

Potentiels basés patchs pour l'extraction interactive de contours

Thoraya BEN CHATTAH, Sébastien BOUGLEUX, Olivier LÉZORAY

Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France

{sebastien.bougleux, olivier.lezoray}@unicaen.fr

Résumé – L'extraction interactive des contours d'objets dans des images est difficile et trouve de nombreuses applications en édition d'images. Les méthodes proposées pour résoudre ce problème ont un objectif commun : requérir un minimum d'effort de l'utilisateur. Pour les techniques de chemins minimaux, cela dépend de la capacité de la fonction de potentielle à capturer les contours. Nous proposons de nouveaux potentiels basés patchs conçus pour avoir de faibles valeurs aux bords de l'objet ciblé. Pour évaluer ceux-ci, nous quantifions leurs capacités en termes de nombre de points germes nécessaires. Les résultats visuels et quantitatifs démontrent la capacité de ces potentiels à réduire l'interaction de l'utilisateur tout en préservant une bonne précision d'extraction.

Abstract – Interactive extraction of object contours from images is difficult and has many applications in image editing. The proposed methods to solve this problem have a common objective: to require a minimum of effort from the user. For minimum path techniques, this depends on the ability of the potential function to capture contours. We propose new patch-based potentials designed to have low values at the edges of the targeted object. To evaluate these, we quantify their capacities in terms of the number of germ points required. The visual and quantitative results demonstrate the ability of the proposed potentials to reduce user interaction while maintaining good extraction accuracy.

1 Introduction

L'extraction interactive de contours, qui consiste à délimiter avec précision le contour d'un objet, reste un problème difficile. Cela est souvent lié à la présence de structures dont les limites sont mal définies, et des outils d'aide au détournage peuvent être très bénéfiques pour les utilisateurs finaux. Pour cela, il est nécessaire que l'utilisateur fournisse des entrées pour faciliter l'extraction de l'objet. Ces entrées peuvent prendre différentes formes qui sont ensuite incorporées dans le processus de segmentation sous forme de contraintes. Comme cela nécessite une interaction avec l'utilisateur, toutes les méthodes tentent d'obtenir des résultats de segmentation précis tout en réduisant l'effort de l'utilisateur. Il y a essentiellement trois méthodes dans la littérature pour permettre à l'utilisateur de spécifier l'objet ciblé. La première est d'étiqueter certains pixels à l'intérieur/à l'extérieur de l'objet [1] et d'utiliser un processus d'étiquetage par coupe de graphe [2]. La deuxième méthode consiste à fournir une sous-région à l'intérieur de l'image qui contient l'objet (une boîte englobante) [3]. La troisième méthode consiste à tracer une courbe ou des points germes près du contour. Fournir des courbes a été largement utilisé dans les méthodes de contours actifs [4]. Fournir des points a été principalement utilisé avec des approches de chemin minimal [5]. Dans la suite, nous nous concentrerons sur ce dernier type d'interaction que nous appelons les méthodes interactives d'extraction de contours à base de points germes.

2 Méthodes interactives d'extraction

2.1 Chemins minimaux

La plupart des méthodes interactives d'extraction de contours à base de germes reposent sur des chemins minimaux. Étant donnée une image $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}_+^n$ et deux germes $p_s, p_e \in \Omega$ situés sur le contour d'un objet, l'objectif est de trouver un chemin entre p_s et p_e qui représente un morceau du contour. Dans cet article, nous considérons le modèle classique présenté dans [5]. Un chemin de p_s à p_e est défini comme une courbe paramétrique $\gamma : [0, 1] \rightarrow \Omega$ pour $\gamma(0) = p_s$ et $\gamma(1) = p_e$, et sa longueur a comme coût $L(\gamma) = \int_0^1 P(\gamma(s)) \|\gamma'(s)\| ds$ avec γ' la dérivée de γ . $P : \Omega \rightarrow \mathbb{R}_+^*$ est une fonction de potentiel dérivée de l'image f qui a de faibles valeurs pour les bords. La longueur $L(\gamma)$ est plus petite lorsqu'elle passe par de faibles valeurs P et l'on peut rechercher des courbes qui minimisent $L(\gamma)$ avec une paire de points de début et de fin. Le but est donc de capturer le chemin minimal γ^* , aussi appelé chemin géodésique, qui minimise globalement la fonction de coût $L(\gamma) : L(\gamma^*) = \min_{\gamma \in \mathcal{A}(p_s, p_e)} L(\gamma)$ avec $\mathcal{A}(p_s, p_e)$ l'ensemble des chemins possibles reliant p_s à p_e . Le coût optimal définit la distance géodésique $d(p_s, p_e) = L(\gamma^*)$. Pour résoudre ce problème de minimisation, la carte de distance géodésique $\mathcal{U}_{\{p_s\}} : \Omega \rightarrow \mathbb{R}_+$ d'un point p_s à tout autre point $p \in \Omega$ est considérée : $\mathcal{U}_{\{p_s\}}(p) = d(p_s, p) = \min_{\gamma \in \mathcal{A}(p_s, p)} L(\gamma)$. La carte de distance est la solution de l'équation Eikonale $\|\nabla \mathcal{U}_{\{p_s\}}(p)\| = P(p)$, $\forall p \in \Omega$ avec $\mathcal{U}_{\{p_s\}}(p_s) = 0$. La valeur $\mathcal{U}_{\{p_s\}}(p)$ peut être interprétée comme le temps d'arrivée à p d'un front se propageant depuis un point p_s avec une vitesse $1/P$. La méthode du Fast Marching [6, 5] est utilisée pour calculer les valeurs de

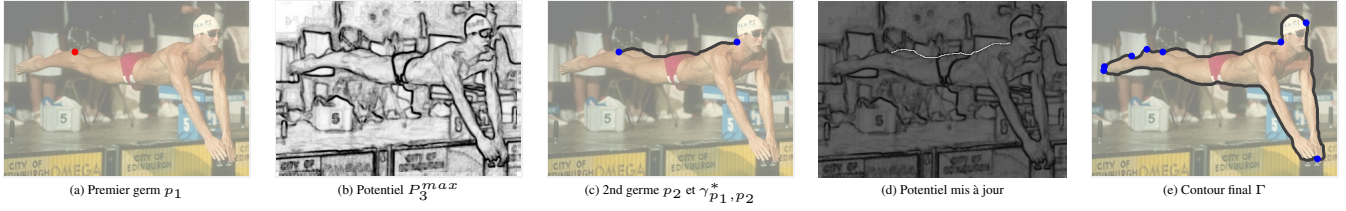


FIGURE 1 – Une extraction de contour interactive avec le framework livewire.

$U_{\{p_s\}}$ en ordre croissant, afin que le processus puisse être arrêté lorsque le point final p_e est atteint. Ensuite, le chemin minimal entre p_s et p_e peut être construit en appliquant une procédure de descente de gradient à partir de p_e le long du gradient de la carte U_{p_s} , pour arriver à p_s (voir [7] pour plus de détails). La courbe de résultat γ^* dépend fortement du potentiel P qui est habituellement construit de telle sorte à prendre des valeurs faibles à la structure d'intérêt désirée et des valeurs plus élevées ailleurs. Certains travaux ont suggéré l'utilisation de nouvelles fonctions de potentiel [8] afin de mieux distinguer les discontinuités entre les objets et l'arrière-plan. Nous nous concentrons sur cette piste de recherche.

2.2 Le framework Livewire

Lorsque les germes sont fournis sous forme de points dans des approches de chemins minimaux pour l'extraction du contour, les méthodes employées adoptent toutes la même stratégie. L'utilisateur place des points germes qui fournissent un étiquetage initial pour certains pixels de l'image. Ensuite, un algorithme est exécuté pour propager les étiquettes des germes dans les régions non étiquetées jusqu'à ce qu'un critère optimal soit atteint. L'utilisateur peut soit placer sur le contour tous les points germes [9] ou bien les fournir l'un après l'autre [10, 11]. Ce dernier type d'interaction est employé dans un framework général nommé livewire. Ils nécessitent une séquence ordonnée de points germes (fournis l'un après l'autre) pour extraire le contour de l'objet. L'avantage de ce type d'interaction est que l'utilisateur peut contrôler le résultat final. L'interaction doit être minimale en ce sens que peu de points germes doivent être utilisées et que le plus grand chemin optimal entre eux doit être obtenu. L'algorithme 2.2 resume le framework livewire et la Figure 1 illustre ce processus interactif pour extraire un contour avec un des potentiels (P_3^{max}) que nous proposons.

3 Potentiels basés patches

Dans cette section, nous présentons de nouveaux potentiels basés patches pour l'extraction de contours avec le framework livewire. Un contour est une courbe Γ qui sépare de deux régions \mathcal{R}^+ et \mathcal{R}^- qui ont des caractéristiques différentes (par exemple, couleur, texture). Notre proposition consiste à utiliser des patches comme descripteurs de texture pour construire un potentiel qui a de petites valeurs pour les pixels de contours.

3.1 Notations

Pour une image f , la couleur à un pixel $p_i = (x_i, y_i)^T$ est donnée par $f(p_i) \in \mathbb{R}^3$. Un patch est une région carrée de taille

Algorithm 1 Framework Livewire pour l'extraction de contour

Fonction de potentiel P , Ensemble de germes $\mathcal{S} = \emptyset$, contour Γ
 Sélectionner un point germe p_1
 Calculer la carte des distances géodésiques $U_{\{p_1\}}$ avec P
 $p_s \leftarrow p_1, \mathcal{S} \leftarrow \mathcal{S} \cup \{p_s\}$
repeat
 Sélectionner un point germe p_e
 Déterminer le plus court chemin γ_{p_s, p_e}^* à partir de $U_{\{p_s\}}$
 $\mathcal{S} \leftarrow \mathcal{S} \cup \{p_e\}, \Gamma \leftarrow \Gamma \cup \gamma_{p_s, p_e}^*$
 Mettre à jour le potentiel $P(q) \leftarrow +\infty, \forall q \in \gamma_{p_s, p_e}^* \setminus \{p_s, p_e\}$
 Calculer la carte des distances géodésiques $U_{\{p_e\}}$ avec P
 Déterminer le plus court chemin γ_{p_e, p_1}^* à partir de $U_{\{p_e\}}$
 $p_s \leftarrow p_e$
until L'utilisateur est satisfait du contour extrait
 $\Gamma = \Gamma \cup \gamma_{p_{i+1}, p_1}^*$

$(2w+1)^2$ autour d'un pixel et est défini par le vecteur $\mathcal{P}(p_i) = (f(p_i + t), \forall t \in [-w, w]^2)^T \in \mathbb{R}^{3(2w+1)^2}$. La distance entre deux patches est fournie par la norme L_2 $d(\mathcal{P}(p_i), \mathcal{P}(p_j)) = \|\mathcal{P}(p_i) - \mathcal{P}(p_j)\|_2$. Étant donné $t_{\theta, \delta}$ un vecteur de translation δ selon l'angle θ avec la ligne horizontale, on peut définir les pixels gauche et droit à une distance δ d'un pixel p_i sur une ligne d'angle θ avec l'horizon comme : $p_i^{l, \theta, \delta} = p_i + t_{\theta, \delta}$ et $p_i^{r, \theta, \delta} = p_i - t_{\theta, \delta}$. Dans la suite, on considérera les lignes d'angle $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$.

3.2 Potentiel basés pixels

Nous proposons deux premiers potentiels qui ne prennent en compte que les pixels entourant un pixel courant p_i dans un 8-voisinage. Un pixel peut être considéré comme étant sur un contour si le patch à sa gauche est différent du patch à sa droite. Nous définissons donc une différence de patch centrée sur p_i comme : $d_c^{\theta, \delta}(p_i) = d(\mathcal{P}(p_i^{l, \theta, \delta}), \mathcal{P}(p_i^{r, \theta, \delta}))$. Cependant, le contour n'est pas nécessairement orienté le long de l'axe y et nous devons tester toutes les orientations possibles dans un 8-voisinage pour définir le potentiel : $P_1(p_i) = \max_{\theta} d_c^{\theta, 1}(p_i)$ avec $\delta = 1$ puisque nous ne considérons que les pixels voisins. Puisqu'un pixel sur un contour n'est pas non plus supposé être similaire aux pixels qui l'entourent, nous pouvons améliorer ce potentiel en calculant les différences à gauche ou à droite entre le pixel et ses voisins. Nous définissons $d_l^{\theta, \delta}(p_i) = d(\mathcal{P}(p_i), \mathcal{P}(p_i^{l, \theta, \delta}))$ et $d_r^{\theta, \delta}(p_i) = d(\mathcal{P}(p_i), \mathcal{P}(p_i^{r, \theta, \delta}))$. Ensuite, nous pouvons agréger toutes ces distances le long d'une direction et obtenir $d_s^{\theta, 1}(p_i) = \{d_l^{\theta, 1}(p_i), d_c^{\theta, 1}(p_i), d_r^{\theta, 1}(p_i)\}$ puis prendre le maximum dans toutes les directions : $P_2^{AF}(p_i) =$

$\max_{\theta} AF(d_s^{\theta,1}(p_i))$ avec AF une fonction d'agrégation parmi les min, max et la moyenne.

3.3 Potentiel basé patches

L'inconvénient des potentiels précédents est qu'ils sont très localisés alors qu'il peut être avantageux d'étudier un voisinage plus vaste. En effet, s'il y a un contour au pixel p_i , les patches à gauche sont sûrement différents des patches à droite. Pour tenir compte de cela, nous généralisons le potentiel précédent P_2 pour opérer sur une ligne plus longue et considérer les pixels qui sont plus éloignés qu'une distance $\delta = 1$. Nous considérons une bande de patches ns sur un côté de la ligne et pour éviter un chevauchement entre le patch à p_i et ses patches gauche et droit, nous considérons les pixels à distance $\delta \in [\epsilon, \epsilon + ns]$ avec ϵ un paramètre. Pour chaque ensemble de distances possibles (gauche d_l , droite d_r ou centré d_c), nous retenons le minimum et agrégeons toutes ces distances sur une direction pour ensuite prendre le maximum dans toutes les directions :

$$P_3^{AF}(p_i) = \max_{\theta} AF(\{\min_{\delta \in [\epsilon, \epsilon + ns]} d_s^{\theta, \delta}(p_i), \forall s \in \{l, c, r\}\}).$$

3.4 Potentiel germe-dépendant

Tous les potentiels précédents que nous avons proposés peuvent être considérés comme aveugles car ils ne prennent pas en compte les germes préalablement positionnées pour calculer le potentiel. Pour surmonter ceci, nous concevons un potentiel germe-dépendant. Étant donné un germe p_s , on s'attend à ce que les régions $\mathcal{R}^+(p_s)$ et $\mathcal{R}^-(p_s)$ sur ses deux côtés soient similaires aux autres régions pour les points suivants du contour. Cependant, pour le premier germe $p_s = p_1$, nous n'avons aucune idée exacte de l'orientation pour définir les régions qu'il sépare. Pour y faire face, nous recherchons l'angle θ_s qui maximise les différences entre l'ensemble des patches des deux régions. Nous considérons des patches de taille 5×5 avec un chevauchement minimum pour chaque région. Cela donne pour p_1 :

$$\theta_1 = \arg \max_{\theta} \max_{\substack{q_1 \times q_2 \\ q_1 \in \mathcal{R}_{\theta}^+(p_1) \\ q_2 \in \mathcal{R}_{\theta}^-(p_1)}} d(\mathcal{P}(q_1), \mathcal{P}(q_2)) \quad (1)$$

Ceci fournit deux régions $\mathcal{R}_{\theta_1}^+(p_1)$ et $\mathcal{R}_{\theta_1}^-(p_1)$ selon l'orientation retenue θ_1 . Pour tout autre germe $p_s \neq p_1$, les régions $\mathcal{R}_{\theta_s}^+(p_s)$ et $\mathcal{R}_{\theta_s}^-(p_s)$ sont définies relativement la normale $N_s = (\gamma_{s-1,s}^{*f}(p_s))^{\perp} / \|\gamma_{s-1,s}^{*f}(p_s)\|$ au chemin le plus court $\gamma_{s-1,s}^{*f}$ par l'angle $\theta_s = \arg \min_{\theta} \mathcal{L}(N_s, p_s p_s^{\perp, \theta, 1})$. Ensuite, pour chaque point p_i de l'image, nous calculons toutes les distances par paires entre les régions $\mathcal{R}_{\theta_s}^+(p_s)$ et $\mathcal{R}_{\theta_s}^-(p_i)$ (et également pour \mathcal{R}^-) pour différentes orientations θ autour p_i . Nous ne conservons que la valeur minimale qui correspond à une configuration très similaire entre p_s et p_i . Ceci peut être formulé par :

$$P_4(p_i) = \min_{\theta} \left(\begin{array}{c} \min_{\substack{q_1 \times q_2 \\ q_1 \in \mathcal{R}_{\theta}^+(p_s) \\ q_2 \in \mathcal{R}_{\theta}^-(p_i)}} d(\mathcal{P}(q_1), \mathcal{P}(q_2)) + \min_{\substack{q_1 \times q_2 \\ q_1 \in \mathcal{R}_{\theta}^-(p_i) \\ q_2 \in \mathcal{R}_{\theta}^-(p_s)}} d(\mathcal{P}(q_1), \mathcal{P}(q_2)) \end{array} \right) \quad (2)$$

Contrairement aux autres potentiels, il doit être calculé à chaque itération de l'Algorithme 2.2.

4 Résultats

Comme présenté dans [12], pour évaluer de manière fiable l'efficacité des approches d'extraction de contour, il faut idéalement éliminer le biais de l'utilisateur, tout en obtenant des mesures de qualité du résultat. Pour cela on peut considérer des agents capables de simuler le comportement de l'utilisateur humain par l'ajout de germes près de la limite de l'objet. Pour ce faire, nous avons conçu notre propre agent utilisateur dédié au framework livewire. Partant d'un germe initial, nous imposons au contour extrait d'être localisé dans un tube \mathcal{T} de rayon r des deux cotés du contour de la vérité terrain. Pour démontrer la capacité des potentiels proposés, nous avons mené des expériences sur la base ISEG [2]. Nous avons lancé l'Algorithme 2.2 pour tous les potentiels avec l'agent simulant l'utilisateur. Pour avoir une comparaison juste entre les potentiels comparés, le premier point de départ est toujours le même sur une image donnée. Les critères utilisés pour comparer deux potentiels sont : i) la moyenne du rapport entre le nombre de germes et la longueur du contour vérité terrain (noté % de germes, moins il y a de germes mieux c'est), ii) l'indice de Jaccard qui mesure la similarité entre le contour extrait et la vérité de sol. Comme nous utilisons des potentiels basés patches, nous avons tout d'abord déterminé quel potentiel donne le pourcentage minimum de germes pour différentes tailles de patches. Ceci nous a permis également de constater que les potentiels basés patches (P_3) étaient toujours meilleurs que ceux basés pixels (P_2) et qu'une taille de patch de 5×5 était adaptée. Une fois cette taille de patch fixée, nous avons comparé le meilleur potentiel basé patch, à savoir P_3^{max} , avec le potentiel classique de [9] basée sur la magnitude du gradient. Quelle que soit la configuration de l'agent (différentes valeurs de r), le potentiel P_3^{max} est toujours meilleur que le potentiel de référence P_E reposant sur un gradient classique (voir [9]) et ce sur la base des deux critères (% de germes et indice de Jaccard). Il est important de noter qu'une différence de 0,1% sur la moyenne du % de germes correspond à 1,5 point). Cependant, si les résultats sont en faveur de P_3^{max} , parfois P_4 est plus performant. Puisqu'aucune règle générale ne peut être conçue pour choisir entre les potentiels P_3^{max} et P_4 , nous les combinons en $P_5 = \alpha P_3^{max} + (1 - \alpha) P_4$. Pour choisir le meilleur α , nous avons testé les valeurs dans $[0, 1]$ et retenu celle qui minimise le % de germes. Nous avons obtenu une valeur de $\alpha = 0.3$. Avec ce nouveau potentiel P_5 , on peut voir que nous dépassons de loin le potentiel pixel classique ainsi que tout potentiel patch que nous avons considéré. De plus, ceci est vrai quel que soit le rayon considéré r de l'agent utilisateur, ce qui renforce la robustesse de notre potentiel proposé P_5 , basé patch et dépendant des germes.

5 Conclusion

Dans cet article, de nouveaux potentiels basés sur des patches ont été proposés pour une extraction des contours. Leur intérêt a été évalué en fonction du nombre de points nécessaires pour

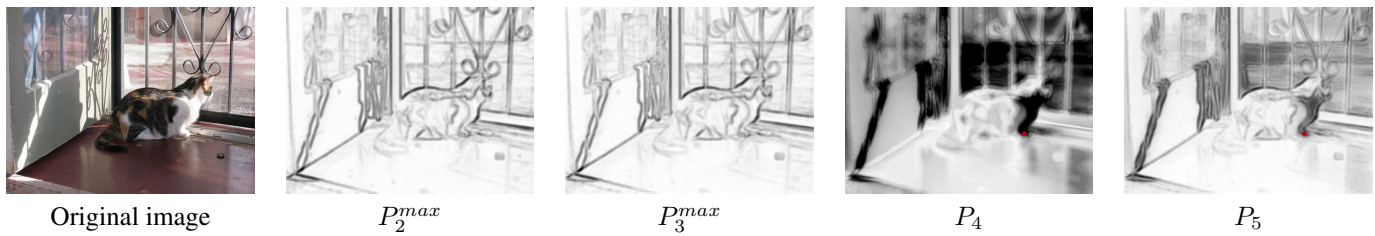


FIGURE 2 – Potentiels avec des patches 5×5 , $AF = \max$ pour P_2 et P_3 , $\epsilon = 2$ et $ns = 5$ pour P_3 , et $\alpha = 0.3$ pour P_5 . Pour P_4 , le point rouge indique le germe p_s .

Potentiel	Rayon	Moyenne du % de germes	Moyenne de l'indice de Jaccard
P_E	2	4.58	0.96
	3	2.96	0.95
	4	2.31	0.94
P_3^{max}	2	3.93	0.97
	3	2.67	0.95
	4	2.12	0.94
P_4	2	4.42	0.95
	3	3.13	0.93
	4	2.52	0.91
P_5	2	3.67	0.97
	3	2.38	0.95
	4	1.86	0.94

FIGURE 3 – Comparaison quantitative entre potentiels.

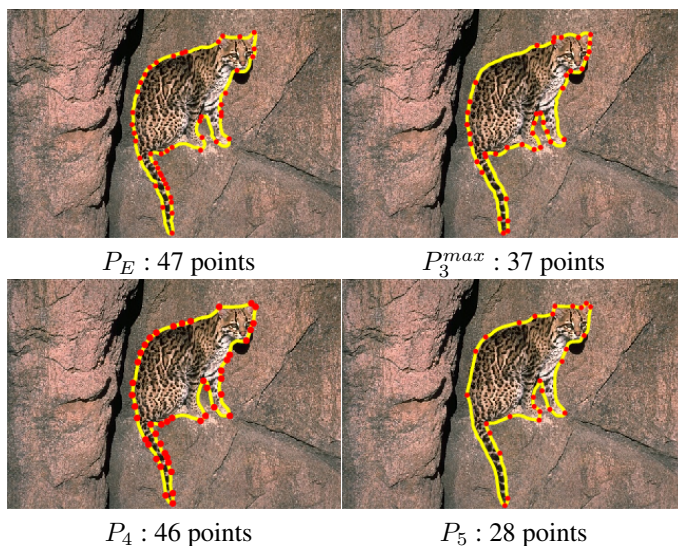


FIGURE 4 – Extraction de contours avec différents potentiels.

effectuer l'extraction du contour, de sorte que l'interaction de l'utilisateur est potentiellement minimisée. Afin d'éliminer le biais de l'utilisateur dans le processus interactif, un agent utilisateur dédié a été conçu. Les résultats ont montré que les potentiels basés sur les patches peuvent être plus efficaces que les potentiels classiques basés sur les pixels et améliorer l'extraction du contour de l'objet.

Références

[1] D. Liu, Y. Xiong, L. Shapiro, and K. Pulli, "Robust interactive image segmentation with automatic boundary refi-

nement," in *ICIP*, 2010, pp. 225–228.

- [2] V. Gulshan, C. Rother, A. Criminisi, A. Blake, and A. Zisserman, "Geodesic star convexity for interactive image segmentation," in *CVPR*, 2010, pp. 3129–3136.
- [3] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut : Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, 2004.
- [4] L. D. Cohen, "On active contour models and balloons," *CVGIP-IMAG UNDERSTAN*, vol. 53, no. 2, pp. 211–218, 1991.
- [5] L. D. Cohen and R. Kimmel, "Global minimum for active contour models : A minimal path approach," *IJCV*, vol. 24, no. 1, pp. 57–78, 1997.
- [6] J. A. Sethian, "A fast marching level set method for monotonically advancing fronts," *Proceedings of the National Academy of Sciences*, vol. 93, no. 4, pp. 1591–1595, 1996.
- [7] G. Peyré, M. Pechaud, R. Keriven, and L. Cohen, "Geodesic methods in computer vision and graphics," *Foundations and Trends in Computer Graphics and Vision*, vol. 5, no. 3-4, pp. 197–397, 2010.
- [8] D. Chen, J.-M. Mirebeau, and L. D. Cohen, "Global minimum for curvature penalized minimal path method," in *BMVC*, 2015, pp. 86–86.
- [9] J. Mille, S. Bougleux, and L. D. Cohen, "Combination of piecewise-geodesic paths for interactive segmentation," *IJCV*, vol. 112, no. 1, pp. 1–22, 2015.
- [10] A. X. Falcão, J. K. Udupa, and F. K. Miyazawa, "An ultra-fast user-steered image segmentation paradigm : live wire on the fly," *IEEE T MED IMAGING*, vol. 19, no. 1, pp. 55–62, 2000.
- [11] P. A. V. Miranda, A. X. Falcão, and T. V. Spina, "Riverbed : A novel user-steered image segmentation method based on optimum boundary tracking," *IEEE T IMAGE PROCESS*, vol. 21, no. 6, pp. 3042–3052, 2012.
- [12] T. V. Spina and A. X. Falcão, "Robot users for the evaluation of boundary-tracking approaches in interactive image segmentation," in *ICIP*, 2014, pp. 3248–3252.