

Etude comparative des stratégies de segmentation non supervisée en régions par champs de Markov

A Comparative Study of Unsupervised Regions Segmentation Strategies by Markov Random Fields

par Catherine ROUQUET*, Pierre BONTON**, Régis TOMCZAK**

* Laboratoire des Instruments et Systèmes
Université Pierre et Marie Curie F-75252 Paris Cedex 05
E-mail : rouquet@moka.ccr.jussieu.fr

** Laboratoire des Sciences et Matériaux pour l'Electronique, et d'Automatique
– URA 1793 CNRS – Université B. Pascal F-63177 Aubière Cedex
E-mail : auteur@lasmea.univ-bpclermont.fr

résumé et mots clés

Cet article aborde le problème de la segmentation non supervisée par champs de Markov d'images de luminance. L'approche développée est de type division-fusion. L'étape de division, qui est une sursegmentation rapide, permet de diminuer de manière importante le nombre de données présentes dans le processus de fusion et, par la suite, les temps de calcul.

La méthode introduite pour la division est basée sur l'extraction de statistiques d'ordre deux à partir des matrices de cooccurrence. L'étude menée montre l'avantage de ces statistiques par rapport à celles, d'ordre un, extraites des histogrammes de niveaux de gris.

Le second point abordé dans cet article concerne la fusion. Elle est réalisée grâce à une modélisation par champs de Markov, à partir du graphe d'adjacence irrégulier de régions issues de la division. Des contributions sont amenées afin d'estimer les hyper-paramètres du système et le nombre d'étiquettes. Plusieurs résultats de segmentation sur différents types d'images réelles sont présentés afin de valider la méthode.

Une étude comparative sur les différentes stratégies d'utilisation de relaxation par champs de Markov est alors menée, de manière rigoureuse, sur des images de synthèse. Cette comparaison est effectuée aussi bien du point de vue qualité des résultats que du point de vue coût algorithmique. Elle permet de montrer les avantages de la méthode proposée : diminuer considérablement les temps de calcul mis pour obtenir les segmentations, tout en n'altérant pas la qualité des résultats de celles-ci.

Segmentation, relaxation, champs de Markov.

abstract and key words

This article deals with the important problem of unsupervised segmentation of luminance images based on a markovian relaxation. As these relaxations need important computation times, we have developed a new approach which considerably decreases these times, without modify quality of segmentation results. This approach is based on a split and merge technique. The splitting step allows to decrease, in a important way, the number of entities present in the merging step and, thereafter, computation times.

The method introduced for the splitting step is original; it is based on extraction of second order statistics from cooccurrence matrices. The study shows advantages of these statistics, and compares them to those of one order extracted from grey levels histograms.

The second point developed in this report concerns the merging step. It is accomplished by a markovian relaxation achieved on irregular adjacency graph of regions coming from the splitting step. Many original contributions are presented to estimate the hyper – parameters of the system.

At this level, several results of segmentation on different real images types are presented in order to validate the method. A comparative study on the different strategies of use of markovian relaxation is then led, in a rigorous manner, on synthesis images. It takes into account quality of results as well as computation times. Interesting results of comparison are presented.

Segmentation, relaxation, Markov random fields.

1. introduction

La segmentation reste un des problèmes fondamentaux de l'analyse d'image. Elle consiste à diminuer considérablement le nombre d'entités présentes dans l'image. Ainsi, le rôle de la segmentation est, à partir de toutes les données, de retrouver l'information utile et de la coder sous une forme de liste de primitives, plus facilement exploitable par un futur processus d'interprétation.

Ici, nous nous intéressons plus particulièrement à la segmentation en régions non supervisée d'images de luminance. Tous les paramètres de l'algorithme, ainsi que le nombre d'étiquettes, sont estimés. Nous utilisons les champs de Markov associés à la théorie bayésienne. Ils offrent un cadre mathématique cohérent pour l'extraction de primitives en introduisant conjointement :

- des connaissances *a priori* sur les primitives;
- une modélisation statistique du lien entre les observations et les primitives.

L'estimation des primitives se fait dans un espace de configurations extrêmement vaste et donne souvent lieu à des optimisations longues et coûteuses. C'est pourquoi, nous nous sommes intéressés à différentes stratégies de relaxation, permettant de réduire les temps de calcul. L'approche proposée dans cet article est de type division-fusion. L'étape de division, réalisée dans un premier temps, permet de réduire de manière importante le nombre d'entités présentes dans le processus de relaxation par champs de Markov. La division amène à une carte de régions largement sur-segmentée. Celles-ci, de formes quelconques, sont modélisées par un graphe d'adjacence irrégulier. Les champs de Markov sont alors définis sur ce graphe pour réaliser la fusion. La relaxation est obtenue rapidement puisqu'un faible nombre d'entités est présent. Cette stratégie d'utilisation des champs de Markov permet de réduire considérablement les temps de calcul mis pour obtenir les segmentations, tout en n'altérant pas la qualité des résultats. Afin d'en évaluer les performances, cette stratégie est comparée à deux autres. Dans un premier temps, la relaxation classique, réalisée directement sur le champ d'observation (les pixels de l'image) est considérée. Cette stratégie de relaxation a été utilisée par de nombreux auteurs, notamment [GGGD90], [GG84], [DE87], [HBG+94] ou [LD89]. Ensuite, la comparaison est effectuée en fonction de l'approche multi-échelle, développée par F. Heitz *et al.* [HPB94], qui consiste à observer l'image à plusieurs résolutions.

L'étude comparative de ces différentes stratégies est menée aussi bien du point de vue qualité des résultats que du point de vue coût algorithmique. Elle permet de tirer des conclusions intéressantes.

L'article se déroule comme suit. Dans un premier temps, l'étape de division est présentée. Elle est basée sur l'extraction de statistiques d'ordre deux des matrices de cooccurrence. Une étude permet de mettre en évidence l'avantage de ces statistiques par rapport à celles, d'ordre un, extraites des histogrammes locaux de niveaux de gris.

La présentation du processus de fusion fait l'objet de la troisième section. Des contributions sont apportées à ce niveau afin d'estimer les hyper-paramètres du système introduits pour rendre la segmentation non supervisée.

La quatrième section aborde la comparaison des différentes stratégies de relaxation. Celle-ci est menée sur des images de synthèse bruitées. La comparaison des segmentations est réalisée en utilisant la mesure de Vinet [VIN91] présentée ultérieurement.

2. la division

L'étape de division est réalisée afin de déterminer les régions initiales du processus de fusion. Elle est basée sur l'utilisation de matrices de cooccurrence.

Dans un premier temps, une présentation de ces matrices, de leur définition mathématique et de leur aptitude à effectuer une division d'image est effectuée. La division et ses différentes étapes sont ensuite expliquées, avant d'introduire les résultats obtenus.

2.1. définition des matrices de cooccurrence

Plusieurs définitions des matrices de cooccurrence existent dans la littérature, nous considérons ici la définition des matrices symétriques; elles déterminent la fréquence d'apparition de paires de niveaux de gris de pixels séparés par une distance δ , dans une direction γ donnée (figure 1).

La définition mathématique des matrices de cooccurrence est la suivante. Soit $f : L_x \times L_y \mapsto G$ une image numérique,

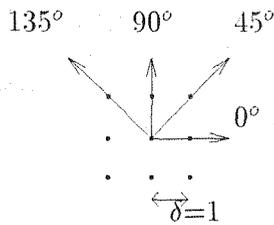


Figure 1. – Calcul des matrices de cooccurrence.

avec des domaines de variations horizontale et verticale : $L_x = \{1, 2, \dots, n_x\}$ et $L_y = \{1, 2, \dots, n_y\}$ respectivement, et avec des niveaux de gris $G = \{0, 1, \dots, n_g\}$. La matrice de cooccurrence $P_{\delta, \gamma}$ pour une distance δ donnée et une orientation γ fixée est définie par : $P_{\delta, \gamma} = [p_{\delta, \gamma}(i, j)]$ avec $0 \leq i \leq n_g$ et $0 \leq j \leq n_g$. Chaque élément de la matrice vaut : $p_{\delta, \gamma}(i, j) = \sum_{x=1}^{n_x} \sum_{y=1}^{n_y} 1_{[f(x, y)=i]} 1_{[f(x', y')=j]}$ où :

– 1_A est la fonction indicatrice pour un ensemble A ;

– x' et y' sont tels que les distances horizontale et/ou verticale entre les pixels de coordonnées (x, y) et (x', y') soient égales à δ . De plus, la condition de direction γ entre les deux pixels doit être vérifiée.

2.2. L'algorithme de division

Beaucoup d'informations peuvent être extraites des matrices de cooccurrence, et plus particulièrement, des informations relatives à la luminance des régions présentes dans la scène. Afin d'introduire les différentes caractéristiques de ces matrices, nous nous reportons aux travaux de Houzelle et Giraudon [HG91] qui étudient leur forme, et plus particulièrement celle de leur diagonale. La méthode décrite ici présente plusieurs originalités par rapport à ces travaux. Tout d'abord, l'utilisation de la diagonale de la matrice n'est pas la même. En effet, Houzelle et Giraudon l'utilisent afin de faire de la croissance de régions. A partir d'un germe, les pixels sont agglomérés pour former des régions, et simultanément, les diagonales des matrices de cooccurrence sont mises à jour (à chaque agglomération de pixels). La croissance a lieu jusqu'à ce que la région déborde sur les régions voisines. Plusieurs pics apparaissent alors, permettant de déterminer les minima locaux de la matrice, caractéristiques de la région. Ici, la diagonale de la matrice de cooccurrence est utilisée comme un histogramme de niveaux de gris : les pics, correspondant aux différentes régions, sont identifiés et permettent de construire des classes dans l'échelle des niveaux de gris. Cette méthode permet d'obtenir une division de l'image très rapide car la diagonale de la matrice est déterminée une seule fois. Dans les travaux de Houzelle et Giraudon, cette matrice devait être calculée autant de fois qu'il y a de pixels. D'où l'avantage de notre approche.

Une autre contribution, présentée ci-dessous, réside dans une étude détaillée des avantages des matrices de cooccurrence. Nous

avons, en particulier, comparé la diagonale aux histogrammes de niveaux de gris. Ceci a permis de mettre en avant les avantages de la première dans le cadre de la division d'image.

Considérons une région de niveau de gris m additionnée d'un bruit gaussien de moyenne nulle et d'écart-type σ . Il est facile de montrer que l'histogramme de cette région possède un pic, centré en m , de la forme $H(x) = K \times e^{-\frac{(x-m)^2}{2\sigma^2}}$, où K est une constante proportionnelle au nombre de pixels de la région.

La diagonale de la matrice de cooccurrence possède un pic, également centré en m , de la forme $D(x) = K' \times e^{-2\frac{(x-m)^2}{2\sigma^2}}$ où K' est une seconde constante. Le pic de la diagonale de la matrice de cooccurrence est moins large que celui de l'histogramme de niveaux de gris. Si la largeur de pic à mi-hauteur est considérée, la matrice de cooccurrence possède un pic de largeur $2\sigma\sqrt{\ln 2}$, tandis que la largeur du pic de l'histogramme des niveaux de gris est de $2\sigma\sqrt{2 \ln 2}$. Soit un rapport $\sqrt{2}$ entre ces deux largeurs. Dans le cadre d'une division de l'image, et en considérant plusieurs régions, il est plus facile de travailler à partir des matrices de cooccurrence car chaque pic est moins étendu, et donc plus distinct que dans l'histogramme.

Ceci peut être vérifié de manière expérimentale dans une fenêtre d'intérêt d'une image réelle (figure 2) dont la matrice de cooccurrence est présentée figure 3. Cette matrice possède deux pics principaux correspondant aux deux régions de la scène étudiée.

Sur la figure 4, où la diagonale de la matrice de cooccurrence ainsi que l'histogramme des niveaux de gris sont présentés, il apparaît que les pics de la diagonale de la matrice sont plus distincts que les pics de l'histogramme des niveaux de gris. D'autre part, dans ce dernier cas, des pics « parasites » interviennent. Pour une division de l'image, il sera donc plus intéressant d'étudier la diagonale de la matrice de cooccurrence. En effet, les éléments de la matrice reposent sur des statistiques d'ordre deux et rendent compte de

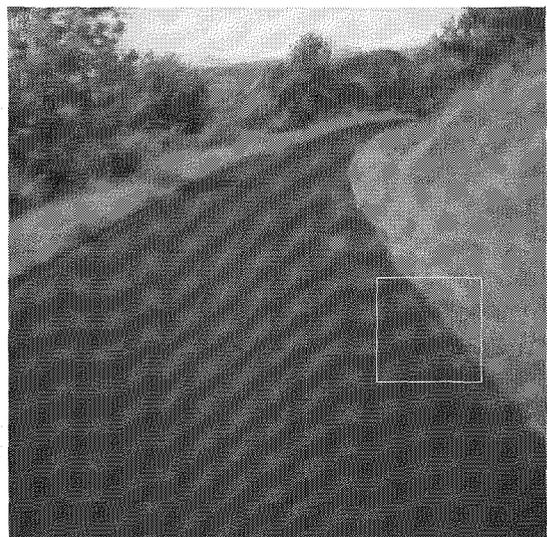


Figure 2. – Image réelle.

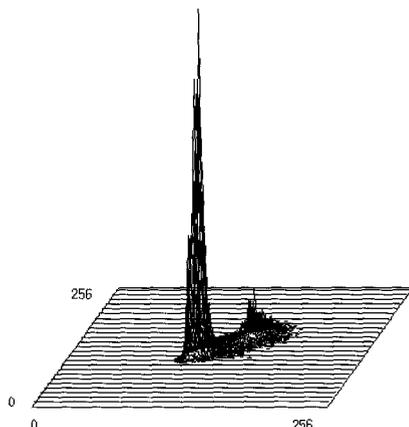


Figure 3. – Représentation 3D de la matrice.

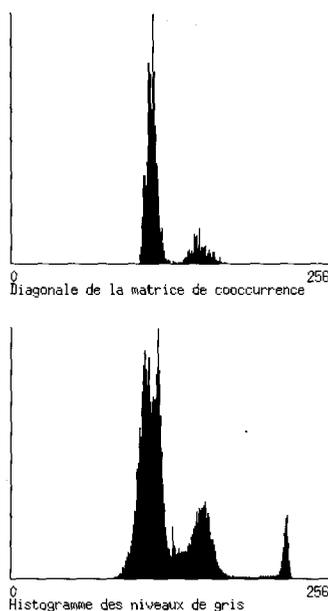


Figure 4. – Diagonale de la matrice et histogramme des niveaux de gris.

la répartition spatiale des niveaux de gris. Une partie du bruit, correspondant par exemple aux transitions de niveaux de gris $(g, g - 1)$, est située en dehors de la diagonale de la matrice; elle n'intervient donc pas lors de la recherche des pics.

L'algorithme de division que nous proposons utilise cette propriété; il se décompose en six phases.

• **Découpage de l'image en secteurs**

Il est réalisé indépendamment des données, et consiste à partitionner l'image en fenêtres rectangulaires. Toutes les étapes suivantes de la division sont réalisées de manière locale, à l'intérieur de chaque fenêtre. La division est donc une étape facilement parallélisable, il suffit de partager les fenêtres sur les différents processeurs.

• **Calcul des matrices de cooccurrence sur chaque secteur**

Pour chaque fenêtre de la grille initiale, la diagonale de la matrice de cooccurrence est déterminée. La matrice est calculée à partir d'une distance $\delta = 1$ pixel. Afin de diminuer les temps de calcul, seuls les pixels correspondant aux orientations 0 et 90 degrés sont considérés (figure 1); ce qui revient à travailler en quatre connexité. D'autre part, l'échelle des niveaux de gris est réduite linéairement à 64. Ceci a pour effet de diminuer considérablement la taille des matrices de cooccurrence, mais aussi de réduire l'effet du bruit.

Une fois la diagonale de la matrice obtenue, un lissage de celle-ci est effectué afin d'éliminer les éventuels pics dus au bruit. Le lissage est réalisé par une moyenne locale sur un voisinage de taille L :

$$d'_{\delta,\gamma}(i) = \frac{1}{L+1} \sum_{j=i-L/2}^{i+L/2} d_{\delta,\gamma}(j)$$

Ce paramètre L est indépendant des images traitées. Seul le rapport signal sur bruit influe sur le choix de celui-ci. Dans notre algorithme, il est fixé à deux et reste inchangé pour toutes les images considérées.

• **Extraction des classes**

Les différentes classes de la diagonale sont extraites, une classe étant définie par un pic et ses deux vallées [GW87].

• **Ré-étiquetage de l'image**

Dans un premier temps, une image d'étiquettes est obtenue dans laquelle, à chaque pixel est affectée l'étiquette de la classe à laquelle il appartient (le niveau de gris du pixel considéré est situé entre les deux vallées définissant la classe). Deux parties non connexes auront donc la même étiquette si elles appartiennent à la même classe de niveaux de gris. Il est nécessaire d'introduire un algorithme de ré-étiquetage en composantes connexes, créant une nouvelle image d'étiquettes dans laquelle chaque étiquette correspond à un ensemble de pixels connexes appartenant à la même classe. Nous appelons région chacun de ces ensembles.

• **Suppression des petites régions**

L'algorithme de division présente l'inconvénient de donner des régions présentant des tailles diverses. Ainsi, plusieurs régions de taille très faible apparaissent (un pixel, deux pixels,...). Nous considérons ces régions comme non significatives. Toutes les régions dont le nombre de pixels est inférieur à un seuil sont donc supprimées par fusion avec la région voisine possédant la moyenne de niveaux de gris la plus proche. Le seuil déterminant l'appartenance d'une région à la catégorie « petite région » est de 25 pixels pour des images de 512×512 pixels. Il faut noter ici que nous ne recherchons pas de minimum global pour réaliser la fusion : les petites régions sont considérées dans leur ordre

d'arrivée et sont fusionnées, dans ce même ordre, avec la région la plus proche au sens des niveaux de gris. Ceci permet d'aboutir à des temps de calcul plus rapides que ceux obtenus par la recherche de minima globaux; par contre, la suppression des petites régions n'est pas optimale.

• Modélisation des résultats

Les régions extraites du processus de division sont de forme quelconque. En effet, certaines fenêtres de l'image originale ont été divisées, ce qui rompt le caractère régulier de la grille initiale. Ces régions sont modélisées par un graphe d'adjacence irrégulier où chaque nœud représente une région et chaque lien représente l'adjacence entre deux régions. C'est ce graphe d'adjacence qui constitue les données d'entrée du processus de fusion.

2.3. résultats expérimentaux de la division

Un résultat de division, obtenu sur une image de route sans marquage au sol, est présenté figure 5. La taille des fenêtres initiales est de 32×32 pixels et nous ne travaillons que sur une trame (fenêtres de 32×64 pixels sur l'image résultat).

Les différentes régions relatives aux objets de l'image ont généralement été retrouvées; d'autre part, peu de régions parasites, ne correspondant pas aux objets réels, sont présentes, ce qui limite par la suite les temps de calcul.

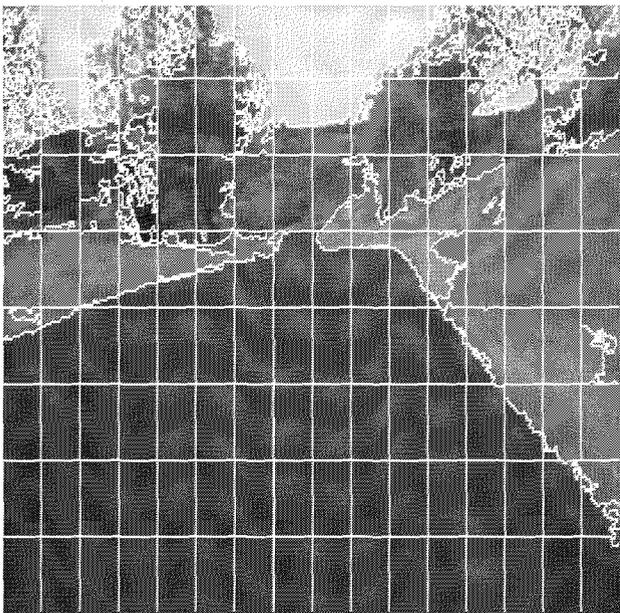


Figure 5. – Résultat de division.

2.4. conclusion

Un algorithme de division de l'image au sens des niveaux de gris a été développé. Le but de celui-ci est de fournir des régions initiales, ultérieurement utilisées par un processus de fusion. Afin d'obtenir une méthode robuste, la division est réalisée en utilisant des statistiques d'ordre deux. Celles-ci, issues des matrices de cooccurrence, sont moins sensibles au bruit que les statistiques d'ordre un, comme montré sur un exemple (paragraphe 2.2). Les résultats de division obtenus sont tout à fait satisfaisants dans le sens où, généralement, les contours des objets de la scène ont été extraits. Ils sont utilisés en donnée d'entrée du processus de fusion.

3. la fusion basée sur une modélisation par champs de Markov

Une fois la division de l'image réalisée, un processus de fusion est engagé pour regrouper certaines régions. Afin d'obtenir de bons résultats de segmentation, notre choix s'est porté vers les techniques basées sur une modélisation par champs de Markov. Connus et étudiés depuis près de vingt ans par les mathématiciens, ils ont progressivement fait leur apparition dans le domaine du traitement d'images [GG84] [AZE87]. Ceux-ci permettent d'introduire de manière simple les variables modélisant les caractéristiques image d'un problème donné. Deux types de variables sont généralement considérés : les variables correspondant à une modélisation des données image et les variables relatives aux champs d'étiquettes. Ces dernières permettent d'éliminer des ambiguïtés, de corriger des erreurs, ou encore de retrouver des informations masquées par le bruit. La segmentation, associée aux champs de Markov et à la théorie bayésienne de l'estimation, se ramène, dans le cas de la maximisation de la probabilité *a posteriori*, à un problème d'optimisation. Le but est de trouver le minimum global d'une fonction représentant l'erreur commise. Cette fonction dépend des deux variables précédemment définies.

L'optimisation se fait dans un espace de configurations très grand, et il n'est généralement pas possible de tester chacune d'entre elles. Des méthodes de relaxation permettent de résoudre ce problème; il en existe deux grandes classes, à savoir, les relaxations stochastiques et les relaxations déterministes. Nous avons choisi d'utiliser les relaxations déterministes, et plus particulièrement l'algorithme *ICM* (Iterated Conditional Modes [BES86]) pour réaliser la fusion des régions. La principale raison de ce choix est que cette méthode conduit à des temps de calcul plus faibles que les approches stochastiques. Par contre, la convergence a souvent lieu dans des minima locaux. Une façon de limiter ce problème consiste à initialiser la relaxation par une

solution proche de la solution finale. Selon les stratégies de relaxation utilisées, des résultats de segmentation de plus ou moins bonne qualité sont obtenus (voir étude § 4.2). Dans cet article, une comparaison des différentes stratégies de relaxation est effectuée. Cette comparaison, menée aussi bien du point de vue coût algorithmique que du point de vue qualité des résultats, est réalisée en utilisant l'algorithme de relaxation *ICM*.

L'organisation de cette section se déroule comme suit. Dans un premier temps, la stratégie de relaxation proposée ainsi que l'estimation des paramètres du système sont définies. Des résultats de segmentation, obtenus de manière non supervisée, sont alors présentés.

3.1. application à la segmentation d'image

Dans le cadre de la segmentation d'images défini ici, les données sont les régions R_i issues de la division. Celles-ci, de formes quelconques, sont modélisées par un graphe d'adjacence irrégulier. Les régions sont élémentaires dans le sens où elles ne pourront pas être redivisées; nous les appelons sites élémentaires. Elles forment le champ de données $R = \{R_1, \dots, R_N\}$. Le problème de la segmentation consiste à mettre en correspondance ce champ de données R et un champ d'étiquettes $E = \{e_1, \dots, e_N\}$ prenant valeurs dans un ensemble fini $\Lambda = \{1, \dots, L\}$. La segmentation est alors assimilée à l'estimation de la réalisation de E , à partir de la réalisation observée R . Plusieurs méthodes de relaxation markovienne définie sur des graphes d'adjacence existent dans la littérature [AG92], [KY94], [MZ92]. L'originalité de notre approche réside dans son aspect entièrement non supervisé. D'autre part, nous avons réalisé une étude permettant de la comparer aux méthodes classiques (relaxation directe sur l'image). Dans notre application, nous nous plaçons dans le contexte de la maximisation de la probabilité *a posteriori* (MAP : Maximum A Posteriori) et cherchons à maximiser la probabilité $P(E/R)$, ce qui revient à minimiser l'énergie $U(E/R)$. Celle-ci varie comme la somme pondérée de deux énergies :

$$U(E/R) = U(R/E) + \alpha U(E) \quad (1)$$

Le paramètre α permet de contrôler l'influence de chacune d'entre elles.

La mise en œuvre de cet algorithme de segmentation requiert la connaissance des paramètres définissant la loi $U(E/R)$. Ceux-ci, formant un vecteur θ , ne sont pas connus comme des *a priori* de l'algorithme. Dans notre cas, les régions issues de la division sont caractérisées par un vecteur de paramètres à une seule composante : la moyenne des niveaux de gris. Celui-ci pourra néanmoins s'étendre et comprendre des caractéristiques telles que la texture ou la couleur. Nous pouvons maintenant définir le vecteur de paramètres θ :

$$\theta = (\alpha, \mu_1, \dots, \mu_L)^T \quad (2)$$

Plusieurs méthodes [DLR77], [CD85], [PIE94], [MC91] ou [LD89] permettent de résoudre le problème de son estimation. Dans notre algorithme, pour des raisons de temps de calcul, une méthode de segmentation courante, où les paramètres sont estimés de manière heuristique au cours de la relaxation (cf. 3.2.2), est utilisée [GGGD90] ou [KH95].

En incluant le vecteur de paramètres cachés, l'énergie *a posteriori* à minimiser peut se mettre sous la forme :

$$U(E/R, \hat{\theta}) = U(R/E, \hat{\theta}) + \alpha U(E) \quad (3)$$

Sa minimisation implique une décomposition de cette fonction de manière locale. Pour cela, un système de voisinage $\mathcal{G} = \{\mathcal{G}_{R_i}, R_i \in R\}$ est défini sur R . Dans notre approche, deux régions R_i et R_j sont voisines si et seulement si les nœuds correspondants dans le graphe d'adjacence sont reliés. Un sous-ensemble c de R est appelé une clique relative au système de voisinage \mathcal{G} si toutes les régions de c distinctes sont voisines deux à deux. Afin de réduire les temps de calcul, seules les cliques d'ordre deux sont considérées (cliques composées de deux régions). Notons \mathcal{C} l'ensemble des cliques de R .

Avec ce système de voisinage, l'équation 3 devient :

$$U(E/R, \hat{\theta}) = \sum_{i=1}^N V(R_i/e_i, \hat{\theta}) + \sum_{c \in \mathcal{C}} V_c(e) \quad (4)$$

où e_i est l'étiquette affectée au site élémentaire R_i et e représente les différentes étiquettes de la clique \mathcal{C} .

Comme nous l'avons justifié précédemment, l'algorithme *ICM* est utilisé pour réaliser la relaxation. Nous minimisons donc, pour chaque région R_i :

$$V(e_i/R_i, e_{\mathcal{G}_{R_i}}, \hat{\theta}) = V(R_i/e_i, \hat{\theta}) + \sum_{e_j \in e_{\mathcal{G}_{R_i}}} V_c(e_i, e_j) \quad (5)$$

$$= V(R_i/e_i, \hat{\theta}) + V(e_i/e_{\mathcal{G}_{R_i}}) \quad (6)$$

où $e_{\mathcal{G}_{R_i}}$ représente les étiquettes des voisins du site R_i .

Le potentiel $V(R_i/e_i, \hat{\theta})$ est modélisé par la distance entre le vecteur de caractéristiques de la régions R_i et celui de l'étiquette e_i , soit :

$$V(R_i/e_i, \hat{\theta}) = \left(\frac{\mu_{R_i} - \mu_{e_i}}{256} \right)^2 \quad (7)$$

Nous définissons le potentiel $V(e_i/e_{\mathcal{G}_{R_i}})$ de telle sorte qu'il favorise l'appartenance de deux régions voisines à la même étiquette. Ainsi, pour ce potentiel, l'étiquette la plus probable pour un site sera celle la plus présente dans son voisinage. La modélisation des connaissances *a priori* s'écrit :

$$V(e_i/e_{\mathcal{G}_{R_i}}) = 1 - \frac{\sum_{R_j \in \mathcal{G}_{R_i}} 1_{e_j=e_i}}{\sum_{R_j \in \mathcal{G}_{R_i}} 1} \quad (8)$$

3.2. algorithme de relaxation

L'algorithme de relaxation se décompose en plusieurs phases. Afin de bien comprendre son déroulement, son synoptique est présenté figure 6. Nous détaillons par la suite chacune de ses étapes.

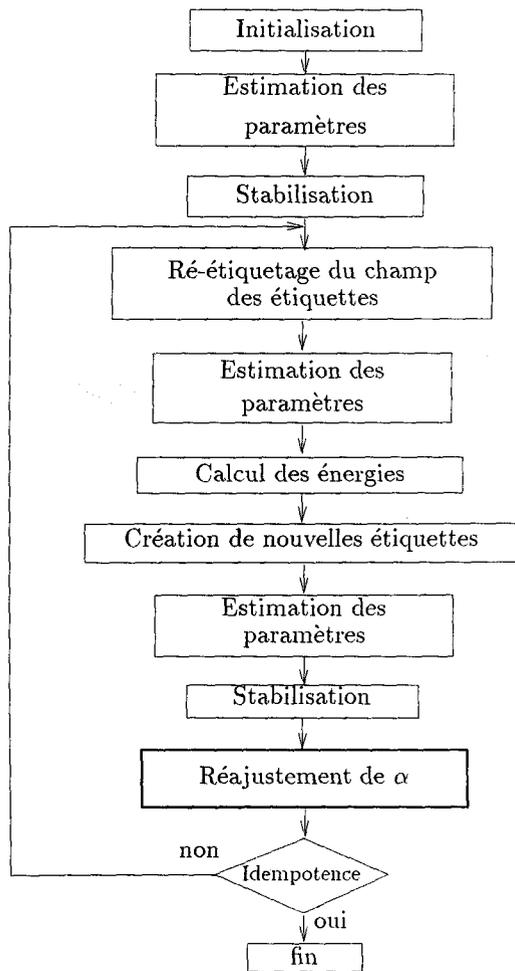


Figure 6. – Synoptique de la relaxation.

3.2.1. initialisation

Au début de cette étape, l'image est composée d'une seule région étiquetée 0. Le prototype de cette étiquette, noté P_0 , est déterminé :

$$P_0 = (\mu_0) \text{ avec } \mu_0 = \frac{1}{N} \sum_{i=1}^N \mu_{R_i}. \text{ Le potentiel moyen sur toute}$$

l'image \bar{V}_0 est alors estimé sans prendre en compte le voisinage puisqu'il est homogène (une seule étiquette 0). D'où :

$$\bar{V}_0 = \frac{1}{N} \sum_{i=1}^N V(R_i/e_i = 0, \hat{\theta}) = \frac{1}{N} \sum_{i=1}^N (\mu_{R_i} - \mu_0)^2 \quad (9)$$

Le graphe d'adjacence est alors balayé et pour chaque région, $V(R_i/e_i = 0, \hat{\theta})$ est déterminé. Parmi tous ces potentiels, nous affectons ceux supérieurs à \bar{V}_0 (et donc correspondant mal à l'étiquette 0) à l'étiquette 1. A la fin de cette étape, le champ d'étiquettes est composé de deux étiquettes : 0 et 1. Il est nécessaire de mettre à jour les prototypes de ces deux étiquettes, c'est-à-dire, d'estimer les paramètres du modèle.

3.2.2. estimation des paramètres du modèle

La remise à jour des vecteurs de caractéristiques correspondant aux différentes étiquettes est réalisée de manière globale, évitant ainsi une nouvelle extraction de paramètres. Pour l'étiquette j correspondant à N_j sites élémentaires R_i , nous aurons par exemple :

$$\mu_j = \frac{\sum_{i=1}^N 1_{e_i=j} \times \mu_{R_i} \times (nb_{pix})_{R_i}}{\sum_{i=1}^N 1_{e_i=j} \times (nb_{pix})_{R_i}} \quad (10)$$

où $(nb - pix)_{R_i}$ représente le nombre de pixels du site élémentaire R_i .

3.2.3. stabilisation du champ des étiquettes

Pour chaque région R_i , $V(e_i = j/R_i, e_{G_{R_i}}, \hat{\theta})$ est déterminé pour toutes les étiquettes j appartenant au voisinage de la région R_i (dans cette étape de stabilisation de l'initialisation, seules les étiquettes 0 ou 1 peuvent être présentes). L'étiquette correspondant au potentiel minimal est affectée à la région considérée, selon le principe de l'algorithme de relaxation ICM. Le graphe d'adjacence est parcouru plusieurs fois, jusqu'à ce que le champ d'étiquettes soit stabilisé.

3.2.4. ré-étiquetage du champ des étiquettes

La stabilisation du champ des étiquettes fait qu'une même étiquette peut être affectée à plusieurs régions non adjacentes. Les prototypes des étiquettes ne sont alors plus représentatifs des objets éventuels de la scène puisqu'ils correspondent à plusieurs zones de l'image. La relaxation peut être continuée avec ce champ d'étiquettes non connexe. Cependant, il est plus rigoureux de mettre en place un processus de ré-étiquetage, permettant de donner des étiquettes différentes à chaque ensemble de régions connexes appartenant à la même classe. Suite à cette étape, les vecteurs de caractéristiques de chaque étiquette sont déterminés.

3.2.5. calcul des potentiels

Le champ des étiquettes est stabilisé avec un certain nombre d'étiquettes. Certaines régions possèdent un potentiel par rapport à leur étiquette $V(e_i/R_i, e_{G_{R_i}}, \hat{\theta})$ très fort; elles ne sont donc pas appropriées à celle-ci, même si elle était la plus probable

dans le champ des étiquettes. De nouvelles étiquettes devront donc être créées. Pour cela, il faut reconnaître les régions mal étiquetées, et donc possédant un fort potentiel. Pour chaque région R_i , le potentiel $V(e_i/R_i, e_{G_{R_i}}, \hat{\theta})$ est déterminé. Sur la figure 7 est représenté l'histogramme des potentiels sur toute l'image.

Cet histogramme possède un pic au voisinage de 0, correspondant aux régions de faible potentiel et donc bien étiquetées. Les valeurs de l'histogramme diminuent ensuite. Une limite doit être fixée, séparant les régions bien étiquetées des régions mal étiquetées, ou encore, les régions de faible potentiel des régions de fort potentiel. On peut, pour cela, utiliser la valeur moyenne : toutes les régions dont le potentiel est supérieur à la moyenne arithmétique des potentiels sont affectées à une nouvelle étiquette. Un problème se pose alors : le nombre de régions rejetées n'est pas contrôlé. L'histogramme figure 7 n'a pas l'allure d'une gaussienne; la moyenne des potentiels ne représente donc pas la valeur telle que 50% des potentiels soient en dessous, et 50% au dessus. Beaucoup de régions peuvent ainsi être rejetées, amenant à une déstabilisation importante du champ des étiquettes. Ceci a pour effet d'augmenter les temps de calcul.

Pour éviter ce problème, nous modélisons les potentiels des différentes régions par une loi log-normale de densité de probabilité

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\ln x - m}{\sigma} \right)^2 \right] \quad (11)$$

Comme les potentiels suivent une loi log-normale, leurs logarithmes népériens suivent une loi normale $\mathcal{N}(m, \sigma)$. La moyenne \bar{U} et la variance V des variables suivant la loi log-normale peuvent se mettre sous la forme :

$$\bar{U} = e^{m + \frac{\sigma^2}{2}} \quad V = (e^{\sigma^2} - 1) e^{2m + \sigma^2} \quad (12)$$

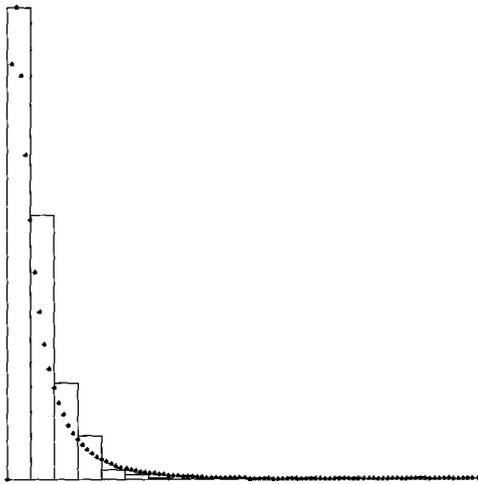


Figure 7. – Histogramme des potentiels et répartition statistique de la loi log-normale correspondante.

Ces deux relations permettent de remonter facilement aux paramètres de la loi log-normale :

$$m = \ln \left(\frac{\bar{U}^2}{\sqrt{U^2 + V}} \right) \quad \sigma^2 = \ln \left(1 + \frac{V}{U^2} \right) \quad (13)$$

Sur la figure 7 sont représentés l'histogramme des potentiels, et la répartition statistique de la loi log-normale correspondante. Ces deux courbes sont bien ajustées et donc, la modélisation de la variation des potentiels sur l'image par une loi log-normale est correcte.

Bien que la distribution des potentiels ait souvent l'allure d'une loi log-normale, dans certains cas, elle est assez éloignée de cette modélisation. Nous avons néanmoins gardé ici la loi log-normale car de manière générale, elle est la plus adaptée aux histogrammes. D'autre part, la qualité de la modélisation a peu d'influence sur la qualité de la segmentation finale.

3.2.6. création de nouvelles étiquettes

Nous supposons que les cinq pour cent des potentiels les plus grands sont trop forts, et par conséquent que l'étiquette qui a été affectée aux régions correspondantes n'est pas adaptée. Cette valeur de potentiel est donnée à partir des tables associées à la loi log-normale par :

$$V_{max} = e^{m+1.645\sigma} \quad (14)$$

Le graphe d'adjacence est alors balayé et, pour chaque région, le potentiel $V(e_i/R_i, e_{G_{R_i}}, \hat{\theta})$ est déterminé.

- Si ce potentiel est supérieur à V_{max} , une nouvelle étiquette est affectée à la région;
- Sinon, l'étiquette de la région est inchangée.

Grâce à la modélisation par une loi log-normale, le nombre de régions rejetées est contrôlé, ce qui permet de ne pas trop déstabiliser le champ des étiquettes, et donc d'aboutir à des temps de calcul plus faibles.

La création de nouvelles étiquettes aurait également pu se faire en rejetant directement les 5 pour cent des régions ayant le plus fort potentiel. Cette méthode possède cependant deux inconvénients. Tout d'abord, elle est plus longue en temps de calcul car elle nécessite le classement de toutes les régions selon leur potentiel. Ensuite, elle est moins robuste. En effet, lors de la relaxation, des régions isolées de fort potentiel (car le potentiel relatif au voisinage est fort) peuvent apparaître. Ces régions vont être à l'origine de l'apparition d'un second pic vers les forts potentiels dans l'histogramme des potentiels. En utilisant la modélisation par une loi log-normale, ce pic est peu pris en compte et des régions de potentiel moins fort vont être rejetées, permettant ainsi à la relaxation de continuer. Par contre, si seules les cinq pour cent des régions de plus forts potentiels sont rejetées (cas de l'utilisation directe de l'histogramme), il se peut que seules les régions isolées le soient, ce qui amènerait à un arrêt de la relaxation et par la suite, à une mauvaise segmentation.

3.2.7. stabilisation

Cette étape de stabilisation est la même que précédemment : l'étiquette correspondant au potentiel minimal est affectée à la région considérée selon le principe de l'algorithme *ICM*.

3.2.8. réajustement du paramètre α

Ce paramètre est introduit pour contrôler l'influence relative des deux énergies : celle due à la modélisation du processus observation-étiquettes et celle due à la modélisation des connaissances *a priori*. La première énergie $U(R/E, \hat{\theta})$ diminue au cours des itérations, tandis que la seconde $U(E)$ augmente avec celles-ci. Il n'est donc pas possible de déterminer un paramètre α tel que ces deux énergies soient toujours équilibrées. Dans beaucoup de travaux [DDB⁺94] [KH95], le paramètre α reste fixe, et ne varie pas en fonction des itérations. Le problème est que l'on ne connaît pas le minimum d'énergie à atteindre, et donc qu'il est impossible de fixer α rigoureusement dès le début de la relaxation. En effet, α doit être assez faible au départ pour permettre la création de nouvelles étiquettes, et assez fort à la fin pour stopper la relaxation. Pour pallier ce problème et obtenir une segmentation adaptative, nous faisons varier le paramètre α au cours des itérations. Dans [GGGD90], une loi de variation linéaire est suggérée. Le problème ici est que la loi de variation n'est pas connue; elle est déterminée *a priori* au début de la relaxation et de plus dépend des images. Dans cet article, nous contournons ce problème en faisant varier α de manière inverse à l'énergie. Nous nous sommes donc intéressés à la variation de l'énergie au cours des itérations et plus particulièrement à la variation du potentiel lié à la vraisemblance de l'observation étant donné une configuration du champ des étiquettes :

$$\overline{V(R/E, \hat{\theta})} = \frac{1}{N} \sum_{i=1}^N V(R_i/e_i, \hat{\theta})$$

Sur la figure 8 est représentée la variation de ce potentiel moyen en fonction des itérations pour une image de route sans marquage au sol. Des figures similaires, avec des décroissances plus ou moins accentuées, sont obtenues pour toutes les images en notre possession.

Le potentiel moyen (représentant également l'énergie $U(R/E, \hat{\theta})$) décroît rapidement au départ pour se stabiliser ensuite. Nous avons également représenté sur la figure l'écart-type des potentiels sur toutes les régions de l'image (barres verticales). Celui-ci diminue très rapidement et devient faible avec les itérations, ce qui est logique puisque les régions correspondant aux forts potentiels sont affectées à une nouvelle étiquette. L'énergie globale $U(R/E, \hat{\theta})$ diminue alors, et l'écart-type relatif aux potentiels devient moins important.

Nous présentons ci-dessous l'évolution de α au cours des itérations :

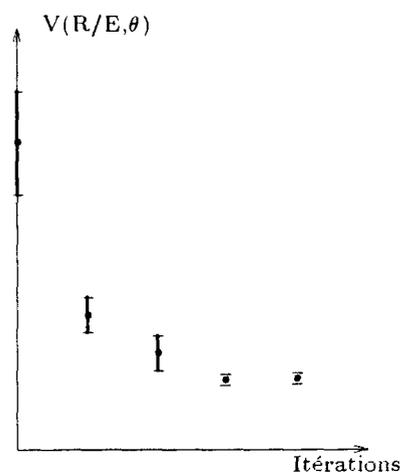


Figure 8. – Valeurs moyennes et écart-types des potentiels. $V(R_i/e_i, \hat{\theta})$ au cours des itérations.

– $k = 0$, $\alpha = 0$. Modélisation des potentiels des différentes régions par une loi log-normale. Extraction des paramètres m et σ définissant cette loi (équation 13) puis de V_{0_max} (équation 14).

– Pour les itérations suivantes, et jusqu'à la stabilisation du processus : modélisation des potentiels par une loi log-normale et extraction de V_{k_max} . La nouvelle valeur de α est fixée à :

$$\alpha = (V_{0_max} - V_{k_max}) \quad (15)$$

De manière générale, nous pouvons conclure que la variation du paramètre α conduit à de meilleurs résultats que ceux obtenus avec α fixe, dans le sens où maintenant, l'algorithme est devenu adaptatif par rapport aux images traitées. De plus, cette variation évite d'avoir à choisir α , choix très difficile à réaliser car ce paramètre dépend des images. D'autre part, un avantage de cette méthode est que la baisse du coefficient α est directement liée à la variation de l'énergie globale $U(R/E, \hat{\theta})$ au cours des itérations. Ainsi, lors du processus de relaxation, lorsque l'énergie est grande au début des itérations, le coefficient α reste voisin de 0, de manière à favoriser la création de nouvelles étiquettes. Par contre, une forte baisse de l'énergie avec les itérations conduit à une grande valeur de α , et favorise ainsi la stabilisation du processus. Cette méthode présente un atout supplémentaire : l'algorithme de relaxation converge moins facilement vers des minima locaux. En effet, une faible valeur de α au début des itérations conduit le système à créer beaucoup d'étiquettes et à ne pas se stabiliser dans des puits d'énergie. La convergence est favorisée par la suite en prenant davantage en compte la notion de voisinage.

Le paramètre α , qui dans un premier temps était fixé par l'utilisateur en fonction des images traitées [ROU96], est devenu non supervisé : il s'adapte au cours des itérations en fonction de la baisse d'énergie (et donc des images étudiées). L'algorithme est entièrement autonome et indépendant du type d'images à segmenter comme nous le montrons dans la section 3.3.

3.2.9. idempotence

Le test de l'idempotence se fait de la manière suivante : le nombre d'étiquettes obtenu après le ré-étiquetage du champ des étiquettes et celui obtenu après la stabilisation sont comparés. Si ces deux nombres sont égaux, l'idempotence est obtenue et le processus de relaxation prend fin. En effet, toutes les nouvelles étiquettes créées ont été supprimées lors de la stabilisation; le champ des étiquettes est identique à celui obtenu avant la création de nouvelles étiquettes. Par contre, si le nombre d'étiquettes a varié, l'idempotence n'est pas réalisée et le processus reboucle sur lui-même, il y a toujours création de nouvelles classes.

3.3. résultats expérimentaux

Les figures 9 à 16 présentent les résultats de segmentation obtenus. De manière générale, les différentes régions de la scène ont été retrouvées, et ce, pour des images de types différents (image de route avec bas-côtés texturés, image de route avec lignes blanches, image de synthèse,...), ce qui montre l'adaptativité de l'algorithme. La méthode de segmentation développée ici présente les avantages suivants. Tout d'abord, l'utilisation de statistiques d'ordre deux (matrice de cooccurrence) lors de l'étape de division permet de rendre celle-ci moins sensible au bruit que les méthodes classiques fondées sur l'étude d'histogrammes locaux (§2.2). Une bonne division est ainsi obtenue, permettant de bien initialiser le processus de fusion.

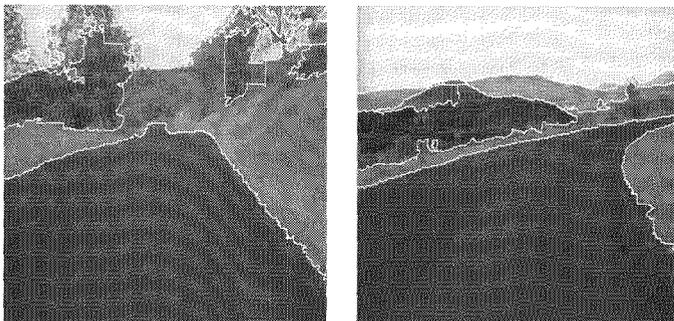


Figure 9 et 10. – Résultat de segmentation.

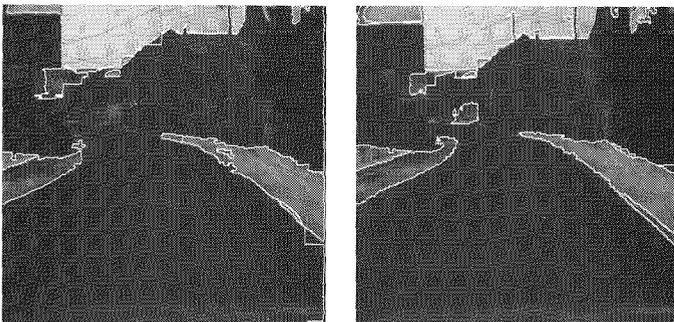


Figure 11 et 12. – Résultat de segmentation.

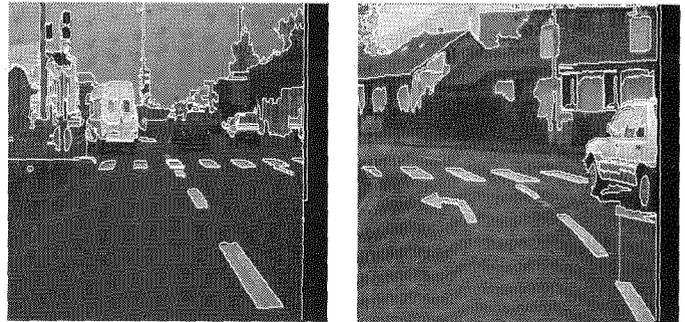


Figure 13 et 14. – Résultat de segmentation.

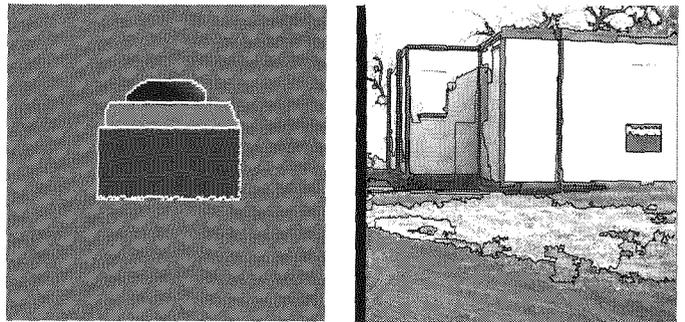


Figure 15 et 16. – Résultat de segmentation.

La fusion est réalisée en utilisant une modélisation par champ de Markov définie à partir du graphe d'adjacence irrégulier des régions issues de la division. Cette technique permet d'une part, de diminuer les temps de calcul et d'autre part, de stabiliser le système : beaucoup de pixels ont déjà été regroupés lors de la division, ce qui permet à la relaxation d'éviter des minima locaux.

Les temps de calcul, mis par la relaxation par champ de Markov, ont fortement diminué en initialisant celle-ci par un graphe d'adjacence. Pour pouvoir estimer les gains de temps obtenus par cette technique, elle doit être comparée aux autres méthodes de relaxation existantes.

4. étude comparative

Il est très difficile de quantifier de manière globale la validité d'une segmentation. En effet, la segmentation optimale est souvent inconnue et de plus, dépend de l'application considérée. Ainsi, pour une même image, plusieurs segmentations différentes pourront être considérées comme correctes. Afin de résoudre en partie ce problème, des études comparatives sont menées sur des images de synthèse simples où la segmentation idéale est connue par construction [CP95], [PP93], [ZG92], [VIN91], [KFB93]. L'évaluation d'un algorithme de segmentation consiste alors à mesurer la différence entre la segmentation obtenue et la segmentation idéale. Il faut noter que cette méthode est très restrictive puisque les images utilisées ne sont pas réelles et que la modélisation du

bruit est souvent éloignée de la réalité. L'algorithme donnant le meilleur résultat sur une image de synthèse ne sera pas nécessairement le mieux adapté à des images réelles. Bien que cette méthodologie présente des défauts, nous l'utilisons afin d'évaluer numériquement les résultats de notre stratégie de relaxation.

Dans un premier temps, nous présentons les outils nécessaires pour réaliser la comparaison, à savoir, une distance entre cartes de régions et une série d'images de synthèse. Le second paragraphe porte sur la comparaison des différentes stratégies de relaxation par champ de Markov. Des tableaux de résultats amèneront à tirer quelques conclusions.

4.1. introduction de la mesure de dissimilarité et des images utilisées

4.1.1. distance entre les cartes de régions

Afin de pouvoir quantifier la distance existant entre deux segmentations, nous utilisons la mesure de Vinet [VIN91]. Considérons une image I contenant N pixels, et deux segmentations en régions de celle-ci notées R et S . La mesure de Vinet consiste à rechercher les couples de régions les plus similaires dans R et dans S , puis à estimer le nombre de pixels n'intervenant pas dans cette superposition. La mesure de Vinet est alors proportionnelle à ce nombre. Dans un premier temps, une table de superposition des étiquettes est déterminée : $T_{i,j} = \text{card}(R_i \cap S_j)$ avec $0 \leq i < m$ et $0 \leq j < n$ si R comporte m régions et S n régions.

La recherche du maximum dans cette table conduit aux deux régions les plus similaires, le critère de similarité étant défini par $C_0 = \max_{\substack{0 \leq i < m \\ 0 \leq j < n}} T_{i,j}$. La recherche du second maximum (sans considérer les deux régions précédentes déjà affectées) conduit au critère de similarité C_1 , et ainsi de suite jusqu'à C_{k-1} où $k = \min(m, n)$.

La mesure de dissimilarité entre les deux segmentations R et S est alors donnée par :

$$D(R, S) = \frac{N - C}{N} \quad \text{avec} \quad C = \sum_{i=0}^{k-1} C_i$$

Un inconvénient de cette mesure est qu'elle passe par la recherche de maxima locaux. Ainsi, un changement de un pixel dans la table de superposition pourra produire un changement de maximum et donc, une forte variation de la mesure de dissimilarité. Pour remédier à cet inconvénient, une modification a été apportée à cette mesure dans [CP95]. Plutôt que de rechercher les maxima de façon itérative, il est suggéré de rechercher la configuration telle que la mesure de dissimilarité soit globalement minimale. En utilisant cette technique, le problème précédemment posé est résolu, et les changements de un pixel auront peu d'influence sur la mesure finale.

4.1.2. présentation des images utilisées

Les images considérées ici sont toutes des images de synthèse, ce qui permet de connaître les segmentations optimales. Elles proviennent de la banque de données du GdR ISIS. Une image contenant sept régions de forme géométrique a été créée (figure 17).

Dans le but de travailler sur des images se rapprochant le plus possible des images réelles, deux bruits blancs différents, de moyenne nulle et d'écart-type dix, ont été superposés aux images originales. Il s'agit de bruits de distributions gaussienne et exponentielle. Deux images bruitées ont ainsi été obtenues : BRUIT1 et BRUIT2. La première image est présentée figure 18, la seconde lui est visuellement très proche.

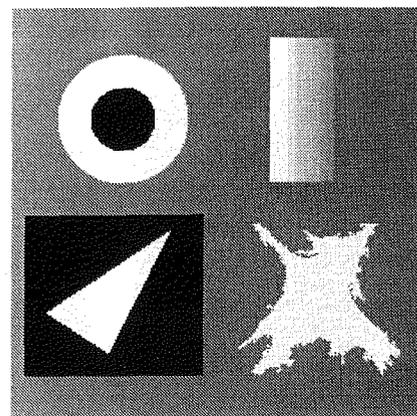


Figure 17. – Image de synthèse.

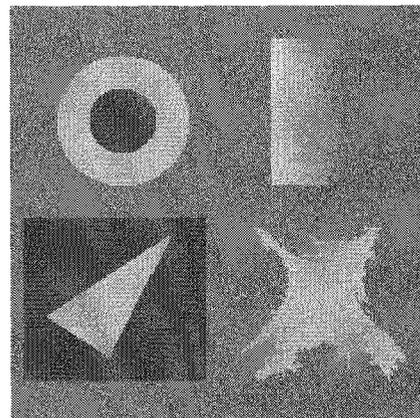


Figure 18. – Image BRUIT1.

4.2. comparaison des stratégies de relaxation par champ de Markov

Le but de cette section est de comparer la stratégie de relaxation par champ de Markov développée avec d'autres méthodes déjà existantes. Les résultats de trois stratégies de relaxation sont ainsi comparés. Il s'agit de :

- la relaxation sur un graphe d'adjacence;
- la relaxation définie sur le champ d'observation;
- la relaxation multi-échelle.

Avant de décrire plus en détails ces différents algorithmes, nous précisons que pour tous les résultats de ce paragraphe, seule une trame de l'image a été considérée. D'autre part, la relaxation est toujours fondée sur l'algorithme *ICM*. Le modèle markovien utilisé est le même pour les trois algorithmes : celui présenté dans la section précédente. De grandes étapes ont été conservées pour les trois méthodes de relaxation, à savoir : l'estimation des paramètres, la modélisation des énergies, la création de nouvelles étiquettes et la variation du paramètre α reliant les deux énergies au cours des itérations.

4.2.1. présentation des trois stratégies de relaxation testées

1) Relaxation définie sur un graphe d'adjacence

L'algorithme considéré est l'algorithme présenté dans cet article. Les paramètres utilisés sont :

- des secteurs initiaux de 32×32 pixels;
- une échelle de niveaux de gris réduite à 64 pour la division.

Les figures 19 et 20 présentent les résultats de segmentation ainsi obtenus. Seuls les contours des régions sont présentés dans un but de clarté.

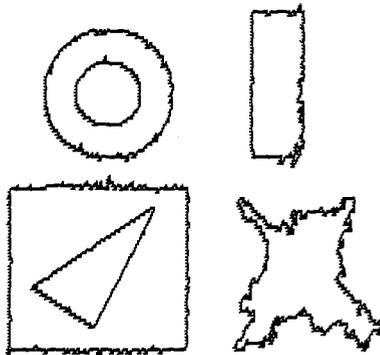


Figure 19. - Résultat sur l'image BRUIT1.

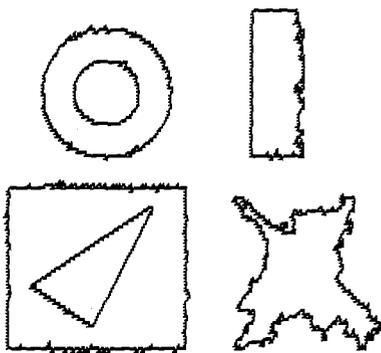


Figure 20. - Résultat sur l'image BRUIT2.

2) Relaxation définie sur le champ d'observation

L'algorithme présenté est un algorithme classique de relaxation par champ de Markov utilisant uniquement les niveaux de gris. Les champs de Markov sont directement définis sur le champ d'observation i.e. les pixels, il n'y a pas d'étape préliminaire de division de l'image. Le nombre de sites lors de la relaxation est donc égal au nombre de pixels de l'image. Les résultats de segmentation de cet algorithme sont présentés figures 21 et 22.

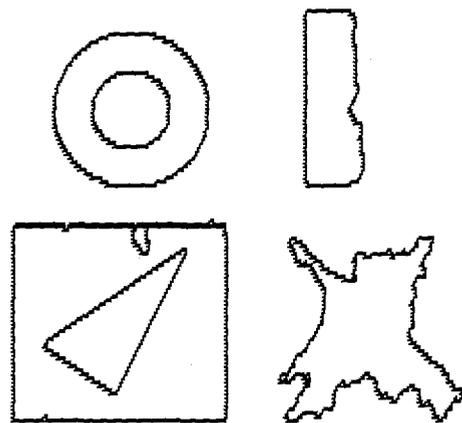


Figure 21. - Résultat sur l'image BRUIT1.

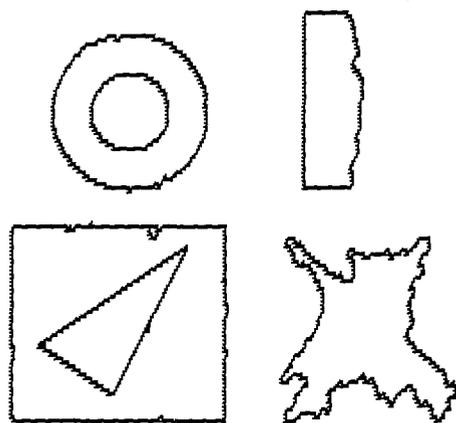


Figure 22. - Résultat sur l'image BRUIT2.

3) Relaxation multi-échelle sur le champ d'observation

De nombreux algorithmes de segmentation, reposant sur une relaxation par champs de Markov, utilisent une approche multi-échelle. Celle-ci a été introduite par Perez et Heitz [PH92] pour réaliser une segmentation en régions au sens du mouvement. L'idée est de résoudre le problème de la segmentation dans des espaces de configurations emboîtés $\Omega^n \subset \Omega^{n-1} \subset \dots \subset \Omega^1 \subset \dots \subset \Omega^0 = \Omega$ correspondant à des configurations de plus en plus

finies. Considérons une grille $S = \{S_1, \dots, S_N\}$ de dimension $2^m \times 2^n$ et un système de voisinage \mathcal{G} défini sur celle-ci. Un champ markovien aléatoire multi-grille peut être engendré à partir de cette grille avec la procédure :

- 1) $B^0 = S, \Omega^0 = \Omega$ et $i = 0$;
- 2) $i = i + 1$; S est divisé en blocs de taille $2^i \times 2^i$. Ces blocs vont former l'échelle $B^i = \{b_1^i, b_2^i, \dots, b_{N_i}^i\}$ et vont être associés à l'espace de configurations Ω^i .

Toutes les étiquettes des pixels appartenant à un bloc sont supposées être identiques. Cette contrainte conduit à un espace de configuration Ω^i qui est un sous-ensemble de l'ensemble original Ω . La démarche employée est illustrée figure 23. L'algorithme de relaxation est le suivant : plutôt que de minimiser la fonction originale U , une séquence de tâches, qui consiste à minimiser les U_i aux différents niveaux de la pyramide, est introduite. Dans un premier temps, le problème est résolu au niveau i en utilisant une relaxation classique mono-échelle. Le niveau $i - 1$ est alors initialisé par projection du champ des étiquettes obtenu sur la nouvelle grille plus fine, et ainsi de suite.

Cette approche multi-échelle a été utilisée par de nombreux auteurs. Bouman et Shapiro [BS94] réalisent ainsi une segmentation d'images texturées, Ackah-Miezan et Gagalowicz [AMG93] une segmentation d'images au sens des niveaux de gris. L'avantage des approches multi-échelle (combinées à des algorithmes de relaxation déterministes pour les différentes optimisations) est qu'elles conduisent à de bons résultats, très proches de ceux obtenus avec des algorithmes stochastiques mono-échelle. Les coûts algorithmiques sont, de plus, inférieurs à ceux des algorithmes déterministes de type *ICM*. Quand aux temps de calcul, ils sont de 60 à 100 fois plus faibles que ceux des méthodes stochastiques mono-échelles.

Les résultats de segmentation sur les deux images précédentes sont présentés figures 24 et 25.

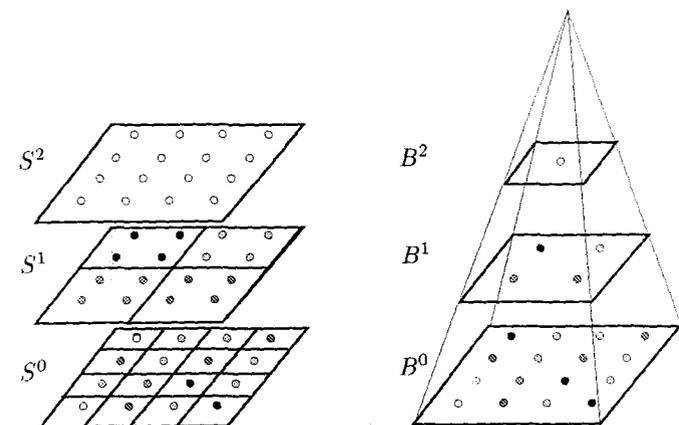


Figure 23. – Les différentes échelles.

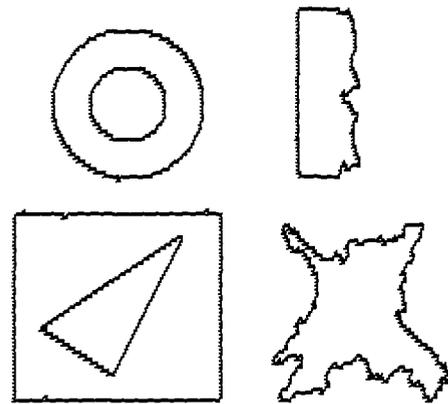


Figure 24. – Image BRUIT1.

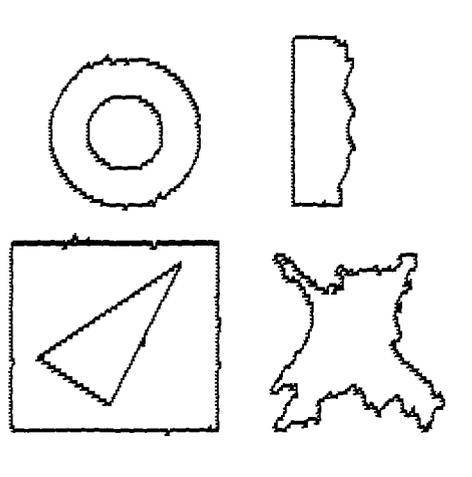


Figure 25. – Image BRUIT2.

4.3. comparaison des différents algorithmes

Nous avons présenté dans les paragraphes précédents les résultats de segmentation des trois algorithmes. Afin de pouvoir les comparer, nous avons introduit § 4.1.1 la mesure de Vinet modifiée. Pour chaque image résultat, nous évaluons la distance qu'il y a entre celle-ci et la segmentation idéale (une faible valeur de la mesure correspond à une bonne segmentation). Par souci de lisibilité, toutes les mesures ont été multipliées par 100. Dans les tableaux I et II sont présentées les différentes mesures pour les trois algorithmes (une mesure sur chaque quadrant et une mesure globale sur toute l'image) ainsi que les temps de calcul obtenus en milli-secondes. Rappelons ici que seule une trame de l'image a été considérée, et que les temps de calcul ont été obtenus sur une station de travail HP9000/735. Les quadrants introduits ici sont ceux représentés figure 26.

Tableau 1. – Image BRUIT1.

	quad. 1	quad. 2	quad. 3	quad. 4	global	Temps
Relaxation sur un graphe	0.64	5.16	1.16	1.51	2.12	450
Relaxation sur le champ d'observation	0.05	7.42	0.67	1.32	2.37	28960
Relaxation multi-échelle sur le champ d'observation	0.04	6.38	0.11	1.12	1.91	3320

Tableau 2. – Image BRUIT2.

	quad. 1	quad. 2	quad. 3	quad. 4	global	Temps
Relaxation sur un graphe	0.44	5.54	0.95	1.14	2.02	410
Relaxation sur le champ d'observation	0.17	7.13	0.37	1.22	2.22	21060
Relaxation multi-échelle sur le champ d'observation	0.22	6.50	0.19	1.03	1.99	4070

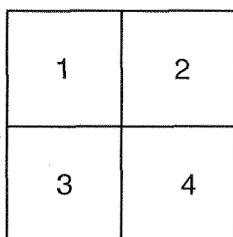


Figure 26. – Définition des quadrants.

Plusieurs conclusions peuvent être tirées de ces tableaux. Tout d'abord, de manière générale, l'algorithme de relaxation défini sur le graphe d'adjacence aboutit à des temps de calcul faibles (≈ 500 ms) pour une meilleure qualité de résultats que celle obtenue avec la relaxation sur le champ d'observation.

Nous remarquons, à partir des tableaux précédents, l'intérêt de diminuer le nombre de sites élémentaires lors de la relaxation :

- tout d'abord, le coût algorithmique est bien plus faible que celui obtenu avec une relaxation sur le champ d'observation;
- d'autre part, l'algorithme ainsi développé est moins sensible aux minima locaux que les méthodes de relaxation sur le champ d'observation. En effet, beaucoup de pixels sont déjà regroupés au début du processus markovien (étape de division). La relaxation est initialisée avec une carte de régions proche de la solution finale, ce qui permet à l'algorithme *ICM* d'éviter un grand nombre de minima locaux.

Notre méthode a également été comparée avec la relaxation multi-échelle. Les résultats sont d'un peu moins bonne qualité, mais les temps de calcul ont fortement diminué.

Nous avons vérifié, figure 27 à 30, ces phénomènes sur deux images réelles segmentées avec :

- a. l'algorithme de relaxation sur le graphe d'adjacence;
- b. l'algorithme de relaxation multi-échelle.

Du point de vue qualité des résultats, il semble que les deux approches donnent des résultats de qualité similaire. Les temps de calcul (présentés entre parenthèses) sont quant à eux très probants puisqu'il y a un facteur 25 entre les deux méthodes. Ces temps dépendent cependant des images et plus particulièrement, de leur complexité.

Il faut néanmoins noter que les résultats de segmentation obtenus avec la méthode présentée dans cet article dépendent de la qualité de la division. Il est bien évident que si des régions ne sont pas détectées lors de cette division, elle ne seront pas retrouvées par la suite. Ce cas de figure ne peut se produire que pour des images fortement bruitées.

De manière générale, la méthode développée ici conduit à des résultats de qualité similaire à ceux obtenus avec d'autres méthodes de relaxation. D'autre part, les temps de calcul sont faibles relativement à ceux obtenus avec les autres stratégies.

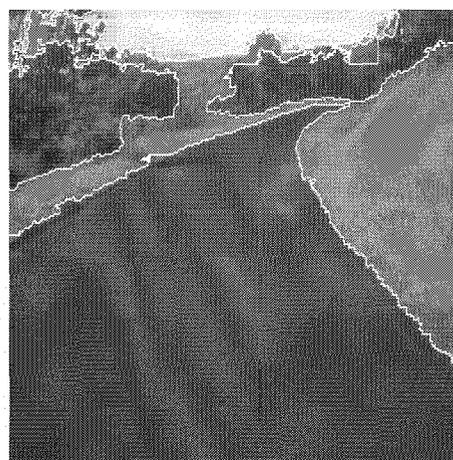


Figure 27. – Relaxation sur le graphe (900ms).

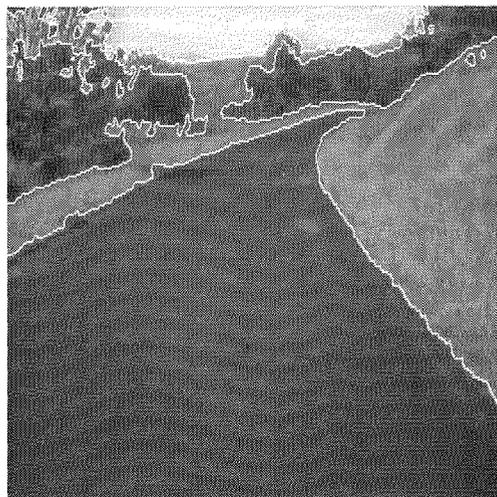


Figure 28. – Relaxation multi-échelle (23380ms).

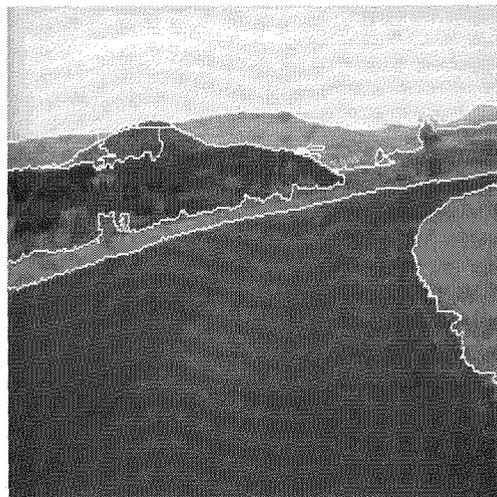


Figure 29. – Relaxation sur le graphe (780 ms).

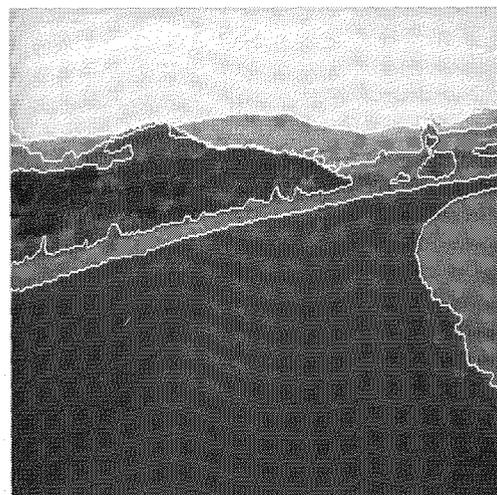


Figure 30. – Relaxation multi-échelle (23670ms).

4.4. conclusion

Nous avons présenté dans cette section une étude comparative entre les résultats obtenus avec l'algorithme développé dans cet article et ceux obtenus par d'autres stratégies. Cette étude permet de tirer plusieurs conclusions.

Etude des temps de calcul

Un des buts principaux de l'algorithme était, dès le départ, de pouvoir être utilisé dans des applications réelles, ce qui impliquait d'obtenir des temps de calcul raisonnables. Après avoir choisi d'utiliser un algorithme de relaxation par champ de Markov pour réaliser la fusion, nous nous sommes confrontés aux forts coûts algorithmiques qu'elle engendre. Pour pallier ceux-ci, une première étape de division de l'image a été mise en place avant de réaliser la relaxation. Elle permet de diminuer le nombre de données initiales présentes lors du processus markovien et par la suite, de diminuer considérablement les temps de calcul. Lors de l'étude comparative (§ 4.2), nous avons constaté que l'algorithme ainsi développé est plus rapide que les autres méthodes de relaxation par champ de Markov (même en considérant la relaxation multi-échelle). Ces gains en temps de calcul permettent d'envisager de manière plus réaliste une application opérationnelle.

Etude qualitative

L'algorithme de relaxation sur un graphe d'adjacence conduit à des résultats de qualité similaire à ceux obtenus avec d'autres relaxations. De manière générale, les différents objets de la scène sont bien retrouvés.

5. conclusion

Nous avons développé dans cet article une méthode de segmentation non supervisée où le nombre d'étiquettes et les paramètres des régions sont estimés. Une préoccupation majeure a été de ne tenir compte d'aucune information *a priori* sur le type d'images traitées. Les résultats sur ce point sont concluants puisque des segmentations aussi bien sur des images de synthèse que sur différents types d'images réelles de scènes d'extérieur sont présentées. Elles ont toutes été obtenues sans changer de seuils lors de l'exécution de l'algorithme.

La méthode de segmentation développée dans cet article est de type division-fusion. La division est réalisée afin d'initialiser un processus de relaxation par champ de Markov. Cette technique permet, comme nous l'avons montré, de diminuer considérablement les temps d'exécution.

La division a été présentée lors d'une première partie; des statistiques d'ordre deux ont été choisies pour rechercher les différentes

régions présentes dans la fenêtre étudiée. En effet, ces régions sont extraites à partir de l'étude de la forme des matrices de co-occurrence. Nous avons montré, lors d'expériences sur des images réelles, que ces statistiques étaient plus robustes au bruit que la simple utilisation d'histogrammes locaux.

La fusion, détaillée dans un second temps, est directement initialisée par la carte de régions issue de la phase de division. Afin d'obtenir des résultats de bonne qualité, une relaxation par champ de Markov est utilisée. Celle-ci, fondée sur une analyse statistique des images, présente l'inconvénient de conduire à des temps de calcul élevés. Pour remédier à ce problème, la relaxation proposée est réalisée de manière particulière. En effet, le champ de Markov est modélisé sur le graphe d'adjacence irrégulier des régions issues de la division. Ceci entraîne une diminution du nombre de sites et permet, comme nous l'avons montré lors d'une étude comparative, de diminuer énormément les temps de calcul tout en gardant des résultats de bonne qualité.

Par l'introduction d'un nouveau paramètre (α) lors de la relaxation et par son estimation, nous avons rendu celle-ci non supervisée et indépendante du type d'image traité. En effet, ce paramètre varie directement en fonction de la baisse d'énergie, permettant ainsi d'équilibrer les différents termes intervenant lors de la relaxation. Une modélisation des énergies, tout au long de la relaxation, par une loi log-normale permet d'autre part de connaître exactement l'évolution de celles-ci au cours des itérations et par la suite, de définir un processus robuste de création de nouvelles étiquettes. Rappelons ici que l'algorithme est entièrement non supervisé et donc que, ni les paramètres des classes, ni leur nombre ne sont connus *a priori*.

Lors d'une dernière partie, nous avons comparé notre méthode de segmentation avec d'autres déjà existantes. Cette étude, menée sur des images de synthèse, a permis de mettre en avant les avantages de notre algorithme (temps de calcul et qualité des résultats).

BIBLIOGRAPHIE

- [1] R. Azencott, C. Graffigne. Non supervised segmentation using multi-level markov random fields, In *Proc. International Joint Conf. on Pattern Recognition*, vol. 3, p. 201–205, The Hague, the Netherlands, septembre 1992.
- [2] A. Ackah-Miezan, A. Gagalowicz. Discrete models for energy minimizing segmentation, In *International Conference on computer vision*, p. 200-207, Berlin, mai 1993.
- [3] R. Azencott. Champs markoviens et imagerie, *Congrès AFCET. Reconnaissance des formes et intelligence artificielle*, p. 1183–1191, 1987.
- [4] J. Besag. *On the statistical analysis of dirty pictures*, vol. 48 of *B. J. Royal Stat. Soc.*, 1986.
- [5] C. Bouman, M. Shapiro. A multiscale random field model for bayesian image segmentation, *IEEE Trans. Image Processing*, 3(2) : 162–177, mars 1994.
- [6] G. Celeux, J. Diebolt. The SEM algorithm : a probabilistic teacher algorithm derived from the EM algorithm for the Mixture Problem, *Computation statistics quarterly*, 2 : 73–82, 1985.
- [7] J.P. Cocquerez, S. Philipp. *Analyse d'images : filtrage et segmentation*, collection Enseignement de la physique, Masson, Paris, 1995.
- [8] M. Derras, C. Debain, M. Berducot, P. Bonton, J. Gallice. Unsupervised regions segmentation : real time control of an upkeep machine of natural spaces, *Ineuropan conference on computer vision*, vol. B, p. 207–212, Stockholm, Sweden, mai 1994, vol. B.
- [9] H. Derin, H. Elliot. Modeling and segmentation of noisy and textured images using gibbs random fields, *IEEE trans. pattern analysis and machine intelligence*, 9(1) : 39–55, janvier 1987.
- [10] A.P. Dempster, N.M. Laird, D.B. Rubin. Maximum likelihood from incomplete Data via the EM Algorithm, *J. Roy. Statist. Soc.*, 39(1) : 1–38, 1977, série B. methodological.
- [11] S. Geman, D. Geman. Stochastic relaxation, gibbs distribution, and the Bayesian restoration of images, *IEEE trans. pattern analysis and machine intelligence*, PAMI 6(6) : 721–741, novembre 1984.
- [12] D. Geman, S. Geman, C. Graffigne, P. Dong. Boundary detection by constrained optimization, *IEEE trans. pattern analysis and machine intelligence*, 12(7) : 609–628, juillet 1990.
- [13] R.C. Gonzalez, P. Wintz. *Digital image processing*, Addison-Wesley publishing comagny, 1987, second edition.
- [14] I.L. Herlin, D. Bereziat, G. Giraudon, C. Nguyen, C. Graffigne. Segmentation of Echocardiographic images with Markov random fields, technical report 2424, INRIA, décembre 1994.
- [15] S. Houzelle, G. Giraudon. Segmentation en régions par modélisation de matrices de cooccurrence. In *AFCET, 8^{ème} congrès de reconnaissance des formes et intelligence artificielle*, vol. 3, p. 1177–1183, Lyon Villeurbanne, novembre 1991, tome 3.
- [16] F. Heitz, P. Perez, P. Boutheymy. Multiscale minimization of global energy functions in some visual recovery problems; *computer vision, graphics and image processing*, 59(1) : 125–134, janvier 1994, image understanding.
- [17] R. Kara-Falah, P. Bolon. Mesure de dissimilarité entre deux segmentations. In *Quatorzième colloque GRETSI*, p. 763–766, Juan les Pins, septembre 1993.
- [18] C. Kervrann, F. Heitz. A markov random field model based approach to unsupervised texture segmentation using local and global spatial statistics, *IEEE trans. image processing*, 4(6) : 856-862, 1995.
- [19] H.Y. Kim, H.S. Yang. A systematic way for region-based image segmentation based on Markov random fields model, *Pattern recognition letters*, 15 : 969–976, 1994.
- [20] S. Lakshmanan, H. Derin. Simultaneous parameter estimation and segmentation of gibbs random fields using simulated annealing, *IEEE trans. pattern analysis and machine intelligence*, 11(8) : 799–813, août 1989.
- [21] B.S. Manjunath, R. Chellappa. Unsupervised texture segmentation using Markov random field models, *IEEE trans. pattern analysis and machine intelligence*, 13(5) : 478–482, mai 1991.
- [22] J.W. Modestino, J. Zhang. A Markov random field model-based approach to image interpretation, *IEEE trans. pattern analysis and machine intelligence*, 14(6) : 606-615, juin 1992.
- [23] P. Perez, F. Heitz. Une approche multi-échelle à l'analyse d'images par champs markoviens, *Traitement du Signal*, 9(6) : 459–472, 1992.
- [24] W. Pieczynski. Champs de Markov cachés et estimation conditionnelle itérative, *Traitement du Signal*, 11(2) : 141–153, 1994.
- [25] N.R. Pal, S.K. Pal. A review of image segmentation technique, *pattern recognition* 26(9) : 1277–1294, 1993.

[26] C. Rouquet, F. Chausse, R. Chapuis, P. Bonton. Segmentation non-supervisée d'images de scènes routières, une approche multi-critères, *Traitement du Signal*, 13(3) : 195–208.

[27] L. Vinet, *Segmentation et mise en correspondance de régions de paires d'images stéréoscopiques*, PhD thesis, Université de Dauphine, Paris IX, juillet 1991.

[28] Y.J. Zhang, J.J. Gerbrands. Segmentation evaluation using ultimate measurement accuracy, *image processing algorithms and techniques III*, SPIE 1657 : 449–460, 1992.

Manuscrit reçu le 5 décembre 1996.

LES AUTEURS

Catherine ROUQUET



Catherine Rouquet est née à Clermont-Ferrand, le 10 décembre 1970. Elle a obtenu le DEA d'Electronique et Système, option Vision pour la Robotique en 1993 puis un doctorat en 1996. Ces deux diplômes ont été préparés au LASMEA, à l'Université Blaise-Pascal de Clermont-Ferrand. Elle est maintenant Maître de Conférences à l'Université Pierre et Marie Curie, Paris 6 et pour la recherche est rattachée au LIS. Ses nouveaux travaux de recherche entrent dans le cadre de l'indexation d'images couleur.

Pierre BONTON



Pierre Bonton est né à Billon (63) le 27 novembre 1947. Professeur des Universités à Clermont-Ferrand (Université Blaise Pascal), il est responsable du DESS « Contrôle Industriel et Gestion de la Qualité ». Sa recherche s'est orientée vers la segmentation d'images depuis quelques années. Les principales applications sont tournées vers la robotique mobile agricole et le médical. Il a cinq ans d'expériences industrielles en conception. Il est aussi responsable de l'action « Théâtre et Sciences » à Clermont-Ferrand.

Régis TOMCZAK



Régis Tomczak est né à Vichy le 11 décembre 1971. Il obtient le DEA d'Electronique et Système, option Vision pour la Robotique en 1994. Actuellement doctorant de 3ème année au sein du LASMEA, son sujet de thèse est la mesure automatique des taux de pollens dans l'air par traitement d'images.