HAL
open science

Une formulation convexe pour la segmentation binaire d'images à partir d'histogrammes.

Romain Yildizoglu, Jean-François Aujol, Nicolas Papadakis

To cite this version:

Romain Yildizoglu, Jean-François Aujol, Nicolas Papadakis. Une formulation convexe pour la segmentation binaire d'images à partir d'histogrammes.. Colloque du Groupe de Recherche et d'Etudes de Traitement du Signal et des Images (GRETSI'13), Sep 2013, Brest, France. pp.1-4. hal-00992086

HAL Id: hal-00992086

https://hal.science/hal-00992086v1

Submitted on 16 May 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Une formulation convexe pour la segmentation binaire d'images à partir d'histogrammes.

Romain YILDIZOĞLU, Jean-François AUJOL, Nicolas PAPADAKIS

Université de Bordeaux, Institut de Mathématiques de Bordeaux (IMB, UMR 5251),
351 Cours de la Libération F-33400 Talence, France
{romain.yildizoglu;jean-francois.aujol;nicolas.papadakis}@math.u-bordeaux1.fr

Résumé – Nous présentons un modèle de minimisation d'une fonctionnelle convexe pour la segmentation binaire à partir d'histogrammes, avec une pénalisation de la longueur de la frontière de la segmentation implémentée par la variation totale. Le terme de données compare les histogrammes estimés sur les deux régions avec deux histogrammes de référence. La distance entre histogrammes choisie dans ce papier est la norme l^1. La fonctionnelle obtenue est minimisée par un algorithme adapté à l'optimisation convexe non lisse. Nous discutons les avantages de la comparaison entre les histogrammes globaux, les modifications entraînées par notre convexification et son indépendance à l'initialisation. Nous donnons une formulation avec les histogrammes cumulés en lieu et place des histogrammes qui est plus robuste dans le cas des images en niveaux de gris. Nous montrons également des résultats obtenus pour des histogrammes de niveaux de gris, de couleur, de Local Binary Pattern et du tenseur de structures.

Abstract – We present a convex minimization model for binary segmentation based on histograms, with a penalization of the length of the frontier given by the total variation regularization. The data term measures the distance between the estimated histograms in the two regions and two given histograms. The chosen distance in this paper is the l^1 one. The functional is minimized with an algorithm adapted to non smooth convex optimization. We discuss the advantages of the comparison between global histograms, the changes induced by our convexification, and the independence to initialization. We give a formulation with cumulative histograms instead of histograms that is more robust for grey level images. We also show results obtained with histograms of gray levels, color, Local Binary Patterns and the structure tensor.

1 Introduction

La segmentation d'image est l'objet de recherches actives depuis plus de vingt ans (voir [2, 11] et les références incluses). On peut par exemple citer l'article fondateur de Mumford et Shah [16] et son approximation par lignes de niveaux développée par Chan et Vese dans [10]. Ce dernier article fournit un algorithme très flexible pour segmenter une image en deux parties homogènes, chacune étant caractérisée par sa valeur moyenne.

Pour les images texturées, un grand nombre d'extensions [10] ont été proposées pour améliorer le modèle reposant sur la valeur moyenne. Par exemple, des filtres de Gabor dans [26] et [21], des paquets d'ondelettes dans [3] ou encore le tenseur de structure dans [6, 25], sont utilisés pour caractériser la texture.

Lorsque l'on considère les histogrammes globaux des régions à segmenter, il existe également une importante littérature [1, 7, 14] qui repose aussi sur les travaux de [10]. Ces approches nécessitent cependant des calculs complexes de dérivées de formes [12] pour l'équation d'évolution de la ligne de niveau. Il est important de remarquer que l'ensemble des approches précédemment décrites repose sur l'utilisation des méthodes par lignes de niveaux [20] et seulement un minimum local associé à l'énergie est calculé. D'autres modèles comme ceux de [27, 4, 13] qui utilisent des algorithmes sur des graphes ainsi que [24] obtiennent de bons résultats sans lignes de niveaux, mais il n'y a pas de preuve que ces algorithmes convergent vers un minimum global de l'énergie originale.

Récemment, des méthodes de convexification ont été proposées pour résoudre ce problème [18, 9, 5, 28]. Le modèle Chan-Vese original [10] est convexifié et une solution globale est calculée efficacement avec des algorithmes de type primal-dual. Un simple seuillage de la solution globale donne alors un minimum global du problème original non convexe.

La contribution de ce papier est de proposer un modèle convexe pour segmenter une image en deux parties, chaque région étant caractérisée par son histogramme. Ce modèle convexe est alors minimisé efficacement par un algorithme d'optimisation convexe non lisse [23]. Notons que des algorithmes de segmentation reposant sur les histogrammes ont été récemment proposés. Ces travaux utilisent la distance de Bhattacharyya [30] ou la distance de Wasserstein [17, 22] pour comparer les histogrammes. Cependant, comme ces méthodes reposent également sur l'évolution d'une ligne de niveau, elles sont non convexes et la solution obtenue dépend de l'initialisation.

Par rapport aux méthodes de segmentation précédemment mentionnées qui reposent sur l'évolution d'une ligne de niveau, l'approche ici est plus simple, et malgré la simplicité de l'approche, les segmentations obtenues sont satisfaisantes tout en

étant indépendantes de l'initialisation.

Le plan du papier est le suivant. Nous introduisons notre modèle pour la segmentation binaire reposant sur les histogrammes dans la section 2. Notre fonctionnelle repose sur une norme l^1 pour le terme de données, et nous proposons un schéma numérique efficace qui repose sur un algorithme d'optimisation convexe non lisse. Nous présentons dans la section 4 des exemples numériques pour montrer l'intérêt de la méthode ainsi que ses limites. Nous concluons dans la section 5 et présentons des pistes futures de recherche.

2 Modélisation

Le problème traité dans cet article est la segmentation d'une image en deux parties à partir de deux histogrammes connus. Pour cela, nous définissons une fonctionnelle dont la segmentation sera un minimiseur. Introduisons tout d'abord les notations. Nous supposons que le domaine de l'image Ω est un ouvert non vide borné de $\mathbb{R}^2$ à bord Lipschtitz. Le produit scalaire sur $L^2(\Omega)$ est noté $\langle .,. \rangle$. Soit $I : \Omega \mapsto \Lambda \subset \mathbb{R}^k$ l'image, Λ fini, h^0 et h^1 deux histogrammes de références : $\Lambda \to [0;1]$, avec $\sum_{\lambda\in\Lambda} h^i(\lambda) = 1$, $i = 0, 1$. Nous souhaitons estimer une segmentation binaire representée par $u : \Omega \to \{0, 1\}$, où l'histogramme calculé sur la région $\Omega_0 := \{x \in \Omega, u(x) = 0\}$ est proche de h^0 (respectivement l'histogramme sur $\Omega_1 := \{x \in \Omega, u(x) = 1\}$ doit être proche de h^1). L'histogramme sur la région Ω_1 est donné par :

$$h_u(\lambda) = \frac{1}{\sum_{x\in\Omega} u(x)} \sum_{x\in\Omega} u(x)\mathbb{1}_{I=\lambda}(x) \tag{1}$$

En notant $TV(u)$ la variation totale de u et $\| h \|_1 = \sum_{\lambda\in\Lambda} |h(\lambda)|$ la norme l^1, nous considérons la fonctionnelle :

$$J(u) = TV(u) + \mu \left|\left|(h_u - h^1)_{\lambda\in\Lambda}\right|\right|_1 + \mu \left|\left|(h_{1-u} - h^0)_{\lambda\in\Lambda}\right|\right|_1 \tag{2}$$

Cette fonctionnelle n'est pas convexe. En introduisant $g^1_\lambda := \mathbb{1}_{I=\lambda} - h^1(\lambda)$, $g^0_\lambda := \mathbb{1}_{I=\lambda} - h^0(\lambda)$, nous obtenons :

$$\| h_u - h^1 \| = \frac{1}{\sum_\Omega u(x)} \left|\left|\left(\left\langle u, g^1_\lambda \right\rangle_\Omega\right)_{\lambda\in\Lambda}\right|\right|_1 \tag{3}$$

En supposant connue la taille de la partition minimale, le facteur $1/\sum_\Omega u(x)$ peut alors être remplacé par une constante $\beta \in [0, 1]$ pour obtenir la nouvelle fonctionnelle :

$$\begin{aligned} J(u) &= TV(u) + \frac{\mu}{\beta} \| (\left\langle u, g^1_\lambda \right\rangle_\Omega)_{\lambda\in\Lambda} \|_1 \\ &\quad + \frac{\mu}{1-\beta} \| (\left\langle 1-u, g^0_\lambda \right\rangle_\Omega)_{\lambda\in\Lambda} \|_1 . \end{aligned} \tag{4}$$

Cette fonctionnelle est bien convexe, mais son domaine de définition ne l'est pas ($u(x) \in \{0; 1\}$). Nous relaxons donc la condition sur u en considérant le domaine convexe $0 \le u \le 1$. Cette nouvelle fonctionnelle peut alors être minimisée globalement. Le minimiseur obtenu est finalement seuillé à 0.5 pour obtenir une segmentation binaire. Même si rien ne garantit que la solution seuillée du problème relaxé est solution du problème original, nous observons en pratique un très bon comportement de notre algorithme. Il est cependant possible de créer des images qui ont de nombreux minima globaux pour le problème original et pour lesquelles nous ne savons pas obtenir de solution intéressante à partir d'un minimiseur de la fonctionnelle convexe.

Notons que la fonctionelle compare les histogrammes globalement, elle ne regarde pas pour chaque pixel dans quel modèle il est plus probable comme dans [15], ni si localement l'histogramme estimé est proche de l'histogramme donné comme dans [17]. Cela permet d'obtenir de bons résultats même si l'histogramme de la région est très différent des histogrammes locaux.

La distance l^1 entre les histogrammes ne tient pas compte d'éventuelles métriques sur l'espace Λ. Ceci peut être gênant si par exemple on connaît l'histogramme d'une image avant que du bruit ne soit rajouté. Dans le cas des images en niveaux de gris, on peut pallier ce problème en remplaçant dans notre fonctionnelle l'histogramme par l'histogramme cumulé. On prend $g_{C,\lambda} = \sum_{\gamma\le\lambda} g^0_\lambda = \mathbb{1}_{I\le\lambda} - h^0_C(\lambda)$, avec $h^0_C(\lambda) = \sum_{\gamma\le\lambda} h^0(\gamma)$, ce qui revient à utiliser la distance de Wasserstein l^1 entre les histogrammes. L'optimisation de cette fonctionnelle suit les mêmes lignes que celle de la précédente.

3 Optimisation

Écrivons la fonctionnelle avec des variables duales. Soient A et B les opérateurs linéaires, $A : u \mapsto (\left\langle u, g^1_\lambda \right\rangle_\Omega)_{\lambda\in\Lambda}$, $B : u \mapsto (\left\langle u, g^0_\lambda \right\rangle_\Omega)_{\lambda\in\Lambda}$. On peut écrire le problème comme une recherche de point selle :

$$\arg \max_{q_1,q_2,q_3} \arg \min_u E(q_1, q_2, q_3, u) \tag{5}$$

où

$$\begin{aligned} E(q_1, q_2, q_3, u) &= \langle q_1, \nabla u\rangle_\Omega + \langle q_2, Au\rangle_\Lambda + \langle q_3, B(1-u)\rangle_\Lambda \\ &\quad -\chi_{B(0,1)}(q_1) - \chi_{[-\mu\frac{1}{\beta},\mu\frac{1}{\beta}]}(q_2) \\ &\quad -\chi_{[-\mu\frac{1}{1-\beta},\mu\frac{1}{1-\beta}]}(q_3) + \chi_{[0,1]}(u) \\ &= \langle q, Ku\rangle_{\Omega\times\Lambda^2} - F^*(q) + G(u) \end{aligned} \tag{6}$$

avec $q = (q_1, q_2, q_3) \in (\mathbb{R}^\Omega \times \mathbb{R}^\Omega) \times \mathbb{R}^\Lambda \times \mathbb{R}^\Lambda$ de dimension $2|\Omega| + 2|\Lambda|$, $K : u \mapsto (\nabla u, Au, -Bu)$, $G(u) = \chi_{[0,1]}(u)$, $F^*(q) = \chi_{B(0,1)}(q_1) + \chi_{[-\mu\frac{1}{\beta},\mu\frac{1}{\beta}]}(q_2) + \chi_{[-\mu\frac{1}{1-\beta},\mu\frac{1}{1-\beta}]}(q_3) - \langle q_3, B1\rangle$, $B(0, 1)$ le disque unité fermé de $\mathbb{R}^2$. χ_C est la fonction caractéristique de l'ensemble convexe C, c'est à dire $\chi_C(x) = 0$ si $x \in C$, $\chi_C(x) = \infty$ sinon.

Optimisation avec un algorithme primal-dual préconditionné (PDP)

Nous donnons la forme générale de l'algorithme [23], puis son application à notre problème.

Soient X, Y deux espaces vectoriels de dimension finie. Notons $\langle .,. \rangle$ le produit scalaire usuel, $K : X \to Y$ un opérateur linéaire et $G : X \to \mathbb{R} \cup \{\infty\}$ et $F^* : Y \to \mathbb{R} \cup \{\infty\}$ deux fonctions convexes (F^* étant la transformée de Legendre Fenchel de F [23]). Nous voulons résoudre :

$$\min_{x \in X} \max_{y \in Y} \langle Kx, y \rangle + G(x) - F^*(y). \tag{7}$$

Dans [23], un algorithme primal-dual préconditionné est proposé pour résoudre ce problème :

Algorithm 1 Algorithme primal-dual préconditionné ([23])

$$u^{k+1} = (I + T\partial G)^{-1} \left(u^k - TK^t z^k\right)$$
$$z^{k+1} = (I + \Sigma \partial F^*)^{-1} \left(z^k - \Sigma K(2u^{k+1} - u^k)\right)$$

avec T et Σ des matrices symétriques définies positives telles que $\| \Sigma^{\frac{1}{2}} K T^{\frac{1}{2}} \|^2 < 1$ et $(I + T\partial G)^{-1}(\hat{u}) := \arg\min_{u \in U} G(u) + \frac{1}{2\tau} \left\langle T^{-1}(u - \hat{u}), u - \hat{u} \right\rangle$. Cet algorithme converge vers un point selle en $O(\frac{1}{k})$ (avec k le nombre d'itérations). Comme il est proposé dans [23], on peut prendre $T = \mathrm{diag}(\tau_1, ..., \tau_n)$ et $\Sigma = \mathrm{diag}(\sigma_1, ..., \sigma_m)$ avec $\tau_j < \frac{1}{\sum_{i=1}^m |K_{i,j}|}$, $\sigma_i < \frac{1}{\sum_{j=1}^n |K_{i,j}|}$.

Application de PDP à la fonctionnelle (5)

Introduisons les notations suivantes :

$$\begin{cases} Tu(x) = \tau_x u(x), & \tau_x < \frac{1}{2 + \sum_{\lambda \in \Lambda} (|g_\lambda^1(x)| + |g_\lambda^0(x)|)}, \\ A^t q_2 = \sum_{\lambda \in \Lambda} q_2(\lambda) g_\lambda^1, & \\ B^t q_3 = \sum_{\lambda \in \Lambda} q_3(\lambda) g_\lambda^0, & \\ \sigma_1 < \frac{1}{4}, & \\ \Sigma_2 q_2(\lambda) = \sigma_{2,\lambda} q_2(\lambda), & \sigma_{2,\lambda} < \frac{1}{\sum_{x \in \Omega} g_\lambda^1(x)}, \\ \Sigma_3 q_3(\lambda) = \sigma_{3,\lambda} q_3(\lambda), & \sigma_{3,\lambda} < \frac{1}{\sum_{x \in \Omega} g_\lambda^0(x)}, \\ b = B1, & \end{cases} \tag{8}$$

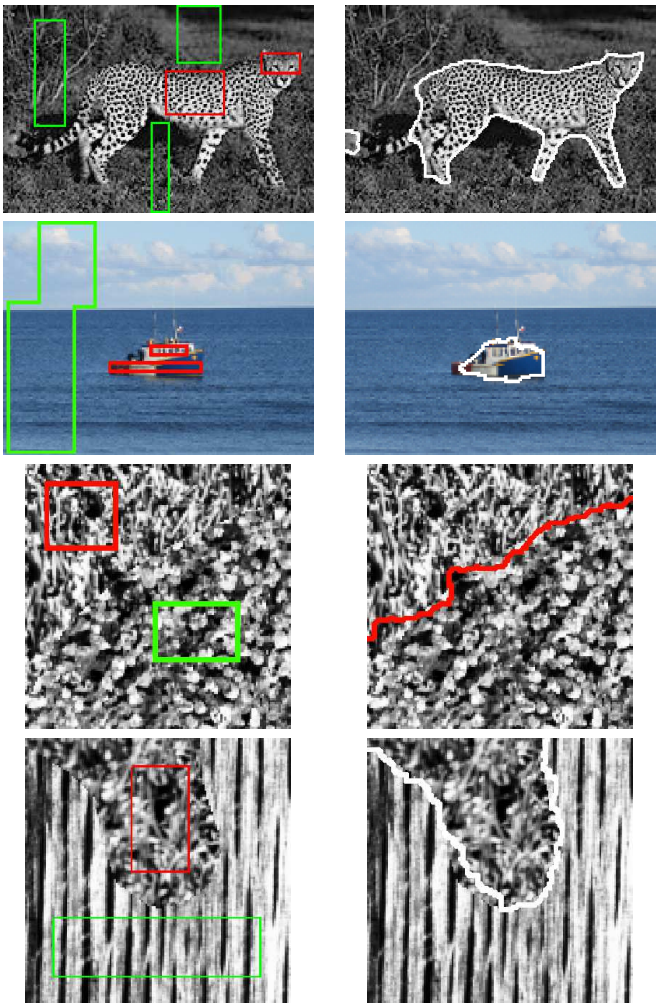
L'algorithme 1 appliqué à la fonctionnelle (5) donne l'algorithme 2, où P_C est la projection orthogonale sur le convexe C :

Algorithm 2

$$u^{k+1} = P_{[0,1]} \left(u_k - T \left(A^t q_2^k - B^t q_3^k - \mathrm{div}(q_1^k)\right)\right)$$
$$q_1^{k+1} = P_{B(0,1)} \left(q_1^k + \sigma_1 \nabla(2u^{k+1} - u^k)\right)$$
$$q_2^{k+1} = P_{[-\mu \frac{1}{\beta}, \mu \frac{1}{\beta}]} \left(q_2^k + \Sigma_2 A(2u^{k+1} - u^k)\right)$$
$$q_3^{k+1} = P_{[-\mu \frac{1}{1-\beta}, \mu \frac{1}{1-\beta}]} \left(q_3^k - \Sigma_3 \left(B(2u^{k+1} - u^k) - b\right)\right).$$

4 Résultats

La Figure 1 présente les résultats de la méthode avec des histogrammes cumulés pour les niveaux de gris, l'histogramme des couleurs pour les image couleurs et le tenseur de structure [25] ou les Local Binary Pattern (LBP) [19] pour la texture. Sur ces images, les zones d'apprentissage des histogrammes de référence sont indiquées en rouge et vert. L'inconnue est systématiquement initialisée comme une constante afin de souligner le caractère global de l'approche. Les segmentations sont satisfaisantes. Notons que le bateau est retrouvé malgré le déséquilibre entre les deux régions, ici une modification du paramètre β peut améliorer le résultat.



(a) Image initiale (b) Segmentation

FIGURE 1 – Segmentations de quelques images avec $\beta = 0.5$. Les types d'histogrammes utilisés sont, de haut en bas : les niveaux de gris, la couleur, le tenseur de structure, les LBP. Les segmentations sont satisfaisantes, une modification du paramètre β permettrait d'améliorer la segmentation du navire [29].

5 Conclusions et perspectives

Nous avons présenté un algorithme de segmentation reposant sur des comparaisons entre histogrammes et nous avons montré son intérêt pour différents types d'images et d'histogrammes. L'utilisation des histogrammes cumulés permet d'être plus robuste pour les images en niveaux de gris. Même si notre convexification fait apparaître un nouveau paramètre β, celui-

ci ne nous a pas paru critique pour les images testées. Il a d'ailleurs été constamment fixé à 0.5 dans les exemples donnés dans ce papier. De futurs travaux pourraient porter sur une version n-aire de la modélisation en utilisant les travaux de [8], sur des tests extensifs pour différents descripteurs sous forme d'histogrammes, ainsi qu'un examen de la relation entre le minimiseur du problème convexe et celui du problème original.

Références

[1] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson. Image segmentation using active contours : Calculus of variations or shape gradients ? *SIAM Applied Mathematics*, 63 :2003, 2002.

[2] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, 2002.

[3] J-F. Aujol, G. Aubert, and L. Blanc-Féraud. Wavelet-based level set evolution for classification of textured images. *IEEE Transactions on Image Processing*, 12 :1634–1641, 2003.

[4] I. Ayed, H. Chen, K. Punithakumar, I. Ross, and S. Li. Graph cut segmentation with a global constraint : Recovering region distribution via a bound of the bhattacharyya measure. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 3288–3295. IEEE, 2010.

[5] E. Brown, T. F. Chan, and X. Bresson. Completely Convex Formulation of the Chan-Vese Image Segmentation Model. *IJCV*, pages 1–19, 2011.

[6] T. Brox, M. Rousson, R. Deriche, and J. Weickert. Unsupervised segmentation incorporating colour, texture, and motion. In *Computer Analysis of Images and Patterns*, 2003.

[7] T. Brox and J. Weickert. Level set segmentation with multiple regions. *IEEE Transactions on Image Processing*, 15(10) :3213–3218, 2006.

[8] A. Chambolle, D. Cremers, and T. Pock. A convex approach to minimal partitions. *SIAM Journal on Imaging Sciences*, 5(4) :1113–1158, 2012.

[9] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *JMIV*, 40 :120–145, 2011.

[10] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2) :266 –277, 2001.

[11] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation : Integrating color, texture, motion and shape. *International Journal of Computer Vision*, 72 :215, 2007.

[12] M. C. Delfour and J.-P. Zolésio. *Shapes and geometries : analysis, differential calculus, and optimization*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

[13] L. Gorelick, Frank R. Schmidt, Y. Boykov, A. Delong, and A. Ward. Segmentation with non-linear regional constraints via line-search cuts. In *European Conference on Computer Vision (ECCV'12)*, pages 583–597. Springer, 2012.

[14] A. Herbulot, S. Jehan-Besson, S. Duffner, M. Barlaud, and G. Aubert. Segmentation of vectorial image features using shape gradients and information measures. *Journal of Mathematical Imaging and Vision*, 25(3) :365–386, October 2006.

[15] J. L. Marroquin, E. A. Santana, and S. Botello. Hidden markov measure field models for image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(11) :1380–1387, 2003.

[16] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42 :577–685, 1989.

[17] K. Ni, X. Bresson, T. F. Chan, and S. Esedoglu. Local histogram based segmentation using the wasserstein distance. *International Journal of Computer Vision*, 84(1) :97–111, August 2009.

[18] M. Nikolova, S. Esedoglu, and T. F. Chan. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal on Applied Mathematics*, 66(5) :1632–1648, 2006.

[19] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7) :971 –987, jul 2002.

[20] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed : Algorithms based on hamilton-jacobi formulations. *JCP*, 79(1) :12 – 49, 1988.

[21] N. Paragios and R. Deriche. Geodesic active regions for supervised texture segmentation. In *ICCV*, 1999.

[22] G. Peyré, J. Fadili, and J. Rabin. Wasserstein active contours. In *IEEE International Conference on Image Processing (ICIP'12)*, 2012.

[23] T. Pock and A. Chambolle. Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In *IEEE International Conference on Computer Vision (ICCV'11)*, pages 1762 –1769, 2011.

[24] K. Punithakumar, J. Yuan, I. Ben Ayed, S. Li, and Y. Boykov. A convex max-flow approach to distribution-based figure-ground separation. *SIAM Journal on Imaging Sciences*, 5(4) :1333–1354, 2012.

[25] M. Rousson, T. Brox, and R. Deriche. Active unsupervised texture segmentation on a diffusion based feature space. In *CVPR*, 2003.

[26] L. A. Vese and T. F. Chan. A multiphase level set framework for image segmentation using the mumford and shah model. *International Journal of Computer Vision*, 50 :271–293, 2002.

[27] S. Vicente, V. Kolmogorov, and C. Rother. Joint optimization of segmentation and appearance models. In *IEEE International Conference on Computer Vision (ICCV'09)*, pages 755–762. IEEE, 2009.

[28] R. Yildizoglu, J-F. Aujol, and N. Papadakis. Active contours without level sets. In *IEEE International Conference on Image Processing (ICIP'12)*, 2012.

[29] R. Yildizoglu, J-F. Aujol, and N. Papadakis. A convex formulation for global histogram based binary segmentation. In *Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR'13)*, 2013.

[30] Y. Yuan, E. Ukwatta, X.C. Tai, A. Fenster, and C. Schnörr. A fast global optimization-based approach to evolving contours with generic shape prior. *In submission in IEEE TPAMI, also UCLA Tech. Report CAM 12-38*, 2012.