

TP7 - Segmentation d'images par contours actifs implicites (*level set*)

Vincent Barra - Christophe Tilmant

2010-2011

1 Partie théorique

1.1 Introduction

Les **contours actifs** sont une méthode relevant de l'analyse d'image qui permet l'extraction de primitives sous contraintes. Il s'agit d'objets de nature mathématique auxquels est conféré un comportement mécanique réaliste par la minimisation d'une énergie potentielle calculée à partir de l'image. La conversion en énergie cinétique amène le contour à converger vers les structures d'intérêt. Cette méthode a pour avantage, entre autre, de contrôler la forme de la solution, qui peut être contrainte à être fermée par la représentation de contour adoptée.



FIGURE 1 – Principe d'un contour actif. Courbe mise en mouvement par minimisation d'énergie afin de converger vers l'objet d'intérêt.

La suite de cette étude s'intéresse à la représentation dite implicite des contours actifs. Par opposition aux représentations explicites (approximation polygonale, courbes paramétrées, décomposition des coordonnées sur une base de fonctions), la représentation implicite, encore appelée représentation par courbe de niveaux ou représentation *level set*, présente moins de contraintes quant à la géométrie des objets qui peuvent être irréguliers (présence de fortes courbures) et dont la topologie peut changer en cours de convergence.

D'un point de vue chronologique, le modèle princeps des contours actifs a été proposé par Kass et al. [KASS-88] et consiste en un modèle déformable intrinséquement fermé qui évolue vers les frontières de la région désirée. La déformation est fondée sur une formulation Lagrangienne de minimisation de l'énergie, exprimée comme la somme d'un terme d'attache aux données de l'image et d'un terme de régularisation. Cette méthode ne permet pas de changements topologiques simples, et l'énergie dépend de la paramétrisation de la courbe.

Caselles et al. [CASE-93] ont introduit un modèle géométrique des contours actifs formulé par des équations aux dérivées partielles (EDPs) d'évolution de courbes. Chaque point de la courbe se déplace dans la direction de son vecteur normal à une vitesse proportionnelle à la courbure. Cette approche est une alternative à la minimisation d'énergie qui peut être vue comme la recherche de la solution d'une équation de Hamilton-Jacobi, et être ainsi résolue de façon efficace par l'utilisation de la méthode des courbes de niveaux d'Osher et de Sethian [OSHE-88]. L'utilisation de cette représentation implicite

explique tout le succès de cette approche.

Caselles et al. [CASE-97] ont, par la suite, proposé un problème équivalent de minimisation d'énergie, fondé sur la recherche d'un chemin géodésique minimal dans un espace de Riemann. Ce problème, qui peut être également résolu numériquement par *level set*, permet en théorie de résoudre le problème d'initialisation des contours actifs. Les résultats théoriques ont montré l'existence, l'unicité et la stabilité de la solution.

De nombreux autres modèles dérivant des approches géométriques et géodésiques ont été proposés. Au lieu de présenter un catalogue des variantes décrites dans la littérature, le paragraphe suivant les unifie dans un même formalisme de minimisation d'énergie. Le lien entre décroissance de l'énergie et champ de vitesse engendrant un déplacement est également présenté, afin d'avoir une approche à la fois énergétique et géométrique.

L'utilisation d'une représentation implicite possède pour principal inconvénient un temps de calcul prohibitif.

1.2 Principe général

Considérons une image, notée I , et définie sur le domaine Ω ,

$$\begin{aligned} I : \quad \Omega \subset \mathbb{R}^2 &\rightarrow [0, 1] \\ \mathbf{m} = (x, y)^T &\mapsto I(x, y) \end{aligned}$$

Cette image peut être appréhendée comme une partition de N régions R_i homogènes par rapport à des caractéristiques données (contraste, texture,...) :

$$\begin{cases} \forall i \in \{1, 2, \dots, N\} : R_i \subset \Omega \\ R_1 \cup R_2 \dots \cup R_N = \Omega \\ \forall (i, j) \in \{1, 2, \dots, N\} \times \{1, 2, \dots, N\} / i \neq j : R_i \cap R_j = \emptyset \end{cases}$$

Notons $\gamma(\mathbf{p})$ une courbe fermée définie sur Ω et paramétrée arbitrairement par un vecteur, noté \mathbf{p} . La région délimitée par ce contour est notée R_γ . La méthode des contours actifs en deux dimensions se fonde sur un modèle déformable en considérant la courbe $\gamma(\mathbf{p})$ dynamique, c'est-à-dire $\gamma(\mathbf{p}) = \gamma(\mathbf{p}, t)$. Ce modèle doit évoluer vers les frontières de la région recherchée R_i . Le problème est donc de trouver le vecteur \mathbf{p} qui vérifie $R_\gamma = R_i$. Afin de contraindre la progression du modèle, une énergie E lui est associée. Elle est composée de la somme d'un terme d'attache aux données et d'un terme de régularisation. Le système minimise cette énergie en la convertissant en énergie cinétique jusqu'à la vérification d'un critère d'arrêt. Ce problème d'énergie peut être formulé par un équilibre de force, correspondant à un problème numérique de descente de gradient :

$$\frac{\partial \gamma(\mathbf{p}, t)}{\partial t} = \mathbf{F} \text{ avec } \begin{cases} \gamma(\mathbf{p}, t=0) = \gamma_0(\mathbf{p}) \\ \mathbf{F} = \frac{\delta E}{\delta \gamma} \end{cases}$$

Dans ce modèle, trois degrés de liberté sont laissés au choix de l'utilisateur :

- la représentation de la courbe, qui peut être explicite, paramétrique ou implicite ;
- la définition de la fonction d'énergie, qui se fonde sur des contraintes géométriques et/ou des données extraites de l'image (gradient, paramètres de texture, information temporelle, ...) ;
- l'initialisation de la méthode qui influe sur la solution trouvée du fait de la non-convexité éventuelle de la fonction d'énergie.

Dans la majorité des cas, l'énergie attachée au modèle peut être décrite comme une combinaison de fonctionnelles intégrales curviligne et surfacique. L'EDP d'évolution se déduit mathématiquement par calcul variationnel.

1.2.1 Formulation énergétique - Cas général

1. énergie intégrale curviligne

Cette énergie de contour peut être exprimée comme une intégrale simple le long du contour d'une fonction f dépendant des caractéristiques de l'image :

$$E_b = \int_{\gamma(\mathbf{p})} f(\mathbf{m}) \, d\mathbf{a}(\mathbf{m})$$

Dans [CASE-97] les auteurs ont démontré que cette forme d'énergie conduit au champ de vitesse :

$$\mathbf{F}_b = [\nabla \cdot (f(\mathbf{m})\mathbf{n})] \mathbf{n} = [f(\mathbf{m})\kappa - \nabla f(\mathbf{m}) \cdot \mathbf{n}] \mathbf{n}$$

où \mathbf{n} est le vecteur normal unitaire, κ la courbure euclidienne et ∇ l'opérateur de gradient.

2. énergie intégrale surfacique

Une énergie régionale peut s'exprimer comme une intégrale double sur un sous-domaine de Ω d'une fonction g des caractéristiques de l'image :

$$E_r = \iint_{R_\gamma} g(\mathbf{m}, R_\gamma) \, d\mathbf{m}$$

Dans la littérature, le cas particulier d'indépendance vis à vis de la région considérée, c'est-à-dire $g(\mathbf{m}, R_\gamma) = g(\mathbf{m})$, est largement utilisé :

$$\mathbf{F}_r = g(\mathbf{m}) \mathbf{n}$$

1.2.2 Formulation énergétique - Utilisation en imagerie

Une grande diversité de termes d'énergie peut être mise en place à partir de la formulation en fonction intégrales curviligne ou surfacique. Ils peuvent être classés en deux catégories : les termes d'énergie interne, provoquant cohérence et régularisation du contour et les termes d'énergie d'attache aux données, qui permettent d'attirer le contour vers les structures d'intérêt.

1. énergies de régularisation

- Lorsque $f(\mathbf{m}) = 1$, l'énergie est équivalente à la longueur euclidienne de la courbe γ . Ce cas est très largement utilisé pour ses propriétés de régularisation et mène au champ :

$$\mathbf{F}_b = \kappa \mathbf{n}$$

Dans ce cas on retrouve l'équation de la chaleur qui est bien connue pour ses propriétés de lissage géométrique. Les points singuliers présentant des valeurs de courbure élevées rattrapent les autres points de la courbe.

- Lorsque $g(\mathbf{m}) = \mu$, l'énergie est proportionnelle à la surface du domaine. Cette énergie correspond à une vitesse largement utilisée et correspondant à une force surfacique homogène à une pression appelée force ballon [COHE-91], qui contraint le modèle à un mouvement de contraction ou dilatation globale. Dans ce cas :

$$\mathbf{F}_r = \mu \mathbf{n}$$

Chaque point du contour subit une force constante dans la direction de la normal au contour, qui se ramène à appliquer un mouvement homothétique à la forme. Cette énergie est couramment utilisée pour pousser la courbe vers la vallée de potentiel.

2. énergies d'attaches aux données

L'information issue de l'image peut découler de données extraites le long du contour γ (intégrale curviligne) ou à l'intérieur de la région R_γ qu'il délimite (intégrale surfacique). Les énergies associées sont qualifiées d'énergie "contour" ou d'énergie "région" respectivement.

Pour le terme contour, la fonction f usuellement utilisée dans l'intégrale curviligne est appelée fonction d'arrêt. Elle dépend d'un opérateur de type dérivatif, par exemple la norme du gradient, qui présente un maximum de réponse sur les sauts d'amplitude dans l'image ; elle décroît lorsque l'opérateur croît et tend théoriquement vers zéro quand il tend vers l'infini. La forme suivante est couramment utilisée en pratique :

$$f(\mathbf{m}) = \frac{1}{1 + \left| \nabla \hat{I}(\mathbf{m}) \right|^p}$$

où \hat{I} est une version régularisée de I (par exemple en appliquant un noyau Gaussien d'écart-type σ) et la puissance p est généralement égale à un ou deux. Caselles et al. [CASE-97] ont nommé le modèle utilisant cette énergie **contour actif géodésique**.

1.3 Représentation implicite d'un contour

Les représentations mathématiques d'un contour peuvent être divisées en trois classes :

- représentation explicite (approximation polygonale, ...),
- représentation paramétrique (descripteurs de Fourier, B-spline, ...),
- représentation implicite (*level set*).

La méthode des *level set* est un cadre de travail analytique travaillant sur l'évolution géométrique d'objets. Au lieu d'employer une représentation Lagrangienne classique pour décrire les géométries, la méthode des *level set* les décrit à travers une fonction scalaire ψ définie sur une grille fixe. L'équation d'évolution, formalisée sous forme d'une EDP, est contrainte par un champ de vitesse imposé dans le sens de la normale au contour. Ce champ est construit de manière à attirer le modèle vers les objets à extraire dans l'image sous contraintes de régularisation géométrique. La normale et la courbure du contour en chaque point sont facilement définies à partir des propriétés différentielles géométriques de cette représentation.

Un des principaux avantages de la description implicite d'un contour est qu'elle gère intrinsèquement les changements topologiques en cours de convergence. Par exemple, si un processus de segmentation est initialisé par des germes multiples, les collisions et fusions des composantes connexes sont gérées sans modification de l'algorithme, ce qui n'est pas le cas par exemple avec des contours actifs classiques. La contrepartie est le coût de calcul important dans sa formulation basique, qui nécessite obligatoirement le recours à des algorithmes améliorés [SETH-99].

Dans le formalisme des *level set*, la courbe γ n'est pas paramétrée, mais implicitement définie par le biais d'une fonction de dimension supérieure :

$$\begin{aligned} \psi : \Omega \subset \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \psi(x, y) \end{aligned}$$

Le front 2D à un instant t est alors défini par la ligne de niveau de valeur zéro de la fonction ψ à cet instant :

$$\gamma(t) = \psi^{-1}(0)(t)$$

Le principe d'évolution des contours sous cette représentation a été introduit par Osher et Sethian dans [OSHE-88]. Généralement, la fonction *level set* initiale ψ_0 est obtenue par l'utilisation de la distance signée à la courbe initiale γ_0 [GOME-00] :

$$\psi_0(x, y) = \begin{cases} +d((x, y), \gamma_0) & \text{if } (x, y) \in R_{\gamma_0} \\ -d((x, y), \gamma_0) & \text{if } (x, y) \notin R_{\gamma_0} \end{cases}$$

où $d(\bullet)$ représente la distance euclidienne, et (x, y) un point de Ω .

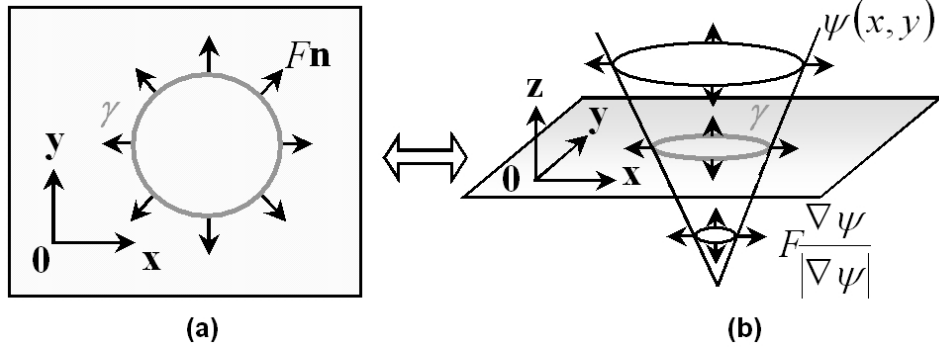


FIGURE 2 – Représentation implicite d'une courbe. (a) Courbe dans un espace à deux dimensions avec définition de la normale extérieure; (b) représentation implicite de la courbe par la ligne de niveau zéro d'une fonctionnelle à deux variables.

Sous ce formalisme, la normale au contour \mathbf{n} et la courbure κ sont accessibles grâce à des opérateurs différentiels appliqués à la carte :

$$\mathbf{n} = \frac{\nabla \psi}{|\nabla \psi|} \text{ et } \kappa = \nabla \cdot \frac{\nabla \psi}{|\nabla \psi|}$$

1.4 équation d'évolution

Il a été montré dans [SETH-99] que si $\gamma(t)$ évolue suivant l'équation d'évolution vue précédemment alors la représentation implicite (c'est-à-dire la fonction ψ) respecte l'équation :

$$\frac{\partial \psi((x, y), t)}{\partial t} = F |\nabla \psi((x, y), t)|$$

où F est la composante normale du champ de vitesse \mathbf{F} : $F = \mathbf{F} \cdot \mathbf{n}$. Le problème d'évolution se pose pour la fonction ψ , puisqu'il est possible de déduire l'évolution de $\gamma(t)$ à partir de celle de $\psi(t)$. Avec cette représentation, il faut noter que seule la composante normale au contour intervient dans la propagation du modèle. Par la suite, il sera fait mention indifféremment de champ de vitesse ou de valeur de vitesse en un point.

Afin d'extraire la région souhaitée R_i de l'image, il est nécessaire de construire un champ de vitesse F adapté à partir de diverses sources. A titre d'exemple, en formulation *level set*, le modèle géodésique de Caselles et al. respecte l'équation :

$$\frac{\partial \psi(t)}{\partial t} = f(I) \cdot |\nabla \psi| \cdot \kappa + \nabla f(I) \cdot \nabla \psi$$

1.5 Discrétisation de l'équation d'évolution

Le coeur de la méthode est un schéma itératif guidé par la fonction de vitesse F , pour lequel un critère d'arrêt doit être défini. Le problème de propagation de front est résolu numériquement en discrétisant les domaines en espace et en temps :

$$\begin{aligned} \mathbf{m} &= (i\Delta x, j\Delta y) \\ t &= n\Delta t \end{aligned}$$

où Δx et Δy sont les pas spatiaux, et Δt le pas temporel. Les valeurs de la fonction ψ et de la vitesse F au point \mathbf{m} et au temps t , sont notées respectivement ψ_{ij}^n et F_{ij}^n . La fonction ψ est initialisée à la distance signée. En utilisant la méthode des différences finies, l'équation d'évolution devient :

$$\psi_{ij}^{n+1} = \psi_{ij}^n + \Delta t F_{ij}^n |\nabla \psi_{ij}^n|$$

Un schéma numérique convexe du premier ordre complet peut être trouvé dans [SETH-99] pour l'estimation des dérivées spatiales de l'équation précédente lorsque la fonction vitesse peut être écrite comme :

$$F = F_{prop} + F_{curv} + F_{adv} = F_0 - \varepsilon\kappa + \mathbf{U}(\mathbf{m}, t) \cdot \mathbf{n}$$

où $F_{prop} = F_0$ est une vitesse de propagation, $F_{curv} = -\varepsilon\kappa$ un terme de vitesse dépendant de la courbure et $F_{adv} = \mathbf{U}(\mathbf{m}, t) \cdot \mathbf{n}$ un terme d'advection avec $\mathbf{U}(\mathbf{m}, t) = (u(\mathbf{m}, t), v(\mathbf{m}, t))^T$. Le calcul numérique des termes de vitesse contour et région utilise ce schéma car il est possible de rapprocher les deux membres de la vitesse géodésique de F_{curv} et F_{adv} et le champ de vitesse région est lui analogue à F_{prop} :

$$\psi_{ij}^{n+1} = \psi_{ij}^n + \Delta t \left[\begin{array}{c} - \left[\max(F_{0ij}, 0) \nabla^+ (\psi_{ij}^n) + \min(F_{0ij}, 0) \nabla^- (\psi_{ij}^n) \right] \\ + \left[\varepsilon\kappa_{ij}^n \sqrt{D^{0x} (\psi_{ij}^n)^2 + D^{0y} (\psi_{ij}^n)^2} \right] \\ - \left\{ \begin{array}{l} \max(u_{ij}^n, 0) D^{-x} (\psi_{ij}^n) + \min(u_{ij}^n, 0) D^{+x} (\psi_{ij}^n) + \\ \max(v_{ij}^n, 0) D^{-y} (\psi_{ij}^n) + \min(v_{ij}^n, 0) D^{+y} (\psi_{ij}^n) \end{array} \right\} \end{array} \right]$$

où $\min(\bullet, \bullet)$ et $\max(\bullet, \bullet)$ correspondent aux opérateurs minimum et maximum, $D^{+\beta}(\bullet)$, $D^{-\beta}(\bullet)$ et $D^{0\beta}(\bullet)$ sont respectivement les dérivées par différences finies à droite, à gauche et centrées par rapport à la variable β . Finalement, $\nabla^+(\bullet)$ et $\nabla^-(\bullet)$ sont les opérateurs dérivatifs suivants :

$$\nabla^+(\bullet) = \sqrt{\max(D^{-x}(\bullet), 0)^2 + \min(D^{+x}(\bullet), 0)^2 + \max(D^{-y}(\bullet), 0)^2 + \min(D^{+y}(\bullet), 0)^2}$$

$$\nabla^-(\bullet) = \sqrt{\max(D^{+x}(\bullet), 0)^2 + \min(D^{-x}(\bullet), 0)^2 + \max(D^{+y}(\bullet), 0)^2 + \min(D^{-y}(\bullet), 0)^2}$$

Il existe également des schémas d'ordre supérieurs s'appliquant à des fonctionnelles convexes ou non pour résoudre itérativement le problème. Cependant, l'application de segmentation d'image ne nécessitant pas une grande précision dans la résolution de la propagation du contour, ces schémas ne s'avèrent pas nécessaires.

2 Partie pratique

Vous trouverez sur <http://ent.univ-bpclermont.fr/> dans la rubrique "Ressources Pédagogiques" ce sujet, ainsi qu'un squelette du TP.

- Il est demandé de mettre en œuvre un contour actif implicite géodésique. Vous utiliserez pour ce faire une fonction d'arrêt comme celle présentée dans la partie théorique.
- Vous testerez votre algorithme sur des images que vous bruiterez.

3 Références

- [KASS-88] Kass M., Witkin A., Terzopoulos D., "Snakes : Active contour models", International Journal of Computer Vision, vol. 1(4), pp. 321-31, 1988.
- [OSHE-88] Osher S., Sethian J.A., "Fronts propagating with curvature-dependant speed : Algorithms based on Hamilton-Jacobi formulations", Journal of Computational Physics, vol. 79(1), pp. 12-49, 1988.
- [COHE-91] Cohen L.D., "On active contour models and balloons", Computer Vision, Graphics, and Image Processing : Image Understanding, vol. 53(2), pp. 211-8, 1991.
- [CASE-93] Caselles V., Catte F., Coll T., Dibos F., "A geometric model for active contours", Numerische Mathematik, vol. 66(1), pp. 1-31, 1993.
- [CASE-97] Caselles V., Kimmel R., Sapiro G., "Geodesic active contours", International Journal of Computer Vision, vol. 22(1), pp. 61-79, 1997.
- [SETH-99] Sethian J.A., "Level set methods; Evolving interfaces in geometry, fluid mechanics, computer vision and material science," Cambridge : Cambridge University Press, second edition, 1999.
- [GOME-00] Gomes J., Faugeras O.D., "Reconciling distance functions and level sets", Journal of Visual Communication and Image Representation, vol. 11(2), pp. 209-23, 2000.