

Segmentation d'images multi-bandes par sélection itérative de la combinaison optimale des bandes chromatiques

MOHAMED ABADI¹, ANNE-SOPHIE CAPELLE-LAIZE¹, MAJDI KHOUDEIR¹, DIDIER COMBES², SERGE CARRE²

¹ Institut XLIM-SIC, UMR CNRS 6172

Université de Poitiers, BP 30179, 86962 Futuroscope-Chasseneuil, Cedex France

² UR4 P3F

INRA Poitou-Charentes, BP 6, 86600 Lusignan, France

¹{abadi, capelle, khoudeir}@sic.univ-poitiers.fr,

²{dcombes, scarre}@lusignan.inra.fr

Résumé - Cet article propose une approche itérative de segmentation d'images multi-bandes fondée sur la détermination des combinaisons adéquates des bandes chromatiques. Elle utilise, à chaque itération, successivement une technique de seuillage ou une méthode de classification pour distinguer les deux classes dominantes dans chaque combinaison et le pouvoir discriminant pour sélectionner la meilleure. 300 images de la base d'images Berkeley sont utilisées pour étudier la pertinence de notre approche. Des évaluations qualitatives et quantitatives basées sur les indices de mesure tels que PRI (0.795) et GCE (0.157) sont employées. Ces mesures montrent l'efficacité de la méthode en comparant les résultats de segmentation obtenus par notre approche à ceux donnés par l'humain et l'algorithme *Mean Shift*.

Abstract - This paper proposes an iterative approach to multi-band image segmentation. It is based on computing the appropriate combinations of the chromatic components. At each iteration, our approach uses, successively a thresholding method or a classification algorithm to distinguish, in each combination, the two dominant classes and the discriminant power to select the best combination. 300 images for Berkeley database are used to study the relevance of our approach. Qualitative and quantitative assessments based on PRI (0.795) and GCE (0.157) indices measuring are used. These measurements show the accuracy of our approach by comparing the segmentation results with those given by the human and the *Mean Shift* algorithm.

1 Introduction

L'information spectrale est un élément important en traitement et analyse d'images. Elle permet d'améliorer la qualité de la segmentation. Cela induit une meilleure compréhension du contenu des images et par conséquent facilite l'interprétation théorique et visuelle de la scène [1, 2]. Différentes approches de segmentation d'images multi-bandes ont été développées [2, 5, 7, 8]. Ce type d'images fournit en effet plus d'informations qu'une seule image en niveaux de gris [1]. Ici, nous supposons que les images couleur sont des images multi-spectrales ; les trois bandes seront traitées les unes indépendamment des autres. Par conséquent, les techniques de segmentation introduites pour les images en niveau de gris deviennent aisément extensibles. Trois types de traitement sont ainsi possibles : (i) les bandes sont utilisées séparément [4], (ii) traitement de deux bandes à la fois (le nombre de cas est $\frac{N(N-1)}{2}$) [6] et (iii) traitement de trois bandes ou plus simultanément [12] (peu d'approches ont été développées à cause de la complexité). Cependant, les performances dépendent forcément des transformations appliquées [8]. Généralement, ces transformations sont obtenues par des estimations empiriques déduites à partir des scènes synthétiques et/ou naturelles. La qualité des transformations est évaluée en fonction des applications, leur capacité à dissocier les différentes

informations spectrales ainsi que leur robustesse face aux variations de la lumière et aux effets naturels.

2 Travaux en relation

La recherche d'un système de représentation adéquat pour tout type d'image est une mission difficile. Des tentatives récentes et originales ont été proposées [9, 10]. Elles nécessitent des connaissances *a priori* de la scène (approches supervisées) où l'idée est de fusionner les informations pertinentes fournies par les différentes bandes, par exemple les composantes des différents espaces couleur classiques (*RVB*, *XYZ*, *IHS*, ...). Les résultats sont des espaces hybrides (*EH*) où le nombre de composantes est inconnu *a priori*. De même ces espaces n'ont pas de fondement qu'il soit physique ou psycho-visuel contrairement aux espaces couleur classiques. Pour remédier aux inconvénients de l'*EH* (algorithme de sélection de type glouton, convergence non assurée, dépendance par rapport au critère d'arrêt), Grandchamp et al. [9] ont proposé un algorithme multi-objectif basé sur la courbe de Pareto pour séparer les couverts forestiers des parcelles agricoles. L'algorithme défini garantit la solution optimale mais il nécessite à son tour l'apprentissage supervisé. Nous constatons que les résultats des différentes méthodes sont principalement influencés par les transformations choisies [1, 7]. Ces transformations sont ajustées en fonction des objectifs des applications [7, 8]. Partant de

ce constat, nous proposons ici une approche basée, d'une part sur la définition d'un modèle qui utilise simultanément toutes les bandes pour éviter le choix des espaces couleur et de la/les bandes à utiliser et d'autre part, sur les techniques de partition hiérarchiques pour remplacer l'apprentissage supervisé. Enfin, un critère de sélection est utilisé pour évaluer les combinaisons. Nous introduisons ainsi une approche qui exclut la notion d'espace couleur car elle cherche à trouver la combinaison optimale des bandes colorimétriques. Cette combinaison est sélectionnée pour mieux identifier les classes à discriminer.

3 Procédure de segmentation

La diversité des approches de segmentation rend leur comparaison difficile car les idées divergent sur le choix des espaces couleur ou la/les bandes à utiliser. Ici nous ne nous préoccupons pas de la notion d'espaces couleur en tant que tel car nous cherchons un moyen simple pour segmenter les images multi-bandes. En effet, notre méthode est une approche itérative qui définit à chaque étape une combinaison hybride optimale permettant la séparation des données en deux classes. Ainsi à chaque itération, l'algorithme construit les combinaisons à partir des N bandes (Figure 1.a). Les pixels de l'image sont alors projetés sur ces combinaisons puis partitionnés en deux classes. La combinaison hybride minimisant le rapport entre les variances inter-classes et intra-classes (CS) est finalement sélectionnée. Il s'agit de la combinaison la plus discriminante selon le critère CS qui sera décrit ci-après. Les classes sont labélisées (Figure 1.b et 1.c) et utilisées à l'itération suivante pour récupérer le contenu des N bandes spécifiques à chaque classe. Enfin, la procédure est arrêtée à l'itération de rang $p \in \mathbb{N}$ défini *a priori*.

3.1 Combinaisons hybrides

Indépendamment de la notion d'espace couleur et quelque soit le nombre de bandes $(b_l)_{l=1,\dots,N}$, nous pouvons modéliser la construction d'une combinaison (C_k) par :

$$(C_k)_{k=1,\dots,K} = \sum_{l=1}^N \alpha_{k,l} \cdot b_l \quad (1)$$

où $\alpha_{k,\cdot} = \{(e_1, \dots, e_l, \dots, e_N) \mid e_l \in E\}$ est le $k^{\text{ième}}$ sous-ensemble de N éléments construits à partir d'un ensemble fini $E \subset \mathbb{R}$ de cardinal n ($n \in \mathbb{N}^*$) défini *a priori*.

Le nombre total des sous-ensembles possibles est $K = n^N$.

La question à résoudre est celle du choix de la meilleure combinaison C_k c'est-à-dire celle qui permettra la meilleure séparation des classes. En examinant l'équation (1), il est évidemment impossible d'explorer toutes les combinaisons pour n grand. Empiriquement nous choisissons $E \in \left\{-\frac{1}{4}, \frac{1}{2}\right\}$. Cette décision est due aux très faibles variations observées sur le comportement du critère CS par rapport aux changements opérés sur les éléments de l'ensemble E .

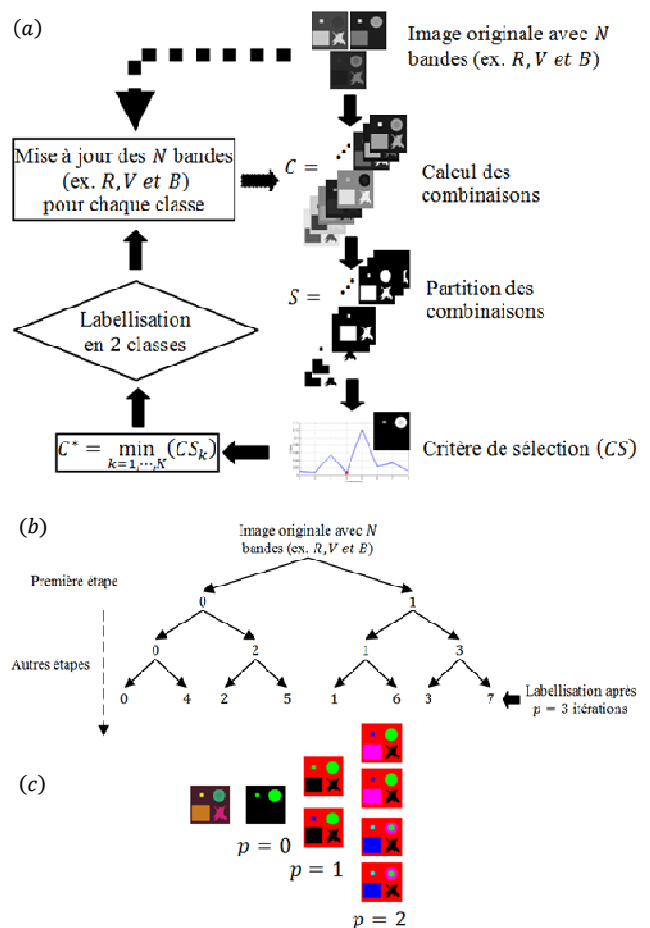


Figure 1 : Schéma de l'algorithme de segmentation automatique des images multi-bandes.

3.2 Choix de la combinaison optimale

Nous cherchons la combinaison C_k qui discrimine au mieux les pixels de l'image en deux classes. Ainsi une approche par seuillage ou par classification non-supervisée est une solution. Les pixels de chaque classe sont ensuite utilisés afin de calculer les critères de sélection CS_k . Ce dernier est défini par sa capacité à sélectionner la combinaison adéquate. Nous l'exprimons via les rapports des variances intra-classes W_k et inter-classes B_k déterminées sur les C_k de la manière suivante :

$$CS_k = \frac{W_k}{W_k + B_k} \quad (2)$$

où

$$W_k = \frac{1}{2N} \sum_{j=1}^2 \sum_{i=1}^{n_j} (x_{i,j} - m_j)^2 \quad (3)$$

et

$$B_k = \frac{1}{2} \sum_{j=1}^2 \sum_{i=1}^{n_j} (m_j - m)^2 \quad (4)$$

avec $x_{i,j}$ et n_j sont, respectivement, la $i^{\text{ème}}$ valeur et le nombre de pixels de la $j^{\text{ème}}$ classe de la combinaison C_k . $m_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{i,j}$ est le vecteur des centres de gravité utilisé pour définir la classe j ($j \in \{1,2\}$). Le vecteur des centres de gravité de toutes les classes est $m = \frac{1}{2} \sum_{j=1}^2 \frac{1}{n_j} \sum_{i=1}^{n_j} x_{i,j}$.

Nous retenons la combinaison qui possède le rapport minimal :

$$C^* = \min_{k=1,\dots,K}(CS_k) \quad (5)$$

Généralement, les combinaisons seront d'autant plus discriminantes que W_k sera élevée par rapport à B_k . En effet, si W_k est élevée, ceci signifie que les centres des classes sont éloignés. Cependant, si B_k est faible, ceci signifie que les pixels d'une même classe sont proches de leur centre. Par contre, si, simultanément W_k est élevée et B_k est faible alors les classes sont éloignées les unes des autres et compactes. La situation idéale pour un problème de discrimination.

4 Résultats

Dans cette section, nous présentons une comparaison quantitative (Figures 2 et 3) et qualitative (Figure 4) entre les résultats obtenus par notre algorithme, la vérité terrain et l'algorithme *Mean Shift*¹. Pour ce dernier nous avons utilisé les paramètres par défaut. Les images choisies appartiennent à la base d'images Berkeley². Ces images sont des extraits des scènes naturelles. Elles témoignent d'une forte variabilité colorimétrique et de la présence de l'information texture pour certaines. Notre approche est testée pour $p = 2$ et $E \in \left\{-\frac{1}{4}, \frac{1}{2}\right\}$. Les résultats obtenus peuvent être considérés comme un compromis entre la sous-segmentation et la sur-segmentation car les régions sont mieux délimitées. Contrairement à l'algorithme *Mean Shift* qui produit une sur-segmentation par rapport à la vérité terrain (Figure 4).

Pour évaluer la complexité de notre algorithme, nous avons mesuré le temps de calcul³ effectué sur l'image synthétique (Figure 1.c), il ressort que notre approche est plus rapide que l'algorithme *Mean Shift*. Le temps de calcul effectué est respectivement 1.19s et 2.87s. L'évaluation quantitative de la segmentation est réalisée en adoptant la même démarche que dans [13]. Sur un ensemble de 300 images issues de la base d'images Berkeley, nous avons utilisé, pour chaque image, les différentes segmentations réalisées par l'humain comme des images références. L'idée est que, même si l'humain produit des segmentations différentes pour la même image, les résultats finaux doit différer essentiellement dans le raffinement local de certaines régions. En effet, Martin et al. [13] et Unnikrishnan et al. [14] ont défini des mesures d'erreur qui ne pénalise pas une segmentation si elle est plus grossière ou plus raffinée que l'autre. Nous montrons dans ce travail les résultats obtenus en utilisant les mesures *GCE* (Global Consistency Error) [13] et *PRI* (Probabilistic Rand Index) [14]. Ces mesures permettent d'évaluer la différence entre les différentes sources de segmentation prises deux à deux. La mesure *PRI* $\in [0,1]$ permet de quantifier le nombre de pixels bien classés. Le critère *GCE* $\in [0,1]$ donne la mesure à laquelle la segmentation d'une image peut-être vue comme un raffinement de l'autre.

Les figures 2 et 3 montrent les histogrammes de ces mesures calculées sur les 300 images de la base d'images Berkeley pour chaque algorithme en fonction des images références. Sur les histogrammes de la mesure *GCE* la valeur 0 signifie qu'il n'y a pas d'erreur entre les paires d'images. Ceci est équivalent à la valeur 1 pour la mesure *PRI*. En effet, la figure 2 montre un pic très proche de la valeur 0 pour notre approche contrairement à l'algorithme *Mean Shift*. De même, sur la figure 3 on observe que l'histogramme produit en utilisant notre approche est moins étalé que celui obtenu par l'algorithme *Mean Shift*. Globalement, les histogrammes montrent que les fréquences d'apparition des faibles valeurs de la mesure *GCE* et des valeurs élevées du critère *PRI* sont plus importantes pour notre approche.

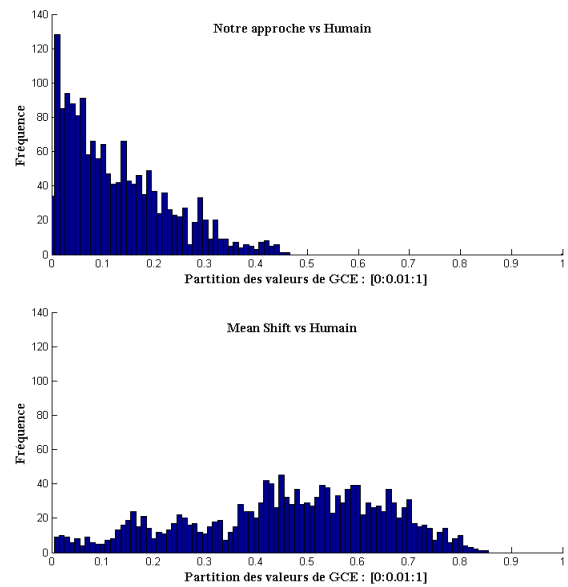


Figure 2 : Histogrammes du critère GCE obtenus en utilisant notre approche (En haut) et l'algorithme *Mean Shift* (en bas).

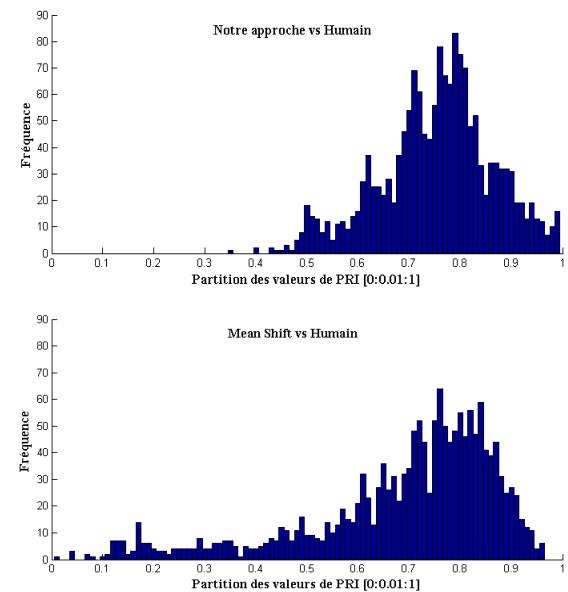


Figure 3 : Histogrammes du critère PRI obtenus en utilisant notre approche (En haut) et l'algorithme *Mean Shift* (en bas).

Le tableau 1 montre la moyenne des mesures GCE et PRI déterminées sur les 300 images de la base d'images Berkeley en utilisant les segmentations obtenues par l'Humain, notre approche et l'algorithme *Mean Shift*.

¹ <http://www.rutgers.edu/riul/research/code/EDISON/index.html>

² <http://www.cs.berkeley.edu/projects/vision/grouping/segbench>

³ Un processeur de 2.00 GHz Intel Core™ 2 et une RAM de 1 Go.

En comparant ces mesures avec celles obtenues par l'Humain, on constate que quantitativement, notre approche retourne des valeurs plus proches des valeurs obtenues par l'humain que l'algorithme *Mean Shift*,

Tab 1 : Performance moyenne des mesures GCE et PRI obtenue sur les 300 images de la base d'images Berkeley.

	GCE	PRI
Humain	0.0797	0.8754
Notre approche	0.1569	0.7950
<i>Mean Shift</i>	0.3717	0.7770

Ces tests nous ont permis de montrer que notre approche a donnée des résultats visuels et quantitatifs, au sens des mesures citées, de bonne qualité.



Figure 4 : Segmentation obtenues sur des images issues de la base d'images Berkeley. De gauche vers la droite : images originales, images segmentées par l'Humain, par notre approche et par l'algorithme *Mean Shift*.

Nous avons présenté une approche originale qui s'étend aisément à la segmentation d'images multi-bandes (ex. les images Landsat : 7 bandes). Elle a été expérimentée pour le cas des images couleur à $N = 3$ bandes (R, V et B). Au lieu d'utiliser la notion d'espaces couleur, nous travaillons sur la définition de

combinaison hybride optimale qui répartie au mieux les données en deux classes. Ainsi nous avons proposé une procédure itérative qui palie au choix délicat de l'espace couleur et de la méthode de sélection. Ainsi nous répondons aux questions du choix et de la sélection de l'espace couleur. En outre, pour des applications spécifiques [11], l'utilisation d'échantillons d'apprentissage est une solution pour estimer plus efficacement les éléments de l'ensemble E . De même, un algorithme de fusion [5] peut aussi être utilisé pour réduire le nombre final de classes.

Références

- [1] H-D. Cheng, X. Jiang, Y. Sun, J. Wang. Color image segmentation: Advances and prospects. *PR* 34(12), 2259–2281, 2001.
- [2] H. Noh, Q. Zhang, S. Han, B. Shin, D. Reum. Dynamic calibration and image segmentation methods for multispectral imaging crop Nitrogen deficiency sensors. *American Society of Agricultural Engineers*. 48(1): 393–401, 2005.
- [3] U. Shapira, I. Herrmann, A. Karnieli, D.J. Bonfil. Weeds detection by ground-level hyperspectral imaging. *ISPRS*, Haifa, Israel, 2010.
- [4] L. Busin, N. Vandenbroucke, L. Macaire, J-G. Postaire. Color space selection for unsupervised color image segmentation by histogram multithresholding. *ICIP*, pp. 203–206, 2004.
- [5] C. Meurie, O. Lezoray, L. Khoudour, A. Elmoataz. Morphological hierarchical segmentation and color spaces, *International Journal of Imaging Systems and Technology*, 20(2), pp. 167-178, 2010.
- [6] O. Lézoray, C. Charrier. Color image segmentation using morphological clustering and fusion with automatic scale selection. *PR Letters*, 30, (4), pp. 397-406, 2009.
- [7] H. Bonesmo, K. Kaspersen, A. Bakken. Evaluating an Image Analysis System for Mapping White Clover Pastures. *Acta Agriculturae Scandinavica, Section B - Soil and Plant Science* 54(2): 76-82, 2004.
- [8] S.D. Noble, R.B. Brown. Plant species discrimination using spectral/spatial descriptive statistics, *Image Analysis for Agricultural Products and Processes*, Potsdam, Germany, 82-92, August 27-28, 2009.
- [9] E. Grandchamp, Abadi, M. Hybrid color space choice: an optimization review for cost-efficiency trade-off. *SITIS*, Marrakesh, 29-4 December 2009.
- [10] N. Vandenbroucke, L. Macaire, J-G. Postaire. Color image segmentation by pixel classification in an adapted hybrid color space. application to soccer image analysis. *CVIU*, 90(2), 190–216, 2003.
- [11] J. Yang, C. Liu, and J. Yang. What Kind of Color Spaces Is Suitable for Color Face Recognition? *Neurocomputing*, 73, 2140-2146, 2010.
- [12] S. Ouattara, A. Clement and B. Vigouroux. Full 3D compact histogram segmentation of color images. *Asian J. Sci. Res.*, 4, 42-52, 2011.
- [13] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. In *Proc. 8th ICCV*, 2, 416–423, July 2001.
- [14] R. Unnikrishnan, C. Pantofaru, and M. Hebert. Toward Objective Evaluation of Image Segmentation Algorithms, *TPAMI*, 29(6), 929-94, June 2007.