

Segmentation d'image couleur par grille de rectangles optimale selon une approche Bayésienne

Marc BOULLÉ, Aymeric LARRUE

France Télécom R&D
2, avenue Pierre Marzin, 23300 Lannion, France
marc.boullé@orange-ftgroup.com

Résumé – Ce papier introduit une nouvelle méthode de segmentation d'image selon une grille de régions rectangulaires, distribuées de façon homogène sur un ensemble de plages de niveaux de couleurs. La segmentation est ici considérée comme un problème de sélection de modèle, ce qui permet d'établir un critère d'évaluation d'une segmentation optimale selon une approche Bayésienne. Une heuristique d'optimisation performante est également proposée afin de rechercher la meilleure segmentation. Des premières expérimentations démontrent la validité de l'approche et ouvrent des perspectives pour la méthode proposée, en tant que technique générique de prétraitement des images.

Abstract – This paper introduces a new image segmentation method according to a grid of rectangular regions, distributed on a set of color level ranges. The segmentation is considered as a problem of model selection, which enables to establish an optimal evaluation criterion based on a Bayesian approach. An efficient search heuristic is proposed as well, in order to retrieve an efficient image segmentation. First experiments demonstrate the validity of the approach and suggest its potential as a generic technique for image preprocessing.

1 Introduction

La segmentation au sens colorimétrique d'une image consiste à la prétraiter en la partitionnant de façon fiable en zones homogènes vis-à-vis de composantes colorimétriques données. Ce prétraitement est très important pour de nombreuses applications de l'image, qui lors de traitements ultérieurs nécessitent de pouvoir travailler sur des informations fiables et pertinentes. De nombreuses approches ont été proposées pour la segmentation d'image. On distingue essentiellement les approches à base de contours et les approches à base de région [1]. Les approches à base de régions partitionnent les pixels en regroupant les pixels adjacents présentant des composantes colorimétriques similaires. Les régions adjacentes sont regroupées selon un critère prenant en compte un compromis entre l'homogénéité des régions et la simplicité des frontières des régions. Notre approche s'apparente à une approche par région.

Dans ce papier, nous proposons une approche générique basée sur une segmentation d'image en grille de rectangles, chaque rectangle étant caractérisé par la distribution de ses pixels sur des plages de niveaux de couleurs. Les dimensions de la grille sont données par une discrétisation des coordonnées d'espace, et les plages de niveaux résultent d'une discrétisation des couleurs de l'image. Chaque rectangle de la grille doit être le plus homogène possible vis-à-vis des couleurs utilisées, et les rectangles doivent être le plus différent possible (vis-à-vis des couleurs) deux à deux. L'approche s'apparente ainsi à une approche à base de régions, pour laquelle la forme des régions serait contrainte à une grille de rectangles. Le problème principal est alors de trouver un critère d'évaluation d'une segmentation réalisant un compromis efficace entre l'homogénéité des régions

(composantes colorimétriques uniformes par rectangle) et la complexité du modèle de segmentation (dimensions de la grille de rectangles et nombre de plages de niveaux par couleur). On propose un critère d'évaluation permettant de trouver la segmentation optimale d'une image en grille de rectangles, résultant de la mise en oeuvre d'une approche Bayésienne de la sélection de modèle.

D'autres travaux apparentés s'appuient sur des segmentations rectangulaires, selon des frontières parallèles aux axes. Par exemple, les segmentations par quadarbre [2] découpent récursivement une image selon les axes horizontaux et verticaux en optimisant un critère d'homogénéité pour les quatre régions ainsi délimitées. En phase de post-traitement, les régions adjacentes similaires sont fusionnées. Par rapport à ces approches, notre méthode recherche une partition en rectangles factorisable sur les deux axes, ce qui fournit un tableau bidimensionnel de régions rectangulaires permettant d'une part une optimisation poussée du critère de segmentation considéré globalement, d'autre part une manipulation aisée des régions lors des traitements ultérieurs.

Dans les approches apparentées, un critère de similitude entre les régions est utilisé afin de décider si deux régions doivent être fusionnées ou rester séparées. Ce critère de similitude nécessite généralement l'ajustement de paramètres, ce qui couplé avec son utilisation répétée pour des décisions locales de fusion ou séparation de régions à différentes échelles, peut entraîner des problèmes de fiabilité de la segmentation obtenue. Notre approche utilise une évaluation MAP globale et sans paramètre de la segmentation, ce qui lui confère à la fois robustesse et simplicité d'utilisation.

Concernant le critère d'évaluation d'une segmentation, d'autres méthodes sont également fondées sur des approches statistiques ou ce qui est équivalent sur des approches de type MDL [3]. Ces approches se basent généralement sur des hypothèses paramétriques de distribution des couleurs, Gaussienne [4] ou Poissonnienne [5] dans le cas des images radar par exemple. Notre approche n'exploite que les connaissances génériques des images, dimensions et nombres de niveaux par couleurs. Elle ne fait d'autre hypothèse que celle d'une distribution uniforme des plages de niveaux de couleurs par région rectangulaire et des niveaux de couleur par plage de niveaux.

2 Modèle d'image couleur en grille

Après avoir présenté notre représentation des images couleur, on introduit les modèles de segmentation en grille. On décrit l'approche Bayésienne permettant d'évaluer la qualité de tels modèles et on présente l'algorithme d'optimisation qui permet de rechercher la meilleure segmentation.

2.1 Représentation d'une image couleur

On considère une image couleur comme un tableau de données bi-dimensionnel dont les instances sont les pixels de l'image et les variables, au nombre de cinq, sont deux variables d'espace et trois variables colorimétriques. Les variables colorimétriques peuvent être (R, G, B) , (H, S, V) , ou même se réduire à une seule variable de niveaux de gris par exemple. Dans le reste du papier, pour des raisons de commodité dans les notations, on se basera sur l'espace colorimétrique (R, G, B) .

Soit $\mathbb{V} = \{X, Y, R, G, B\}$ l'ensemble des variables d'une image, se décomposant en un sous-ensemble $\mathbb{V}_S = \{X, Y\}$ de variables d'espace et un sous-ensemble $\mathbb{V}_C = \{R, G, B\}$ de variables de couleur. En utilisant la connaissance a priori du domaine des images et de leur codage au format bitmap, on sait que chacune des variables peut prendre un nombre fini de valeurs entières, selon les constantes suivantes : M_X, M_Y : largeur et hauteur en pixels de l'image, $N = M_X M_Y$: nombre de pixels, M_R, M_G, M_B : nombres de niveaux pour les variables de couleurs R, G, B .

2.2 Modélisation d'une image couleur

On définit un modèle d'image comme un modèle statistique de régression des couleurs, plus précisément comme un modèle d'estimation de la densité des variables colorimétriques conditionnellement aux variables d'espace.

De façon similaire à l'objectif classique des méthodes de segmentation d'image, on recherche des régions d'espace homogènes vis à vis de la distribution des couleurs. Pour définir les régions, on discrétise chaque variable d'espace en un ensemble d'intervalles adjacents. Le produit cartésien des discrétisations des variables X et Y définit un ensemble de régions rectangulaires, sur lesquelles on définit la distribution des couleurs. Le nombre de niveaux étant important, on discrétise également chaque variable de couleur en plages de niveaux adjacentes, en supposant

que la répartition des niveaux est homogène localement à chaque plage de niveaux. Pour chaque rectangle d'espace, on décrit la distribution des pixels sur les plages de niveaux. L'ensemble des rectangles d'espace, des plages de niveaux, et des distributions des pixels de chaque rectangle sur les plages de niveaux fournit alors un modèle synthétique de la densité des composantes colorimétriques conditionnellement à la position des pixels.

Cette approche est formalisée dans la définition 1.

Définition 1 *Un modèle d'image en grille est défini par la discrétisation de chaque variable d'espace et de couleur et sur chaque région rectangulaire de la partition de l'image en grille ainsi définie par la distribution des pixels sur les plages de niveaux de couleur.*

Les paramètres d'un modèle d'image en grille sont entièrement définis par :

- I_X, I_Y : nombre d'intervalles de largeur et de hauteur,
- I_R, I_G, I_B : nombre de plages de niveaux pour R, G, B ,
- N_x^X : largeur en pixels de l'intervalle x de X ,
- N_y^Y : hauteur en pixels de l'intervalle y de Y ,
- N_i^V : nombre de niveaux pour la plage i de la variable de couleur V ($V \in \mathbb{V}_C$).
- N_{xyi}^V : nombre de pixels de la région (x, y) pour la plage de niveaux i de la couleur V .

Soient enfin N_{xy} le nombre de pixels de la région (x, y) et $N_i^{[V]}$ le nombre total de pixels de l'image pour la plage de niveaux i de la variable de couleur V . Ces quantités sont déduites de la connaissance du format bitmap des images et des paramètres d'un modèle d'image en grille selon les équations $N_{xy} = N_x^X N_y^Y$ et $N_i^{[V]} = \sum_{x=1}^{I_X} \sum_{y=1}^{I_Y} N_{xyi}^V$.

2.3 Evaluation d'un modèle en grille

Les modèles d'image en grille sont très expressifs, puisqu'ils peuvent passer d'une description très grossière de l'image en une seule région et une seule plage de couleur (modèle nul) à une description très fine contenant autant de régions élémentaires que de pixels et autant de plages que de niveaux potentiels (modèle complet).

Nous proposons ici de formuler le choix de la meilleure granularité d'un modèle en grille comme un problème de sélection de modèle. Une approche Bayésienne est appliquée pour choisir le meilleur modèle en grille, qui est recherché en maximisant la probabilité $p(\text{Modèle}|\text{Données})$ du modèle en grille sachant les données, c'est à dire l'image. En utilisant la règle de Bayes, et puisque la probabilité des données $p(\text{Données})$ ne dépend pas du modèle, il s'agit alors de maximiser $p(\text{Modèle})p(\text{Données}|\text{Modèle})$, c'est-à-dire le produit d'un terme d'a priori sur les modèles et d'un terme de vraisemblance des données connaissant le modèle.

On propose une distribution a priori des modèles en grille exploitant la hiérarchie des paramètres : les nombres d'intervalles sont d'abord choisis sur chaque variable d'espace et de couleur, puis les bornes des intervalles et enfin les effectifs par plage de couleur dans chaque région. Le choix est uniforme à chaque étage de cette hiérarchie. De plus, les distributions des plages de niveaux de couleurs par région sont supposées indépendantes entre elles.

En utilisant la définition de la famille de modèles en grille et sa distribution a priori, la formule de Bayes permet de calculer explicitement les probabilités a posteriori des modèles connaissant les données. En prenant le log négatif de ces probabilités, cela conduit au critère d'évaluation d'un modèle en grille M fourni dans l'équation 1 :

$$\begin{aligned}
C(M) = & \sum_{V \in \mathbb{V}} \log M_V \\
& + \sum_{V \in \mathbb{V}} \log \binom{M_V + I_V - 1}{I_V - 1} \\
& + \sum_{x=1}^{I_X} \sum_{y=1}^{I_Y} \sum_{V \in \mathbb{V}_C} \log \binom{N_{xy} + I_V - 1}{I_V - 1} \quad (1) \\
& + \sum_{x=1}^{I_X} \sum_{y=1}^{I_Y} \sum_{V \in \mathbb{V}_C} \left(\log N_{xy}! - \sum_{i=1}^{I_V} N_{xyi}^V! \right) \\
& + \sum_{V \in \mathbb{V}_C} \sum_{i=1}^{I_V} N_i^{[V]} \log N_i^V
\end{aligned}$$

La première ligne de l'équation 1 correspond au choix des nombres d'intervalles, la deuxième ligne au choix des bornes de ces intervalles. La troisième ligne correspond au choix des paramètres de la distribution multinomiale des pixels de chaque région sur les plages de niveaux de couleurs, indépendamment pour chacune des variables de couleur. La quatrième ligne est un terme du multinôme pour chaque région, représentant la vraisemblance d'observer les pixels distribués suivant les paramètres de la multinomiale définie sur la troisième ligne. Le terme de la dernière ligne se base sur le nombre de niveaux possibles N_i^V de chaque plage et sur l'effectif cumulé $N_i^{[V]}$ des pixels de l'image sur cette plage. Il s'agit également d'un terme de vraisemblance du niveau des pixels sur leur plage, en se basant sur une hypothèse de distribution uniforme des niveaux dans chaque plage.

Il est à noter que comme le log négatif d'une probabilité représente une longueur de codage [6], la formule 1 s'interprète comme une longueur d'encodage d'une image, avec la partie d'a priori encodant les paramètres du modèle en grille et la partie de vraisemblance encodant la position et le niveau colorimétrique exact des pixels connaissant le modèle d'image. On peut vérifier que les deux modèles "extrêmes" (nul et complet) ont une longueur de codage approximativement égale à celle du format bitmap natif, qui utilise 24 bits par pixel. Les modèles intermédiaires sont à même d'identifier des régions homogènes vis à vis des couleurs, ce qui leur permet un codage plus efficace.

2.4 Optimisation d'une segmentation

On résume dans l'algorithme 1 une heuristique d'optimisation consistant à optimiser alternativement la discrétisation de chaque variable de l'image, en figeant la discrétisation des autres variables. L'algorithme de discrétisation utilisé est celui présenté dans [7] dans le cadre de la discrétisation des variables numériques pour les problèmes de classification supervisée du Data Mining. Les optimisations sont répétées tant qu'il y a amélioration du critère

d'évaluation. En pratique, comme le montre la figure 1, l'optimum est vite atteint indépendamment du modèle initial en entrée de l'algorithme, en moins de trois itérations sur l'ensemble des variables selon nos expérimentations.

Algorithme 1 Optimisation d'une image en grille

Entrée: M {Modèle initial}

Sortie: M^* ; $c(M^*) \leq c(M)$ {Modèle final amélioré}

- 1: **Tant que** amélioration **répéter**
 - 2: **Pour tout** $V \in \mathbb{V}$ **répéter**
 - 3: Figer la discrétisation des autres variables
 - 4: Optimiser la discrétisation de la variable V
 - 5: **Fin pour**
 - 6: **Fin tant que**
-

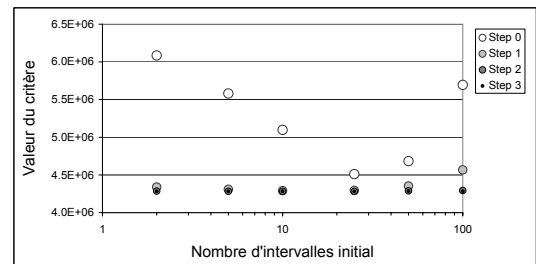


FIG. 1 – Valeur du critère après 0 à 3 étapes d'optimisation, en partant de modèles initiaux pour 2, 5, 10, 25, 50 et 100 intervalles équidistants par variable.

3 Evaluation

On présente dans cette section des résultats d'expérimentations préliminaires sur la segmentation d'une image en grille et ses applications potentielles pour la recherche de zone d'intérêt, la compression d'image ou la détection de mouvement.

3.1 Exemple de segmentation en grille

La méthode de segmentation en grille est évaluée en utilisant l'image Lena de dimension $512 * 512$ au format RGB, dont le fichier bitmap a une taille de 768 ko.

La segmentation obtenue est représentée sur la gauche de la figure 2. Cette grille comprend 57 intervalles horizontaux et 27 intervalles verticaux, soient 1539 régions à comparer avec les 262144 pixels de la représentation initiale. Les couleurs sont discrétisées en 21 plages pour le rouge, 25 pour le vert et 20 pour le bleu. Dans la droite de la figure 2, les régions de la grille sont dessinées en prenant leur couleur moyenne, ce qui permet de définir un modèle simplifié de l'image de taille inférieure à 5 ko, soit environ 0.5% de la taille de la bitmap. Ce format est très aisément manipulable puisque les régions rectangulaires sont indexables directement dans un tableau bidimensionnel. Les motifs de l'image sont préservés à un niveau macroscopique, comme le montre la figure 2. La segmentation en grille s'avère ainsi potentiellement intéressante en tant que prétraitement pour la recherche de zones d'intérêt.

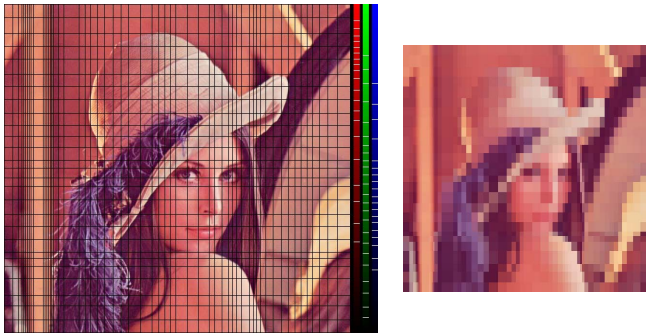


FIG. 2 – Grille de la segmentation et plages de niveaux de couleur pour l'image Lena ; à droite, le modèle en grille avec couleur moyenne par région.

3.2 Application à la compression d'image

La valeur du critère d'évaluation de la grille pour l'image Lena est 4281750 en prenant la base 2 du logarithme dans la formule (1). Selon la section 2.3, cette valeur correspond au coût de codage théorique de l'image. Ce coût de codage permettrait en théorie de comprimer l'image Lena sans perte d'information en utilisant environ 69% de la taille de l'image au format bitmap initial. A titre de coût de comparaison, le format PNG permet une compression sans perte pour un facteur d'environ 78%.

Ce résultat préliminaire permet d'envisager l'utilisation des modèles d'image en grille pour la compression d'image. Il montre surtout que la démarche Bayésienne de choix de la granularité de la grille conduit à une solution particulièrement performante : le compromis est optimal entre la complexité de la grille et la fidélité de la description des régularités de l'image.

3.3 Application à la détection de mouvement

On propose ici d'évaluer l'utilisation potentielle des modèles d'image en grille pour la détection de mouvement. Dans un flux vidéo, on effectue la différence entre deux trames successives afin de détecter les mouvements en se basant sur les régions ayant changé de couleur. La figure 3 présente un tel exemple dans le cadre d'un flux vidéo en niveaux de gris.



FIG. 3 – Détection de mouvement dans un flux vidéo : à gauche, image résultant de la différence entre deux trames successives ; à droite, segmentation en grille de cette image avec visualisation des zones de forte entropie

Le traitement de ce type d'image pose le problème de l'identification des zones de mouvement, à distinguer des

zones de bruits vidéo liée par exemple aux variations locales de luminosité ou aux caractéristiques des capteurs de la caméra. Du fait de la robustesse de l'approche Bayésienne utilisée, la technique de segmentation en grille est appropriée pour le prétraitement de ce type d'image. Dans le cas de l'image présentée sur la gauche de la figure 3, la grille obtenue comporte $72 = 13 \times 4$ régions distribuées sur 6 plages de niveaux de gris. La droite de la figure 3 représente la grille obtenue en coloriant chaque région selon son entropie. Les régions d'entropie nulle ou faible (en blanc ou gris clair) sont celles où l'image est restée constante ou a légèrement varié en raison du bruit vidéo. Les régions de forte entropie (gris foncé) correspondent aux zones ayant le plus changé, ce qui permet d'identifier les zones candidates pour la détection de mouvement.

4 Conclusion

Les modèles d'image en grille proposés ici se basent sur un partitionnement de l'image en une grille bidimensionnelle de régions rectangulaires et sur la discrétisation des couleurs en plages. L'apport de la méthode réside principalement dans sa généralité, qui la rend indépendante du domaine d'application, et dans l'approche Bayésienne utilisée pour l'évaluation globale des modèles d'image, qui permet d'obtenir une granularité optimale des grilles de segmentation.

Des expérimentations préliminaires confirment la validité de l'approche et illustrent son potentiel en tant que méthode générique de prétraitement d'image, dans le cadre de la segmentation, de la compression sans perte d'information ou de la détection de mouvement. Les travaux à venir viseront à finaliser l'approche présentée et à développer son potentiel pour le prétraitement d'image.

Références

- [1] H. Maître. *Le traitement des images*. Lavoisier, 2003.
- [2] S.L. Horowitz and T. Pavlidis. Picture segmentation by a tree traversal algorithm. *JACM*, 23(2) :368–388, April 1976.
- [3] J. Rissanen. Modeling by shortest data description. *Automatica*, 14 :465–471, 1978.
- [4] U. Ndili, R. Nowak, and M. Figueiredo. Coding theoretic approach to image segmentation. In *IEEE International Conference on Image Processing*, October 2001.
- [5] V. Venkatachalam, R.D. Nowak, R.G. Baraniuk, and M.A. Figueiredo. Unsupervised sar image segmentation using recursive partitioning. In E.G. Zelnio, editor, *Proc. SPIE, Algorithms for Synthetic Aperture Radar Imagery VII*, volume 4053, pages 121–129, 2000.
- [6] C.E. Shannon. A mathematical theory of communication. Technical report, Bell systems technical journal, 1948.
- [7] M. Boullé. MODL : a Bayes optimal discretization method for continuous attributes. *Machine Learning*, 65(1) :131–165, 2006.