

Segmentation d’Images par Sélection de Courbes de Niveau

Charles KERVRANN, Mark HOEBEKE, Alain TRUBUIL

INRA – Biomérie

Domaine de Vilvert, 78352 Jouy en Josas, France

Charles.Kervrann@jouy.inra.fr, Mark.Hoebeke@jouy.inra.fr, Alain.Trubuil@jouy.inra.fr

Résumé – Nous étudions une énergie de segmentation d’image pour laquelle la solution minimisante peut être déterminée explicitement. L’approche estime à la fois le nombre inconnu d’objets et les frontières associées en sélectionnant un ensemble restreint de lignes de niveau de l’image. Dans notre cas, aucune minimisation d’énergie est requise et l’algorithme résultant est non-itératif. Des techniques de filtrage anisotropique sont introduites pour lisser les lignes de niveau des images bruitées.

Abstract – We study an image segmentation energy, which minimizer can be determined. The approach estimates the unknown number of objects and draws object boundaries by selecting the “best” level lines computed from level sets of the image. As a consequence, no energy minimization methods is necessary, yielding to a fast and non-iterative segmentation algorithm. Finally, anisotropic diffusion is used to smooth level lines in noisy images.

1 Introduction

Nous considérons ici le problème de la segmentation d’image lorsque l’image reconstruite est supposée constante par morceaux. L’analyse bayésienne d’image [5] procure des solutions élégantes à ces problèmes de segmentation. Elle conduit en définitive à la formulation du problème comme celui de la minimisation globale d’une fonctionnelle d’énergie mêlant variables observables et paramètres du modèle [5, 3, 7, 11, 10, 13]. Cette fonctionnelle se décompose elle-même en deux termes distincts : un premier terme mesurant la qualité de l’approximation et un second terme qui encourage *a priori* l’émergence de régions délimitées par des bords réguliers. Les modèles mis en jeu dans ce contexte peuvent être discrets (modèles markoviens [5]) ou continus (contours actifs [4], fonctionnelle de Mumford et Shah [11, 10, 13]). Cependant, la mise en œuvre de ces méthodes peut s’avérer délicate en raison de la charge calculatoire nécessaire pour mener la minimisation globale des énergies [5, 3, 7].

Nous nous intéressons ici aux questions de performance et de rapidité des schémas de segmentation non-supervisés s’appuyant sur une minimisation de fonction d’énergie. On peut remarquer qu’il n’est pas toujours aisé d’expliquer le minimum global de fonctions non-convexes utiles dans les applications [10]. Il apparaît donc pertinent d’étudier les fonctionnelles pour lesquelles les solutions peuvent être déterminées à l’avance. On peut alors s’affranchir de recourir à des procédures itératives de minimisation lors de la mise en œuvre de l’algorithme de segmentation. Nous avons choisi ici des énergies simples qui se déclinent par le choix judicieux de termes de régularisation adjoints à un modèle de log-vraisemblance des observations. Pour ces choix de modèles de segmentation, il s’avère que les frontières des régions extraites se superposent à des lignes ou courbes de niveau de l’image. Ces modèles d’énergie introduits par Beaulieu et Goldberg [1] dans un cadre discret

ne permettent pas cependant de contrôler la régularité des frontières des régions. Nous introduisons donc un filtrage anisotropique [2] de manière à lisser les lignes de niveau et réduire le bruit ambiant. L’algorithme complet de segmentation réalise ainsi d’une part une sélection non-itérative d’un sous-ensemble de lignes de niveau de l’image qui minimise l’énergie globale de segmentation et d’autre part un filtrage anisotropique de l’image contraint par la segmentation. Quelques itérations suffisent pour obtenir la stabilité de la carte de segmentation.

2 Formulation du problème

On suppose l’existence d’une fonction inconnue f , i.e. l’image originale définie de $[0, 1]^2$ sur \mathbb{R} . Nous observons un vecteur $y = f + \varepsilon$ discret de variables bruitées où ε est un bruit blanc gaussien. Les observations sont réparties régulièrement sur une grille 2D :

$$y(s) = f(s) + \varepsilon(s) \quad (1)$$

où $y(s)$ est l’intensité observée au pixel $s = (x, y) \in [0, 1]^2$ et N est le nombre de pixels de l’image. On suppose que la minimisation de l’énergie sur f permet de définir un partitionnement de l’image en régions.

Soit $\Omega_i \subset [0, 1]^2$, $i = 1, \dots, P$ une région ou objet de l’image. Nous associons un modèle énergétique inspiré de [1, 6], aux régions dont les supports sont inconnus :

$$E(f, \Omega_1, \dots, \Omega_P) = \sum_{i=1}^P \int_{\Omega_i} (f(x, y) - \overline{f_{\Omega_i}})^2 dx dy + \lambda J(P, \Omega_i) \quad (2)$$

où $\overline{f_{\Omega_i}}$ désigne la luminance moyenne calculée sur le support de la région Ω_i :

$$\overline{f_{\Omega_i}} = \frac{1}{|\Omega_i|} \int_{\Omega_i} f(x, y) dx dy, \quad |\Omega_i| = \int_{\Omega_i} dx dy, \quad (3)$$

λ est un réel positif et $J(P, \Omega_i)$ est une fonctionnelle *a priori* qui permet d'imposer des contraintes sur la solution.

Notre objectif est de sélectionner une collection C_P de régions ($\{\Omega_1, \dots, \Omega_{P-1}\}$, $P \geq 1$) disjointes, strictement incluses dans $[0, 1]^2$, simplement connexes, se superposant sur le fond Ω_P de l'image et telles que $\cup_{i=1}^P \Omega_i = [0, 1]^2$. Si $P = 1$, l'image ne contient pas de régions. Une segmentation optimale de l'image f (si elle existe) est par définition un minimum global de l'énergie :

$$(\Omega_1^*, \dots, \Omega_{P^*}^*) = \operatorname{argmin}_{P \geq 1} \operatorname{argmin}_{(\Omega_1, \dots, \Omega_P) \in C_P} E(f, \Omega_1, \dots, \Omega_P). \quad (4)$$

Une minimisation directe de cette énergie sur les domaines inconnus des régions Ω_i et les paramètres $\overline{f_{\Omega_i}}$, s'avère extrêmement complexe. Dans le paragraphe suivant, nous prouvons que les frontières des régions qui minimisent ce critère correspondent à des lignes de niveau de l'estimateur non paramétrique de f . Pour cela, nous devons prendre par exemple $J(P, \Omega_i) = P$ ou $J(P, \Omega_i) = \sum_{i=1}^{P-1} |\Omega_i|$. Ce choix de fonctionnelle permet de pénaliser uniquement l'émergence d'un trop grand nombre de régions dans l'image et λ peut être interprété comme un paramètre d'échelle [1, 10].

3 Estimateur

Notre estimateur \hat{f} est un estimateur défini de la manière suivante :

$$(\widehat{\Omega}_1, \dots, \widehat{\Omega}_{\widehat{P}}) = \operatorname{argmin}_{P \geq 1} \operatorname{argmin}_{(\Omega_1, \dots, \Omega_P) \in C_P} E(\hat{f}, \Omega_1, \dots, \Omega_P). \quad (5)$$

La question de l'existence d'un minimum global admissible pour des modèles de segmentation défini par Beaulieu and Goldberg [1] ou Mumford et Shah [11] a été abordée dans [10] mais ne sera pas discutée ici.

3.1 Description du minimiseur

Nous faisons l'hypothèse *ad-hoc* assurant l'existence d'un minimum global de l'énergie pour des fonctions prises dans le voisinage de l'image originale f [6]. Dans ces conditions, nous proposons le lemme suivant :

Lemme *S'il existe un minimum global admissible pour chaque fonction prise dans un voisinage de f et qu'aucun minimum pathologique existe [6], alors*

$$\hat{f}|_{\partial\Omega_i} \equiv \mu_i, \quad i = 1, \dots, P-1.$$

i.e. le bord $\partial\Omega_i$ de chaque objet Ω_i est une frontière d'un ensemble de niveau de la fonction f .

Preuve Sans perte de généralité, nous prouvons ce lemme pour un objet Ω et un fond Ω^c , où Ω^c désigne l'ensemble fermé complémentaire de Ω . Pour deux ensembles A et B , on notera $\int_{A-B} \hat{f} \equiv \int_A \hat{f} - \int_B \hat{f}$. Soit également Ω_δ une faible perturbation de Ω , i.e. la distance de Hausdorff $d_\infty(\Omega_\delta, \Omega) \leq \delta$. On définit alors

$$\int_{\Omega_\delta - \Omega} \mathbb{I} = |\Omega_\delta| - |\Omega|, \quad (6)$$

$$\left(\int_{\Omega_\delta} \hat{f} \right)^2 - \left(\int_{\Omega} \hat{f} \right)^2 = 2 \int_{\Omega} \hat{f} \int_{\Omega_\delta - \Omega} \hat{f} + \left(\int_{\Omega_\delta - \Omega} \hat{f} \right)^2.$$

La différence entre les énergies est égale à $E(f, \Omega_\delta, \Omega_\delta^c) - E(f, \Omega, \Omega^c) = T_1 + T_2 + T_3 + T_4$, avec

$$T_1 = \int_{\Omega_\delta} \hat{f}^2 - \int_{\Omega} \hat{f}^2, \quad T_2 = -\frac{1}{|\Omega_\delta|} \left(\int_{\Omega_\delta} \hat{f} \right)^2 + \frac{1}{|\Omega|} \left(\int_{\Omega} \hat{f} \right)^2,$$

$$T_3 = \int_{[0,1]^2 - \Omega_\delta} \hat{f}^2 - \int_{[0,1]^2 - \Omega} \hat{f}^2,$$

$$T_4 = -\frac{1}{1 - |\Omega_\delta|} \left(\int_{[0,1]^2 - \Omega_\delta} \hat{f} \right)^2 + \frac{1}{1 - |\Omega|} \left(\int_{[0,1]^2 - \Omega} \hat{f} \right)^2.$$

En utilisant (6), on obtient

$$T_1 = \int_{\Omega_\delta - \Omega} \hat{f}^2, \quad (7)$$

$$T_2 = -\frac{2}{|\Omega|} \int_{\Omega_\delta - \Omega} \hat{f} \int_{\Omega} \hat{f} - \frac{1}{|\Omega|} \left(\int_{\Omega_\delta - \Omega} \hat{f} \right)^2 + \frac{1}{|\Omega|^2} \int_{\Omega_\delta - \Omega} \mathbb{I} \left(\int_{\Omega} \hat{f} \right)^2 + o \left(\int_{\Omega_\delta - \Omega} \mathbb{I} \right)^2$$

et des expressions similaires pour T_3 et T_4 . On définit

$$m_0 = \int_{\Omega} \mathbb{I}, \quad m_1 = \int_{\Omega} \hat{f}, \quad K_1 = \int_{[0,1]^2} \hat{f}, \quad (8)$$

$$P_0 = -\frac{2m_1}{m_0} + \frac{2(K_1 - m_1)}{1 - m_0}, \quad P_1 = \frac{m_1^2}{m_0^2} - \frac{(K_1 - m_1)^2}{(1 - m_0)^2}.$$

Soit M un point fixe appartenant au bord $\partial\Omega$ et Ω_δ tel que $\partial\Omega_\delta = \partial\Omega$ excepté sur un petit voisinage de M . L'énergie admettant un minimum pour Ω , $f(M)$ doit être solution de l'équation

$$P_0 \hat{f}(M) + P_1 = 0 \quad (9)$$

qui admet une solution unique. P_0 and P_1 ne dépendent pas du choix de M et $P_0 \neq 0$. Donc, $\hat{f}(M)$ est constant lorsque M couvre $\partial\Omega$. \square

Nous venons de montrer que le minimiseur est une famille de courbes d'iso-intensité sur l'image pouvant être identifiées par les frontières d'ensembles de niveau.

3.2 Représentation par lignes de niveau

Matheron a proposé une représentation «morphologique» des images par des ensembles de niveaux \mathcal{S}_η [8] :

$$\mathcal{S}_\eta = \{s \in [0, 1]^2 : f(s) \geq \eta\} \quad (10)$$

pour des valeurs η situées dans la gamme de f . Cette représentation morphologique permet une reconstruction complète et invariante aux changements de contraste, des images [8]. Les ensembles de niveau peuvent être identifiés à partir des composantes connexes de l'image qui décrivent les relations spatiales entre les intensités lumineuses. Une opération de seuillage permet de construire une carte binaire de l'ensemble \mathcal{S}_η , contenant plusieurs composantes.

Une représentation analogue consiste à considérer seulement les frontières de ces ensembles de niveau, i.e. les lignes de niveau [9]. On supposera que ces lignes sont déterminées par les frontières des composantes connexes

pour des objets simplement connexes. Une région Ω_i correspondra à une composante connexe et les lignes de niveau apparaissent comme des lignes continues d'iso-valeurs situées aux bords des composantes. Notons que ces lignes (frontières des objets) n'ont aucun propriété de régularité. Une filtrage anisotropique est donc préconisé pour lisser l'image au préalable tout en préservant les structures.

3.3 Régularisation des lignes de niveau

Black *et al.* ont montré que la diffusion anisotropique pouvait être interprétée comme une procédure d'estimation robuste d'une image constante par morceaux [2]. Il est notamment établi que, formulée dans un cadre discret, l'équation de diffusion construite autour de l'estimateur de Tukey [12] préserve davantage les discontinuités et évite un lissage trop important des structures de l'image. Les équations de diffusion correspondantes s'écrivent :

$$f_{t+1}(s) = f_t(s) + \frac{\tau}{|\mathcal{G}_s|} \sum_{p \in \mathcal{G}_s} z_{sp} (f_t(p) - f_t(s)) \quad (11)$$

$$z_{sp} = \begin{cases} [1 - ((f_t(p) - f_t(s))/\sigma)^2] & |f_t(p) - f_t(s)| \leq \sigma \\ 0 & \text{sinon} \end{cases}$$

où t désigne le nombre d'itérations, τ est une constante positive qui contrôle le taux de diffusion, \mathcal{G}_s représente le voisinage au 1^{er} ordre de s , $|\mathcal{G}_s|$ est le nombre de voisins de s , $z_{sp} \in [0, 1]$ est un processus analogue à un processus ligne qui indique la présence (z_{sp} proche de 0) ou l'absence (z_{sp} proche de 1) de discontinuités et σ est un paramètre d'échelle.

Le paramètre d'échelle σ influence les résultats du filtrage et peut être estimé de manière robuste [12, 2]. Dans notre approche, nous privilégions une estimation robuste de $\sigma_i = k \tilde{\sigma}_i$ adaptée à chaque objet Ω_i où k est un réel positif,

$$\tilde{\sigma}_i = 2,6477 \sqrt{\frac{2}{|\Omega_i|} \sum_{t=0}^{|\Omega_i|/2} r_{0(t)}^2}, \quad (12)$$

et $r_{0(1)}^2 \leq r_{0(2)}^2 \leq \dots \leq r_{0(|\Omega_i|/2)}^2$ est une suite ordonnée de $|\Omega_i|/2$ résidus croissants [12], calculés à partir de l'approximation discrète des gradients de l'image [2].

Le lien entre l'extraction des lignes de niveau et le filtrage adaptatif est ainsi clairement établi. À chaque itération t , les données sont lissées en fonction de σ_i calculé sur le support de l'objet Ω_i . Il est alors possible de contrôler le nombre d'itérations nécessaire pour obtenir la stabilité de la carte de segmentation. Invariablement, les mêmes lignes de niveau sont extraites si l'image restaurée est suffisamment constante par morceaux, ce qui advient au bout de cinq à dix itérations. Notons que le contrôle de la stabilité de la restauration d'image par filtrage anisotropique effectué indépendamment de la segmentation est délicat [2].

4 Implémentation numérique

Soit les observations f_t obtenues à l'itération t après filtrage par diffusion anisotropique. Quatre étapes sont nécessaires pour construire une carte de segmentation.

CONSTRUCTION DES ENSEMBLES DE NIVEAU Soit $f_t(s)$ appartenant à $[f_{min}, f_{max}]$ partitionné en $K = 4, 8$ ou 16 intervalles disjoints de même longueur $\{[l_1, h_1], \dots, [l_K, h_K]\}$. Un pixel s est affecté à $[l_j, h_j[$ si $l_j \leq f_t(s) < h_j$.

EXTRACTION DES OBJETS On procède à un étiquetage de l'image en composantes connexes permettant d'identifier chaque objet Ω_i de taille significative et le fond Ω_P . Ceci revient à construire K images g_t^j , $j = 1, \dots, K$ pour lesquelles $g_t^j(s) = \eta$, $\forall s \in \mathcal{S}_\eta$ et $g_t^j(s) = 0$, $\forall s \notin \mathcal{S}_\eta$. La liste des composantes connexes pour chaque image permet d'établir la liste complète des objets $\{\Omega_1, \dots, \Omega_{P-1}\}$.

DÉTERMINATION DES CONFIGURATIONS À partir de la liste $\{\Omega_1, \dots, \Omega_{P-1}\}$, des configurations sont construites en énumérant toutes les combinaisons possibles d'objets, soit 2^{P-1} configurations. Chaque configuration peut être indexée par une suite de bits exprimant la présence ou l'absence d'un objet pour la configuration étudiée.

CALCUL DES ÉNERGIES Les moyennes empiriques (3) et les erreurs d'approximation (2) sont calculées une seule fois pour tous les objets Ω_i . Les quantités (2) et (3) associées au fond Ω_P sont actualisées pour chaque configuration traitée. La configuration qui minimise globalement l'énergie (2) détermine la carte de segmentation optimale. La complexité de l'algorithme dépend essentiellement de la longueur de la liste des objets.

Ces étapes sont enchainées au sein d'un algorithme :

- INITIALISATION : Soit l'image originale f_{t_0} et un choix de paramètres K, λ, τ et k .
- TANT QUE $f_{t+1}(s) \neq f_t(s)$, $\forall s \in [0, 1]^2$ RÉPÉTER
 - Segmentation d'images:
 1. Construction des ensembles de niveau
 2. Extraction des objets
 3. Détermination des 2^{P-1} configurations
 4. Calcul des énergies et sélection de la configuration optimale.
 - Diffusion anisotropique: voir (11) et (12).
 - Actualiser σ_i et k : $k \leftarrow 0.9 k$ et incrémenter t ;
- FIN TANT QUE.

5 Résultats expérimentaux

La méthode décrite a été validée sur des images d'origine variées: imagerie satellitaire, médicale et de microscopie confocale. Les paramètres de l'algorithme précédent sont les suivants: $K = 8$, $J(P, \Omega_i) = \sum_{i=1}^{P-1} |\Omega_i|$, les régions dont la taille $|\Omega_i|$ est inférieure à $0.01 \times N$ sont écartées, $\tau = 1$ et k est fixé à 5 puis abaissé linéairement avec un coefficient de 0.9 à chaque itération. Les temps de calcul sont mesurés en secondes sur une station de travail 296MHz.

Nous présentons des résultats de segmentation d'images (Figs.1-2) acquises à l'INSERM-Rouen en microscopie confocale. La microscopie confocale est une technique 2D ou 3D de visualisation optique non-invasive qui permet l'analyse

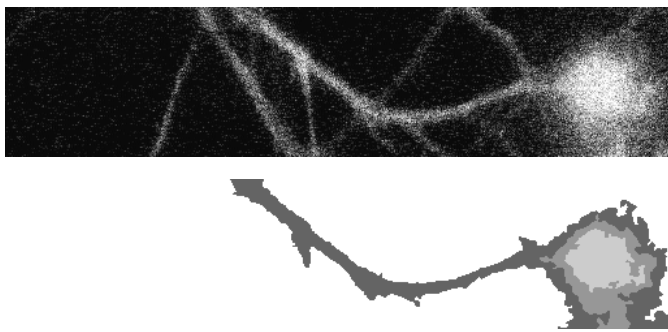


FIG. 1: Segmentation d'image (115×512) en microscopie confocale obtenue au bout de 61s et 10 itérations ($\lambda = 650$).

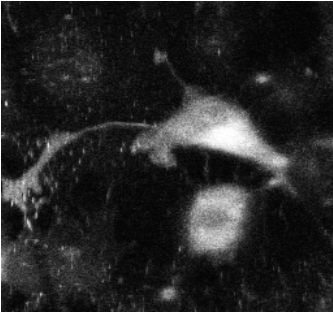


FIG. 2: Segmentation d'image (256×239) en microscopie confocale obtenue au bout de 51s et 8 itérations ($\lambda = 3250$).

en profondeur de tissus. La première série de figures présente des cellules neuronales 2D observées en culture pour lesquelles on mesure la fluorescence (Fig. 1) voire leur migration (Fig. 2). On s'intéresse ici à quelques régions d'intérêt superposées sur un fond étiqueté en blanc sur les figures.

La seconde étude de cas concerne l'extraction du cortex cérébral d'une section IRM (Fig. 3). Au préalable, la tête a été isolée du fond par la méthode que nous avons décrit. L'image étant peu bruitée, nous n'avons pas procédé ici à un filtrage de l'image et les régions d'intérêt ont été extraites au bout d'une seule itération.

Enfin, nous présentons des résultats de segmentation d'imagerie satellitaire (Fig. 4). L'algorithme a isolé les couches nuageuses significatives et l'océan (fond). Les étiquettes correspondent ici aux valeurs moyennes estimées sur chaque support des régions Ω_i .

Références

- [1] J. Beaulieu et M. Goldberg. «Hierarchy in picture segmentation: a stepwise optimization approach». *IEEE on Trans. on Patt. Anal. and Mach. Int.*, 11(2): 150–163, 1989.
- [2] M.J. Black, G. Sapiro, D. Marimont et D. Heeger. «Robust anisotropic diffusion». *IEEE on Trans. Image Process.*, 7(3): 421–432, 1998.
- [3] A. Blake et A. Zisserman. «Visual reconstruction». MIT Press, 1987.
- [4] V. Caselles, R. Kimmel et G. Sapiro. «Geodesic active contours». *Int J. of Comp. Vis.*, 22(1): 61–79, 1997.
- [5] S. Geman et D. Geman. «Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images». *IEEE on Trans. on Patt. Anal. and Mach. Int.*, 6(6): 721–741, 1984.
- [6] J. Istas. «Statistiques des processus, segmentation du signal et de l'image». HdR, Université de Paris VII, 1997.

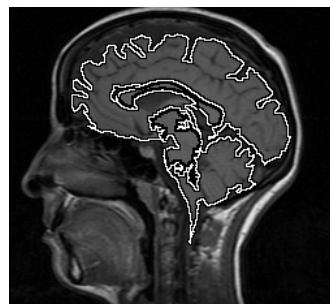


FIG. 3: Segmentation 2D d'image IRM (256×256) obtenue au bout de 4s et une seule itération ($\lambda = 32500$).

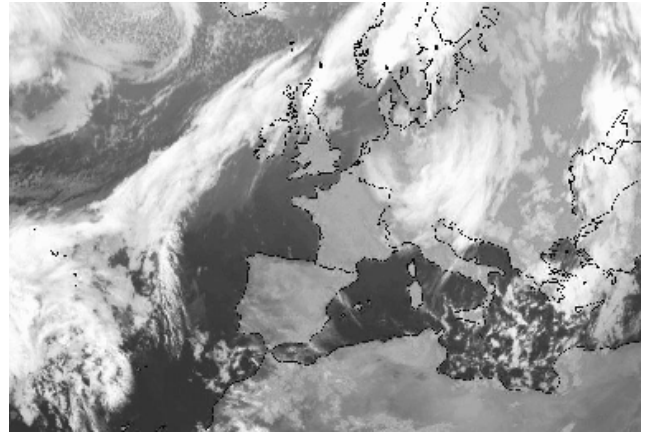


FIG. 4: Segmentation d'image météorologique (383×260) obtenue au bout de 18s et d'une seule itération ($\lambda = 2000$).

- [7] Y.G. Leclerc. «Constructing simple stable descriptions for image partitioning». *Int J. of Comp. Vis.*, 3: 73–102, 1989.
- [8] G. Matheron. «Random sets and integral morphology». John Wiley, New York, 1975.
- [9] P. Monasse et F. Guichard. «Fast computation of a contrast-invariant image representation». *IEEE on Trans. on Image Process.*, à paraître.
- [10] J.M. Morel et S. Solimini. «Variational methods in image segmentation». Birkhauser 1994.
- [11] D. Mumford et J. Shah. «Optimal approximations by piecewise smooth functions and variational problems». In *Communication on Pure and applied Mathematics*, 42(5): 577–685, 1989.
- [12] P. Rousseeuw et A. Leroy. «Robust regression and outlier detection». *Wiley Series in Probability and Mathematical Statistics*, Wiley, 1987.
- [13] S.C. Zhu et A. Yuille. «Region competition: unifying snakes, region growing, and bayes/MDL for multiband image segmentation». *IEEE on Trans. on Patt. Anal. and Mach. Int.*, 18(9): 884–900, 1996.