

Segmentation d'images couleur par coalescence non supervisée d'histogrammes 2D et fusion de régions selon la théorie de Dempster-Shafer

Color image segmentation by unsupervised 2D histogram clustering and Dempster-Shafer region merging

Olivier Lezoray, Christophe Charrier

LUSAC EA 2607, Équipe Vision et Analyse d'Image, IUT SRC, 120 rue de l'exode, 50000 Saint-Lô (France),
{o.lezoray,c.charrier}@chbg.unicaen.fr

Manuscrit reçu le 15 juin 2004

Résumé et mots clés

Dans cet article nous proposons une méthode de segmentation d'images couleur selon une nouvelle approche que nous appelons bi-marginale. Afin de pallier les défauts des approches marginales classiques, nous considérons les composantes couleur deux à deux afin d'avoir une vue partielle de leur corrélation. Travaillant selon cette vision bi-composante, nous considérons les trois combinaisons possible comme trois sources d'informations indépendantes. Chaque information bi-composante est tout d'abord analysée selon un schéma de coalescence morphologique non supervisé qui recherche les couleurs dominantes d'un histogramme bidimensionnel. Cela permet de construire trois cartes de segmentation distinctes qui sont combinées par intersection après avoir été simplifiées. L'intersection produisant une sur-segmentation, une fusion des régions deux à deux est opérée selon un critère de similarité et selon la combinaison de Dempster-Shafer jusqu'à un critère de terminaison. Afin d'automatiser la méthode de segmentation, une mesure d'énergie est proposée afin de quantifier la qualité d'une segmentation, celle-ci sert tout au long de la méthode proposée comme indicateur de performance de la segmentation afin d'en régler les différents paramètres.

Couleur, segmentation, coalescence, fusion, Dempster-Shafer, multi-échelle, qualité.

Abstract and key words

In this paper, a color image segmentation method based on a new approach called bimarginal is proposed. To overcome the drawbacks of the classical marginal approaches, color components are considered in pairs in order to have a partial view of their inner correlation. Working with color images, the three possible combinations are considered as three independent information sources. Each pairwise component combination is firstly analyzed according to an unsupervised morphologic clustering which looks for the dominant colors of a 2D histogram. This leads to obtain three segmentation maps combined by intersection after being simplified. The intersection process itself producing an over-segmentation of the image, a pairwise region merging is done according to a similarity criterion with the Dempster-Shafer theory up to a termination criterion. To fully automate the segmentation, an energy function is proposed to quantify the segmentation quality. The latter acts as a performance indicator and is used all over the segmentation to tune its parameters.

Color, segmentation, clustering, fusion, Dempster-Shafer, multi-scale, quality.

Remerciement

Les auteurs remercient les relecteurs dont les commentaires ont permis d'améliorer l'article. Ce travail a été financé par les fonds FEDER et la Fnadt.

1. Introduction

La segmentation d'images est une étape essentielle car elle conditionne la qualité de l'interprétation future des éléments composant l'image. C'est un problème essentiel, difficile et pour lequel il n'existe pas de solution universelle. Beaucoup de méthodes de segmentation d'images couleur donnent de très bons résultats et l'on peut les classer en deux familles selon qu'elles utilisent ou non la répartition spatio colorimétrique [TFMB04]. Quelles que soient les méthodes utilisées, lorsqu'elles travaillent sur la notion de couleur, se pose le problème de la dimension des données. On trouve par conséquent trois types d'approches de la segmentation couleur (chaque méthode de segmentation appartenant à l'une de ces approches) [TFMB04]: l'approche marginale, l'approche scalaire, l'approche vectorielle. L'approche marginale de la segmentation opère une segmentation sur chaque composante couleur puis fusionne les cartes de segmentation en une seule. L'approche scalaire de la segmentation fusionne tout d'abord les composantes couleur en une seule et segmente celle-ci. On se ramène alors à un problème de segmentation scalaire (en niveau de gris). La segmentation vectorielle segmente directement l'image couleur en utilisant une information vectorielle. Le principal désavantage des approches marginales et scalaires est la non-prise en compte des interactions colorimétriques inter compo-

santes. Leur avantage est d'être généralement plus simples et rapides que les techniques vectorielles. Un certain nombre d'auteurs utilisent néanmoins des approches mixtes qui combinent les propriétés de plusieurs approches. Dans cet article nous prenons ce point de vue de la segmentation d'images couleur. Travaillant sur des images couleur, l'approche marginale considère chaque composante couleur comme différentes sources d'informations dont les segmentations sont ensuite fusionnées. L'aspect marginal de l'approche est restreignant car elle ne prend pas en compte l'aspect vectoriel des données, nous proposons donc une approche intermédiaire que nous appellerons bi-marginale qui considère les composantes couleur deux à deux afin d'avoir une vue partielle de la corrélation entre les composantes. L'avantage de cette approche réside dans cette vision partielle de la corrélation entre les composantes tout en conservant une approche marginale. Le schéma général de l'approche de segmentation proposée est résumé par la figure 1. Nous travaillons sur les trois combinaisons bi-composantes possibles, chacune de ces combinaisons est traitée suivant un schéma marginal de coalescence et de simplification, les segmentations bi-marginales sont ensuite fusionnées par intersection et celle-ci est simplifiée par regroupement de régions en utilisant les informations bi-composantes comme trois sources d'informations indépendantes. La première étape bi-marginale repose sur une méthode de coalescence morphologique non supervisée qui recherche les couleurs dominantes d'un histogramme 2D. Cela

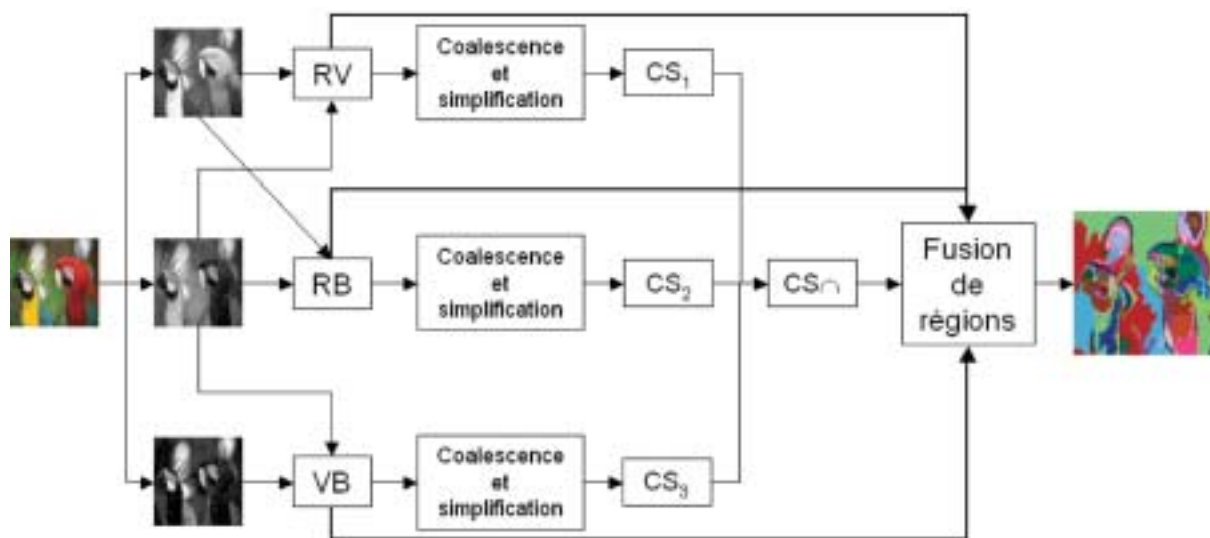


Figure 1. Schéma général de l'approche bi-marginale proposée de la segmentation d'images couleur.

permet de construire automatiquement trois cartes de segmentation distinctes. Dans une deuxième étape, les trois cartes obtenues sont tout d'abord combinées par intersection. Or le résultat du processus d'intersection des segmentations conduit à une sursegmentation de l'image, une étape de fusion des régions est alors opérée jusqu'à une condition d'arrêt qui peut être par exemple un prédicat de qualité de segmentation, un nombre de régions finales à atteindre, etc. Quelle que soit la technique retenue, force est de constater que la fusion est très souvent opérée selon le paradigme du choix forcé : sélection de la région la plus proche de la région à fusionner selon un critère de similarité. Or le choix de ce critère est drastique, *i.e.*, un critère de piètre performance peut induire une qualité de segmentation médiocre. Qui plus est, même si ce critère est de très bonne qualité, le résultat final peut être décevant d'un point de vue qualité. Ce choix d'affectation d'une région par rapport à une autre est souvent pris relativement tôt dans le processus de fusion. Dès lors que l'on a sélectionné la région la plus proche au sens d'un critère de similarité, la fusion est opérée. Or plusieurs recherches menées depuis quelques années, ont permis de mettre en évidence qu'une prise de décision tardive en élargissant le champ de sélection des régions candidates peut singulièrement améliorer les résultats de la segmentation. Par exemple Lefevre *et al.* [LVC00] ont ainsi développé un algorithme de segmentation basé sur la combinaison de Dempster-Shafer pour la segmentation de mélanomes. Cependant, même si les taux de classification sont très bons, le nombre de classes utilisées reste très faible. Avant de procéder à l'étape de fusion de régions selon la combinaison de Dempster-Shafer, une élimination des petites régions (*i.e.*, dont la taille est inférieure à la taille d'un élément structurant) est opérée. Cette élimination est contrôlée par deux critères de qualité qui empêchent la disparition d'une petite région si celle-ci est d'une importance certaine dans l'évaluation de la qualité de la segmentation. À l'issue de ce processus de fusion des petites régions, une nouvelle carte de segmentation est générée. Pour chaque région, au lieu de sélectionner seulement le plus proche candidat, un ensemble de candidats potentiellement intéressants est construit en fonction de plusieurs critères de similarité sur lesquels un relâchement est effectué (ce qui permet de sélectionner k régions candidates à la fusion). La sélection du candidat est ainsi repoussée et la prise de décision n'est effectuée qu'après combinaison d'un certain nombre de prédicats. À l'issue de cette combinaison, le meilleur candidat à la fusion avec la région sélectionnée sera déterminé. Les désavantages majeurs des approches avec fusion sont d'une part qu'elles sont généralement supervisées et d'autre part qu'elles nécessitent un certain nombre de paramètres difficiles à régler. L'approche que nous proposons ici utilise une fonction d'énergie permettant de quantifier la qualité d'une segmentation. Cette fonction d'énergie est utilisée afin de régler les paramètres de chaque méthode (la coalescence et la fusion) et permet d'obtenir une approche totalement non supervisée qui n'utilise aucune information *a priori*.

L'article s'articule de la façon suivante. Dans la section 2, après une introduction sur la coalescence d'histogrammes, nous pro-

posons une méthode morphologique non supervisée de coalescence opérant sur des histogrammes bi-dimensionnels. Afin de déterminer automatiquement la meilleure coalescence, une mesure de comparaison de segmentations basée sur une fonction d'énergie est introduite. Cette dernière étant utilisée tout au long de l'article comme un indicateur de performance de la segmentation [BLC03] afin d'en régler les différents paramètres. La section 3 présente le principe de la fusion des cartes de segmentation selon la théorie de Dempster-Shafer. L'intersection des trois cartes de segmentation obtenues *via* l'approche bi-marginale est simplifiée tout d'abord par l'élimination des régions dont la taille est inférieure à un élément structurant. Le principe de la fusion des régions par regroupement est ensuite détaillé à savoir la construction de la fonction de masse et le critère de terminaison. Nous concluons finalement sur l'apport de la méthode.

2. Coalescence d'histogrammes

Pour réaliser un clustering d'une image couleur plusieurs stratégies peuvent être employées. Dans tout cet article nous utiliserons le terme de coalescence pour désigner le terme anglais de clustering. Les stratégies de coalescence diffèrent généralement par la dimensionnalité des données utilisées pour diviser l'image. Pour le cas unidimensionnel, les histogrammes de chaque composante couleur sont pris séparément et la méthode de coalescence consiste à déterminer des seuils découpant chaque histogramme [Cel90, LC02]. Cette méthode est basée sur la supposition que les différents objets d'une image appartiennent à différents pics de chaque histogramme. Pour le cas tridimensionnel la nature vectorielle des données est prise en compte et la coalescence opère directement sur des données multivariées (les couleurs réparties dans l'espace couleur 3D). C'est la répartition spatiale des couleurs dans l'espace couleur associé qui est utilisée pour diviser l'histogramme 3D [Soi96, GSD01a, PYL98, PZLB93]. On suppose alors que les couleurs des différents objets sont réparties en différents groupes autour de couleurs dominantes qui représentent les centres des classes de l'histogramme. Les méthodes de coalescence uni et tri dimensionnelles nécessitent toutes deux la donnée d'un paramètre : le nombre de classes (*i.e.* le nombre de couleurs dominantes) de la coalescence finale. La méthode unidimensionnelle est handicapée par le fait qu'un regroupement de couleurs présent dans l'histogramme 3D n'est pas toujours présent dans chacun des histogrammes des composantes couleur. La méthode tridimensionnelle est quant à elle handicapée par la répartition spatiale des données qui peuvent être très éparpillées mais également par la complexité de l'algorithme de recherche des regroupements. Un autre problème inhérent cette fois-ci aux images couleur est lié à la représentation de celles-ci. Une image peut-être

représentée dans différents espaces couleur, chaque espace ayant des propriétés plus ou moins intéressantes. Ce choix est généralement difficile. Nous prenons ici le parti pris d'une méthode de coalescence devant fonctionner sans *a priori* sur les propriétés de l'espace couleur et qui doit fonctionner quel que soit celui-ci. Bien entendu ce choix a une influence, on consultera [LC03] pour une étude de la méthode de coalescence proposée dans différents espaces couleur. Par soucis de simplicité, nous travaillerons ici uniquement dans l'espace RVB .

2.1. Histogrammes bidimensionnels

Une alternative intéressante aux deux précédentes méthodes de coalescence 1D ou 3D réside dans l'utilisation d'histogrammes bidimensionnels [KSH01, XGSDL03] (histogrammes 2D). Ceux-ci utilisent conjointement deux composantes couleur, par exemple RV , RB , VB dans l'espace couleur RVB . Ceci a plusieurs avantages. Tout d'abord l'aspect très dispersé des couleurs dans l'histogramme 3D est partiellement réduit et surtout la complexité de l'algorithme de coalescence est grandement réduite. L'utilisation d'histogrammes 2D permet de disposer d'une méthode intermédiaire à celles uni et tri dimensionnelles. Une autre atout majeur de l'utilisation d'histogrammes 2D réside dans le fait qu'un histogramme 2D est représenté par une image en niveaux de gris et par conséquent les algorithmes classiques de traitement d'images en niveaux de gris peuvent être directement appliqués afin de diviser l'histogramme 2D.

2.2. Coalescence morphologique non supervisée

Dans cette section, nous proposons une méthode de coalescence non supervisée d'histogrammes 2D basée sur une stratégie morphologique afin de diviser une image en un certain nombre de classes. La classification par morphologie mathématique n'est pas nouvelle. Elle a été initialement proposée par Soille [Soi04]. Tous les classificateurs morphologiques considèrent un histogramme comme une fonction de densité de probabilité générale qui est représentée par une image numérique. L'avantage d'utiliser des classificateurs morphologiques réside

dans l'absence d'hypothèse de modes multi gaussiens qui s'avère généralement très éloignée de la réalité [GSD01b]. L'utilisation comme classificateur de la Ligne de Partage des Eaux (LPE) a ainsi été proposée par Soille [Soi96, Soi04] : les minima de plus forte dynamique sont sélectionnés comme marqueurs, une reconstruction morphologique de l'histogramme permet une imposition des minima et finalement une LPE est calculée sur l'histogramme reconstruit. La méthode de Soille présente un paramètre *a priori* dont il est préférable de s'affranchir à savoir le nombre de classes. Nous proposons donc une méthode non supervisée où la LPE opère sur une fonction de distance aux marqueurs des classes.

Revenons tout d'abord sur les histogrammes 2D. Ceux-ci sont obtenus par la projection de l'histogramme 3D sur les plans bi-composantes. La figure 2(1) présente l'histogramme 3D de la figure 4(3) et les figures 2(2)-(4) les trois histogrammes 2D obtenus par projection bi-composante. On remarque que les projections ne contiennent évidemment pas les mêmes informations. Les images des histogrammes 2D des figures 2(2)-(4) sont inversées pour une meilleure visualisation. Pour diviser un histogramme 2D, nous supposons (comme pour les autres méthodes de coalescence) que les différents objets de l'image couleur à diviser sont présents au niveau des pics dominants de l'histogramme. Ceci qui est caractérisé par des régions de niveaux de gris homogènes dans l'image représentant l'histogramme 2D. Les principaux pics (*i.e.* les principales régions connexes) de l'histogramme 2D sont considérés comme les centres des classes à extraire.

La principale difficulté d'une stratégie de coalescence réside dans la détermination du nombre de pics (*i.e.* du nombre de classes) sans aucun *a priori* sur leur nombre. Ce problème se retrouve quel que soit la dimensionnalité de l'histogramme. La méthode que nous proposons permet de s'affranchir de ce problème. Nous utilisons une stratégie basée sur des opérations morphologiques qui sont rapides, efficaces et qui ont prouvé leur grande robustesse sur les images à niveaux de gris [Ser82]. La coalescence d'un histogramme 2D se fait en plusieurs étapes. Puisque les histogrammes sont généralement bruités (à cause de l'éparpillement des couleurs), ceux-ci sont lissés par un filtre exponentiel symétrique (qui prend un paramètre $\beta \in [0, 1]$ qui permet de lisser plus ou moins l'image, à β faible, lissage fort).

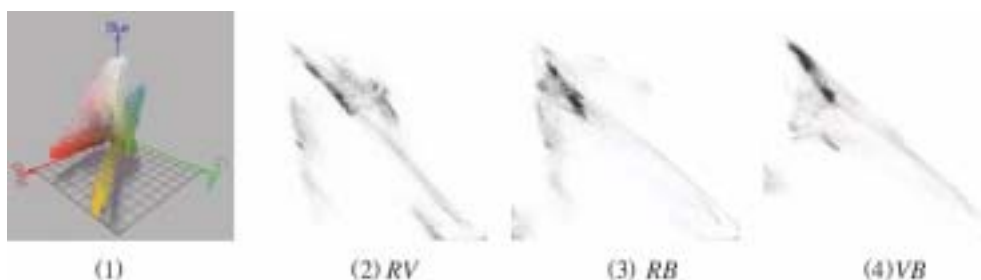


Figure 2. Un histogramme 3D et ses trois projections 2D bi-composantes (inversées).

Le résultat de ce lissage est reconstruit dans l'image originale afin d'obtenir une version simplifiée et régularisée de l'histogramme. À partir de l'image régularisée, nous cherchons à extraire les principaux pics. Les pics d'un histogramme 2D se traduisent par des pixels de forte intensité par rapport à leur voisinage. Dans l'histogramme beaucoup de pixels ont une valeur nulle statuant de la non présence d'une couleur dans l'image originale. Une opération d'érosion appliquée directement sur l'image de l'histogramme permet d'éliminer certaines couleurs de l'histogramme, ces couleurs éliminées étant considérées comme non dominantes [GBLP00]. Les couleurs dominantes correspondent quant à elles à des maxima dans l'image, des opérations d'érosions répétées permettent de réduire l'histogramme à ses principales couleurs dominantes. Itérer ce principe permet d'extraire les couleurs dominantes et l'ensemble des érodés ultimes permet donc d'extraire automatiquement de l'histogramme toutes les couleurs dominantes correspondant à des pics dans l'histogramme. L'ensemble des érodés ultimes correspond aux ensembles connexes de pixels qui disparaissent en dernier lors d'un procédé itératif d'érosions. Cet ensemble d'érodés ultimes représente les centres des classes de l'histogramme 2D. Ces centres sont labellisés et utilisés comme marqueurs pour une opération de ligne de partage des eaux réalisée sur une fonction de distance aux marqueurs. Ceci permet d'obtenir une division de l'histogramme 2D en un certain nombre de classes (autant que d'érodés ultimes). Toutes ces étapes sont décrites par la figure 3. L'histogramme 2D RV original (figure 3(1)) est simplifié par reconstruction (figure 3(2)), ses érodés ultimes sont extraits (figure 3(3)), labellisés et une ligne de partage des eaux

est calculée sur une fonction de distance (figure 3(4)) aux marqueurs (les érodés ultimes). La coalescence finale de l'histogramme est donnée par la figure 3(5). Un élément structurant carré 3×3 est utilisé ici.

À partir de cet histogramme 2D divisé, une carte de segmentation est obtenue car à chaque région de l'histogramme correspond un ensemble de couleurs de l'image originale : à chaque vecteur RVB on associe la classe correspondante dans l'histogramme 2D divisé. Ceci est illustré par la figure 4. La figure 4(2) donne la carte de segmentation obtenue à partir de la coalescence de l'histogramme 2D RV . Ces résultats ont été obtenus pour une valeur de $\beta = 0.5$. La figure 4(1) est l'image originale et la figure 4(3) représente une image quantifiée obtenue en affectant à chaque région sa moyenne couleur. On peut constater que le résultat est proche de l'image originale attestant de la qualité de la coalescence. La méthode de coalescence est appliquée à chaque histogramme 2D ce qui permet d'obtenir trois cartes de segmentation. Les cartes de segmentation sont relabellisées afin d'obtenir une partition en régions connexes et non juste une classification des pixels. Le nombre de régions de chaque carte de segmentation est donc supérieur au nombre de classes car les pixels d'une même classe ne sont pas forcément connexes dans l'image.

2.3. Aspect multi-échelle et minimisation d'énergie

La méthode de coalescence morphologique non supervisée permet de découper un histogramme 2D en un nombre de classes

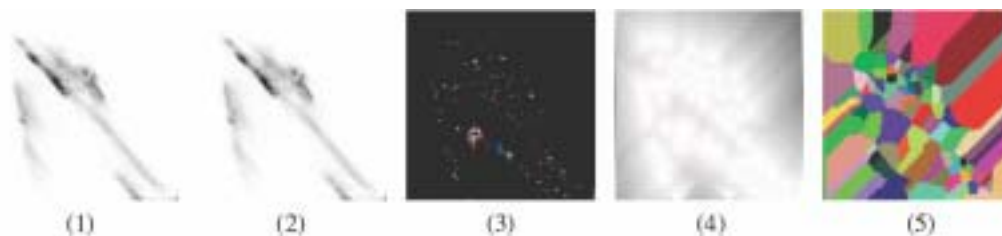


Figure 3. Les étapes de la coalescence d'un histogramme 2D.



Figure 4. (1) Une image couleur; (2) la coalescence de l'histogramme 2D RV , (3) l'image quantifiée des régions.

automatiquement déterminé mais cette méthode utilise un unique paramètre β qui précise le degré du lissage. Plus le lissage est fort, plus l'image correspondant à l'histogramme est simplifiée et moins on obtiendra de régions par la coalescence. De même moins le lissage est fort, plus on obtiendra de régions finales. Le paramètre β agit donc comme un paramètre multi-échelle qui permet de régler la complexité de la carte de segmentation finale car il permet de régler le nombre de couleurs dominantes qui seront extraites par la phase de coalescence. Ce paramètre peut être réglé par l'utilisateur afin de régler le niveau de précision de la segmentation. Cependant il serait plus intéressant de disposer d'une méthode automatique permettant de régler automatiquement ce paramètre afin de trouver le niveau de la stratégie multi-échelle qui soit le plus satisfaisant possible. L'ensemble des segmentations obtenues à différents niveaux d'échelle correspond à une hiérarchie de partitions, le paramètre β servant à générer les différents niveaux de la hiérarchie. La difficulté pour régler le paramètre β réside dans le fait que nous devons disposer une méthode de comparaison de deux cartes de segmentation.

Ceci rejoint un principe exprimé par Koepler, Lopez, et Morel [KLM94] sous le nom de principe de comparaison :

« Nous adoptons un principe sans lequel aucune discussion sur la segmentation ne peut avoir lieu, et que nous appelons le principe de comparaison. Selon ce principe, étant données deux segmentations différentes d'une même donnée, nous pouvons toujours décider laquelle des deux doit être considérée comme meilleure que l'autre (ou équivalente). Nous supposons donc l'existence d'un ordre total sur l'ensemble des segmentations possibles, ce qui se traduit par le fait qu'il existe une fonctionnelle réelle E telle que si $E(M_1) < E(M_2)$, alors la segmentation M_1 doit être considérée comme meilleure que la segmentation M_2 ».

Selon ce principe, la formulation générale d'un problème de segmentation revient à spécifier une fonction d'énergie [GMC01], $E : M \times I \rightarrow R^+$ sur l'ensemble des segmentations M possibles et sur l'image initiale I et à définir une solution du problème de segmentation d'une image I comme la segmentation $M^*(I)$ qui minimise cette énergie, c'est-à-dire par

$$M^*(I) = \underset{M_k \in M}{\operatorname{argmin}} E(M_k, I)$$

Ceci soulève donc un problème de compromis car il n'existe certainement pas d'unique meilleure segmentation d'une image. Il va donc nous falloir faire un compromis entre la qualité de la segmentation (par rapport à l'image initiale) et sa complexité. La qualité de la segmentation décrit la capacité à décrire fidèlement les données initiales et la complexité de la segmentation décrit la simplicité du modèle [MS89]. L'énergie peut alors se mettre sous la forme suivante [GMC01] :

$$E(M, I) = C(M, I) + D(M, I)$$

On peut faire un parallèle de la modélisation de cette énergie avec deux notions : celle de classification et celle liée aux cou-

pures dans les graphes. Pour la classification $D(M, I)$ représente l'erreur intra-classe et $C(I, M)$ représente l'erreur inter-classe. Pour les coupures de graphe [BVZ01], $D(M, I)$ représente un terme d'attache aux données et $C(I, M)$ représente un terme de régularisation. Nous proposons une formulation d'énergie basée sur ces modèles et sur les travaux de Liu [LY94] et Borsotti [BSC98] permettant d'évaluer la qualité d'une segmentation. La valeur d'énergie $E(M, I)$ permet de disposer d'une mesure de qualité de la segmentation et l'on cherche à minimiser cette mesure afin de trouver la meilleure segmentation possible. L'énergie que nous proposons de minimiser est la suivante :

$$E(M, I) = C(M, I) + D(M, I) = \alpha_d \times \left(\sum_{j=1}^{\Theta(G)} \sum_{k>j}^{\Theta(G)} L_{j,k} \right) + \alpha_c \times \left(\sum_{i=1}^{\Theta(G)} \frac{e_i^2}{1 + \log(N_i)} \right) \quad (1)$$

avec $\alpha_c = \frac{\sqrt{\Theta(G)}}{1000 \times h \times w}$ et $\alpha_d = \frac{1}{\operatorname{card} L}$ qui sont des poids de normalisation.

Les notations sont les suivantes : G désigne le graphe d'adjacence des régions obtenu à partir du modèle de segmentation M , $\Theta(G)$ donne le nombre de régions du modèle, L désigne l'ensemble des arêtes, $\operatorname{card} L$ donne le nombre d'arêtes du graphe G , h et w donnent la hauteur et la largeur de l'image I , $N_i = \operatorname{card} \mathcal{R}_i$ donne le nombre de pixels de la région numéro i notée \mathcal{R}_i , $L_{j,k}$ est le poids de l'arête reliant les deux régions j et k dans le graphe d'adjacence, e_i^2 quantifie l'erreur couleur moyenne de la région i (la somme des écarts quadratiques entre les vecteurs couleur de l'image originale et la moyenne couleur de la région numéro i).

Ceci est différent du formalisme proposé par Guigues [GMC03, GMC01, MPF04] qui utilise un critère d'énergie basé sur la fonctionnelle constante par morceaux de Mumford et Shah [MS89] et dont le terme de complexité dépend uniquement du modèle M par l'expression de la longueur des frontières des régions. Cela repose sur un *a priori* fort qui sous-entend que les frontières doivent être simples et de plus la longueur des frontières n'est pas forcément significative de la simplicité des régions. Nous préférons donc nous baser sur le formalisme de Boykov [BVZ01] qui prend en compte la structure topologique du graphe d'adjacence dans le terme de complexité $C(M, I)$. Nous avons modélisé dans la fonction d'énergie deux quantités qui permettent d'exprimer le compromis entre fidélité aux données et complexité. Tout d'abord, $D(M, I)$ correspond à un terme d'attache aux données et est une extension de la mesure de Borsotti [LY94, BCS98] afin de prendre en compte la distribution des couleurs des régions en fonction de leur taille ainsi que le nombre de régions, cela correspond à une énergie interne aux régions comme cela a été proposé initialement par Mumford et Shah [MS89], l'image étant considérée constante par morceaux. $C(M, I)$ décrit la complexité du graphe en tenant compte non seulement de sa structure topologique mais également des valeurs associées aux arêtes. Les arêtes du graphe sont pondérées par une distance entre deux régions, ce qui permet de décrire la similarité entre elles. Nous avons repris ici celle proposée par Salembier exprimée par [SG00, SGG97] :

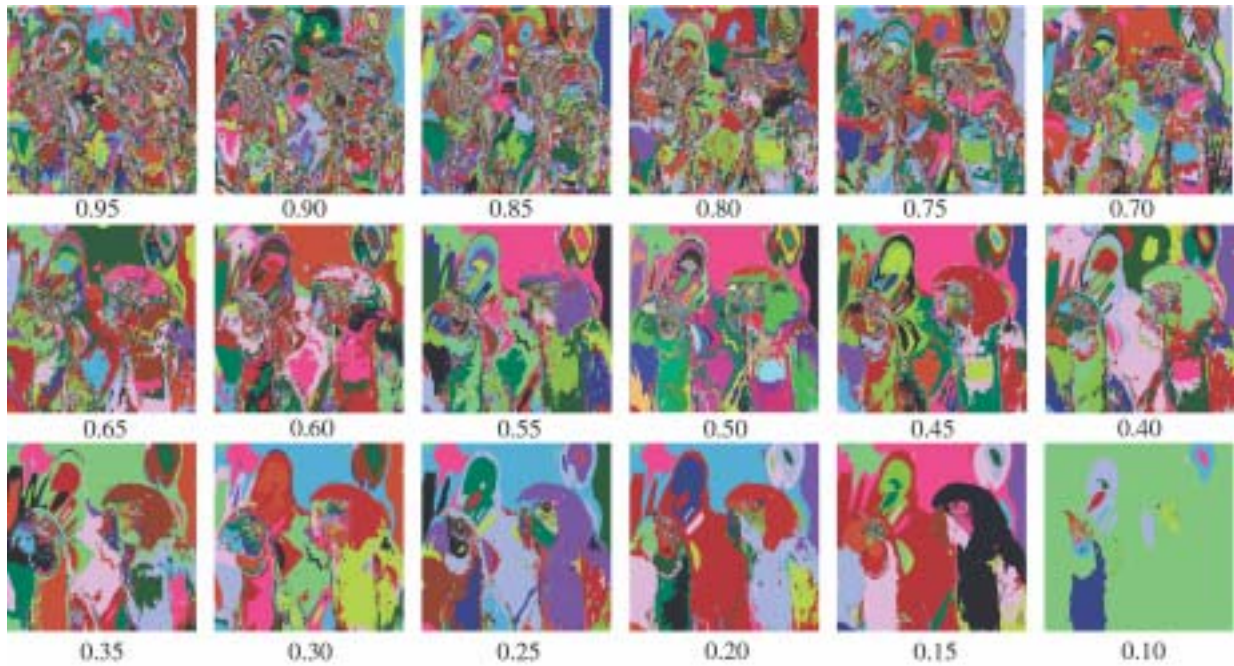


Figure 5. Cartes de segmentation obtenues par coalescence d'un histogramme 2D à différents niveaux d'échelle (différentes valeurs de β)

$$L_{i,j} = N_i d_E(\mathcal{R}_i, \mathcal{R}_i \cup \mathcal{R}_j) + N_j d_E(\mathcal{R}_j, \mathcal{R}_i \cup \mathcal{R}_j)$$

où d_E désigne la norme L_2 entre les couleurs moyennes des régions considérées.

Maintenant que nous disposons d'une énergie permettant de quantifier la qualité d'une segmentation en vertu du principe de comparaison cité plus haut, nous pouvons utiliser ceci afin de déterminer automatiquement le niveau d'échelle approprié à notre segmentation multi-échelle. Comme la mesure d'énergie s'applique sur une carte de segmentation, le résultat de la coalescence n'est donc pas utilisé directement. En effet une même classe de pixels peut se trouver à différents endroits dans une image constituant donc des régions différentes. Les cartes de

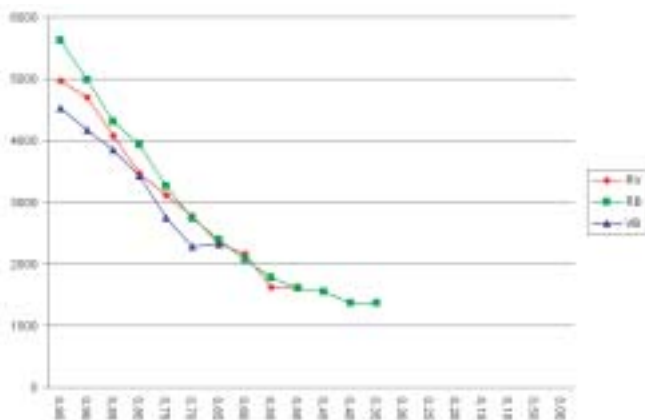


Figure 6. Variation d'énergie de la coalescence des trois histogrammes 2D (E en fonction de β).

segmentations que nous utilisons correspondront donc à une labélisation des composantes connexes de la coalescence, le nombre de régions est alors généralement supérieur au nombre de classes extraites par la coalescence. On obtient donc une hiérarchie de partitions dont le nombre de régions décroît lorsque β augmente. L'énergie E décroissant également en fonction du paramètre d'échelle β , il existe donc toujours une solution triviale qui ne nous intéresse pas à savoir l'image elle-même qui est toujours solution du problème de minimisation d'énergie. De plus la solution n'est pas forcément unique: on cherchera donc un minimum local correspondant à une description simplifiée de l'image permettant un compromis entre attache aux données et régularisation car le terme $C(M, I)$ avantage les modèles avec peu de régions et $D(M, I)$ ceux avec beaucoup de régions. Pour déterminer automatiquement le niveau d'échelle, il suffit donc de chercher un minimum local de la fonction d'énergie. Les différents modèles de segmentations sont obtenus en faisant varier le paramètre multi-échelle β par pas de 0.05 à partir de 1.0. On réalise une segmentation de l'histogramme 2D à chaque niveau d'échelle jusqu'à tomber sur un minimum local d'énergie. Grâce à cette méthode, nous pouvons nous affranchir du paramètre β dont la valeur « optimale » est trouvée par rapport à l'énergie modélisée. La figure 5 donne les segmentations obtenues par coalescence de l'histogramme 2D RG à 18 niveaux différents d'échelle (de 0.95 à 0.10). On remarque que les premiers niveaux sont sur-segmentés et que les derniers sont sous-segmentés. La minimisation d'énergie permet de trouver un compromis entre une sur et une sous segmentation en tenant compte de l'attache aux données initiales et de la complexité de la segmentation obtenue. La figure 6 donne la variation d'énergie

pour les trois histogrammes 2D de la figure 4(1). La variation d'énergie n'est donnée que jusqu'à l'obtention d'un minimum local qui correspond à la fin de la recherche de la partition « optimale » par rapport à la coalescence et la fonction d'énergie considérée. Évidemment ce minimum local est différent pour les trois énergies calculées pour les trois histogrammes 2D car ils ne contiennent pas les mêmes informations. On constate donc que minimiser l'énergie E permet de trouver facilement la valeur souhaitée de β pour chacun des trois histogrammes. La segmentation finale obtenue pour l'histogramme 2D RV de la figure 4(1) est donnée par la figure 4(2): le meilleur niveau est obtenu pour $\beta = 0.50$. Une constatation s'impose: ceci n'est le meilleur niveau qu'au vu de la fonction d'énergie que nous avons définie. Une segmentation plus simplifiée aurait pu être préférée mais le terme d'attache aux données $D(M, I)$ permet de ne pas trop s'éloigner de l'image originale. On peut le constater par la figure quantifiée 4(3) très proche de la figure 4(1). Le but premier de l'étape de coalescence automatique n'est pas d'obtenir directement une segmentation finale, mais une partition fine qui soit à la fois suffisamment compacte et proche de l'image originale. Cette partition fine sera simplifiée par la suite.

2.4. Simplification par fusion de régions

À partir du schéma multi-échelle précédent, nous obtenons de façon automatique non supervisée et sans paramètres, trois segmentations à partir des trois histogrammes 2D couleur. Les segmentations obtenues possèdent cependant trop de régions pour décrire de façon compacte l'image originale. En effet la coalescence néglige le contexte spatial car seule la cohérence des nuages de points-couleurs est utilisée. Cela a une conséquence importante : des nuages de points peuvent ne pas correspondre à des régions significatives dans l'image que l'on veut segmenter et un même nuage de points extrait comme une seule classe peut très bien être réparti sur plusieurs zones non connexes de l'image. Comme on ne peut pas garantir que les classes de pixels extraites par la coalescence correspondent effectivement aux régions de l'image, une étape supplémentaire est nécessaire afin d'introduire la prise en compte de l'information spatiale. Ce type

d'approche en deux étapes à savoir, sélectionner les couleurs dominantes de l'espace couleur et construire des régions spatialement cohérentes a été également proposé par Macaire qui prend en compte à la fois les couleurs dominantes et les degrés de connexité et d'homogénéité de celles-ci [Mac04]. Comme le remarque Macaire [Mac04], pour regrouper des pixels appartenant à des mêmes régions extraites à partir des classes de pixels de la coalescence, il est intéressant de construire un graphe d'adjacence des régions afin de construire des régions tenant compte des propriétés colorimétriques et spatiales des pixels. Fontaine [Fon04] a proposé une analyse floue du graphe d'adjacence des régions pour réaliser ceci. Nous proposons ici une autre approche qui se focalise sur la simplification des cartes de segmentation obtenues par coalescence en fusionnant deux à deux les régions de couleurs similaires. Pour cela nous réutilisons la fonction d'énergie $E(M, I)$ définie ci-dessus pour simplifier à nouveau les trois segmentations obtenues. Nous pouvons réutiliser la même fonction d'énergie car le schéma multi-échelle induit par la fusion de régions est différent de celui induit par la coalescence et la hiérarchie de partitions est donc différente. La réglage de β procède à une minimisation qui permet de découvrir les classes de couleurs dominantes et la fusion des régions permet ensuite de prendre en compte la cohérence spatiale dans l'image. De plus on peut dire que la minimisation se fait dans une nouvelle direction de recherche qui ne peut pas être atteinte par le réglage de β . Pour chaque segmentation, une opération itérative de fusion de régions est réalisée. Une implémentation efficace du processus de fusion basé sur une file hiérarchique de priorité et un graphe d'adjacence de régions est utilisée [GSG98]. Dans notre approche, la segmentation obtenue à l'étape précédente est considérée comme la partition la plus fine. Lors du processus de fusion, chaque région est décrite par sa moyenne couleur et le critère de terminaison est basé sur la fonction d'énergie $E(M, I)$. Les régions sont fusionnées deux à deux itérativement jusqu'à ce que l'on obtienne un certain nombre de valeurs consécutives identiques de la fonction d'énergie (à ε près). Ici nous considérons que le processus de fusion doit se terminer lorsque l'on obtient $\sqrt{\Theta(G)}$ valeurs consécutives identiques. Il est à noter que $\Theta(G)$ varie dynamiquement tout au long du processus de fusion car on procède à



Figure 7. Cartes de segmentation et graphes d'adjacence de régions avant et après fusion des régions similaires.

une décimation du graphe. Pour plus de détails sur ce processus de fusion de régions, nous renvoyons le lecteur intéressé aux articles suivants [LC03, Lez04] qui décrivent en détail le principe de fusion utilisé. Chaque segmentation est donc simplifiée en fusionnant les régions identiques jusqu'à une minimisation de l'énergie statuant que la qualité de la segmentation est suffisante tout en étant plus compacte. Le schéma de segmentation d'histogramme 2D que nous proposons consiste donc à :

- Trouver pour chaque histogramme 2D par coalescence automatique une partition fine fidèle à l'image qui soit suffisamment compacte,
- Simplifier cette partition fine en fusionnant les régions similaires.

Ces deux étapes minimisent la même fonction d'énergie. La première étape cherche une partition fine satisfaisante et la deuxième cherche une simplification satisfaisante de la partition fine. La figure 7 présente la simplification par fusion de régions réalisée. Les images de la première ligne donnent respectivement les trois partitions fines et les trois graphes d'adjacence. On remarque tout de suite pourquoi nous utilisons le terme de partition fine : les graphes permettent de reconnaître facilement les données initiales attestant de la finesse de la partition obtenue. Les images de la seconde ligne présentent les résultats obtenus après fusion pour respectivement les cartes de segmentation et les graphes d'adjacence. On constate que les graphes ont bien été simplifiés car la segmentation finale obtenue est une représentation plus compacte de la partition fine.

Il est évident de noter que segmenter une projection 2D d'un histogramme 3D permet d'obtenir moins de classes que lors d'une coalescence dans l'espace 3D : les classes observables en 3D se chevauchent en 2D. En conséquence, aucune des segmentations 2D obtenues n'est suffisante en tant que telle, un procédé de fusion des cartes de segmentation obtenues est donc requis. La figure 10 illustre ce fait sur une image comportant peu de régions et dont certaines couleurs n'apparaissent pas dans tous les plans bi-composantes (les deux zones vertes dans les disques). Une étape de fusion est donc nécessaire [XGDL03] : l'intersection des trois cartes de segmentation (Figure 10) permet de construire une carte compatible avec les trois cartes obtenues. La figure 10 présente sur la première ligne les résultats de la coalescence pour les trois histogrammes, sur la deuxième ligne les résultats de la fusion des régions, sur la troisième ligne l'intersection des cartes simplifiées, non simplifiées et l'image originale. Les deux zones vertes des disques ne peuvent être discriminées que dans l'histogramme *BV* mais cela produit une sur-segmentation car l'image est bruitée. La simplification par fusion de régions rétablit la cohérence spatiale des régions et l'intersection est grandement simplifiée par rapport au résultat obtenu sans fusion de régions. Les couleurs présentées sur les cartes de segmentation de la figure 10 sont des fausses couleurs et n'ont aucune signification particulière. On remarque donc d'une part que le réglage de β permet d'obtenir des segmentations fines sur-segmentées car ne prenant pas en compte la cohérence spatiale des régions et d'autre part que la fusion au

niveau du graphe d'adjacence permet de rétablir la cohérence spatiale des régions et d'en minimiser le nombre. On aurait très bien pu s'affranchir de cette étape de simplification par fusion des régions dans le graphe d'adjacence mais cela augmenterait considérablement la complexité de la stratégie de fusion ultérieure des cartes de segmentation : de nombreuses fusions évitables sont éliminées et le nombre de régions diminué ce qui permet de réduire le nombre d'hypothèses ultérieures du cadre de discernement. Cela est facilement observable sur la figure 10 où sont présentées les cartes sans et avec simplification ainsi que l'intersection avec et sans simplification : le nombre de régions est grandement réduit et les régions sont spatialement cohérentes.

3. Fusion de cartes de segmentation selon la théorie de Dempster-Shafer

L'objectif à présent est de fusionner de manière non supervisée les cartes de segmentation obtenues lors de l'étape précédente afin d'obtenir une carte finale de segmentation dont la qualité est contrôlée selon un paradigme. Cette approche a été appliquée de nombreuses fois. On peut citer, par exemple, les travaux de Mascle *et al.* qui ont développé une approche similaire permettant d'obtenir une classification multisource non supervisée d'images radar [MBVM97].

Dans cette étude, le processus, qui est appliqué sur des images couleur, se décompose en deux étapes : 1) le calcul de l'intersection des diverses cartes de segmentation avec élimination des régions de taille inférieure à un seuil λ_T , 2) fusion des régions deux à deux selon un critère de similarité et selon la combinaison de Dempster-Shafer jusqu'à un critère de terminaison.

3.1. Intersection des images segmentées et élimination des petites régions

À partir des trois cartes de segmentation obtenues $CS_{i \in [1,2,3]}$, une quatrième carte CS_{\cap} est générée en tant que résultat de l'intersection des trois premières cartes (Figure 8),

$$CS_{\cap} = \bigcap_{i=1}^3 CS_i. \quad (2)$$

Dès lors la carte ainsi générée contient N_s classes $CS_{\cap} = \{C_i, i \in [1, \dots, N_s]\}$, où C_i représente un ensemble i de pixels. L'un des inconvénients majeurs inhérent à la construction de la carte de segmentation CS_{\cap} est une sur-segmentation

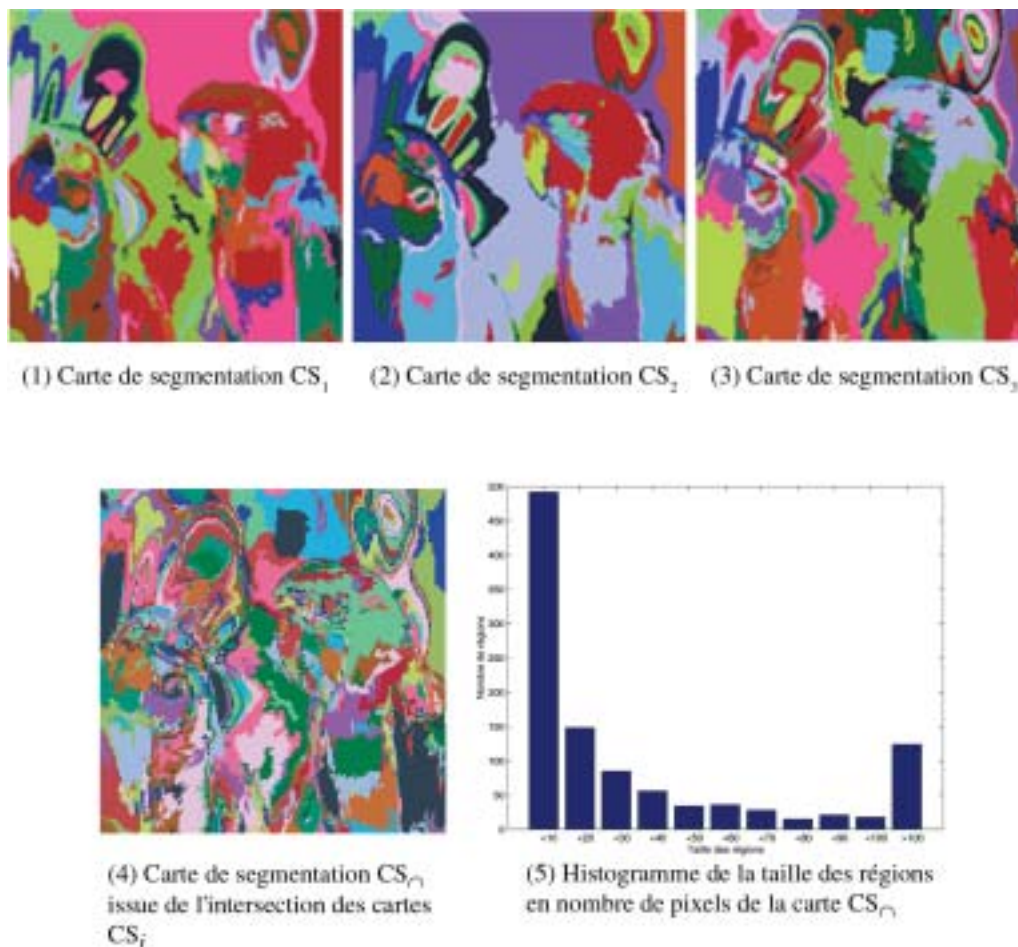


Figure 8. Image résultante de l'intersection des cartes de segmentation.



Figure 9. Image après élimination des régions de taille inférieure au seuil λ_T .

nécessairement induite par ce processus, comme il est possible de l'observer sur la Figure 8(4). Ainsi, de nombreuses régions dont la taille en pixels est inférieure à celle d'un élément structurant 3×3 soit $\lambda_T = 10$ (Cf. Figure 8(5)) sont créées.

Une fusion des régions dont la taille est inférieure au seuil λ_T est opérée selon un critère de similarité d . Ce critère est basé sur une combinaison selon la norme L_2 de la distance colorimé-

trique d_E entre le barycentre couleur de la région adjacente \mathcal{R}_j avec celui de la région considérée \mathcal{R}_i et de la mesure de la dispersion spatio-colorimétrique D_{SC} du nuage couleur de la région \mathcal{R}_i par rapport à la région adjacente \mathcal{R}_j , tel que :

$$d(\mathcal{R}_i, \mathcal{R}_j) = \sqrt{d_E^2(\mathcal{R}_i, \mathcal{R}_j) + D_{SC}^2(\mathcal{R}_i, \mathcal{R}_j)} \quad (3)$$

La mesure d_{SC} utilisée est définie par :

$$D_{SC}(\mathcal{R}_i, \mathcal{R}_j) = \min_{(p,q) \in (\mathcal{R}_i, \mathcal{R}_j)} d_{sc}(p, q) \quad (4)$$

avec

$$d_{SC}(p, q) = \frac{\| \sum_{k=1}^3 \text{cov}_{c_k}^{\mathcal{R}_i, \mathcal{R}_j}(p, q) \|}{\sqrt{\sum_{k=1}^3 \sigma_{c_k}^{\mathcal{R}_i}(p)^2} \cdot \sqrt{\sum_{k=1}^3 \sigma_{c_k}^{\mathcal{R}_j}(q)^2}} \quad (5)$$

c_i représente la composante couleur considérée. $\sigma_{c_k}^{\mathcal{R}_i}(p)$ représente la variance du voisinage d'étude $V(p)$ du pixel spécifié sur la région \mathcal{R}_i , tandis que $\text{cov}_{c_k}^{\mathcal{R}_i, \mathcal{R}_j}(p, q)$ représente la covariance du voisinage d'étude $V(p)$ du pixel spécifié sur la région \mathcal{R}_i par rapport au voisinage d'étude $V(q)$ du pixel spécifié sur la région \mathcal{R}_j .

Cette formule n'a de sens qu'à partir du moment où les valeurs de variance $\sum_{k=1}^3 \sigma_k^{\mathcal{R}_i}(p)^2$ et $\sum_{k=1}^3 \sigma_k^{\mathcal{R}_j}(q)^2$ sont suffisamment grandes par rapport à une valeur seuil δ donnée. Si les deux variances sont inférieures à δ , on considérera qu'il n'y a pas de différence de corrélation entre les deux régions aux points p et q spécifiés, i.e. $d_{SC}(p,q) = 0$. A contrario, si l'une des deux variances est supérieure à ε , on considérera que la différence est maximale, i.e. $d_{SC}(p,q) = 1$.

La mesure $d(\mathcal{R}_i, \mathcal{R}_j)$ permet ainsi d'aboutir à une « nouvelle » distance entre barycentres couleur en introduisant une pondération relative à la dispersion colorimétrique de chacune des régions.

Durant cette étape toutes les régions dont la taille est inférieure à λ_T pixels ne sont pas nécessairement éliminées. En effet il se peut que l'élimination d'une région de petite taille modifie la qualité de l'image (ce qui peut être le cas lorsqu'une petite région intégrée dans une région homogène de taille plus grande est éliminée). Dans ce cas, la disparition de la petite région est tout de suite détectable et va influencer sur la qualité de la segmentation finale.

Afin de prévenir l'introduction de tels artefacts, un prédicat de fusion des régions [GSG98] a été introduit : à partir du graphe d'adjacence de la carte de segmentation initiale CS_\cap deux régions $(\mathcal{R}_i, \mathcal{R}_j)$ telles que $\text{card } \mathcal{R}_i \leq \lambda_T$ ou $\text{card } \mathcal{R}_j \leq \lambda_T$ sont fusionnées si et seulement si

$$d(\mathcal{R}_i, \mathcal{R}_j) < \min(d(\mathcal{R}'_i, \mathcal{R}'_j))$$

tel que $\text{card } \mathcal{R}'_i > \lambda_T$ et $\text{card } \mathcal{R}'_j > \lambda_T$. (6)

À l'issue de cette étape d'élimination, une nouvelle carte de segmentation est générée avec un nombre réduit de régions $N'_s < N_s$ (Cf. Figure 9).

3.2. Fusion des régions

Dans la plupart des cas, l'étape de fusion repose sur la règle du plus proche voisin : une région \mathcal{R}_i sera fusionnée à une des autres régions $(\mathcal{R}_j)_{j \in [1, \dots, M]}$ de la carte de segmentation CS_\cap sous la condition que cette fusion engendre la plus petite distorsion possible, la plupart du temps, au sens de la distance euclidienne seule. Or, ce type de sélection implique un choix binaire. Une telle règle de fusion peut se révéler trop drastique dans les cas où les distances entre la région \mathcal{R}_i et deux autres régions \mathcal{R}_j et \mathcal{R}_k sont très proches. Dans ce cas, l'une ou l'autre des régions pourrait convenir, selon une mesure de distance intégrant un ou plusieurs critères basés sur quelques propriétés du SVH (Système Visuel Humain) ou sur des critères statistiques autres que la seule distance euclidienne.

Une des améliorations possibles pour éviter cette décision dure, serait de considérer chacun des plans couleur de manière indépendante, pour obtenir ainsi une classification par composante. Les couleurs représentées dans l'espace RVB sont fortement corrélées. Ne travailler que sur les plans R , V et B (sous l'hypothèse d'indépendance) ne permet pas de tenir suffisamment compte de la corrélation initiale. Dès lors, en utilisant les plans RV , VB et BR , une partie de la corrélation initiale est conservée, et permettra donc de leur appliquer des traitements intégrant divers critères perceptuels tels que la dispersion colorimétrique, la corrélation, l'émergence, etc. afin d'améliorer la qualité du traitement final. Ainsi dans cette étude, les composantes colorimétriques sont matérialisées par les plans RV , VB et BR . Ces trois plans sont considérés comme trois sources d'information qui sont supposées être indépendantes. Une définition pratique et précise de l'indépendance est relativement difficile à trouver. Dans notre cas d'étude, on supposera l'indépendance comme cela se fait couramment en statistique.

Dans un premier temps, une étape de sélection de k régions candidates à la fusion est appliquée sur chacun des trois plans, générant alors trois séries de classes potentielles candidates à la fusion. Cette étape, tenant compte de k voisins et non pas d'un, permet de considérer l'incertitude selon chacun des plans couleur et de repousser la prise de décision. Ainsi pour une région donnée, au lieu d'obtenir une et une seule région candidate, on obtient un ensemble de régions candidates à la fusion selon un critère de similarité. Pour chacun des plans, on obtient trois ensembles de régions qu'il est nécessaire de combiner pour sélectionner la région finale qui sera fusionnée avec la région candidate.

Cette technique fait référence aux méthodes de fusion de données. Parmi tous les outils qui s'offrent à nous dans ce domaine, nous décidons d'utiliser la théorie de l'évidence (aussi appelée théorie de Dempster-Shafer ou théorie des fonctions de croyance) [Dem67, Sha76], qui permet d'une part de traiter des informations incertaines (un maximum de k classes possible sur chacun des plans couleur), et d'autre part de combiner des informations provenant de plusieurs sources (chacun des plans couleur). Le processus de segmentation par coalescence d'histogramme aboutit à des données qui peuvent être considérées comme imparfaites, puisque la qualité de la segmentation, en l'absence de vérité terrain, n'est pas optimale. L'avantage de la théorie de Dempster-Shafer repose sur la règle de combinaison qui est un outil particulièrement efficace et adapté à la fusion de ce type de données.

3.2.1. Éléments de la théorie de l'évidence

Soit $\Omega = \{\omega_1, \dots, \omega_N\}$, l'ensemble des $N \ll N'_s$ classes finales possibles pour une région \mathcal{R}_i , appelé cadre de discernement et correspondant à la carte de segmentation CS_\cap obtenue après élimination des petites régions dans notre application. Initialement, l'ensemble Ω est constitué des N'_s classes qui correspondent aux N'_s régions de la carte de segmentation. Les N classes finales sont obtenues selon le processus suivant : lorsqu'une région est sélectionnée pour être candidate à la fusion, sa classe associée est retirée du cadre de discernement ($\text{card}(\Omega) = N'_s - 1$). Ce processus est itéré jusqu'à ce que $\text{card}(\Omega) = N$. Cette étape implique alors une redéfinition du

cadre de discernement après chaque fusion d'une région. On émet alors l'hypothèse que la définition initiale de ce cadre est erronée. Au lieu de restreindre ses mesures à cet ensemble (comme le ferait la théorie des probabilités contrainte par son axiome d'additivité), la théorie de l'évidence s'étend sur l'ensemble puissance de Ω , noté 2^Ω , l'ensemble des 2^N sous-ensembles de Ω .

On définit alors une fonction de masse m qui représente la croyance que l'on accorde aux différents états du système, à un instant donné. Cette fonction est également connue en tant que masse de croyance élémentaire. Elle est définie de 2^Ω dans $[0,1]$ et vérifie les conditions suivantes :

$$\sum_{A \subseteq \Omega} m(A) = 1 \quad (7)$$

où $m(A)$ représente la croyance que l'on place dans la proposition A . Ici, $m(A)$ quantifie la croyance que l'on accorde au fait que la classe cherchée appartienne au sous-ensemble A de Ω (et à aucun autre sous-ensemble de A). Les propositions A telles que $m(A) > 0$ sont appelées éléments focaux.

Dans le cas où l'ensemble des hypothèses est exhaustif et exclusif, la masse de l'ensemble vide \emptyset est nulle. Dans ce cas, on parle de monde fermé, et la fonction de masse m est définie par

$$\sum_{A \subseteq \Omega} m(A) = 1 \text{ et } m(\emptyset) = 0 \quad (8)$$

Deux fonctions de masse initiales m_1 et m_2 représentant les informations respectives de deux sources différentes, peuvent être combinées selon la règle de Dempster [Dem67] :

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - K}, \quad \forall A \in 2^\Omega, \quad A \neq \emptyset. \quad (9)$$

K est appelé le facteur de conflit et représente le désaccord entre les deux sources. Il correspond à la masse de l'ensemble vide $K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$. Notons que la combinaison de Dempster, appelée aussi somme orthogonale et notée $m = m_1 \oplus m_2$ est commutative et associative.

Smets propose d'abandonner la normalisation et de conserver la masse de conflit qui peut être redistribuée lors d'un conditionnement. Ainsi, dans le cadre du monde fermé, qui implique que la solution se trouve dans l'ensemble Ω , la masse placée sur l'ensemble vide \emptyset est répartie sur l'ensemble des éléments focaux. Il suppose ainsi que l'origine du conflit provient d'une définition erronée du cadre de discernement, ce qui nécessite une redéfinition de ce cadre. Smets propose alors une nouvelle formulation de la combinaison :

$$\begin{cases} m_S(A) = m_\cap(A) & \forall A \subseteq \Omega \\ m_S(\emptyset) = K \end{cases} \quad (10)$$

où S représente la source considérée. Cette approche a été adoptée à de nombreuses reprises. Par exemple, dans [LCV02], Lefevre *et al.* ont développé un cadre de travail générique uni-

fiant les règles de combinaison tout en développant une famille d'opérateurs de combinaison.

Après combinaison, il reste à prendre une décision quant à l'élément le plus « probable » de Ω . Plusieurs règles de décision sont alors possibles, mais une des plus utilisées reste celle du maximum de probabilité « pignistique ». Cette règle de décision présentée par Smets [Sme90] utilise la transformation pignistique, qui permet de répartir uniformément la masse associée à un sous-ensemble de Ω sur chacun de ses éléments. La définition de la probabilité pignistique se fonde ainsi sur le principe de raisonnement insuffisant, à savoir qu'en l'absence de toute raison de privilégier une hypothèse particulière, il convient de les considérer équiprobables. Cette opération permet de transformer une masse m_i en une distribution de probabilité permettant la prise de décision telle que :

$$\text{BetP}(\omega) = \sum_{\omega \in A \subseteq \Omega} \frac{m(A)}{|A|}, \quad \forall \omega \in \Omega \quad (11)$$

$|A|$ étant le cardinal de A . La décision va alors vers l'élément de Ω dont cette valeur est la plus grande :

$$\omega^* = \text{Arg} \left\{ \max_{\omega \in \Omega} [\text{BetP}(\omega)] \right\}. \quad (12)$$

3.2.2. Construction des fonctions de masse

La croyance d'un observateur sur l'état d'un système à un instant t influe sur la forme des fonctions de masse. La plupart du temps la fonction de masse est intrinsèque à l'application, même s'il existe de nombreuses formulations. Par exemple, Denœux [Den95] a proposé un modèle basé sur la notion de distance aux voisins, tandis que Shafer [Sha76] et Appriou [App91] ont tous les deux proposé des modélisations basées sur le maximum de vraisemblance.

Dans cette étude, on propose de représenter l'information fournie par chaque classification indépendante (*i.e.* selon chacun des trois plans (RV, VB, BR)) par une fonction de masse initiale basée sur le critère de similarité défini dans l'équation 3. Ces trois fonctions de masse initiale $(m_i)_{i \in \{RV, VB, BR\}}$ sont donc créées après calcul des k régions candidates à la fusion et avant décision du vote majoritaire. Elles permettent de prendre en compte l'incertitude associée à chacun des plans en intégrant une mesure de sensibilité associée. Ainsi, on regroupe dans le même élément focal les classes qui se trouvent très proches les unes des autres sur un même plan, et la prise de décision ne s'effectue qu'après avoir combiné les résultats des deux autres projections.

Pour chaque plan $i \in \{RV, VB, BR\}$ et pour une région donnée \mathcal{R}_j , on repère donc les k éléments les plus significatifs selon la mesure de similarité $d_i(\mathcal{R}_j, \mathcal{R}_i)$. La fonction de masse initiale construite selon le plan i comporte alors trois éléments focaux A_i, \bar{A}_i et Ω où \bar{A}_i est le complément de $A_i \subseteq \Omega$. On construit ainsi [CJ04]

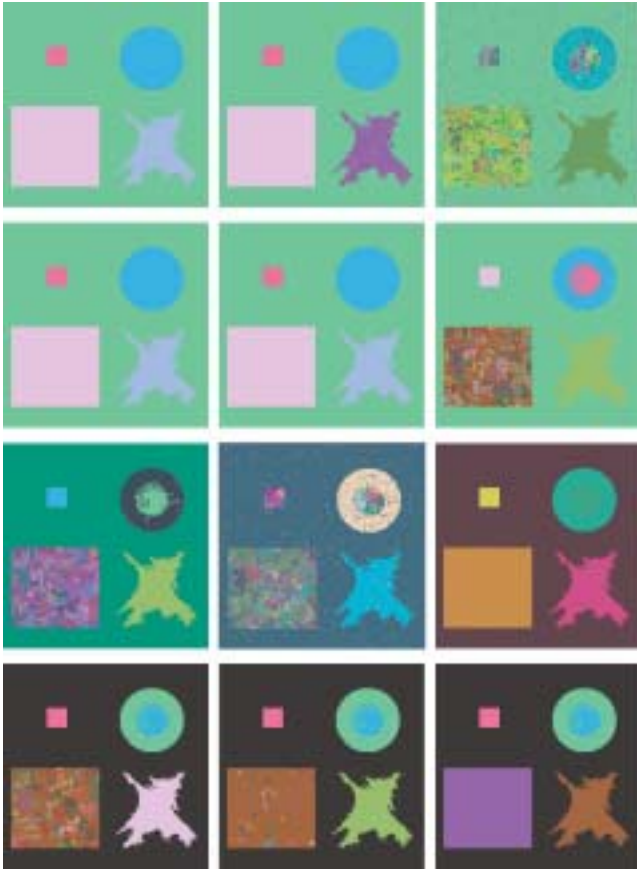


Figure 10. Illustration des différentes étapes de la méthode proposée.

$$A_i = \left\{ \omega \in \Omega, \omega = \text{Classe}(\mathcal{R}_j) \mid d_i(\mathcal{R}_j, \mathcal{R}_1) \leq \varepsilon_i d_i(\mathcal{R}_1, \mathcal{R}_j), \forall \mathcal{R}_j \right\}, \quad (13)$$

pour $i \in \{RV, VB, BR\}$. d_i représente l'application de la formule D (eq. 3) sur le plan i , au lieu de l'espace spatio-colorimétrique 3D en opérant une projection des couleurs sur le plan bi-dimensionnel i . Dans ce cas, la valeur k de l'équation 3 varie de 1 à 2. \mathcal{R}_1 est la région que l'on cherche à fusionner et \mathcal{R}_1 son plus proche voisin (selon la mesure de similarité d_i) dans un contexte spatio-colorimétrique. ε_i est une constante supérieure à 1 permettant de prendre en compte la sensibilité du système visuel humain selon le plan chromatique i : plus la sensibilité sur le plan i est importante, plus ε_i est élevée et plus le critère d'association de deux éléments doit être strict. À la limite, si $\varepsilon_i = 1$, alors $A_i = \{\text{Classe}(\mathcal{R}_1)\}$ est un singleton correspondant au plus proche voisin. La figure 11 illustre ce processus. Dans cet exemple la région \mathcal{R}_2 est plus proche de la région \mathcal{R}_6 que des autres régions candidates $\{\mathcal{R}_3, \mathcal{R}_4, \mathcal{R}_5, \mathcal{R}_7\}$. Dans ce cas, l'élément focal A_i se résume au singleton $\{\mathcal{R}_6\}$ si l'on ne prend en compte que le critère du plus proche voisin. En relâchant la contrainte de sélection par la valeur ε_i sur le plan i , l'élément focal A_i devient dans ce cas l'ensemble $\{\mathcal{R}_5, \mathcal{R}_6, \mathcal{R}_7\}$. Ceci signifie que l'on ne peut pas, à cet instant, décider de la région parmi l'ensemble des k voisins, qui sera fusionnée avec

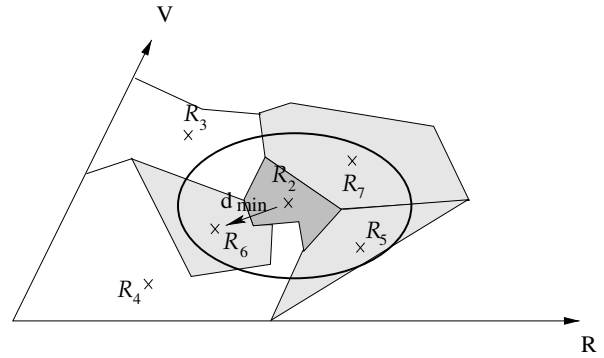


Figure 11. Construction de l'élément focal A_i sur le plan i .

la région \mathcal{R}_2 . Ainsi pour chacun des trois plans colorimétriques, un ensemble de régions contenant les k régions les plus proches est construit. Une combinaison de ces trois ensembles est ensuite réalisée afin de sélectionner la région qui sera finalement fusionnée avec la région \mathcal{R}_1 selon le maximum de probabilité pignistique.

L'une des difficultés restantes réside dans le choix de la modélisation des fonctions de masse. Dans [CCFM04], Capelle *et al.* ont comparé trois modèles d'évidence (Denøeux, Shafer et Appriou) dans le cadre de la segmentation d'images obtenues par résonance magnétique. Les résultats tendent à prouver que les meilleurs taux de classification sont induits par l'utilisation de la modélisation proposée par Denøeux. Néanmoins, dans l'approche proposée par Appriou, l'aspect multi-niveau de la méthode de construction des fonctions de masse est des plus intéressantes. En effet, il est possible de considérer chaque plan couleur (RV , RB et BV) comme des sources indépendantes portant des informations complémentaires et redondantes. En se basant sur ces deux constats, une modélisation de la fonction de masse a été réalisée en utilisant les deux modèles cités précédemment.

Tout d'abord, afin d'éviter un conflit total entre deux sources (deux plans), la masse accordée au cadre de discernement est fixée à une valeur constante. Ceci permet d'accorder une croyance minimale à l'ensemble Ω . Cette valeur a été ajustée de manière empirique suite à plusieurs tests réalisés sur une base composée de 50 images couleur allant de faiblement à fortement texturées. Ensuite les masses sont assignées aux ensembles A_i en tenant compte de la répartition des éléments dans l'ensemble A_i que l'on représente par la distance moyenne (au sens du critère de similarité) entre deux de ces éléments. La fonction de masse initiale pour le plan i est alors définie par :

$$m_i(A_i) = \alpha_i e^{-\gamma_i \bar{d}} \quad (14)$$

$$m_i(\bar{A}_i) = 1 - m_i(A_i) - m_i(\Omega) \quad (15)$$

$$m_i(\Omega) = 0.01 \quad (16)$$

où α_i est une constante permettant d'ajuster la pente de la fonction $m_i(A_i)$ et $\gamma_i = 1/d_{\max}$. d_{\max} est la similarité maximale



Figure 12. Cartes de segmentation obtenues par fusion de Dempster-Shafer.

entre \mathcal{R}_i et les éléments de A_i dans l'espace (R,V,B) et \bar{d} est la similarité moyenne entre chacun des éléments de A_i dans l'espace (R,V,B), au sens de l'équation 3. Ainsi, plus \bar{d} est grand (plus les éléments sont éloignés les uns des autres), plus la masse de A_i est petite.

Soient m_{RV} , m_{VB} et m_{BR} , les trois fonctions de masse initiale issues des projections sur les plans RV , VB et BR respectivement. La fonction de masse résultante de la combinaison des trois fonctions m_{RV} , m_{VB} et m_{BR} est obtenue à partir de l'équation (10):

$$m = m_{RV} \oplus m_{VB} \oplus m_{BR} \tag{17}$$

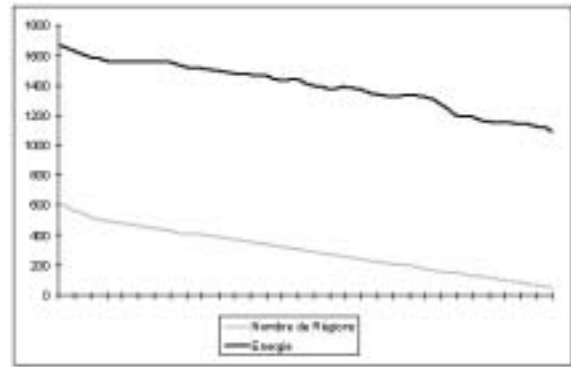
Finalement, la classe d'affectation de \mathcal{R}_i (c'est-à-dire la région avec laquelle elle va fusionner) est sélectionnée à partir de m (équation 17) sur la base du maximum de probabilité pignistique (équation 11). Dans le cas où les régions candidates sont équidistantes (même probabilité pignistique), la région ayant le cardinal le plus petit sera sélectionnée pour fusion.

Cependant, il se peut que dans certain cas très particulier, la règle de combinaison de Dempster-Shafer puisse engendrer des valeurs de conflit très importante. La figure 10 illustre ce propos. Une région n'est pas présente dans deux plans, alors qu'elle est présente dans le troisième. Ainsi, une règle d'abandon de la fusion de la région courante avec les régions candidates est

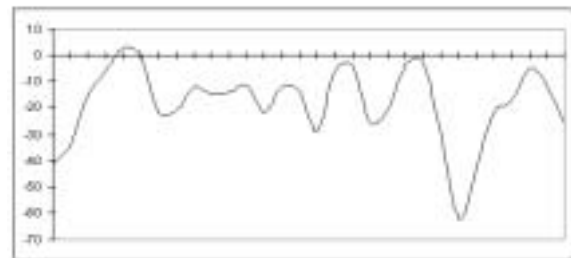
appliquée, si le facteur de conflit est très élevé, avec mise à jour appropriée du cadre de discernement. Néanmoins, dans le cas d'images naturelles, un faible taux d'abandon des fusions est constaté (cela reste marginal). La dernière ligne de la figure 10 présente les résultats obtenus à différents niveaux de fusion, la fusion optimale étant obtenue pour l'image la plus à droite de la dernière ligne, ce qui correspond bien aux régions de l'image originale.

3.2.3. Critère de terminaison

La combinaison de cartes de segmentation par la méthode de Dempster-Shafer est utilisée afin de fusionner les régions deux à deux suivant les règles de fusion précédemment énoncées. Cette combinaison permet d'obtenir une carte finale de segmentation à partir de la carte d'intersection simplifiée. Le processus de fusion des régions permet de contrôler le nombre de régions de la carte de segmentation finale. La figure 12 présente les 30 cartes de segmentation obtenues en appliquant la fusion sur l'intersection des trois cartes de segmentation obtenues par coalescence non supervisée. On remarque que les premières cartes obtenues sont sur-segmentées tandis que les dernières sont sous-segmentées. Comme il est difficile de trouver la meilleure segmentation, il nous faut établir un compromis entre sur et sous segmentation afin de contrôler le nombre de fusions de régions à réaliser. Ceci nous ramène directement au principe de comparaison énoncé au paragraphe 2.3. Nous pouvons donc réutiliser la fonction d'énergie que nous avons définie afin de trouver un compromis entre sur et sous segmentation. Si l'on analyse la variation d'énergie sur l'ensemble des segmentations de la figure 12, on remarque que celle-ci n'est pas régulière (Figure 13(1)). La recherche d'un minima local n'est donc pas directement possible à moins de le faire à ε près. Pour trouver automatiquement le meilleur niveau de simplification, nous procédons à une recherche automatique en nous basant sur l'énergie associée à chaque carte. L'énergie étant décroissante à quelques irrégularités près sur les bornes de l'intervalle choisi (nombres maximum et minimum de régions finales), on ne peut choisir un minimum local. Ce que l'on cherche est en fait une chute dans l'énergie traduisant une grande modification de la carte de segmentation. Cela revient à déterminer les endroits où la fonction d'énergie présente des transitions 1D prononcées. On utilise pour cela la dérivée première de l'énergie sur les bornes de l'intervalle choisi. Le minimum global de la dérivée première nous précisant la zone de plus grande variation d'énergie (figure 13(2)). On obtient ainsi la carte de segmentation finale la mieux simplifiée au niveau d'un compromis entre sur et sous segmentation. Dans le cas des images de la figure 12, une chute d'énergie est détectée autour des cartes de segmentation ayant 161 régions (Figure 12(24)) et 145 régions Figure 12(25)). De manière visuelle, on constate qu'une région entière a disparu entre les deux figures (région située à gauche de l'image et qui représente le feuillage). Cette disparition provoque ainsi une forte dégradation de la qualité de la segmentation, ce qui explique la différence de variation d'énergie observée entre les deux cartes.



(1) Énergie et nombre de régions.



(2) Dérivée première de l'énergie.

Figure 13. Critère d'arrêt du processus de fusion.

4. Conclusion

Dans cet article, nous avons proposé une nouvelle méthode mixte bi-marginale de segmentation d'images couleur. Tout d'abord l'image couleur est considérée selon la répartition spatiale de ses couleurs à travers ses trois histogrammes bidimensionnels. Ceux-ci sont segmentés automatiquement par une méthode rapide de coalescence morphologique non supervisée. Le réglage du paramètre multi-échelle de la coalescence se fait via la minimisation d'une fonction d'énergie modélisant la « qualité » d'une segmentation en vertu d'un principe de comparaison. Ceci permet d'arriver à un compromis entre la qualité et la complexité de la segmentation. Chacune des cartes de segmentation obtenues est simplifiée par fusion de régions connexes jusqu'à stabilisation de la même fonction d'énergie. Les trois cartes de segmentation obtenues doivent ensuite être fusionnées. Pour cela, leur intersection est calculée afin d'obtenir une unique carte de segmentation. Sur cette dernière, une pré-étape d'élimination contrôlée des régions dont la taille est inférieure à celle d'un élément structurant est opérée. Afin d'éviter l'introduction d'artefacts liés à ce processus d'élimination (élimination d'une petite région dans une région homogène), un prédicat de fusion des régions [CSG98] a été mis en place, ce qui permet de générer une nouvelle carte de segmentation avec un nombre de régions réduit. Une étape de fusion est ensuite

opérée sur cette nouvelle carte. Les plans RV , VB et BR sont alors considérés comme trois sources d'information indépendantes. Dans un premier temps, une étape de sélection des k régions les plus proches est appliquée sur chacun des trois plans, générant alors trois séries de classes potentielles. Cette étape, tenant compte de k voisins et non pas un, permet de considérer l'incertitude selon chacun des plans couleur et de repousser la prise de décision. Enfin, la décision de la région finale se fait après avoir combiné ces trois classifications, selon la règle de combinaison de Dempster-Shafer.

Les résultats des simulations de la méthode de segmentation sont prometteurs et indiquent sa robustesse et sa flexibilité. L'utilisation d'informations bi-composantes tout au long de la méthode permet d'obtenir un nouveau approche de la segmentation d'images couleur que nous avons appelée bi-marginale. Grâce à la fonction d'énergie définie, tous les paramètres sont estimés automatiquement. Les principes présentés dans cet article peuvent en outre être utilisés séparément à savoir la coalescence, la fusion par la théorie de l'évidence ou bien encore la quantification de la qualité d'une segmentation.

Références

- [App91] A. APPRIOU, Probabilités et incertitude en fusion de données multisenseurs. *Revue scientifique et technique de la défense*, 11:27-40, 1991.
- [BCS98] M. BORSOTTI, P. CAMPADELLI, and R. SCHETTINI, Quantitative evaluation of color image segmentation results. *Pattern recognition letters*, 19:741-747, 1998.
- [BLC03] N. BOULLOUDANI, P. LAMBERT, and D. COQUIN, Segmentation automatique des images couleur à base d'indicateurs de performance. In *CORESA'03*, 2003.
- [BVZ01] Y. BOYKOV, O. VEKSLER, and R. ZABIH, Fast approximate energy minimization via graph cuts. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222-1239, Novembre 2001.
- [CCFM04] A.-S. CAPELLE, O. COLOT, and C. FERNANDEZ-MALOIGNE, Evidential segmentation scheme of multi-echo mr images for the detection of brain tumors using neighborhood information. *Information Fusion*, 5(3):203-216, 2004.
- [Cel90] M. CELENK, A color clustering technique for image segmentation. *Computer Vision Graphics and Image Processing*, 52:145-170, 1990.
- [CJ04] C. CHARRIER and A.-L. JOUSSELME, Color space combination and approximation for vector quantization. In *ISIVC'04*, pages 231-234, Brest, September 2004.
- [Dem67] A. DEMPSTER, Upper and Lower Probabilities Induced by Multivalued Mapping. *Ann. Math. Statist.*, 38:325-339, 1967.
- [Den95] T. DENOUEX, A k-nearest neighbor classification rule based on dempster-shafer theory. *IEEE Trans. on Systems, Man and Cybernetics*, 25(5):804-813, 1995.
- [Fon04] M. FONTAINE, *Segmentation non supervisée d'images par analyse de la connexité des pixels*. PhD thesis, Université des Sciences et Techniques de Lille, 2004.
- [GBLP00] A. GILLET, L. MACAIRE, C. BOTTE-LECOCQ, and J.-G. POSTAIRE, Fuzzy unsupervised color image segmentation. In *Proceedings of CGIP*, pages 141-146, 2000.
- [GMC01] L. GUIGUES, H. LE MEN, and J.-P. COCQUEREZ, Segmentation d'images par minimisation d'un critère mdl dans une pyramide de segmentations. In *Proceedings of GRETSI'2001*, 2001.
- [GMC03] L. GUIGUES, H. LE MEN, and J.-P. COCQUEREZ, Analyse et représentation ensembles-échelle d'une image, in *Proceedings of GRETSI'2003*, 2003.
- [GSD01a] T. GÉRAUD, P.Y. STRUB, and J. DARBON, Color image segmentation based on automatic morphological clustering. In *IEEE International Conference on Image Processing*, volume 3, pages 70-73, 2001.
- [GSD01b] T. GÉRAUD, P.Y. STRUB, and J. DARBON, Segmentation d'images en couleur par classification morphologique non supervisée. In *International Conference on Image and Signal Processing (ICISP)*, pages 387-394, 2001.
- [GSG98] L. GARRIDO, P. SALEMBIER, and D. GARCIA, Extensive operators in partition lattices or image sequence analysis. *Signal Processing*, 6(2):157-180, 1998.
- [KLM94] G. KOEPFLER, C. LOPEZ, and J.-M. MOREL, A multiscale algorithm for image segmentation by variational method. *SIAM Journal on Numerical Analysis*, 31(1):282-299, 1994.
- [KSH01] F. KURUGOLLU, B. SANKUR, and A. HARMANCI, Color image segmentation using histogram multithresholding and fusion. *Image and Vision Computing*, 19(13):915-928, 2001.
- [LC02] O. LEZORAY and H. CARDOT, Histogram and watershed based segmentation of color images. In *Proceedings of CGIV'2002*, pages 358-362, 2002.
- [LC03] O. LEZORAY and H. CARDOT, Hybrid color image segmentation using 2d histogram clustering and region merging. In *Proceedings of ICISP'2003*, volume 1, pages 22-29, 2003.
- [LCV02] E. LEFEVRE, O. COLOT, and P. VANNOORENBERGHE, Belief function combination and conflict management. *Information Fusion*, 3(2):149-162, 2002.
- [Lez04] O. LEZORAY, An unsupervised color image segmentation based on morphological 2d clustering and fusion. In *Proceedings of CGIV'2004*, 2004.
- [LVC00] E. LEFEVRE, P. VANNOORENBERGHE, and O. COLOT, About the use of dempster-shafer theory for color image segmentation. In *Proceedings of CGIP'2000*, pages 164-169, 2000.
- [LY94] J. LIU and Y.-H. YANG, Multiresolution color image segmentation. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 16(7):689-700, 1994.
- [Mac04] L. MACAIRE, *Exploitation de la couleur pour la segmentation et l'analyse d'images*, Habilitation à diriger des recherches, Université des Sciences et Techniques de Lille, 2004.
- [MBVM97] S. MASCLE, I. BLOCH, and D. VIDAL-MADJAR, Application of dempster-shafer evidence theory to unsupervised classification in multisource remote sensing. *IEEE Trans. on Geoscience and Remote Sensing*, 34(5):1018-1031, 1997.
- [MS89] D. MUMFORD and J. SHAH, Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. On Pure and Applied Math.*, 17(4):577-685, 1989.
- [PYL98] S.H. PARK, I.D. YUN, and S.U. LEE, Color image segmentation based on 3-d clustering: morphological approach. *Pattern Recognition*, 31(8):1061-1076, 1998.
- [PZLB93] J.G. POSTAIRE, R.D. ZHANG, and C. LECOCQ-BOTTE, Cluster analysis by binary morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(2):170-180, 1993.
- [Ser82] J. SERRA, *Image Analysis and mathematical morphology*. Academic Press, London, 1982.
- [SG00] P. SALEMBIER and L. GARRIDO, Binary partition tree as an efficient representation for image processing, segmentation and information retrieval. *IEEE transactions on Image Processing*, 9(4):561-576, 2000.
- [SGG97] P. SALEMBIER, L. GARRIDO, and D. GARCIA, Image sequence analysis and merging algorithm. In *International Workshop on Very Low Bit-rate Video*, pages 1-8, 1997.
- [Sha76] G. SHAFER, *A mathematical theory of evidence*. Princeton University Press, 1976.

- [Sme90] P. SMETS, Constructing the pignistic probability function in a context of uncertainty. *Uncertainty in Artificial Intelligence*, 5:29-39, 1990. Elsevier Science Publishers.
- [Soi96] P. SOILLE, Morphological partitioning of multispectral images. *Journal of Electronic Imaging*, 18(4):252-265, 1996.
- [Soi04] P. SOILLE, *Morphological Image Analysis: Principles and Applications*. Springer-Verlag, 2004.
- [TFMB04] A. TRÉMEAU, C. FERNANDEZ-MALOIGNE, and P. BONTON, *Image numérique couleur*. Dunod, 2004.
- [XGDL03] H. XUE, T. GÉRAUD, and A. DURET-LUTZ, Multi-band segmentation using morphological clustering and fusion application to color image segmentation. In *Proceedings of the IEEE International Conference on Image Processing (ICIP'03)*, volume 1, pages 353-356, 2003.



Olivier **Lezoray**

Olivier Lezoray a obtenu une thèse de l'Université de Caen en Informatique en 2000. Depuis Septembre 2000, il est Maître de Conférences de l'Université de Caen à l'IUT Cherbourg-Manche, Département Services et Réseaux de communication. Il est membre du groupe Vision et Analyse d'Image du Laboratoire Universitaire des Sciences Appliquées de Cherbourg (LUSAC). Ses recherches portent sur les techniques des segmentation d'images couleur (les graphes, la morphologie mathématique et les méthodes variationnelles) ainsi que sur l'extraction des connaissances à partir des données et l'apprentissage (réseaux de neurones et machines à support de vecteurs). Il est membre du Groupe Français de l'Imagerie Numérique Couleur (GFINC).



Christophe **Charrier**

Christophe Charrier a obtenu une thèse de l'Université Jean Monnet (Saint-Etienne) en informatique en 1998. De novembre 1998 à septembre 2001, il a été stagiaire postdoctoral puis chercheur associé au Laboratoire de Radiocommunications et de Traitement du Signal à l'Université Laval (Québec, Canada). Depuis septembre 2001, il est Maître de Conférences de l'Université de Caen à l'IUT Cherbourg-Manche, Département Services et Réseaux de communication. Il est membre du groupe Vision et Analyse d'Image du Laboratoire Universitaire des Sciences Appliquées de Cherbourg (LUSAC). Ses recherches portent sur la compression des images couleur, les mesures de qualité des images, la modélisation du système visuel humain et la fusion de données. Il est membre du Groupe Français de l'Imagerie Numérique Couleur (GFINC). Depuis 2003, il est éditeur associé de la revue IASTED International Journal of Robotics and Automation.



