (i, j) dans l'ordre de parcours lexicographique IJ des composantes.*

Le parcours séquentiel des C P -uplets se fait dans l'ordre lexicographique des composantes. La spécificité de l'histogramme compact est que les voisins d'un P -uplet ne

sont pas situés à des positions connues d'avance (voir paragraphe 6.2.2.2.2). Comme illustré par le tableau 6.4, ces positions précèdent néanmoins toujours le P -uplet dans l'ordre de parcours et par conséquent sont déjà étiquetées. Pour P composantes, le nombre théorique des prédécesseurs d'un P -uplet vaut $(3^P - 1)/2$. Dans la pratique, le nombre réel des prédécesseurs est limité par le nombre de P -uplets présents dans l'image.

Composante I	Composante J
0	0
...	...
i-1	j-1
i-1	j
i-1	j+1
...	...
i	j-1
i	j
...	...
$2^Q - 1$	$2^Q - 1$

Tableau 6.4. Coordonnées en 2 dimensions des voisins théoriques du point (i,j) dans l'histogramme compact IJ d'une image codée sur Q bits par composante.

L'étiquetage en composantes connexes de l'histogramme compact permet aux méthodes de classification de pixels d'opérer sans re-quantification ni sous-échantillonnage de l'espace multidimensionnel de l'image. Un exemple de mise en œuvre est décrit au paragraphe suivant (voir paragraphe 6.2.3). Le lecteur pourra aussi se référer à [TOR 10] pour un exemple d'utilisation des histogrammes multidimensionnels compacts pour la classification spatio colorimétrique d'images couleur abordée dans la suite de ce chapitre.

6.2.3. Classification de pixels par analyse d'histogrammes compacts

Ouattara et Clément [OUA 07] ont développé un processus non supervisé de classification de pixels par analyse d'histogrammes multidimensionnels compacts. Ce processus fournit une image étiquetée où à chaque pixel est associée une étiquette correspondant à la classe d'assignation. Les pixels connexes assignés à une même classe sont alors regroupés en une région par une analyse en composantes connexes de l'image étiquetée.

Initialement introduit dans [CLÉ 03], le principe d'analyse des modes d'histogrammes consiste à décomposer hiérarchiquement l'histogramme multidimensionnel de l'image en construisant une arborescence des modes représentatifs des classes. L'algorithme parcourt, dans l'ordre croissant, les différentes valeurs d'effectifs de

l'histogramme compact, en effectuant à chaque étape un étiquetage en composantes connexes de l'histogramme. Les modes de l'histogramme identifiés par les composantes connexes sont ainsi récursivement décomposés en sous-modes de moindre dispersion. Un mode sera considéré comme significatif s'il présente un effectif total supérieur à un seuil de population S dont le choix conditionne *in fine* le nombre des classes. Une fois construite l'arborescence des modes, la classification des valeurs colorimétriques non assignées aux modes significatifs (noyaux des classes) est effectuée par minimisation d'une distance colorimétrique aux centres de gravité de ces différents modes. Les régions de l'image segmentée sont ensuite constituées des zones de pixels connexes dont les couleurs appartiennent à une même classe. Le faible encombrement mémoire des histogrammes compacts autorise la méthode de segmentation à traiter vectoriellement un nombre P quelconque de composantes colorimétriques, sans requantification, ni sous-échantillonnage. Un exemple de classification est donné dans la figure 6.1 pour une image couleur naturelle. Les auteurs montrent que la méthode est capable de distinguer des classes colorimétriquement très proches et identifiées dans l'histogramme par des modes de faible dispersion. L'inconvénient avec des images texturées ou ne présentant que peu de régions de couleurs homogènes est une tendance à la sur-segmentation de l'image. Les distributions des deux régions verticales de l'image 6.2(a) (la composante bleue est fixée à 0) sont si proches dans le plan chromatique RV qu'elles ne peuvent pas être distinguées par une analyse de l'histogramme couleur représenté par la figure 6.2(b). La figure 6.2 montre que la seule analyse de la distribution des couleurs ne permet pas de retrouver les deux régions. Cet exemple illustre le besoin d'analyser conjointement la distribution des couleurs au sein d'un espace couleur et la répartition spatiale des couleurs au sein du plan image.



Figure 6.1. Classification des pixels de l'image "Lena" 6.1(a) pour différents nombres de classes. Les résultats sont représentés par les images d'étiquettes correspondant aux couleurs moyennes des classes.

6.3. Classification spatio colorimétrique

6.3.1. Introduction

Tenir compte conjointement de l'arrangement spatial des couleurs au sein du plan image et de la distribution des couleurs au sein de l'espace couleur peut permettre

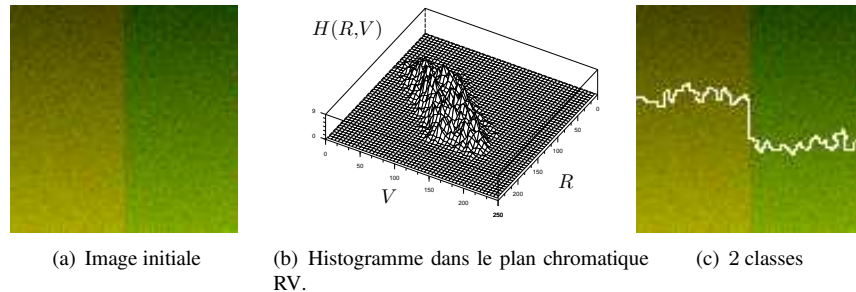


Figure 6.2. Classification des pixels de l'image de synthèse 6.2(a) par analyse de l'histogramme. Les résultats sont représentés par les contours des régions reconstituées à partir de l'image d'étiquettes en fausses couleurs. Une version en couleur de cette figure est disponible à la page 346.

de partitionner l'espace couleur de telle sorte que les classes de pixels correspondent effectivement aux régions de l'image. Nous divisons ces méthodes dites "spatio colorimétriques" en deux groupes, selon qu'elles analysent conjointement ou successivement les informations spatiales et colorimétriques décrivant l'image.

6.3.2. Analyse conjointe

Cette analyse associe les pixels à des points caractérisés par leur coordonnées spatiales et leurs composantes couleur et donc projetés dans un espace de dimension 5. Nous présentons la méthode "mean shift" qui s'appuie sur une estimation de la densité de probabilité de ces points dans cet espace, tandis que le degré de compacité spatio colorimétrique présenté ensuite, couple une mesure de dispersion des couleurs dans l'espace couleur avec une mesure de la connexité des pixels sur le plan spatial de l'image.

6.3.2.1. Mean shift

Comaniciu et Meer construisent une fonction spatio colorimétrique décrivant l'image, et proposent de détecter les modes en considérant conjointement les distributions spatiales et couleur. Ils appliquent l'algorithme "mean shift" dans cet espace spatio colorimétrique pour segmenter l'image [COM 02]. La méthode du "mean shift" s'appuie sur un noyau utilisé par l'estimateur de densité de Parzen, et d'un second noyau défini à partir de sa dérivée. En chaque point de l'espace couplant l'espace couleur avec le plan image, le "mean shift" est calculé comme la différence entre la moyenne pondérée des observations des voisins et le centre du noyau. Pour un ensemble de points, la procédure converge vers les maxima de densité sous-jacente sans estimer cette densité. Par ailleurs, la translation du noyau dans l'espace s'adapte

automatiquement à la densité locale des points : des densités basses provoquent un pas important de déplacement du “mean shift”.

Une image couleur peut être représentée par N points \mathbf{x}_i de dimension 5 représentant les N pixels de l'image dans le domaine spatial (\mathbf{x}_i^s : coordonnées spatiales) et couleur (\mathbf{x}_i^c : point couleur dans l'espace couleur c).

L'étape initiale affecte j à 1 et \mathbf{x}_i à $\mathbf{y}_{i,1}$. Ensuite, l'algorithme répète le calcul de $\mathbf{y}_{i,j+1}$ jusqu'à convergence par l'équation :

$$\mathbf{y}_{i,j+1} = \frac{\sum_{k=1}^N \mathbf{y}_{k,j} \cdot \frac{\delta}{h_s^2 h_c^3} g\left(\left\|\frac{\mathbf{y}_{k,j}^s - \mathbf{y}_{i,j}^s}{h_s}\right\|_2^2\right) g\left(\left\|\frac{\mathbf{y}_{k,j}^c - \mathbf{y}_{i,j}^c}{h_c}\right\|_2^2\right)}{\sum_{k=1}^N \frac{\delta}{h_s^2 h_c^3} g\left(\left\|\frac{\mathbf{y}_{k,j}^s - \mathbf{y}_{i,j}^s}{h_s}\right\|_2^2\right) g\left(\left\|\frac{\mathbf{y}_{k,j}^c - \mathbf{y}_{i,j}^c}{h_c}\right\|_2^2\right)} \quad (6.3)$$

où $\|\cdot\|_2$ est la norme \mathcal{L}_2 , g est une fonction noyau de type Gaussien, δ un facteur de normalisation, h_s et h_c définissant la taille des domaines spatiaux et couleur pris en compte pour l'estimation des densités de probabilité. Soient les points $\hat{\mathbf{y}}_i$ représentant les pixels après convergence, l'image de labels est composée des points $\mathbf{z}_i = (\mathbf{x}_i^s, \hat{\mathbf{y}}_i^c)$. Les points \mathbf{z}_i dont la distance dans le plan image est inférieure à h_s et dont la distance dans l'espace couleur est inférieure à h_c sont regroupés pour former les régions. Chaque région dont la surface est inférieure à un seuil est fusionnée avec la région adjacente dont la couleur moyenne est la plus proche de celle de la région considérée.

Losson *et al.* ont montré que cette méthode a tendance à fournir une sur-segmentation notamment quand les objets de la scène observée ne sont pas éclairés de manière uniforme [LOS 08]. L'image 6.4(a), qui montre la segmentation fournie par une procédure “mean shift”¹ appliquée à l'image de synthèse 6.2(a), illustre ce phénomène. Dans cette image, les contours des régions reconstruites sont indiqués en blanc. Les résultats sont généralement très sensibles à l'ajustement des paramètres h_s et h_c et aboutissent à une sur-segmentation des images. C'est la raison pour laquelle il est nécessaire de fusionner les régions adjacentes dont les couleurs moyennes sont proches par une des méthodes décrites dans le chapitre segmentation de [TRÉ 04].

Plusieurs méthodes de segmentation s'inspirent du “mean shift”, telle que celle de Paris *et al.* qui proposent de simplifier l'algorithme du “mean shift” en détectant les germes des régions par une simple recherche de maxima locaux de la densité de probabilité alors estimée [PAR 07]. Les régions ensuite reconstruites par “mean shift” étant en sur-nombre, elles sont fusionnées par une analyse itérative d'un graphe d'adjacence des régions, ce qui débouche sur une segmentation hiérarchique de l'image.

1. Cette image a été obtenue grâce à l'exécutable disponible à l'adresse <http://www.caip.rutgers.edu/riul/research/code/EDISON/>.

Par ailleurs, Shapira *et al.* proposent de remplacer la moyenne pondérée des observations des voisins utilisée pour calculer la différence avec le centre du noyau par une valeur médiane des observations [SHA 09]. Comme pour le filtrage médian des couleurs, la difficulté se concentre alors sur la définition d’une relation d’ordre entre ces observations multi-dimensionnelles.

La méthode “mean shift” juxtapose les informations spatiales et colorimétriques dans un même espace de dimension 5. Ces informations étant hétérogènes, il est nécessaire de séparer l’ajustement de la taille h_s du domaine spatial de celle h_c du domaine couleur pris en compte pour estimer les densité de probabilité. Néanmoins, il est difficile de définir des seuils sur une distance calculée dans cet espace qui intègre des attributs de nature hétérogène. C’est la raison pour laquelle Macaire *et al.* proposent un degré de compacité couleur qui mesure séparément la distribution des couleurs dans l’espace couleur et leur répartition spatiale sur le plan image pour construire des classes de pixels [LOS 08, MAC 06, BUS 05].

6.3.2.2. Analyse des degrés de compacité spatio colorimétrique

Dans l’espace couleur RVB, soit $\mathcal{D}_l(\mathbf{x}_i^c)$ un cube centré en $\mathbf{x}_i^c = [x_i^R, x_i^V, x_i^B]^T$ et dont les côtés de longueur l (un nombre impair) sont parallèles aux axes de cet espace. Le cube $\mathcal{D}_l(\mathbf{x}_i^c)$ contient toutes les couleurs $\mathbf{c}' = [c'^R, c'^V, c'^B]^T$ adjacentes à \mathbf{x}_i^c , de telle sorte que $x_i^k - \frac{l-1}{2} \leq c'^k \leq x_i^k + \frac{l-1}{2}$, $k = R, V, B$. Soit $\mathcal{S}_l(\mathbf{x}_i^c)$ le sous-ensemble constitué de tous les pixels \mathbf{x}_j dont la couleur \mathbf{x}_j^c appartient au cube couleur $\mathcal{D}_l(\mathbf{x}_i^c)$.

Le degré de connexité $CD_l(\mathbf{x}_i^c)$ est défini comme le nombre moyen de voisins de chaque pixel de $\mathcal{S}_l(\mathbf{x}_i^c)$ qui appartiennent également à $\mathcal{S}_l(\mathbf{x}_i^c)$ [BUS 05]. Quand le sous-ensemble $\mathcal{S}_l(\mathbf{x}_i^c)$ est vide, $CD_l(\mathbf{x}_i^c)$ est fixé à 0. Sinon, en considérant un voisinage spatial de $nb \times nb$ pixels (nb étant un nombre impair fixé à 3 ici), $CD_l(\mathbf{x}_i^c)$ est défini comme :

$$CD_l(\mathbf{x}_i^c) = \frac{\sum_{\mathbf{x}_j \in \mathcal{S}_l(\mathbf{x}_i^c)} \text{Card}\{N[\mathcal{S}_l(\mathbf{x}_i^c)](\mathbf{x}_j)\}}{(nb^2 - 1) \cdot \text{Card}\{\mathcal{S}_l(\mathbf{x}_i^c)\}}, \quad (6.4)$$

où $N[\mathcal{S}_l(\mathbf{x}_i^c)](\mathbf{x}_j)$ représente les voisins du pixel $\mathbf{x}_j \in \mathcal{S}_l(\mathbf{x}_i^c)$, parmi ses $(nb^2 - 1)$ voisins, qui appartiennent également à $\mathcal{S}_l(\mathbf{x}_i^c)$. Ce degré $CD_l(\mathbf{x}_i^c)$ varie entre 0 et 1 ; une valeur proche de 0 indique que les pixels de $\mathcal{S}_l(\mathbf{x}_i^c)$ sont dispersés sur le plan image tandis qu’une valeur proche de 1 signifie que la plupart des pixels de ce sous-ensemble sont connexes sur l’image.

Le degré d’homogénéité $HD_l(\mathbf{x}_i^c)$ est défini comme le ratio entre la dispersion locale moyenne des couleurs dans le voisinage de chaque pixel appartenant à $\mathcal{S}_l(\mathbf{x}_i^c)$, et la dispersion globale couleur des pixels de $\mathcal{S}_l(\mathbf{x}_i^c)$ [MAC 06] :

$$HD_l(\mathbf{x}_i^c) = \frac{\sigma_{local}(S_l(\mathbf{x}_i^c))}{\sigma(S_l(\mathbf{x}_i^c))} \quad (6.5)$$

avec

$$\sigma(S) = \frac{1}{\text{Card}\{S\}} \cdot \sqrt{\sum_{\mathbf{x}_j \in S} \|\mathbf{x}_j^c - \mathbf{m}_S^c\|_2^2}, \quad (6.6)$$

et

$$\sigma_{local}(S) = \frac{1}{\text{Card}\{S\}} \cdot \sum_{\mathbf{x}_j \in S} \sigma(N[S](x_j)). \quad (6.7)$$

où \mathbf{m}_S^c est le barycentre des couleurs représentant les pixels qui appartiennent à S .

Quand les couleurs des pixels appartenant à $S_l(\mathbf{x}_i^c)$ donnent naissance à un nuage compact dans l'espace couleur, le degré d'homogénéité estimé en \mathbf{x}_i^c est proche de 1. Au contraire, quand ces couleurs forment plusieurs nuages distincts, le degré d'homogénéité est proche de 0.

La fonction de compacité spatio colorimétrique $sccf_l$ est une fonction dont la valeur en chaque couleur \mathbf{x}_i^c est définie comme le produit entre les degrés de connexité et d'homogénéité [LOS 08] :

$$sccf_l(\mathbf{x}_i^c) = CD_l(\mathbf{x}_i^c) \cdot HD_l(\mathbf{x}_i^c), \quad (6.8)$$

Une valeur élevée de $sccf_l(\mathbf{x}_i^c)$ indique que les pixels du sous-ensemble $S_l(\mathbf{x}_i^c)$ sont fortement connexes dans l'image ($CD_l(\mathbf{x}_i^c)$ proche de 1) et que leurs couleurs sont concentrées dans l'espace couleur ($HD_l(\mathbf{x}_i^c)$ proche de 1). À l'opposé, une faible valeur de $sccf_l(\mathbf{x}_i^c)$ signifie que les pixels appartenant à $S_l(\mathbf{x}_i^c)$ sont dispersés dans l'image ($CD_l(\mathbf{x}_i^c)$ proche de 0) et/ou que leurs couleurs ne forment pas un nuage compact dans l'espace couleur ($HD_l(\mathbf{x}_i^c)$ proche de 0). La longueur l du cube couleur doit être fixée avec soin par l'utilisateur en fonction du contenu de l'image. La complexité algorithmique de construction de $sccf_l$ atteint $N.l^3.C$ où C est le nombre de couleurs différentes présentes dans l'image de taille N pixels.

La figure 6.3(a) montre la fonction $sccf_7$ de l'image 6.2(a) représentant deux régions dont les distributions couleur sont mélangées dans le plan chromatique RV. La longueur l est fixée à 7 afin que les sous-ensembles de pixels $S_l(\mathbf{x}_i^c)$ soient suffisamment peuplés. Examinons les trois couleurs différentes \mathbf{c}_1 , \mathbf{c}_2 et \mathbf{c}_3 mises en

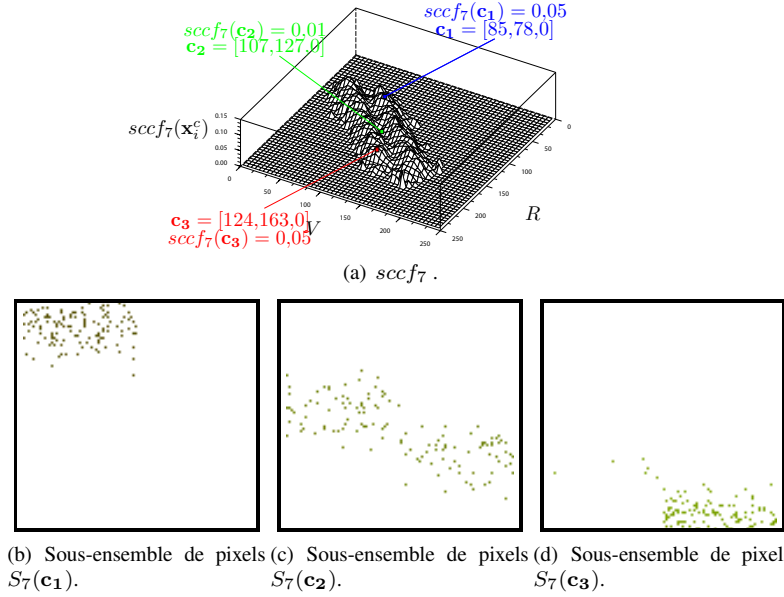


Figure 6.3. Fonction de compacité spatio colorimétrique de l'image 6.2(a), avec trois couleurs et leurs sous-ensembles de pixels correspondants.

évidence par des symboles couleur dans la figure. Nous pouvons voir que les valeurs de $sccf_7$ sont élevées pour les couleurs c_1 et c_3 , puisque la plupart des pixels des sous-ensembles $S_7(c_1)$ et $S_7(c_3)$ sont connexes dans l'image et que leurs couleurs sont proches les unes des autres (voir figures 6.3(b) et 6.3(d)). Notons que ces deux couleurs sont situées sur des modes à détecter. À l'opposé, la valeur de $sccf_7$ est faible pour la couleur c_2 située dans une vallée de l'histogramme car les pixels du sous-ensemble $S_7(c_2)$ sont dispersés dans l'image (voir figure 6.3(c)). Cet exemple illustre que le degré de compacité spatio colorimétrique met évidence les deux modes qui n'auraient pas été distinguables par une analyse de la simple distribution des couleurs (voir figure 6.2).

Comme pour l'histogramme compact, il reste ensuite à détecter les modes de la fonction de compacité spatio colorimétrique afin de partitionner l'espace couleur en différentes classes. Les pixels connexes dont les couleurs appartiennent à un même mode forment les régions retrouvées dans l'image segmentée. L'image 6.4(b) montre un exemple de segmentation de l'image de synthèse 6.2(a) après détection des modes par des outils morphologiques flous décrits dans [LOS 08]. Cette image montre que cette analyse du degré de compacité spatio colorimétrique permet de retrouver en grande partie les deux régions dont les distributions couleur sont mélangées.

La qualité de la segmentation dépend fortement de la fonction utilisée pour représenter conjointement la distribution des couleurs au sein de l'espace couleur et sa répartition spatiale dans le plan image. La méthode "mean shift" ainsi que l'analyse des degrés de compacité spatio colorimétriques supposent que les modes de cette fonction correspondent aux régions à retrouver dans l'image. Ces approches atteignent leurs limites quand la scène observée est si complexe qu'il est difficile d'associer des modes distincts aux régions à retrouver. C'est la raison pour laquelle certains auteurs procèdent par une analyse successive dans l'espace couleur et dans le plan image.

6.3.3. Analyse successive

JSEG est l'une des méthodes de segmentation les plus connues qui effectue une étape de classification dans l'espace couleur suivie d'une analyse spatiale [DEN 01]. Une première étape quantifie de manière non-uniforme les couleurs présentes dans l'image en N_c couleurs pour fournir une image appelée 'class-map' où aux pixels représentés par les points bi-dimensionnels \mathbf{x}_i^s est assigné un label scalaire \mathbf{x}_i^l , $\mathbf{x}_i^l = 1, \dots, N_c$. Soit ω_k le sous-ensemble des pixels \mathbf{x}_i dont le label \mathbf{x}_i^l est égal à k .

Considérons que l'image est partitionnée en n^A régions \mathcal{R}_n^A , ($n = 1, 2, \dots, n^A$). Pour chaque région \mathcal{R}_n^A , est évalué le critère J_n :

$$J_n = \frac{\sum_{\mathbf{x}_i^s \in \mathcal{R}_n^A} \sum_{k=1}^{N_c} \left\| \mathbf{x}_i^s - \mathbf{m}_{k,n}^s \right\|_2^2}{\sum_{\mathbf{x}_i^s \in \mathcal{R}_n^A} \left\| \mathbf{x}_i^s - \mathbf{m}_n^s \right\|_2^2} \quad (6.9)$$

où

$$\mathbf{m}_{k,n}^s = \frac{1}{\text{Card}\{\mathcal{R}_n^A \cap \omega_k\}} \cdot \sum_{\mathbf{x}_i^s \in \mathcal{R}_n^A \cap \omega_k} \mathbf{x}_i^s \quad (6.10)$$

et

$$\mathbf{m}_n^s = \frac{1}{\text{Card}\{\mathcal{R}_n^A\}} \cdot \sum_{\mathbf{x}_i^s \in \mathcal{R}_n^A} \mathbf{x}_i^s. \quad (6.11)$$

Le critère J_n mesure le rapport entre la dispersion spatiale des pixels de \mathcal{R}_n^A assignés à un même label et la dispersion spatiale des pixels de toute la région. Deng *et al.* ont montré par des exemples que plus ce rapport est bas, plus les pixels assignés à des labels identiques sont connexes dans la région [DEN 01].

Une mesure \bar{J} globale sur l'image tient compte de l'ensemble des régions par :

$$\bar{J} = \frac{1}{n^A} \sum_{n=1}^{n^A} \text{Card}\{\mathcal{R}_n^A\} J_n. \quad (6.12)$$

La minimisation de ce critère permet d'obtenir une partition où dans chacune des régions, les pixels assignés aux mêmes labels sont les plus connexes possible. En utilisant le critère \overline{J} , une procédure de croissance de régions suivie d'une étape de fusion de régions est appliquée à la 'class-map' afin d'éviter une sur-segmentation.

Wang *et al.* montrent que la classification grossière fournie par l'étape de quantification des couleurs dégrade la flexibilité du critère JSEG [WAN 06b]. Ainsi, les régions voisines dont la transition couleur est douce ont tendance à être séparées en plusieurs classes distinctes. Ce phénomène est illustré par l'image 6.4(c) résultant du programme disponible sur le site web², et configuré avec les valeurs par défaut des paramètres.

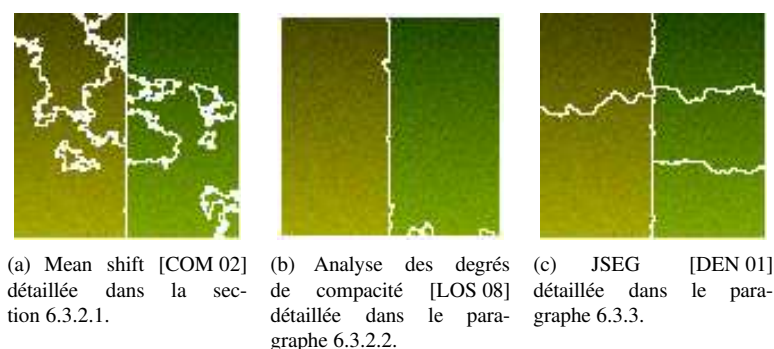


Figure 6.4. Résultats de segmentation de l'image 6.2(a) fournis par différentes méthodes de classification spatio colorimétrique. Les résultats sont représentés par les contours des régions reconstruites.

Wang *et al.* montrent, sur un exemple simple, que le critère \overline{J} appliqué à la 'class-map', ne permet pas de fournir une information sur le niveau des transitions couleur entre deux régions [WAN 06c]. Par conséquent, il ne permet pas de distinguer des régions avec des textures similaires mais dont les distributions couleur sont différentes. JSEG a donc tendance à sur-segmenter les images en cas d'éclairage non uniforme. Ceci est probablement lié au fait que les informations couleur et spatiale sont prises en considération séparément. Wang *et al.* proposent alors de combiner le critère \overline{J} basé sur une mesure locale d'homogénéité des textures avec une mesure de discontinuité couleur H utilisée par la méthode HSEG [JIN 03].

2. <http://vision.ece.ucsb.edu/segmentation/jseg/>

6.3.4. Conclusion

Pour conclure sur cette partie, nous proposons d'étudier les temps de calcul nécessaires pour segmenter une image de taille 1000x100 pixels pour les trois approches présentées qui ont été implémentées sur un PC dont le micro-processeur est cadencé à 2 GHz. Le temps requis par l'approche "mean shift" est le plus faible à savoir 5 s, suivi de celui consommé par JSEG qui est de 8s. Le temps requis pour calculer les degrés de compacité spatio-couleur est plus important car atteignant environ 60 secondes. Ce surcoût pourrait être réduit par une implémentation similaire à celle de l'histogramme compact. Les résultats issus de ces trois méthodes seront présentés et comparés à la fin de ce chapitre.

Les méthodes dites de classification spatio colorimétrique sont généralement bien adaptées aux images contenant des régions homogènes en couleur, l'analyse spatiale permettant surtout de distinguer des régions dont les distributions couleur sont proches ou se mélangent dans l'espace couleur. Elles fournissent une première segmentation qui doit être complétée par l'analyse d'un graphe formalisant les relations spatiales entre pixels adjacents, décrite au paragraphe suivant.

6.4. Segmentation par analyse de graphes

Un graphe est une structure tout naturellement adaptée à la représentation des données au sein des images car il modélise les propriétés de voisinage entre ces données. Dans cette partie, nous allons donc décrire comment formaliser une image par un graphe et les opérations à y appliquer pour segmenter une image couleur. Après avoir décrit la structure d'un graphe et les opérations à y appliquer, nous détaillerons leur exploitation pour une segmentation dite semi supervisée, à savoir à partir de marqueurs fournis par un utilisateur. Les arêtes d'un graphe peuvent être pondérées à partir de la similarité entre données adjacentes. Nous verrons ensuite que ces similarités peuvent être regroupées au sein d'une matrice dont l'analyse spectrale permet de couper le graphe afin de classifier les données de l'image.

6.4.1. Graphes et images couleur

6.4.1.1. Généralités sur les graphes

En traitement d'images, on utilise habituellement le support de l'image comme une grille où les relations de connexité spatiale entre les pixels déterminent le voisinage considéré. En conséquence, sur une grille 2D, le graphe de voisinage associé à une image correspond par exemple à l'utilisation d'une 4 ou d'une 8 connexité au niveau pixel (voir figures 6.5(a) et 6.5(b)). On appelle ceci un graphe grille où les nœuds sont les pixels et les arêtes les relations de connexité entre pixels. Une grille est une tessellation régulière dont les éléments sont des carrés, mais il existe d'autres tessellations

constituées de polygones dont la plus célèbre est la tessellation de Voronoï. A cette tessellation est associée également un graphe de voisinage nommé le graphe de Delaunay. La figure 6.5(c) présente une partition de Voronoï avec son graphe de Delaunay superposé. Plus spécifiquement, lorsque que l'on segmente une image en régions, on peut associer à cette partition un graphe de voisinage nommé le graphe d'adjacence de régions où chaque nœud représente une région et chaque arête représente une relation de voisinage entre ces régions. La figure 6.5(d) présente une partitions en régions d'une image avec son graphe de régions superposé. Finalement, un graphe peut être associé à n'importe quelle représentation de données d'images pour peu que l'on se munisse d'une mesure de distance ou de similarité. Dans ce cadre, traiter des données images revient à traiter des graphes dont les propriétés attachées aux nœuds et aux arêtes dépendent du problème considéré.

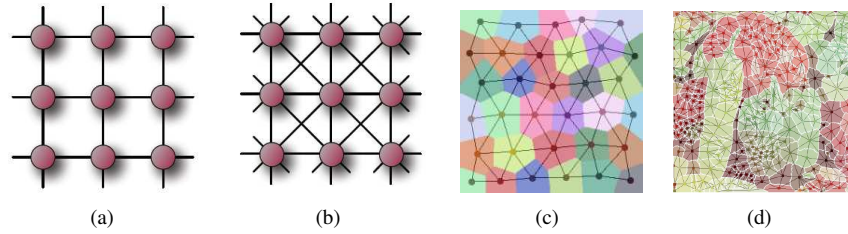


Figure 6.5. Exemples de graphes de type grille en 4 (a) ou 8 voisinage (b), de Delaunay (c) ou d'adjacence de régions (d).

6.4.1.2. Éléments de base sur les graphes

Un graphe est donc une structure de données utilisée pour décrire un ensemble d'objets et les relations entre ceux-ci. Les objets sont des *nœuds* du graphe et les relations des *arêtes*. Un $G = (\mathcal{V}, \mathcal{E}, w)$ est constitué d'un ensemble $\mathcal{V} = \{v_1, \dots, v_N\}$ de N nœuds et d'un ensemble $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ d'arêtes pondérées. Un graphe G est supposé non dirigé, sans boucles ni arêtes multiples [DIE 05]. Soit (i, j) une arête de \mathcal{E} qui connecte deux nœuds i et j de \mathcal{V} . Son poids, noté $w(i, j)$ représente la similarité entre ces deux nœuds. Ces similarités sont obtenues par le biais d'une fonction symétrique et positive $w : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}_+$ satisfaisant $w(i, j) = 0$ si $(i, j) \notin \mathcal{E}$.

À un graphe G , on associe les espaces de fonctions sur arêtes et nœuds. Soit $\mathcal{H}(\mathcal{V})$ l'espace de Hilbert des fonctions à valeurs réelles définies sur les nœuds d'un graphe $G = (\mathcal{V}, \mathcal{E}, w)$. Une fonction $f : \mathcal{V} \rightarrow \mathbb{R}^3$ de $\mathcal{H}(\mathcal{V})$ assigne une couleur $f(j) = \mathbf{x}_j^c = (f_k(j))_{k=1,2,3}^T$ à chaque nœud $j \in \mathcal{V}$. L'espace $\mathcal{H}(\mathcal{V})$ est muni du produit scalaire usuel $\langle f, h \rangle_{\mathcal{H}(\mathcal{V})} = \sum_{j \in \mathcal{V}} f(j)h(j)$, où $f, h : \mathcal{V} \rightarrow \mathbb{R}^3$. De manière similaire, soit $\mathcal{H}(\mathcal{E})$ l'espace de Hilbert des fonctions à valeurs réelles définies sur les arêtes de G . Il est également muni du produit scalaire $\langle F, H \rangle_{\mathcal{H}(\mathcal{E})} = \sum_{i \in \mathcal{V}} \sum_{j \sim i} F(i, j)H(i, j)$, où

$F, H : \mathcal{E} \rightarrow \mathbb{R}^3$ sont deux fonctions de $\mathcal{H}(\mathcal{E})$. On note $i \sim j$, le fait que deux nœuds i et j sont adjacents. Enfin, habituellement dans un graphe pondéré, les poids suivants sont considérés : $w(i, j) = \exp(-\|f(i) - f(j)\|_2^2 / 2\sigma^2)$ avec σ un paramètre d'échelle et $\|\cdot\|_2$ la norme \mathcal{L}_2 . L'ensemble des poids peut être rassemblé au sein d'une matrice de similarité w de dimension $N \times N$.

6.4.1.3. Opérateurs de différences sur graphes

Nous présentons ici un formalisme basé sur des équations aux différences partielles sur graphes pour la segmentation d'images couleur³. Un ensemble d'opérateurs de différences pondérées sur graphes a été introduit dans [ELM 08b, ELM 08a, LEZ 07, TA 09]. Soit $G = (V, \mathcal{E}, w)$ un graphe pondéré, et soit $f : V \rightarrow \mathbb{R}^3$ une fonction de $\mathcal{H}(\mathcal{V})$. L'opérateur de différence de f , noté $d : \mathcal{H}(\mathcal{V}) \rightarrow \mathcal{H}(\mathcal{E})$, est défini sur une arête $(i, j) \in \mathcal{E}$ par :

$$(df)(i, j) = \sqrt{w(i, j)}(f(j) - f(i)). \quad (6.13)$$

L'opérateur adjoint $d^* : \mathcal{H}(\mathcal{E}) \rightarrow \mathcal{H}(\mathcal{V})$ de l'opérateur de différence est défini, pour toutes les fonctions $f \in \mathcal{H}(\mathcal{V})$ et $H \in \mathcal{H}(\mathcal{E})$, par :

$$\langle df, H \rangle_{\mathcal{H}(\mathcal{E})} = \langle f, d^* H \rangle_{\mathcal{H}(\mathcal{V})}, \quad (6.14)$$

En utilisant les définitions des produits scalaires, l'expression de l'opérateur adjoint d'une fonction $H \in \mathcal{H}(\mathcal{E})$ peut s'exprimer en un nœud $i \in \mathcal{V}$ par :

$$(d^* H)(i) = \sum_{j \sim i} \sqrt{w(i, j)}(H(j, i) - H(i, j)). \quad (6.15)$$

L'opérateur de divergence d'une fonction $H \in \mathcal{H}(\mathcal{E})$ est défini par $-d^*$. L'opérateur gradient pondéré $\nabla_w : \mathcal{H}(\mathcal{V}) \rightarrow \mathbb{R}^{3 \times \text{Card}\{\mathcal{E}\}}$ d'une fonction $f \in \mathcal{H}(\mathcal{V})$ en un nœud $i \in \mathcal{V}$ est la matrice de toutes les différences pour l'ensemble des arêtes de \mathcal{E} et pour toutes les composantes couleur : $\nabla_w f(i) = [(df)(i, j)]_{(i, j) \in \mathcal{E}}^T$. La norme \mathcal{L}_2 de ce vecteur représente la *variation locale* de la fonction f en un nœud du graphe et est définie par :

$$\|\nabla_w f(i)\|_2 = \sqrt{\sum_{j \sim i} ((df)(i, j))^2} \stackrel{(6.13)}{=} \sqrt{\sum_{k=1}^3 \sum_{j \sim i} w(i, j)(f_k(j) - f_k(i))^2}. \quad (6.16)$$

Le p -Laplacien pondéré, $p \in \mathbb{R}^+$, d'une fonction $f \in \mathcal{H}(\mathcal{V})$, noté $\Delta_w^p : \mathcal{H}(\mathcal{V}) \rightarrow \mathcal{H}(\mathcal{V})$, est défini par :

$$\Delta_w^p f(i) = \frac{1}{2} d^* (\|\nabla_w f(i)\|_2^{p-2} df(i, j))(i). \quad (6.17)$$

3. En collaboration avec Abderrahim Elmoataz et Vinh-Thong Ta.

Le p -Laplacien pondéré de $f \in \mathcal{H}(\mathcal{V})$, en un nœud $i \in \mathcal{V}$, peut être obtenu par :

$$\Delta_w^p f(i) = \frac{1}{2} \sum_{j \sim i} \gamma_w^f(i, j) (f(i) - f(j)), \quad (6.18)$$

avec

$$\gamma_w^f(i, j) = w(i, j) (\|\nabla_w f(j)\|_2^{p-2} + \|\nabla_w f(i)\|_2^{p-2}). \quad (6.19)$$

Le p -Laplacien pondéré est un opérateur non-linéaire sauf pour $p = 2$ (il correspond dans ce cas au Laplacien classique [CHU 97]). En pratique pour éviter, une division par zéro dans l'équation (6.19) quand $p \leq 1$, la variation locale dans l'équation (6.16) est remplacée par une version régularisée : $\|\nabla_w f(j)\|_2 \approx \sqrt{\|\nabla_w f(j)\|_2^2 + \epsilon^2}$, où $\epsilon \rightarrow 0$ est une valeur constante faible [CHA 01].

6.4.1.4. Régularisation discrète sur graphes

Soit $f^0 : \mathcal{V} \rightarrow \mathbb{R}^3$ une fonction définie sur les nœuds d'un graphe pondéré $G = (\mathcal{V}, \mathcal{E}, w)$. On considère que la fonction f^0 représente une mesure d'une fonction initiale $g : \mathcal{V} \rightarrow \mathbb{R}^3$ corrompue par un bruit additif n (de moyenne nulle et de variance σ^2) tel que $f^0 = g + n$. Afin de retrouver la fonction originale g , une méthode usuelle est de chercher une fonction $f : \mathcal{V} \rightarrow \mathbb{R}^3$ qui soit suffisamment régulière sur G mais également proche de f^0 . Ce problème inverse peut être formalisé par la minimisation d'une fonctionnelle qui contient un terme de régularisation et un terme d'approximation. Cela amène à considérer le problème variationnel suivant :

$$g \approx \min_{f: \mathcal{V} \rightarrow \mathbb{R}^3} \{E_w^p(f, f^0, \lambda) = R_w^p(f) + \frac{\lambda}{2} \|f - f^0\|_2^2\}, \quad (6.20)$$

où la fonctionnelle de régularisation $R_w^p : \mathcal{H}(\mathcal{V}) \rightarrow \mathbb{R}$ est la forme discrète de p -Dirichlet de la fonction $f \in \mathcal{H}(\mathcal{V})$:

$$R_w^p(f) = \frac{1}{p} \sum_{i \in \mathcal{V}} \|\nabla_w f(i)\|_2^p \stackrel{(6.16)}{=} \frac{1}{p} \sum_{i \in \mathcal{V}} \left(\sum_{k=1}^3 \sum_{j \sim i} w(i, j) (f_k(j) - f_k(i))^2 \right)^{\frac{p}{2}}. \quad (6.21)$$

Le second terme d'attache aux données est pondéré par un paramètre de fidélité $\lambda \geq 0$ qui spécifie le compromis entre les deux termes de régularisation et d'approximation. Les deux termes de $E_w^p(f, f^0, \lambda)$ dans l'équation (6.20) sont des fonctions strictement convexes [CHA 01]. Ce problème d'optimisation admet une unique solution pour $p = 1$ et $p = 2$ qui satisfait pour tout $i \in \mathcal{V}$:

$$\frac{\partial}{\partial f(i)} E_w^p(f, f^0, \lambda) = \Delta_w^p f(i) + \lambda(f(i) - f^0(i)) = 0, \quad (6.22)$$

en utilisant le fait que $\frac{\partial R_w^p(f)}{\partial f(i)} = 2\Delta_w^p f(i)$, $\forall i \in \mathcal{V}$ (voir [ELM 08a] pour une preuve). L'équation (6.22) peut être interprétée comme un analogue discret sur graphes de

l'équation d'Euler-Lagrange. En utilisant la formulation du p -Laplacien de l'équation (6.18) dans l'équation (6.22), la solution du problème d'optimisation est également la solution du système d'équations suivant. Pour tout $i \in \mathcal{V}$,

$$\left(\lambda + \sum_{j \sim i} \gamma_w^f(i, j) \right) f(i) - \sum_{j \sim i} \gamma_w^f(i, j) f(j) = \lambda f^0(i). \quad (6.23)$$

Afin d'approximer la solution de la minimisation de l'équation (6.20), le système d'équations est linéarisé par la méthode de Gauss-Jacobi pour obtenir l'algorithme itératif suivant pour chaque composante [ELM 08b, ELM 08a, LEZ 07] :

$$\begin{cases} f_k^{(0)} = f_k^0 \\ f_k^{(t+1)}(i) = \frac{\lambda f_k^0(i) + \sum_{j \sim i} \gamma_w^{f_k^{(t)}}(i, j) f_k^{(t)}(j)}{\lambda + \sum_{j \sim i} \gamma_w^{f_k^{(t)}}(i, j)}, \forall i \in \mathcal{V}. \end{cases} \quad (6.24)$$

où $\gamma_w^{f_k^{(t)}}(i, j)$ est la fonction γ (dans l'équation (6.19)). Dans ce processus, les régularisations sont effectuées sur chaque composante séparément. Ceci n'est pas adapté au traitement d'images couleur et la géométrie locale doit prendre en compte des informations vectorielles [TSC 05]. Pour pallier ceci, la norme de l'opérateur gradient étant définie pour des vecteurs couleur, cela permet de coupler les régularisations et de prendre en compte l'aspect vectoriel de la couleur (lorsque $p = 1$). Lorsque $p = 2$, le gradient couleur n'intervient pas, mais la fonction de poids w sert également de terme de couplage permettant ainsi de prendre à nouveau en compte l'aspect vectoriel des données à traiter. Enfin, ce formalisme peut être utilisé directement pour le débruitage et la simplification d'images couleur, avec l'avantage de permettre naturellement l'intégration d'informations non locales exploitant des patches [BUA 05] (le lecteur intéressé pourra consulter [ELM 08b, ELM 08a, LEZ 07]).

6.4.2. Classification semi supervisée à partir d'un graphe

De nombreux schémas de segmentation automatique ont été proposés [COM 02, DEN 01, MAC 06] dans la littérature. Récemment, les approches semi automatiques ont connu un regain d'intérêt. Celles-ci reformulent le problème de segmentation en un problème d'apprentissage semi supervisé [ZHO 05, BEL 06] basé sur des stratégies de propagation de labels [WAN 06a, GRA 06, SIN 07]. Le précédent formalisme de régularisation sur graphes peut être adapté à la segmentation d'images couleur comme une méthode de classification semi supervisée. Pour permettre cette adaptation, le problème doit être réécrit comme [TA 09] :

– Soit $\mathcal{V} = \{v_1, \dots, v_N\}$ un ensemble de nœuds auxquels sont associés des vecteurs couleur ;

– Soit $G = (\mathcal{V}, \mathcal{E}, w)$ un graphe pondéré associé aux nœuds de \mathcal{V} et qui sont connectées par des arêtes de \mathcal{E} .

La classification semi supervisée de \mathcal{V} consiste à grouper l'ensemble \mathcal{V} en N_ω classes, le nombre de classes étant supposé connu à l'avance. Le but est alors d'estimer les labels des données non labellisées à partir des données labellisées.

Soit $\Omega = \{\omega_m\}_{m=1, \dots, N_\omega}$ l'ensemble initial des nœuds labellisés et soit $\mathcal{V} \setminus \Omega$ l'ensemble initial des nœuds non labellisés. L'appartenance d'un nœud à une classe est modélisé par N_ω fonctions de labels initiales (une par classe) $f_m^0 : \mathcal{V} \rightarrow \mathbb{R}$; avec $m = 1, \dots, N_\omega$. Pour un nœud $i \in \mathcal{V}$ donné, si i est initialement labellisé, alors

$$f_m^0(i) = \begin{cases} +1, & \text{si } i \in \omega_m \\ -1, & \text{sinon.} \end{cases} \quad (6.25)$$

Si i est initialement non labellisé (*i.e.* $i \in \mathcal{V} \setminus \Omega$) alors $f_m^0(i) = 0$. La classification semi supervisée est obtenue en effectuant en parallèle N_ω processus de régularisation afin d'estimer les fonctions $f_m : \mathcal{V} \rightarrow \mathbb{R}$ pour chaque classe ω_m . En utilisant le précédent formalisme de régularisation discrète, ceci est formalisé ainsi :

$$\min_{f_m \in \mathcal{H}(\mathcal{V})} \left\{ R_w^p(f_i) + \frac{\lambda}{2} \|f_m(i) - f_m^0(i)\|_2^2 \right\}, \quad (6.26)$$

où le terme $R_w^p(f_m)$ est le terme de régularisation défini dans l'équation (6.21). Le processus de diffusion discret décrit par l'équation (6.24) est utilisé pour chaque minimisation. À la fin des processus, les probabilités d'appartenance à chaque classe peuvent être estimées et la classification finale peut être obtenue pour un nœud $i \in \mathcal{V}$ par :

$$\arg \max_{m \in 1, \dots, N_\omega} \left\{ f_m(i) / \sum_m f_m(i) \right\} \quad (6.27)$$

La démarche pour segmenter les images par classification semi supervisée se situe dans le plan image représenté par un graphe. L'objectif global visé est d'obtenir une segmentation grossière dans un premier temps et de la raffiner par l'introduction de nouveaux marqueurs. En plus des marqueurs initiaux fournis par l'utilisateur, quatre paramètres sont requis par la méthode : la topologie du graphe, la fonction de pondération des arêtes w , le paramètre de fidélité λ et la valeur de p . Trois paramètres sont généralement fixes : p est choisi égal à 2 sauf pour des images présentant beaucoup de contours, $\lambda = 1$ afin d'éviter de modifier les labels initiaux, la fonction de pondération des arêtes est classiquement un noyau Gaussien. Une topologie de graphe simple (graphe grille en 4-adjacence) est suffisante pour des images simples, mais dès lors que les images sont texturées, une topologie plus large est nécessaire ainsi qu'une prise en compte de la texture à travers l'utilisation de patches. La taille de ces derniers est fixée par l'utilisateur et dépend des textures présentes dans les images. La robustesse de la

méthode dépend donc de trois facteurs principaux : les germes initiaux, la topologie et la pondération du graphe. Pour un graphe présentant $|\mathcal{V}|$ arêtes, la complexité requise est en $O(nk|\mathcal{V}|)$ où n est le nombre d'itérations et k le nombre de voisins de chaque nœud. La figure 6.6 présente des résultats pour l'image 6.4(c) quand l'utilisateur fixe les germes initiaux en plein centre des deux régions. La comparaison de cette figure avec la figure 6.4 montre que cette technique semi supervisée fournit des résultats probants même quand les distributions couleur des régions sont mélangées dans l'espace couleur.

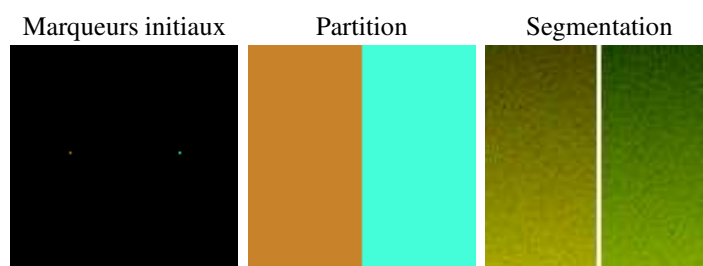


Figure 6.6. Segmentation de l'image 6.4(c) par classification semi supervisée sur un graphe grille en 4-adjacence. Les arêtes sont pondérées par un noyau Gaussien ($\sigma = 5$), $\lambda = 1$ et $p = 2$. La distance entre deux nœuds est calculée sur les couleurs des pixels.

Un autre exemple d'utilisation de la méthode de segmentation semi supervisée est présenté à la figure 6.7 pour la segmentation en trois classes d'une image cytologique (fond, noyaux et cytoplasme). La figure 6.7(a) présente l'image initiale ainsi que trois marqueurs initiaux (un seul par classe). La segmentation semi supervisée opère ensuite sur le graphe complet d'une partition initiale de l'image couleur originale (voir figure 6.7(b)), ce qui signifie que toutes les régions sont connectées par une arête. La fonction de poids des arêtes utilise la couleur moyenne des régions. Du résultat final (voir figure 6.7(c)), nous pouvons dresser deux conclusions intéressantes. D'une part, peu de marqueurs sont nécessaires pour obtenir une segmentation totale des objets non connexes, rendant la méthode semi automatique moins fastidieuse pour l'utilisateur. D'autre part, une convergence rapide de l'algorithme et une diminution du temps d'exécution sont obtenues du fait de la réduction du nombre de données à traiter (avec la partition initiale, le nombre de sommets à traiter est réduit de plus de 90% comparativement au nombre de pixels de l'image). L'utilisation du graphe complet n'est pas envisageable si les sommets du graphe sont associés aux pixels de l'image, mais dans le cas où l'on considère une partition initiale, le nombre de nœuds du graphe est grandement réduit.

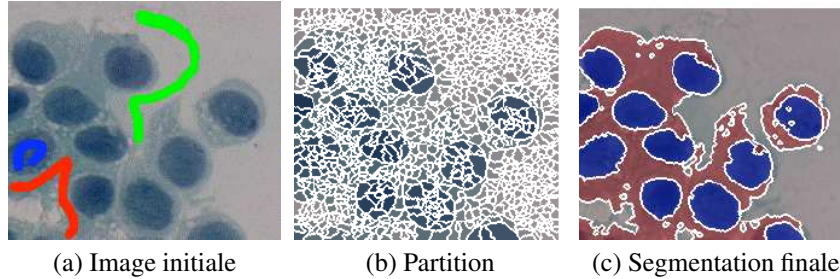


Figure 6.7. Segmentation semi supervisée par diffusion de labels.

6.4.3. Classification spectrale appliquée à la segmentation d'images couleur

6.4.3.1. Introduction

Plutôt que de régulariser itérativement les fonctions caractérisant chacun des nœuds d'un graphe, les méthodes de classification spectrale s'attachent à exploiter directement les poids pondérant les arêtes entre les nœuds afin de les regrouper en classes ou régions. Le problème de la classification de données à partir de graphes de similarité peut être redéfini de la manière suivante : Comment partitionner un graphe de manière à ce que les poids de toutes les arêtes entre nœuds appartenant à différents sous-graphes soient faibles et que ceux des arêtes au sein d'un même sous-graphe soient élevés ? Pour ce faire, les poids sont rassemblés dans une matrice dite d'affinité ou de similarité à partir de laquelle une analyse spectrale (au sens décomposition des vecteurs propres) est menée. Bien qu'il existe un large éventail d'algorithmes différents de classification spectrale, ces derniers reposent souvent sur le schéma suivant :

- La construction d'un graphe de similarité entre les données qui sera formalisée par une matrice d'affinité à partir des données à analyser. La similarité entre les données est généralement calculée à partir de fonctions noyaux de type gaussienne.
- L'analyse spectrale des données projetées dans une nouvelle base formée par les vecteurs propres de cette matrice d'affinité.
- L'application d'une méthode de classification (généralement un simple k -means) sur les données à analyser dans cette nouvelle base.

Soit $G = (\mathcal{V}, \mathcal{E}, w)$ un graphe pondéré non-orienté où $\mathcal{V} = \{v_1, \dots, v_N\}$ représente l'ensemble des nœuds constituant le graphe. Le poids entre deux nœuds v_i et v_j est noté $w(i, j)$ et la matrice de similarité notée w est une matrice carrée de taille $N \times N$ symétrique puisque le graphe est non-orienté. Le degré $d_i = \sum_{j=1}^N w(i, j)$ d'un nœud v_i correspond à la somme des poids associés aux nœuds qui lui sont connectés par une

arête. La matrice D des degrés est une matrice diagonale où les degrés d_i sont placés à la $i^{\text{ème}}$ ligne et $i^{\text{ème}}$ colonne de la matrice.

Les matrices D et w sont utilisées pour construire la matrice d'affinité L aussi appelée matrice Laplacienne. Il existe plusieurs matrices Laplaciennes utilisées dans la littérature dont les principales sont la matrice Laplacienne non-normalisée définie par la relation $L = D - w$ et la matrice Laplacienne normalisée $L_{sym} = D^{-1/2} L D^{-1/2}$. L'exploitation de l'une ou l'autre de ces matrices permet de résoudre un problème de coupure de graphe différent et doit être utilisée suivant le contexte du travail à réaliser [LUX 07]. Les démonstrations pour résoudre ce problème suivant la matrice Laplacienne considérée peuvent être trouvées dans [LUX 07] de manière générale tandis que Shi et Malik [SHI 00] résolvent ce problème pour la matrice L_{sym} appliquée à la segmentation d'images.

Les méthodes de classification spectrale peuvent être appliquées soit dans le plan image (voir paragraphe 6.4.3.2) soit dans l'espace couleur (voir paragraphe 6.4.3.3).

6.4.3.2. Segmentation par classification spectrale dans le plan image

Shi et Malik [SHI 00] introduisent une méthode de classification spectrale appliquée à la segmentation par une analyse dans le plan image. Chaque pixel de l'image correspond à un nœud du graphe et ces nœuds sont reliés par des arêtes dont les poids $w(i, j)$ sont exprimés par :

$$w(i, j) = \exp \left(- \left(\frac{\|\mathbf{x}_i^c - \mathbf{x}_j^c\|_2}{\sigma^c} \right)^2 \right) \cdot \begin{cases} \exp \left(- \left(\frac{\|\mathbf{x}_i^s - \mathbf{x}_j^s\|_2}{\sigma^s} \right)^2 \right) & \text{si } \|\mathbf{x}_i^s - \mathbf{x}_j^s\|_2 < r, \\ 0 & \text{sinon.} \end{cases} \quad (6.28)$$

où \mathbf{x}_i^s représente les coordonnées spatiales du pixel et donc du nœud i et \mathbf{x}_i^c est un vecteur caractéristique calculé à partir de mesures d'intensité, de couleur et de texture. Notons que l'équation (6.28) nécessite de fixer la taille r du voisinage et les paramètres des noyaux Gaussiens σ^c , σ^s .

Une fois que le graphe représentant l'image à segmenter est créé, la méthode proposée consiste à extraire itérativement et une par une les régions de l'image I en coupant en deux le graphe. Cet algorithme appelé "Ncut" mesure les dissimilarités entre deux sous-graphes ainsi que la similarité au sein d'un même sous-graphe [SHI 00]. Il exploite la matrice Laplacienne L_{sym} afin de prendre en compte la population de chacun des sous-graphes. Ceci permet d'éviter une partition en sous-graphes composés uniquement de pixels isolés. La méthode de segmentation proposée pour segmenter l'image peut être décomposée en quatre étapes :

1) Construction d'un graphe pondéré non-orienté $G = (\mathcal{V}, E, w)$ à partir d'une mesure de similarité entre deux nœuds.

2) Résoudre $L_{sym}x = \lambda x$ pour les vecteurs propres avec les plus petites valeurs propres.

- 3) Utiliser le vecteur propre avec la deuxième valeur propre la plus petite (nommé vecteur de Fiedler) pour partitionner le graphe.
- 4) Décider si la partition courante doit être à nouveau segmentée ou pas.

Yu et Shi [YU 03] remarquent que les calculs de valeurs et vecteurs propres peuvent être instables lorsque les matrices Laplaciennes ne sont pas bloc diagonales, c'est-à-dire lorsque les coupures du graphes pour le partitionner en sous-graphe ne sont pas franches et proposent de couper le graphe en plusieurs sous-graphes plutôt que de le faire itérativement. En outre, cette approche permet de partitionner directement l'image en régions avec une seule itération et de réduire ainsi le temps de traitement.

Les méthodes de classification spectrale nécessitent l'ajustement de paramètres comme l'écart-type des noyaux gaussiens pour calculer la similarité entre deux nœuds ainsi que le nombre de classes. Pour cette raison, L. Zelnik-Manor et P. Perona proposent d'ajuster automatiquement ces paramètres [ZEL 04] pour obtenir une méthode non supervisée.

Les méthodes effectuant l'opération de segmentation dans le plan image considèrent chaque pixel de l'image comme un nœud d'un graphe qu'il faut couper. Une image de N pixels implique la construction d'une matrice d'affinité avec N^2 éléments, ce qui peut provoquer des problèmes d'allocation mémoire. Ceci limite par conséquent son usage à des images de petites tailles ou à des images considérées à sous-résolution.

Afin de limiter l'espace mémoire requis par la matrice d'affinité, Cour *et al.* proposent un algorithme de classification multi-échelle en remplaçant le paramètre de voisinage r par un paramètre d'échelle déterminé empiriquement [COU 05]. Par ailleurs, Les auteurs soulignent l'importance de la fonction noyaux pour la qualité du résultat de la segmentation. Pour cette raison, ils construisent un graphe pondéré non-orienté $G = (\mathcal{V}, \mathcal{E}, w)$ à partir de la mesure de similarité $w(i, j)$ entre deux nœuds suivante :

$$w(i, j) = \sqrt{w^I(i, j) \times w^e(i, j) + \alpha w^e(i, j)}. \quad (6.29)$$

Comme la mesure $w^I(i, j)$ basée sur l'intensité des pixels ne permet pas de discriminer des zones texturées, elle est couplée avec la mesure $w^e(i, j)$ analysant l'amplitude des contours détectés dans l'image. La mesure d'intensité est calculée par l'équation

$$w^I(i, j) = \exp \left(-\frac{\|\mathbf{x}_i^s - \mathbf{x}_j^s\|_2^2}{2\sigma^s^2} - \frac{\|\mathbf{x}_i^c - \mathbf{x}_j^c\|_2^2}{2\sigma^c^2} \right).$$

La mesure basée sur les contours est calculée par l'équation $w^e(i, j) = \exp \left(-\frac{\max_{\mathbf{x} \in \text{line}(i, j)} \|\text{Edge}(\mathbf{x})\|_2^2}{2\sigma^e^2} \right)$ où $\text{line}(i, j)$ est une droite joignant les pixels i et j , $\text{Edge}(\mathbf{x})$ est l'amplitude d'un gradient au pixel \mathbf{x} et σ^e est le paramètre d'un noyau Gaussien.

Des essais effectués sur les images de taille 481×321 pixels nécessitent un temps de calcul important car proche de la minute avec un ordinateur équipé d'un micro-processeur cadencé à 2.4 GHZ.

6.4.3.3. Segmentation par classification spectrale dans l'espace couleur

Appliquer une classification spectrale sur le plan image nécessite la construction d'une matrice d'affinité de grande taille. Pour réduire cette taille, il peut être opportun d'appliquer la classification spectrale à un espace couleur puisque le nombre des couleurs présentes au sein de l'image est *a priori* plus faible que celui des pixels. Les méthodes de segmentation d'images couleur par classification des points couleur cherchent à déterminer les nuages de points dans l'espace couleur caractérisant les différentes régions de l'image. Tandis que les méthodes de classification classique recherchent des formes spécifiques comme l'algorithme des *k*-means qui recherche des nuages de forme globulaire, l'emploi des méthodes de classification spectrale permet de s'affranchir des formes des nuages de points caractérisant les régions de l'image dans l'espace couleur. Le problème clé réside dans la définition de similarité entre les nœuds i et j représentant les couleurs présentes au sein de l'image considérée. Dans un premier temps, les couleurs de l'image sont sous-quantifiées en N_C couleurs afin de réduire au maximum la taille de la matrice de similarité sans toutefois altérer la représentation de l'image.

Busin *et al.* considèrent plusieurs espaces couleur pour effectuer l'opération de segmentation [BUS 09]. Pour chaque espace couleur examiné, un graphe constitué de N_C nœuds est construit et les poids entre ces nœuds sont :

$$w^c(i,j) = \min\{H^c(\mathbf{x}_i^c), H^c(\mathbf{x}_j^c)\} \exp^{-\frac{\|\mathbf{x}_i^c - \mathbf{x}_j^c\|_2^2}{2\sigma^c}}, \quad (6.30)$$

où $H^c(\mathbf{x}_i^c)$ est le nombre d'occurrences de la couleur \mathbf{x}_i^c dans l'espace couleur c et σ^c est l'écart-type du noyau Gaussien fixé par l'utilisateur. Un autre poids basé sur la fonction de compacité spatio colorimétrique (voir équation (6.8)) a été proposé par Hébert *et al.* [HEB 08]. Le graphe est alors coupé en sous-graphes à partir de l'algorithme des "Ncut" et la qualité des résultats obtenus pour chaque espace couleur candidat est évaluée grâce à un critère basé sur l'écart entre deux valeurs propres successives [LUX 07]. L'espace candidat avec l'écart le plus élevé est alors retenu pour partitionner les couleurs. Les régions composées des pixels connexes dont les couleurs sont assignées à une même classe, sont itérativement divisées par ce même algorithme jusqu'à satisfaction d'un critère d'arrêt.

La réduction du nombre C de couleurs à 1024 dans l'image à analyser permet d'obtenir des temps de calcul d'environ 15 secondes pour les images de taille 481×321 pixels avec un ordinateur équipé d'un micro-processeur cadencé à 2,4 GHZ. Il est à noter que le temps de calcul pour cette méthode varie en fonction du nombre de couleurs dominantes choisi pour quantifier l'image plutôt qu'en fonction de la taille de l'image.

En conclusion, l'algorithme de L. Zelnik-Manor et P. Perona [ZEL 04] est l'algorithme le plus simple d'utilisation puisqu'il ne nécessite de fixer aucun paramètre par l'utilisateur. Cependant, la quantité d'espace mémoire requis par cet algorithme ne permet de l'utiliser que pour des images de petites tailles et limite considérablement son usage dans la pratique. L'algorithme de T. Cour *et al.* [COU 05] et de L. Busin *et al.* [BUS 09] permettent de s'affranchir de cette limite. Cependant, ces deux algorithmes nécessitent de fixer le nombre de régions à retrouver dans l'image pour le premier et le nombre de classes de pixels pour le deuxième. Ainsi, le choix vers l'un ou l'autre de ces deux algorithmes se fera suivant que l'utilisateur souhaite effectuer une segmentation à partir d'une approche région ou à partir d'une approche par classification des couleurs.

6.5. Évaluation des méthodes de segmentation à partir de vérités-terrain

Les procédures d'évaluation de la qualité d'une segmentation sont grossièrement divisées en deux groupes, celles qui ne sont pas basées sur une vérité terrain, dont le lecteur trouvera une description approfondie dans [PHI 06], et celles basées sur une vérité terrain. L'objectif de cette partie n'est pas de comparer les performances des différentes méthodes de segmentation présentées au sein de ce chapitre, mais plutôt de montrer comment évaluer la qualité de segmentation à partir de plusieurs vérités terrain. Mesurer la qualité d'une segmentation permet de comparer les performances de méthodes entre elles, mais également d'ajuster au mieux les valeurs des paramètres requis par chacune de ces méthodes.

Pour ce faire, Martin *et al.* ont construit une base d'images naturelles segmentées par des observateurs humains [MAR 01].⁴ Chaque observateur a détourné les objets les plus importants selon lui présents dans l'image de telle sorte que chaque image segmentée contienne entre 2 et 30 contours fermés. Cette vérité terrain est toute relative car la partition a été effectuée une fois par chaque observateur. Pour juger de la qualité d'une segmentation fournie par une procédure, les auteurs proposent de comparer la partition automatique avec chacune des partitions fournies par un observateur [MAR 04].

De nombreux critères d'évaluation ont été proposés dont le plus récent est le NPR-index [UNN 07]. Toutefois, l'indice de Jaccard reste l'un des critères les plus utilisés [GE 06] pour mesurer la correspondance des régions de chaque segmentation humaine $\mathbf{S}^{\mathcal{H}}$ (composée de $n^{\mathcal{H}}$ régions $\mathcal{R}_i^{\mathcal{H}}, i = 1, 2, \dots, n^{\mathcal{H}}$) et de la segmentation automatique $\mathbf{S}^{\mathcal{A}}$ (composée de $n^{\mathcal{A}}$ régions $\mathcal{R}_j^{\mathcal{A}}, j = 1, 2, \dots, n^{\mathcal{A}}$), la partition humaine étant considérée comme la vérité-terrain. Pour deux régions $\mathcal{R}_i^{\mathcal{H}}$ et $\mathcal{R}_j^{\mathcal{A}}$, l'indice de Jaccard est défini comme :

4. La base de données est disponible à l'adresse web <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>.

$$J(\mathcal{R}_i^{\mathcal{H}}, \mathcal{R}_j^{\mathcal{A}}) = \frac{\text{Card}\{\mathcal{R}_i^{\mathcal{H}} \cap \mathcal{R}_j^{\mathcal{A}}\}}{\text{Card}\{\mathcal{R}_i^{\mathcal{H}} \cup \mathcal{R}_j^{\mathcal{A}}\}}. \quad (6.31)$$

Le numérateur de l'équation (6.31) mesure le niveau auquel la région $\mathcal{R}_j^{\mathcal{A}}$ issue de la segmentation automatique coïncide avec la région $\mathcal{R}_i^{\mathcal{H}}$ de la vérité terrain. Le dénominateur normalise l'indice de Jaccard de telle sorte qu'il varie entre $[0,1]$.

Par chaque paire de segmentations humaine-automatique, une comparaison deux à deux de toutes les régions respectives est menée avec l'indice de Jaccard. La mesure de similarité résultante est :

$$\text{Sim}(\mathbf{S}^{\mathcal{H}}, \mathbf{S}^{\mathcal{A}}) = \frac{1}{\max(n^{\mathcal{H}}, n^{\mathcal{A}})} \sum_{i=1}^{n^{\mathcal{H}}} \sum_{j=1}^{n^{\mathcal{A}}} J(\mathcal{R}_i^{\mathcal{H}}, \mathcal{R}_j^{\mathcal{A}}). \quad (6.32)$$

Cette mesure varie entre 0, ce qui signifie que les segmentations humaines et automatique sont fortement différentes, et 1, correspondant au cas où les segmentations sont identiques.

Afin de montrer un cas pratique, nous proposons de segmenter les images de référence (#15088 et #37073) extraites de la Base Berkeley [MAR 01], respectivement présentes dans les images 6.8(a) et 6.9(a), et référencées *Boat* et *Plane*. Chacune de ces images montre une partition humaine parmi les cinq disponibles. Par ailleurs, comme ces images contiennent des objets pouvant être extraits sans ambiguïté, les cinq partitions humaines sont concordantes entre elles.

Pour les comparaisons des performances, nous avons retenu trois méthodes spatio colorimétriques, à savoir la méthode "mean shift" décrite dans le paragraphe 6.3.2.1, l'analyse des degrés de compacité décrite dans le paragraphe 6.3.2.2 et la méthode JSEG décrite dans le paragraphe 6.3.3, ainsi que trois analyses de graphes, à savoir la classification spectrale de pixels décrite dans le paragraphe 6.4.3.2, la classification spectrale des couleurs décrite dans le paragraphe 6.4.3.3 et la segmentation semi supervisée par diffusion de labels décrite dans le paragraphe 6.4.2. Les cinq premières méthodes opèrent avec un apprentissage non supervisé (le nombre de classes ou de région peut être cependant introduit) tandis que la sixième méthode requiert la définition par l'utilisateur des marqueurs initiaux dans l'image (voir figures 6.8(g) et 6.9(g)). Les résultats automatiques montrés dans les figures 6.8 et 6.9, sont comparés avec chacune des cinq segmentations humaines grâce à la mesure de similarité. Les faibles indices de Jaccard (valeur max 0,26) montrent bien que les méthodes de segmentation peinent à reproduire la partition des observateurs humains. Par ailleurs, ces figures montrent bien que l'indice de Jaccard reflète de manière relativement satisfaisante

l'évaluation de la qualité de segmentation par comparaison visuelle entre la partition humaine et automatique.

Pour l'image 'Boat' de la figure 6.8, on peut remarquer que l'analyse des degrés de compacité spatio colorimétrique et la segmentation semi supervisée parviennent à bien distinguer la zone représentant l'eau du bateau, ce qui explique la supériorité des indices de Jaccard obtenus par rapport à ceux des autres méthodes. Pour l'image 'Plane' de la figure 6.9, seules les approches par classification spectrale parviennent à respecter en partie la partition humaine du fond de l'image en plusieurs régions. Comme la méthode JSEG réussit à bien distinguer les différents éléments de l'avion que la partition humaine a mis en évidence, elle obtient le meilleur indice de Jaccard qui est très sensible à la sur-segmentation. Les marqueurs initiaux n'ayant pas respecté complètement la finesse de la partition humaine, l'indice de Jaccard obtenu par la segmentation semi supervisée est plus faible que celui obtenu par JSEG. Cela montre qu'obtenir une partition fine par la segmentation semi supervisée requiert le placement de nombreux marqueurs décrivant différentes classes, ce qui peut se révéler fastidieux pour l'utilisateur.

Ces résultats corroborent ceux montrés dans les paragraphes précédents, à savoir que les méthodes de classification spatio colorimétrique qui analysent une fonction représentant conjointement la distribution des couleurs et leur répartition sur le plan image (voir paragraphe 6.3.2) fournissent des résultats grossiers de segmentation qui méritent d'être affinés par une analyse ultérieure spatiale. Cette démarche suivie par la méthode JSEG (voir paragraphe 6.3.3) permet d'obtenir des résultats satisfaisants, au prix d'un surcoût calculatoire dû à ces deux étapes successives, et d'un ajustement délicat des paramètres requis.

Ces figures montrent également l'intérêt d'utiliser une connaissance *a priori*, sous la forme des marqueurs présentés dans le paragraphe 6.4.2, utilisée par un algorithme de segmentation semi supervisée. La partition finale respecte en très grande partie la connaissance *a priori* fournie sous forme de marqueurs, ce qui permet de suivre la segmentation souhaitée par l'utilisateur.

6.6. Conclusion

Ce chapitre, dédié aux méthodes de segmentation d'images couleur en régions, avait pour objectif de décrire les méthodes de segmentation les plus récentes et qui n'étaient donc pas décrites dans le chapitre segmentation de l'ouvrage [TRé 04]. Sans vouloir dresser une liste exhaustive des approches, les méthodes les plus connues ont été présentées, à savoir les méthodes de classification spatio colorimétrique et les méthodes d'analyse de graphes. Le lecteur trouvera un complément sur la segmentation couleur dans le chapitre 1 consacré notamment à la segmentation couleur par

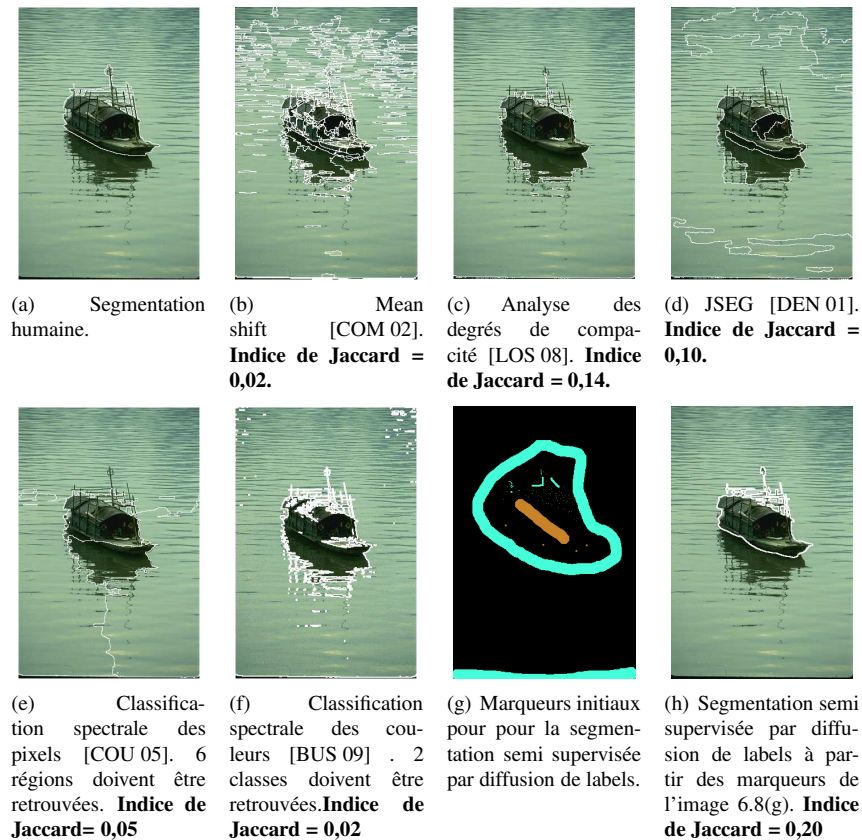


Figure 6.8. Segmentations de l'image 'Boat'.

outils morphologiques et dans le chapitre 5 consacré à la segmentation d'images présentant des textures couleur.

Il est à noter que nous n'avons pas présenté de méthodes de détection de contours couleur, car cette voie est actuellement très peu approfondie, du fait notamment de la nécessité d'obtenir des contours fermés pour une reconnaissance ultérieure des objets [TRÉ 04].

Les méthodes de classification spatio colorimétrique qui analysent conjointement la distribution des couleurs dans l'espace couleur ainsi que leur répartition spatiale dans le plan image nécessitent souvent l'utilisation de structures de données lourdes à manipuler. L'histogramme compact, présenté dans la première partie de ce chapitre, offre une solution intéressante permettant de réduire les temps de calcul et le

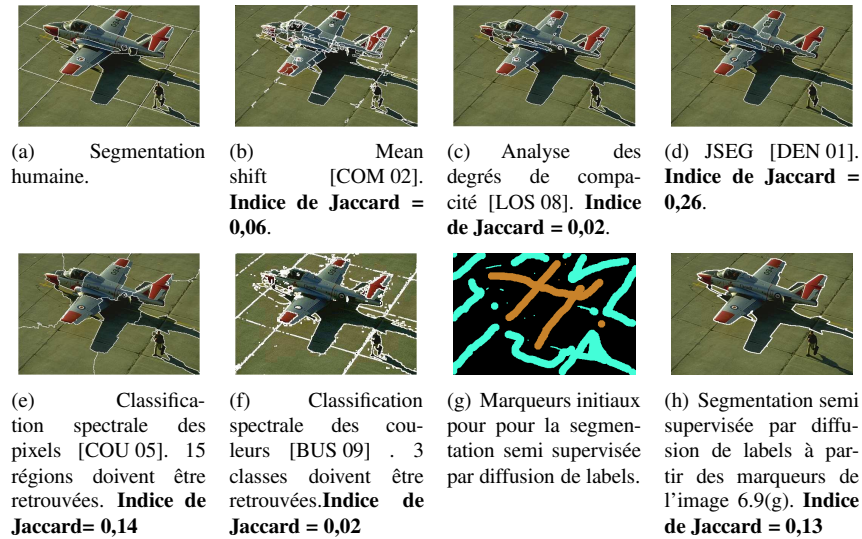


Figure 6.9. Segmentations de l'image 'Plane'.

coût de stockage. Les méthodes de classification spatio colorimétrique sont basées sur des hypothèses fortes d'homogénéité des couleurs au sein des régions à retrouver. Elles ne nécessitent pas d'information *a priori* sur l'image et fournissent des résultats satisfaisants pour les images qui respectent cette hypothèse. Cette hypothèse étant peu valide pour des scènes naturelles complexes, elles provoquent souvent une sur-segmentation qui devra être compensée par une fusion ultérieure des régions. Par ailleurs, ces méthodes basées sur l'analyse des distributions couleur se révèlent être sensibles au choix des espaces couleur retenus pour établir la segmentation. Outre la recherche de méthodes de segmentation dont l'ajustement des paramètres est adapté à un grand nombre de familles d'images couleur, la détermination d'espaces couleur qui permettent de fournir une segmentation satisfaisante reste un problème ouvert, et totalement dépendant de l'application visée.

Ces méthodes de classification fournissent donc une première segmentation qui peut être complétée par l'analyse d'un graphe formalisant les relations spatiales entre régions adjacentes. La formalisation des éléments d'une image sous la forme d'un graphe permet entre autres de représenter une connaissance *a priori* d'un expert sous la forme de marqueurs initiaux. Ces marqueurs, simples à mettre en œuvre, permettent de guider efficacement la segmentation dite alors semi supervisée.

Peu de marqueurs sont nécessaires pour obtenir une segmentation satisfaisante, rendant la méthode semi supervisée peu fastidieuse pour l'utilisateur. Une convergence rapide de l'algorithme et une diminution du temps d'exécution sont obtenues du fait de la partition initiale fournie par la classification. Enfin, l'analyse d'un graphe complet portant sur les pixels n'étant pas envisageable, on peut associer les sommets du graphe aux couleurs présentes dans l'image afin de les classifier.

6.7. Bibliographie

- [BEL 06] BELKIN M., NIYOGI P., SINDHWANI V., BARTLETT P., « Manifold Regularization : A Geometric Framework for Learning from Labeled and Unlabeled Examples », *Journal of Machine Learning Research*, vol. 7, p. 2399-2434, 2006.
- [BUA 05] BUADES A., COLL B., MOREL J.-M., « A review of image denoising algorithms, with a new one », *Multiscale Modeling and Simulation*, vol. 4, n° 2, p. 490-530, 2005.
- [BUS 05] BUSIN L., VANDENBROUCKE N., MACAIRE L., POSTAIRE J.-G., « Colour space selection for unsupervised colour image segmentation by analysis of connectedness properties », *International Journal of Robotics and Automation*, vol. 20, n° 2, p. 70-77, 2005.
- [BUS 09] BUSIN L., SHI J., VANDENBROUCKE N., MACAIRE L., « Color space selection for color image segmentation by spectral clustering », *IEEE International Conference on Signal and Image Processing Applications*, Kuala Lumpur, Malaisie, p. 262-267, 2009.
- [CHA 01] CHAN T., OSHER S., SHEN J., « The digital TV filter and nonlinear denoising », *IEEE Transactions on Image Processing*, vol. 10, n° 2, p. 231-241, 2001.
- [CHE 01] CHENG H. D., JIANG X. H., SUN Y., WANG J., « Color image segmentation : advances and projects », *Pattern Recognition*, vol. 34, n° 12, p. 2259-2281, 2001.
- [CHU 97] CHUNG F. R., « Spectral Graph Theory », *CBMS Regional Conference Series in Mathematics*, vol. 92, p. 1-212, American Mathematical Society, 1997.
- [CLÉ 01] CLÉMENT A., VIGOUROUX B., « Un histogramme compact pour l'analyse d'images multicomposante », *Proc. 18e Colloque GRETSI sur le Traitement du Signal et des Images*, p. 305-307, 2001.
- [CLÉ 03] CLÉMENT A., VIGOUROUX B., « Unsupervised segmentation of scenes containing vegetation (Forsythia) and soil by hierarchical analysis of bi-dimensional histograms », *Pattern Recognition Letters*, , n° 24, p. 1951-1957, 2003.
- [COC 95] COCQUEREZ J.-P., PHILIPP S., *Analyse d'images : filtrage et segmentation*, Masson, 1995.
- [COM 02] COMANICIU D., MEER P., « Mean shift : a robust approach toward feature space analysis », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, n° 5, p. 603-619, mai 2002.
- [COU 05] COUR T., BENEZIT F., SHI J., « Spectral Segmentation with Multiscale Graph Decomposition », *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2*, Washington, DC, USA, IEEE Computer Society, p. 1124-1131, 2005.

- [DEN 01] DENG Y., MANJUNATH B. S., « Unsupervised segmentation of color-texture regions in images and video », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, n° 8, p. 800-810, 2001.
- [DIE 05] DIESTEL R., *Graph Theory*, vol. 173 de *Graduate Texts in Mathematics*, Springer-Verlag, août 2005.
- [ELM 08a] ELMOATAZ A., LÉZORAY O., BOUGLEUX S., « Nonlocal Discrete Regularization on Weighted Graphs : a Framework for Image and Manifolds Processing », *IEEE transactions on Image Processing*, vol. 17, n° 7, p. 1047–1060, 2008.
- [ELM 08b] ELMOATAZ A., LEZORAY O., BOUGLEUX S., TA V., « Unifying local and nonlocal processing with partial difference operators on weighted graphs », *International Workshop on Local and Non-Local Approximation in Image Processing (LNLA)*, p. 11-26, 2008.
- [GE 06] GE F., WANG S., LIU T., « A New Benchmark for Image-Segmentation Evaluation », *Journal of Electronic Imaging*, vol. 16, n° 3, page033011 (16 pages), 2006.
- [GRA 06] GRADY L., « Random walks for image segmentation », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, n° 11, p. 1768–1783, 2006.
- [HEB 08] HEBERT P.-A., MACAIRE L., « Spatial-color pixel classification by spectral clustering for color image segmentation », *3rd International Conference on Information & Communication Technologies : from Theory to Applications*, p. 1 - 5, 2008.
- [JIN 03] JING F., LI M., ZHANG H.-J., ZHANG B., « Unsupervised image segmentation using local homogeneity analysis », *Proceedings of the 2003 International Symposium on Circuits and Systems (ISCAS'03)*, vol. 2, Bangkok, Thailand, p. 456-459, mai 2003.
- [LEZ 07] LEZORAY O., ELMOATAZ A., BOUGLEUX S., « Graph regularization for color image processing », *Computer Vision and Image Understanding (CVIU)*, vol. 107, n° 1-2, p. 38-55, 2007.
- [LOS 08] LOSSON O., BOTTE-LECOCQ C., MACAIRE L., « Fuzzy mode enhancement and detection for color image segmentation », *EURASIP Journal on Image and Video Processing*, vol. Special issue "Color in Image and Video processing", p. 1-19, 2008.
- [LUX 07] VON LUXBURG U., « A tutorial on spectral clustering », *Statistics and Computing*, vol. 17, n° 4, p. 395–416, Kluwer Academic Publishers, 2007.
- [MAC 06] MACAIRE L., VANDENBROUCKE N., POSTAIRE J.-G., « Color image segmentation by analysis of subset connectedness and color homogeneity properties », *Computer Vision and Image Understanding*, vol. 102, n° 1, p. 105–116, 2006.
- [MAR 01] MARTIN D., FOWLKES C., TAL D., MALIK J., « A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics », *Proceedings of the 8th International Conference on Computer Vision (ICCV'01)*, vol. 2, Vancouver, BC, Canada, p. 416–423, juillet 2001.
- [MAR 04] MARTIN D. R., FOWLKES C., MALIK J., « Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, n° 5, p. 530–549, IEEE Computer Society, 2004.

- [OUA 07] OUATTARA S., CLÉMENT A., « Étiquetage d'histogrammes multidimensionnels compacts pour l'analyse d'images multicomposantes », *Proc. 21e Colloque GRETSI sur le Traitement du Signal et des Images*, 2007.
- [PAL 07] PALUS H., « *Color image processing : methods and applications*, R. Lukac and K.N. Plataniotis », Chapitre Color image segmentation : selected techniques, p. 103-128, CRC Press, 2007.
- [PAR 07] PARIS S., DURAND F., « A Topological Approach to Hierarchical Segmentation using Mean Shift », *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*, Minneapolis, p. 1-8, 2007.
- [PHI 06] PHILIPP-FOLIGUET S., GUIGUES L., « Evaluation de la segmentation : état de l'art, nouveaux indices et comparaison », *Traitement du Signal*, vol. 23, n° 2, p. 109-125, 2006.
- [PRE 92] PRESS W., TEUKOLSKY S., VETTERLING W., FLANNERY B., *Numerical Recipes in C. The Art of Scientific Computing*, Second Edition. Cambridge University Press, 1992.
- [SHA 09] SHAPIRA L., AVIDAN S., SHAMIR A., « Mode-detection via median-shift », *Computer Vision, 2009 IEEE 12th International Conference on*, p. 1909-1916, septembre 2009.
- [SHI 00] SHI J., MALIK J., « Normalized cuts and image segmentation », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, n° 8, p. 888-905, 2000.
- [SIN 07] SINOP A. K., GRADY L., « A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm », *International Conference on Computer Vision*, p. 1-8, 2007.
- [TA 09] TA V.-T., LEZORAY O., ELMOATAZ A., SCHÜPP S., « Graph-based tools for microscopic cellular image segmentation », *Pattern Recognition*, vol. 42, n° 6, p. 1113-1125, 2009.
- [TOR 10] TORRES C., CLÉMENT A., « Unsupervised Hierarchical Spatio-Colorimetric Classification For Color Image Segmentation », *Fifth European Conference on Color in Graphics, Image and Vision*, Joensuu (Finlande), p. 235-239, 2010.
- [TRé 04] TRÉMEAU A., FERNANDEZ-MALOIGNE C., BONTON P., *Image numérique couleur : de l'acquisition au traitement*, Dunod, Paris, 2004.
- [TSC 05] TSCHUMPERLÉ D., DERICHE R., « Vector-Valued Image Regularization with PDEs : A Common Framework for Different Applications », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, n° 4, p. 506-517, 2005.
- [UNN 07] UNNIKRISHNAN R., PANTOFARU C., HEBERT M., « Toward Objective Evaluation of Image Segmentation Algorithms », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, n° 6, p. 929-944, 2007.
- [VAN 03] VANDENBROUCKE N., MACAIRE L., POSTAIRE J.-G., « Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis », *Computer Vision and Image Understanding*, vol. 90, n° 2, p. 190-216, 2003.
- [WAN 06a] WANG F., WANG J., ZHANG C., SHEN H. C., « Semi-Supervised Classification Using Linear Neighborhood Propagation », *IEEE Computer Society Conference on Computer Vision and Pattern Recognition - (CVPR'06)*, vol. 1, p. 160-167, 2006.

- [WAN 06b] WANG Y., YANG J., PENG N., « Unsupervised color-texture segmentation based on soft criterion with adaptive mean-shift clustering », *Pattern Recognition Letters*, vol. 27, n° 5, p. 386-392, avril 2006.
- [WAN 06c] WANG Y.-G., YANG J., CHANG Y.-C., « Color-texture image segmentation by integrating directional operators into JSEG method », *Pattern Recognition Letters*, vol. 27, n° 16, p. 1983-1990, décembre 2006.
- [YU 03] YU S., SHI J., « Multiclass Spectral Clustering », *International Conference on Computer Vision*, p. 313–319, 2003.
- [ZEL 04] ZELNIK-MANOR L., PERONA P., « Self-tuning spectral clustering », *Advances in Neural Information Processing Systems 17*, MIT Press, p. 1601–1608, 2004.
- [ZHO 05] ZHOU D., SCHÖLKOPF B., « Regularization on Discrete Spaces », *LNCS 3663, Proc. of the 27th DAGM Symp.*, Springer-Verlag, p. 361–368, août 2005.