

Segmentation d'images couleur par méthode variationnelle

Françoise Dibos⁽¹⁾ et Georges Koepfler^(1,2)

⁽¹⁾ CEREMADE, Université de Paris 9 Dauphine,
Place de Lattre de Tassigny, 75775 Paris cedex 16, France

⁽²⁾ PRISME, Université de Paris 5 René Descartes,
45, rue des Saints Pères, 75270 Paris cedex 06, France

Résumé

On présente un algorithme de segmentation d'images couleurs. L'algorithme est basée sur un modèle variationnel proposé par Mumford et Shah dans le cas des images noir et blanc. Cette approche permet de comparer diverses représentations de l'espace des couleurs.

Abstract

In this paper we present a method for the segmentation of color data. It is based on the minimization of the functional proposed by Mumford and Shah for grey level images. We combine this variational approach with several representations used for color space and compare the experimental results.

1 Introduction

Pour segmenter une image, Mumford et Shah ont proposé de minimiser la fonctionnelle (voir [MS]).

$$E(u, K) = \int_{\Omega \setminus K} \|u - g\|^2 + \mu \int_{\Omega \setminus K} \|\nabla u\|^2 + \lambda \mathcal{H}^1(K), \quad (1)$$

où g est l'image originale définie sur $\Omega \subset \mathbb{R}^2$, u l'image reconstruite, régulière par morceaux, K l'ensemble des singularités de u et μ, λ deux réels positifs. Le premier terme de cette fonctionnelle mesure la qualité de l'approximation, le second la régularité de u à l'intérieur d'une région et le dernier la longueur de l'ensemble des bords de la segmentation. $\mathcal{H}^1(\cdot)$ est la mesure de Hausdorff unidimensionnelle.

Quand g est une image noir et blanc, la norme $\|u - g\|$ est la valeur absolue de la différence des deux niveaux de gris. Si g est une image couleur, une norme pondérée sur les différents canaux est utilisée. Cette norme diffère pour chaque application. Dans cet article, nous présentons une application de la méthode variationnelle de Mumford et Shah à la segmentation d'images couleur. L'image g peut, par exemple, être donnée par ses composantes RGB. D'autres applications de segmentation d'images multicanaux peuvent être trouvés dans [KLR].

Si l'image reconstruite u est supposée constante par morceaux, la fonctionnelle générale (1) se simplifie en

$$E(K) = \int_{\Omega \setminus K} \|u - g\|^2 + \lambda \mathcal{H}^1(K). \quad (2)$$

L'image originale g est alors découpée en régions O_i ($\Omega = \cup O_i$, $O_i \cap O_j = \emptyset$) sur lesquelles u est constante et l'ensemble des bords K est l'union $\cup \partial O_i$. Lorsque K est donné, on montre que si u qui réalise le minimum de

la fonctionnelle (2) alors sa valeur dans la région O_i est $u^{O_i} = \frac{1}{|O_i|} \int_{O_i} g$, où $|O_i|$ est la surface de la région O_i .

Dans [MoS] on a montré que la minimisation de la fonctionnelle (2) est liée aux méthodes de croissance de régions. Ici, deux régions O_i et O_j sont concaténées si

$$E(K \setminus \partial(O_i, O_j)) - E(K) \leq 0,$$

où $\partial(O_i, O_j)$ est la frontière commune aux régions O_i et O_j , [KMS]. Le minimum obtenu en utilisant cet algorithme vérifie des estimations a priori sur la taille des régions (pour plus de détails voir [MoS, chapter 5]).

Le coefficient λ de (2) peut être interprété comme un paramètre d'échelle. En effet, plus grand est λ , plus la segmentation est grossière (peu de régions); si λ est petit on obtient de nombreuses régions et donc une segmentation plus fine.

Considérons maintenant une image couleur. En chaque pixel la couleur est usuellement donnée par ses trois composantes R (rouge), G (vert) et B (bleu), $(R, G, B) \in [0, 1]^3$. Des études de colorimétrie ont montré que pour tenir compte de la réalité visuelle on devait munir l'espace des couleurs d'une distance riemannienne. En effet, on a pu constater qu'autour de chaque point $u \in [0, 1]^3$ de l'espace des couleurs l'on pouvait trouver un ellipsoïde (E) tel que l'oeil ne fasse pas de différence entre la couleur c et les autres couleurs incluses dans (E) (voir [WS, p.306] et [J]). La distance infinitésimale ds dans l'espace des couleurs a donc été écrite, sous sa forme la plus générale,

$$(ds)^2 = \varphi_{11}(u)(du_1)^2 + 2\varphi_{12}(u)du_1du_2 \\ + \varphi_{22}(u)(du_2)^2 + 2\varphi_{23}(u)du_2du_3 \\ + \varphi_{33}(u)(du_3)^2 + 2\varphi_{31}(u)du_3du_1,$$

où $u = (u_1, u_2, u_3) = (u_R, u_G, u_B)$.

$(ds)^2$ est une forme quadratique définie positive et les coefficients $\varphi_{ij}(u)$ sont des fonctions continues de u .

La distance entre deux couleurs u et v est alors donnée par

$$d(u, v) = \min \int_u^v ds = \inf_p \left(\int_0^1 \sqrt{\sum \varphi_{ij} \frac{dp_i}{dt} \frac{dp_j}{dt}} dt \right),$$

où p est un chemin de u à v , c'est à dire une fonction régulière $p : [0, 1] \rightarrow [0, 1]^3$ telle que $p(0) = u$ et $p(1) = v$.

On peut alors écrire la fonctionnelle simplifiée (2) pour des images couleur

$$E(K) = \sum_{\alpha} \int_{O_{\alpha}} d(g, u^{O_{\alpha}})^2 dx + \lambda \ell(K),$$

où $\Omega = \bigcup O_{\alpha}$, $O_{\alpha} \cap O_{\beta} = \emptyset$ ($\alpha \neq \beta$), $u^{O_{\alpha}}$ est la valeur de u dans la région O_{α} et $K = \bigcup \partial O_{\alpha}$. À la place de la longueur de Hausdorff \mathcal{H}^1 nous utilisons la longueur classique ℓ ce qui est équivalent pour des courbes rectifiables.

2 Le choix de la distance dans l'espace des couleurs

Afin de rendre compte le mieux possible de la perception des couleurs, Helmholtz (1896), Schrödinger (1920), Stiles (1946) et plus récemment Vos and Walraven (1972) ont proposé différentes distances des espaces de couleurs.

La distance que nous avons utilisé est la distance de Stiles. Elle a l'avantage de combiner fidélité aux résultats expérimentaux en perception des couleurs et facilité d'implémentation. La distance infinitésimale ds est alors définie, si $(R, G, B) \in [0, 1]^3$, par

$$(ds)^2 = \left(\frac{\xi(R)}{\rho} dR \right)^2 + \left(\frac{\xi(G)}{\gamma} dG \right)^2 + \left(\frac{\xi(B)}{\beta} dB \right)^2,$$

$$\text{où } \xi(R) = \frac{9}{1+9R}, \xi(G) = \frac{9}{1+9G}, \xi(B) = \frac{9}{1+9B}.$$

Les constantes de normalisation $\rho = 1.28$, $\gamma = 1.65$ et $\beta = 7.25$ ont été définies expérimentalement et vérifient

$$\frac{1}{\rho^2} + \frac{1}{\gamma^2} + \frac{1}{\beta^2} = 1.$$

Cette distance est sensible aux variations relatives des composantes de la couleur et pondère l'importance des canaux rouge, vert et bleu.

Grâce au changement de coordonnées $\tilde{R} = \log(1+9R)$, $\tilde{G} = \log(1+9G)$ and $\tilde{B} = \log(1+9B)$, cette distance peut être vue comme une distance euclidienne,

$$(ds)^2 = \frac{(d\tilde{R})^2}{\rho^2} + \frac{(d\tilde{G})^2}{\gamma^2} + \frac{(d\tilde{B})^2}{\beta^2},$$

et

$$d(u(x), g(x))^2 = \frac{(\tilde{u}_R(x) - \tilde{g}_R(x))^2}{\rho^2} + \frac{(\tilde{u}_G(x) - \tilde{g}_G(x))^2}{\gamma^2} + \frac{(\tilde{u}_B(x) - \tilde{g}_B(x))^2}{\beta^2}.$$

Wyszecki et Stiles ont montré [WS, p.660] que cette distance était compatible avec de nombreuses données expérimentales, voir [WS, p.672]. Elle a cependant quelques défauts, Wyszecki et Stiles ont pensé que cela pouvait être du au fait que les trois composantes étaient considérées comme indépendantes. On pourrait y remédier en introduisant des termes rectangles mais cela compliquerait de manière sensible les calculs.

De nombreuses autres approches pour la segmentation couleur ont été proposées. Citons le travail de G. Klinker (voir [KSK]) qui propose un modèle basé sur les propriétés physiques de réflectance de la surface d'un objet. Les applications étant dans le domaine de la robotique elle tient en plus compte des caractéristiques des caméras utilisées.

Dans notre approche on n'a pas utilisé les standards de la CIE, adaptés aux caractéristiques de la vision humaine. Les représentations non linéaires $L^*u^*v^*$ et $L^*a^*b^*$ (voir [P]) sont choisies de façon à obtenir une représentation uniforme pour la perception, c'est-à-dire des variations identiques des coordonnées résultent en variations perceptibles de même amplitude.

Mais le rôle des coordonnées n'est pas le même, on introduit une coordonnée de luminosité et deux coordonnées de chromacité. Il est difficile de trouver les bonnes pondérations dans ce cas et d'utiliser ces trois coordonnées dans le modèle multicanaux.

Par contre dans la distance de Stiles, qui est aussi basée sur des résultats expérimentaux, on ne particularise aucune des coordonnées. Dans ce sens elle est plus proche de la représentation physique RGB de la couleur et se prête parfaitement à être utilisée dans la fonctionnelle (2).

3 Description de l'algorithme

Dans ce paragraphe nous donnons une description de l'algorithme de segmentation, plusieurs distances pouvant être utilisées. Des résultats expérimentaux seront donnés au paragraphe suivant avec différentes distances.

Cet algorithme combine une méthode classique de croissance de régions et le modèle de minimisation d'énergie proposé par Mumford et Shah. Présentons d'abord le critère de fusion (angl. merging).

Le critère de fusion. La décision de fusionner deux régions voisines, O_i et O_j , dépend du signe de

$$E(K \setminus \partial(O_i, O_j)) - E(K),$$

où $\partial(O_i, O_j)$ est la frontière commune aux régions O_i et O_j .

L'algorithme cherche à faire diminuer l'énergie globale (2) en fusionnant éventuellement deux régions voisines. Si pour tout couple de régions voisines O_i et O_j , la segmentation obtenue K' en fusionnant ces deux régions vérifie $E(K') > E(K)$ on dit que K est une segmentation 2-normale à l'échelle λ .

Le critère s'écrit

$$E(K \setminus \partial(O_i, O_j)) - E(K) = \frac{|O_i| \cdot |O_j|}{|O_i| + |O_j|} \cdot d(u^{O_i}, u^{O_j})^2 - \lambda \cdot \ell(\partial(O_i, O_j)),$$

où $|\cdot|$ désigne la surface et u^{O_i} la valeur de l'approximation de g sur O_i .

Les données utilisées dans le calcul du critère sont obtenues de la manière suivante:

Soit $g = (g_R, g_G, g_B)$, à chaque région O on associe sa surface $|O|$ et les 3 canaux $c_l^O = \int_O g_l$, ($l = R, G, B$).

On peut alors en déduire les valeurs de u sur O :

$$u^O = (u_R^O, u_G^O, u_B^O) \text{ où } u_l^O = \frac{c_l^O}{|O|}.$$

Les canaux de la région O_{new} obtenue en fusionnant O_i et O_j sont donnés par $|O_{new}| = |O_i| + |O_j|$ et $c_l^{O_{new}} = c_l^{O_i} + c_l^{O_j}$, ($l = R, G, B$).

Ainsi la fusion de deux régions se fait en additionnant les canaux correspondants et en actualisant les données.

L'algorithme.

1. Soit (u_0, K_0) la segmentation triviale (en pixels) de l'image.
2. Pour tout couple de régions voisines, calculer la valeur de λ pour laquelle une fusion est possible. Ordonner par valeurs croissantes de λ les couples de régions.
3. Effectuer la fusion du couple correspondant à la plus petite valeur de λ et mettre à jour la structure de données.
4. Répéter l'étape précédente jusqu'à obtenir le nombre de régions ou la valeur de λ demandés.

4 Expériences et conclusion

Comme nous l'avons vu précédemment, un des problèmes de la segmentation d'images couleur est la distance choisie dans l'espace des couleurs. Nous présentons ici plusieurs expériences faites en utilisant le logiciel de segmentation multicanaux `msegct` distribué avec `Megawave2` ([MW]).

La première consiste à transformer dans un premier temps l'image couleur en une image noir et blanc. Le niveau de gris d'un pixel est alors donné par: $(R, G, B) \mapsto 0.33 \cdot R + 0.5 \cdot G + 0.17 \cdot B$.

On obtient ainsi une représentation de l'intensité lumineuse de l'image.

Pour obtenir une caractérisation de la couleur on va utiliser les coordonnées (x, y) calculées par

$$x = \frac{2R - G - B}{2} \quad \text{et} \quad y = \frac{\sqrt{3}}{2}(G - B).$$

On obtient ainsi une représentation indépendante de l'intensité, qui contient l'information sur la teinte et la saturation des couleurs de l'image. On utilisera ces canaux avec les poids $(0.5, 0.5)$.

Enfin on présentera les résultats obtenus avec la distance de Stiles.

On ne présente pas des résultats de segmentations séparées des trois canaux *RGB*: en effet on obtient dans ce cas des couleurs inexistantes dans l'image original.

Dans notre approche multicanaux de la segmentation on se propose d'intégrer tous les canaux qui caractérisent une région.

Commentaire des figures.

L'intensité de l'image est représentée dans la figure 1.

La figure 2 montre des segmentations 2-normales avec 100 régions, obtenues avec les trois représentations présentées plus haut. L'image du haut montre une segmentation basée sur l'intensité lumineuse de la figure 1. L'image du centre utilise les coordonnées chromatiques (x, y) et l'image de bas utilise la distance de Stiles.

En comparant les images de 2, on remarque que la segmentation basée sur l'intensité est moins satisfaisante que celle basée sur l'information chromatique. La distance de Stiles donne des bords plus réguliers que la représentation (x, y) , réputée instable (voir [PK]).

Remerciements: Nous remercions le Professeur Pierre-Louis Lions de nous avoir introduit aux espaces de couleurs.

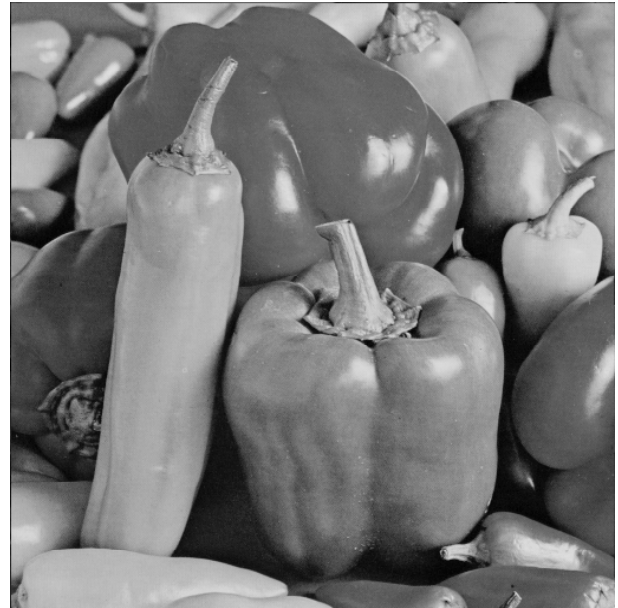


Figure 1 : Intensité lumineuse de l'image RGB "peppers".

Bibliographie

- [J] A.K. Jain. *Color Distance and Geodesics in Color 3 Space*. JOSA, vol.62, num.11, pp.1287-1291, Nov.1972.

- [KMS] G. Koepfler, J.M. Morel et S. Solimini. *Segmentation by minimizing a functional and the “merging” methods*. Proceedings ‘GRETSI Colloque’, pp.1033–1036, Juan-les-Pins, September 1991 (France).
- [KLR] G. Koepfler, C. Lopez et L. Rudin. *Data fusion by segmentation. Application to texture discrimination*. Proceedings ‘GRETSI Colloque’, pp.707–710, Juan-les-Pins, September 1993 (France).
- [KSK] G.J. Klinker, S.A. Shafer et T. Kanade. *A physical approach to color image understanding*. Int.J.Comput.Vision, 4, pp.7-38, 1990.
- [MS] D. Mumford and J. Shah. *Boundary detection by minimizing functionals*. Image Understanding, 1988, Ed. S. Ullman and W. Richards.
- [MoS] J.M. Morel and S. Solimini. *Variational Methods in Image Segmentation*. Birkhäuser Boston, 1995.
- [MW] Megawave2 sur le web
www.ceremade.dauphine.fr/~mw/
- [P] C.A. Poynton. *Frequently Asked Questions about Color*. Sur le web
www.inforamp.net/~poynton/Poynton-colour.html
où par ftp “anonymous” :
[ftp.inforamp.net/pub/users/poynton/doc/color](ftp://inforamp.net/pub/users/poynton/doc/color),
May 1995.
- [PK] F. Perez et C. Koch. *Toward Color Segmentation in Analog VLSI: Algorithm and Hardware*. Int.J.Comput.Vision, 12:1,pp.17-42, 1994.
- [WS] G. Wyszecki and W.S. Stiles. *Color Science : Concepts and Methods, Quantitative Data and Formulae*. J. Wiley and Sons, 1982.

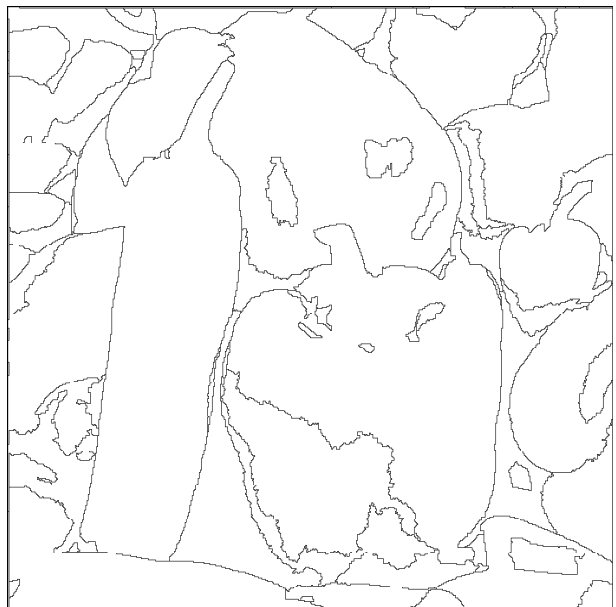
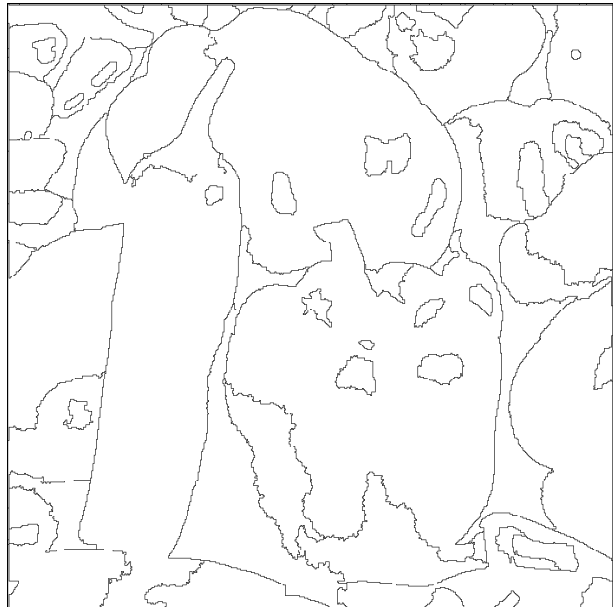
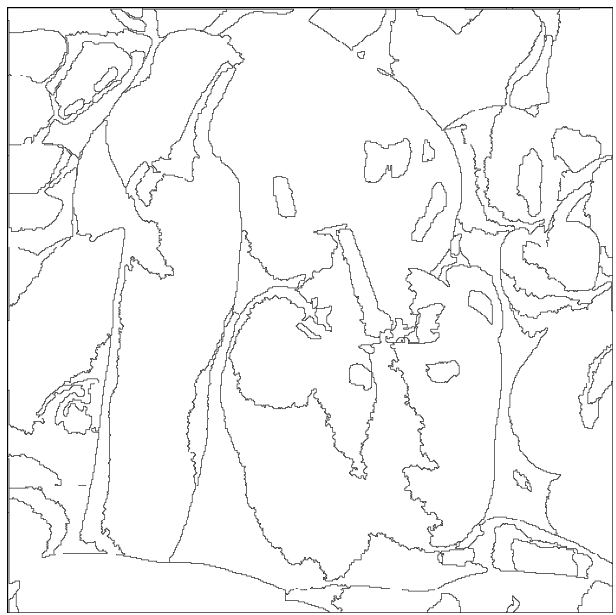


Figure 2 : Segmentation de l'image “peppers” en 100 régions (voir texte).