

Segmentation en régions non supervisée par relaxation markovienne, une étude comparative

Catherine ROUQUET et Pierre BONTON

Laboratoire des Sciences et Matériaux pour l'Electronique, et d'Automatique
UMR 6602 du CNRS - Université Blaise Pascal
24,avenue des Landais - 63110 Aubière Cedex

RÉSUMÉ

Nous présentons un algorithme de segmentation en régions non supervisée qui utilise la théorie des champs de Markov. Ces techniques de relaxation markoviennes amènent à de bons résultats mais présentent cependant les inconvénients, dans leur implantation classique, de conduire à des temps de calcul élevés et de nécessiter l'introduction de plusieurs seuils. Nous nous sommes donc proposé de résoudre ces deux problèmes majeurs et avons développé une nouvelle technique de relaxation markovienne.

Une étude comparative des différentes stratégies d'utilisation des relaxations markoviennes est ensuite menée, de manière rigoureuse, aussi bien du point de vue qualité des résultats que du point de vue coût algorithmique. Des conclusions intéressantes en découlent.

La segmentation reste un des problèmes fondamentaux de l'analyse d'images. Son rôle consiste, à partir des 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. Afin de résoudre ce problème, nous utilisons des méthodes de relaxation markovienne, particulièrement réputées pour la qualité des résultats auxquels elles amènent. Celles-ci possèdent cependant les inconvénients de demander de grands temps de calcul et de nécessiter l'introduction de plusieurs seuils. Afin de remédier à ces inconvénients, nous avons mis en place une méthode de relaxation particulière. En effet, elle est réalisée, non pas sur des pixels comme usuellement, mais sur des régions déterminées antérieurement lors d'un processus de division. Ceci permet de diminuer considérablement le nombre d'entités présentes dans la relaxation et par la suite, les temps de calcul. D'autre part, certains éléments, comme par exemple une étude de la variation d'énergie au cours de la relaxation, permettent de fixer, de manière adaptative, tous les seuils de l'algorithme, rendant celui-ci entièrement non supervisée.

Dans le premier chapitre, nous développons la technique de division utilisée. La fusion, fondée sur une relaxation markovienne définie sur le graphe d'adjacence, est décrite dans le second chapitre. Elle amène à présenter des résultats expérimentaux. Afin d'estimer la qualité de ceux-ci, nous menons lors du troisième chapitre une étude comparative entre

ABSTRACT

We present a non-supervised region based segmentation algorithm which uses Markov Random Fields. These markovian relaxation techniques lead to good results but present some inconvenients in their classic implementation. Indeed, they lead to important computation times and need using some thresholds. So, we proposed to resolve these problems and have developed a new markovian relaxation technique.

A comparative study on the different strategies of use of markovian relaxation is then led on synthesis images. It takes into account quality of results as well as computation times. Interesting results of comparison are presented.

les résultats obtenus par différentes techniques de relaxation markovienne.

1 La division

L'algorithme de division est fondé sur l'étude de la forme des matrices de cooccurrence définies par Haralick [HSD73] et plus particulièrement, sur celle de leur diagonale [HG91]. En effet, cette diagonale peut être utilisée comme un histogramme de niveaux de gris : recherche des pics et des vallées, puis extraction des classes. L'avantage des matrices de cooccurrence dans ce contexte est que celles-ci sont basées sur des statistiques d'ordre deux et utilisent donc la notion de voisinage ; ce qui les rend beaucoup moins sensible au bruit. Ceci a été montré, de manière théorique sur des modèles de bruit simples, et de manière pratique sur plusieurs images réelles bruitées dans le cadre de la division [ROU96].

L'algorithme de division se décompose en cinq phases :

- Découpage de l'image en secteurs.
L'image est découpée en fenêtres rectangulaires à l'intérieur desquelles sera effectuée la division.
- Calcul de la diagonale de la matrice de cooccurrence.
La matrice de cooccurrence est déterminée à partir d'une distance de 1 pixel et des orientations 0 et 90 degrés. De plus, une échelle de niveaux de gris réduite de façon linéaire à 64 est utilisée.

- Extraction des classes.
Une classe est définie à partir d'un pic de la diagonale et de ses deux vallées.
- Ré-étiquetage de l'image.
L'image est ré-étiquetée afin de donner une étiquette différente à chaque ensemble de pixels connexes appartenant à la même classe.
- Suppression des petites régions.
Toutes les régions dont la taille est inférieure à 25 pixels sont supprimées par fusion avec la région voisine ayant la moyenne de niveaux de gris la plus proche.

Cet algorithme de division, basé sur des statistiques d'ordre deux, permet d'obtenir de bons résultats qui maintenant, initialisent le processus de fusion.

2 La fusion

2.1 Les données

Les données d'entrée du processus de fusion sont les régions issues de la division. Celles-ci sont modélisées par un graphe d'adjacence. Le vecteur de caractéristiques des régions est composé de trois paramètres :

- la moyenne des niveaux de gris, notée μ ;
- l'énergie, notée E ;
- l'homogénéité, notée H .

Les deux derniers paramètres sont liés à la texture. Ils sont définis par Haralick [HSD73] à partir des matrices de cooccurrence.

2.2 Approche statistique

Afin de résoudre le problème de la fusion, nous nous plaçons dans le cadre des champs de Markov. Les régions forment le champ de données $R = \{R_1, \dots, R_N\}$. Le problème est de mettre en correspondance ce champ de données et un champ d'étiquettes $E = \{e_1, e_2, \dots, e_N\}$. Nous nous plaçons dans le contexte de la maximisation *a posteriori* et des champs de Gibbs et nous ramenons à la minimisation de la somme de deux énergies :

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

où α est utilisé pour ajuster les deux énergies.

Pour minimiser cette somme, nous utilisons l'algorithme *ICM* (Iterated Conditional Mode) [DJ89] qui consiste à minimiser l'énergie de manière locale, en chaque région. De plus amples détails sur la relaxation peuvent être trouvés dans [RCCB96]. D'autres méthodes de relaxation sur des régions sont exposées dans [AG92] [MZ92]. Pour une région R_i , d'étiquette e_i , l'énergie affectée à cette région est :

$$U(e_i/R_i, \mathcal{G}_{R_i}) = U(R_i/e_i) + \alpha \times U(e_i/e_{\mathcal{G}_{R_i}}) \quad (2)$$

où :

- \mathcal{G}_{R_i} représente le voisinage de la région R_i .
- $e_{\mathcal{G}_{R_i}}$ représente les étiquettes des régions voisines à R_i .

Nous modélisons l'énergie de transposition $U(R_i/e_i)$ par la distance euclidienne entre le vecteur de caractéristiques de la région R_i et celui de l'étiquette e_i :

$$U(R_i/e_i) = 4 \times (\mu_{R_i} - \mu_{e_i})^2 + (H_{R_i} - H_{e_i})^2 + (E_{R_i} - E_{e_i})^2 \quad (3)$$

L'énergie a priori $U(e_i/e_{\mathcal{G}_{R_i}})$ est modélisée par une fonction linéaire :

$$U(e_i/e_{\mathcal{G}_{R_i}}) = 1 - \frac{\sum_{R_j \in \mathcal{G}_{R_i}} \delta(e_j = e_i)}{\sum_{R_j \in \mathcal{G}_{R_i}} 1} \quad (4)$$

où $\delta(x) = 1$ si x est vraie et est nulle dans les autres cas.

La première énergie, $U(R/E)$, diminue au cours des itérations, tandis que la deuxième, $U(E)$, augmente avec celles-ci. Une valeur de α inférieure à 1 favorise l'énergie $U(R/E)$ au départ, ce qui tend à créer de nouvelles étiquettes. A la fin du processus, les deux énergies sont équilibrées de manière à stabiliser le système. Le problème est que l'on ne connaît pas le minimum d'énergie à atteindre et donc, qu'il est impossible de déterminer la valeur de α rigoureusement, dès le début de la relaxation. Pour pallier ce problème et obtenir une segmentation adaptative, nous faisons varier α au cours des itérations. L'idée est de le faire varier de manière inverse à l'énergie. Une faible valeur de α au début de la relaxation (forte énergie) permettra la création de nouvelles étiquettes, alors qu'une forte valeur à la fin des itérations (faible énergie) stabilisera le processus. Nous présentons ci-dessous l'évolution de α au cours des itérations :

- $k = 0, \alpha = 0$; *Stabilisation du processus et extraction de l'énergie $U(R_i/e_i)$ maximale : U_{0_max} ;*
- *Pour les itérations suivantes, et jusqu'à la stabilisation du processus : extraction de U_{k_max} . La nouvelle valeur de α est fixée à : $\alpha = U_{0_max} - U_{k_max}$*

Nous avons implanté cette méthode de variation du paramètre α dans notre algorithme et avons obtenu de bons résultats. L'algorithme est devenu adaptatif.

2.3 Résultats expérimentaux

Afin de montrer la généralité de l'algorithme, nous présentons des résultats de segmentation sur des images de taille 512×512 et de types différents (figures 1, 2).

Pour ces images, de bons résultats de segmentation sont obtenus, sans changer aucun seuils. Nous avons fortement diminué les temps de calcul mis par la relaxation markovienne en initialisant celle-ci par le graphe d'adjacence des régions obtenu par la méthode de division. Afin d'estimer les gains de temps réalisés, nous devons comparer cette méthode aux autres méthodes de relaxation existantes.

3 Etude comparative

Afin de comparer plusieurs algorithmes de segmentation, nous utilisons une image de synthèse provenant de la banque



FIG. 1 — Résultat de segmentation.



FIG. 2 — Résultat de segmentation.

de donnée du GdR TDSI (figure 3). Les différentes segmentations sont évaluées par rapport à la segmentation idéale, connue a priori, en utilisant la mesure de Vinet. Celle-ci, présentée dans [VIN91], mesure la distance entre deux segmentations. Une faible valeur de la mesure correspond à deux segmentations très proches, tandis qu'une forte valeur représente deux cartes de régions assez éloignées.

Dans le premier paragraphe, nous définissons les algorithmes que nous comparons. Les résultats expérimentaux sont présentés dans un second paragraphe et amènent à tirer quelques conclusions.

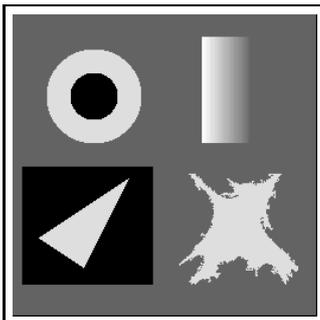


FIG. 3 — Image de synthèse.

3.1 Présentation des algorithmes utilisés

Quatre algorithmes de relaxation markovienne sont étudiés. Pour tous, la relaxation est fondée sur l'algorithme *ICM* et le modèle markovien utilisé est :

$$U(s_i/e_i) = (\mu_{s_i} - \mu_{e_i})^2 \quad (5)$$

$$U(e_i/e_{\mathcal{G}_{s_i}}) = 1 - \frac{\sum_{s_j \in \mathcal{G}_{s_i}} \delta(e_j = e_i)}{\sum_{s_j \in \mathcal{G}_{s_i}} 1} \quad (6)$$

Nous détaillons ci-dessous chacun de ces algorithmes.

1 - Relaxation définie sur un graphe d'adjacence. L'algorithme considéré ici est une version légèrement modifiée de celui décrit dans ce document car la texture n'intervient plus.

2 - Relaxation double-échelles définie sur un graphe d'adjacence. Cet algorithme contient deux niveaux ou échelles. Le premier est l'algorithme présenté ci-dessus où les champs de Markov sont définis sur un graphe d'adjacence. Le second niveau de relaxation est initialisé par le résultat de segmentation du premier. Ici, la modélisation des champs de Markov est réalisée directement sur le champ d'observation (les pixels).

3 - Relaxation définie sur le champ d'observation. C'est l'algorithme de relaxation markovienne classique où les champs de Markov sont directement définis sur le champ d'observation.

4 - Relaxation multi-échelles définie sur le champ d'observation. L'algorithme repose sur l'approche multi-échelles introduite par Heitz et al. [HPB94] où la minimisation des énergies est fondée sur une structuration pyramidale des étiquettes. Cette méthode amène à de bons résultats, avec des temps de calcul plus faibles que ceux obtenus par une relaxation classique

3.2 Comparaison des résultats

Nous avons appliqué les quatre algorithmes précédents à deux images de synthèse. Celles-ci correspondent à l'image présentée ci-dessus additionnée par des bruits blancs de distributions gaussienne et exponentielle. Nous appelons *BRUIT1* et *BRUIT2* ces deux images.

Dans le tableau ci-dessous, nous présentons la mesure de Vinet estimée à partir des segmentations déterminées et de la segmentation idéale. Les temps (en milli-secondes) ont été obtenus sur une station de travail *HP9000/735* en considérant une seule trame de l'image. Les différents algorithmes sont référencés en fonction du numéro qui leur a été attribué.

Algorithme	1	2	3	4
Image BRUIT1				
Mesure de Vinet	2.12	1.58	2.37	1.91
Temps (ms)	450	2800	28960	3320
Image BRUIT2				
Mesure de Vinet	2.02	1.76	2.22	1.99
Temps (ms)	410	1870	21060	4070

Plusieurs conclusions peuvent être tirées de ce tableau. Tout d'abord, de manière générale, l'algorithme de relaxation défini sur un graphe d'adjacence (1) mène à des temps de calcul faibles pour une qualité de résultats comparable à celle obtenue avec la relaxation sur le champ d'observation (3).

De plus, l'algorithme proposé 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. 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.

Cependant, un des inconvénients de cette méthode est que les régions issues de l'étape de division sont figées, elles ne peuvent pas être changées lors de la relaxation. Afin de résoudre ce problème, nous avons développé l'algorithme double-échelles (2) : suite à l'algorithme précédent, une seconde relaxation, initialisée par le résultat du premier algorithme, est réalisée directement sur le champ d'observation. Cette étape permet d'affiner les contours des différentes régions et donc, d'amener à de meilleurs résultats. Les temps de calcul eux sont bien évidemment augmentés. Ils restent cependant inférieurs à ceux obtenus avec une relaxation sur le champ d'observation (cinq fois plus faibles). Du point de vue qualité des résultats, nous voyons que ceux obtenus avec cette méthode sont meilleurs que les segmentations obtenues avec une relaxation sur le champ d'observation (3).

Nous avons également comparé notre méthode avec la relaxation multi-échelles définie sur le champ d'observation (4). Nous avons pu conclure que les résultats de notre méthode sont obtenus plus rapidement et de plus, sont de meilleure qualité.

De manière générale, la méthode développée conduit à des résultats de meilleure qualité que ceux obtenus avec d'autres méthodes de relaxation. De plus, les temps de calcul sont plus faibles.

3.3 Conclusion

Nous proposons une méthode de segmentation en régions non supervisée, de type division-fusion. La division est réalisée afin d'initialiser un processus de relaxation markovienne. Cette technique permet, comme nous l'avons montré, de diminuer considérablement les temps d'exécution mis pour obtenir les segmentations.

La division, utilisant uniquement les niveaux de gris, a été présentée lors d'une première partie. Son originalité réside dans le fait que des statistiques d'ordre deux ont été choisies pour rechercher les différentes régions et rendre ainsi la division plus robuste.

La fusion est détaillée dans un second temps. Afin d'obtenir des résultats de bonne qualité, nous utilisons une relaxation markovienne. Celle-ci présente cependant l'inconvénient de conduire à des temps de calcul élevés. Pour pallier ce problème, la relaxation proposée est réalisée de manière particulière. En effet, le champ de Markov est modélisé sur le graphe d'adjacence 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 lors de la relaxation markovienne, nous avons rendu celle-ci entièrement autonome et indépendante du type d'images traitées. En effet, celui-ci varie directement en fonction de la baisse d'énergie, permettant ainsi de réguler les différents termes intervenant lors de la relaxation.

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

Références

- [AG92] AZENCOTT (R.) et GRAFFIGNE (C.). – Non Supervised Segmentation using Multi-Level Markov Random Fields. *Pattern Recognition*, vol. 3, Août 1992, pp. 201–205. – The Hague.
- [DJ89] DUBES (R.C.) et JAIN (A.K.). – Random Field Models in Image Analysis. *Journal of Applied Statistics*, vol. 16, n2, 1989, pp. 131–164.
- [HG91] HOUZELLE (S.) et GIRAUDON (G.). – Segmentation en régions par modélisation de matrices de cooccurrence. *Reconnaissance des Formes et Intelligence Artificielle*, vol. 3, 1991, pp. 1177–1183. – Lyon Villeurbanne.
- [HPB94] HEITZ (F.), PEREZ (P.) et BOUTHEMY (P.). – Multiscale Minimization of Global Energy Functions in some Visual Recovery Problems. *Computer Vision, Graphics and Image Processing*, vol. 59, n1, Janvier 1994, pp. 125–134.
- [HSD73] HARALICK (R.M.), SHANMUGAM (K.) et DINSTEN (I.). – Textural Features for Image Classification. *IEEE Trans. Systems, Man, and Cybernetics*, vol. 3, n6, 1973, pp. 610–621.
- [MZ92] MODESTINO (J.W.) et ZHANG (J.). – A Markov Random Field Model-Based Approach to Image Interpretation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, n6, 1992, pp. 606–615.
- [RCCB96] ROUQUET (C.), CHAUSSE (F.), CHAPUIS (R.) et BONTON (P.). – Segmentation non-supervisée d'images de scènes routières, une approche multicritères. *Traitement du Signal*, vol. 3, n13, 1996, pp. 195–208.
- [ROU96] ROUQUET (C.). – *Segmentation en régions non supervisée par relaxation markovienne*. – LASMEA, Thèse de PhD, Université Blaise-Pascal, Clermont-Ferrand, Juin 1996.
- [VIN91] VINET (L.). – *Segmentation et mise en correspondance de régions de paires d'images stéréoscopiques*. – Thèse de PhD, Université de Dauphine, Paris IX, 1991.