Segmentation d'Energie Minimale

Laurent D. COHEN
CEREMADE, Université Paris - Dauphine
Place du Maréchal de Lattre de Tassigny
75775 Paris cedex 16
Tel. 01 44 05 46 78 Fax 01 44 05 45 99
Cohen@ceremade.dauphine.fr
http://www.ceremade.dauphine.fr/∼cohen
(mes principales publications sont disponibles sur le web)
Septembre 1999

Un module intensif de DEA sur les modèles déformables aura lieu à l'ENS Cachan vers Janvier 2000. Plusieurs stages, eventuellement poursuivis par une thèse, seront possibles en Traitement d'images, au CEREMADE, à l'INRIA ou en entreprise. Contactez moi par courrier électronique ou envoyer moi un CV.

1 Reconstruction avec Régularisation

Nos travaux traitent de l'utilisation des modèles déformables et modèles paramétriques pour le traitement d'images, en particulier pour le problème de segmentation et reconstruction avec régularisation. Le problème général de la reconstruction d'une courbe ou surface u à partir de données d présegmentées se formalise par la minimisation par rapport à u de:

$$E(u,d) = \int R(u(s))ds + \int V(u(s),d(s))ds \tag{1}$$

où

- $R(u)$ mesure le lissage de la reconstruction u
  (courbe ou surface). On impose la régularité de u
  - soit à travers le terme $R(u)$ composé alors de dérivées de u dans le cas de courbe ou surface libre, il s'agit des modèles déformables les plus généraux;
  - soit en restreignant u à un ensemble de forme d'un certain type, $R(u)$ est alors vide, c'est le cas de modèles paramétriques (Bsplines, superquadriques, hyperquadriques,...).

- V mesure la fidélité de u à la donnée d. Il peut être
  - explicite: $\int V(u,d) = \int \|u(s) - d(s)\|^2$
  - implicite: $\int V(u,d) = \int P(u)$ où P est un Potentiel d'attraction par les données d.

Cette formulation s'applique à de nombreux domaines du traitement d'images incluant extraction de contours, segmentation, amélioration, restoration et mise en correspondance de signaux [Coh96, Coh91, Coh95, PVY88, Gri81, Ter88, MS85, BZ87, KWT88, TWK88, Wei90, GY93, GY91].

Nous nous interessons ici à la segmentation d'images par minimisation d'énergie.

2 Energie de Mumford et Shah

On définit une segmentation comme une fonction régulière par morceaux. On cherche à minimiser une fonctionnelle dépendant de la différence entre l'image originale donnée d et l'image lissée reconstruite à déterminer u et de la forme des bords des régions B. La fonction u est régulière partout sauf aux points de B où elle peut être discontinue. On définit donc une énergie à minimiser du type :

$$E(u,B) = \int_R (d-u)^2 + \lambda^2 \int_{R-B} \|\nabla u\|^2 + \alpha \, l(B) \tag{2}$$

R représente le domaine de l'image. λ désigne un facteur d'échelle et α un facteur de détection de contraste. $l(B)$ est la longueur de l'ensemble des bords.
On minimise simultanément trois termes qui traduisent des propriéés différentes rechercheées:

- Le premier terme traduit la fidélité aux données. u doit rester proche de d.
- Le second est un terme régularisant. Cela modélise une membrane élastique qui est posée sur les données.
- Le dernier terme à un effet régularisant sur les lignes de contours. En minimisant leur longueur, ces lignes se comportent elles aussi comme des sortes d'élastiques.

Le troisième terme est l'apport essentiel de la méthode. En effet, sans lui, u serait une simple régularisation des données. Un lissage simple permet de supprimer le bruit mais émousse aussi les contours des objets. On introduit donc un ensemble B de points en lesquels u peut admettre une discontinuité. Lorsqu'un point appartient à B, on paye une pénalité dûe au dernier terme. Un point de l'image contribue soit au second terme soit au dernier. Sa contribution la plus faible pour l'énergie determine si on introduit ou pas une discontinuité en ce point. Cela sera plus clair lors de la discrétisation du problème dans la section suivante.

Une version simplifiée de ce problème a été étudiée de manière théorique dans [MS85] en dimension 1 pour la première fois, puis dans [MS88]. Elle concerne l'énergie:

$$E_1(u, B) = \int_R (d - u)^2 + \alpha\, l(B)$$

où la fonction u recherchée est supposée constante à l'intérieur de chaque région.

Lorsque B est fixée la fonction qui minimise E est simple à déterminer. Elle est égale sur chaque région A à la moyenne de d sur A.

En associant à chaque frontière B la fonction u correspondante il suffit de minimiser E par rapport à B. On essaie de minimiser cette énergie par une méthode de croissance de régions. Partant d'une segmentation très fine (chaque pixel est une région), on fusionne deux régions adjacentes dès qu'il y a une diminution de l'énergie après fusion.

On obtient ainsi un critère d'agglomération de 2 régions. Cette méthode de croissance de régions peut englober la plupart des autres méthodes par la définition d'une énergie adaptée au critère proposé.

3 Algorithme de Blake et Zisserman

On minimise $E(u, B)$ en remarquant que l'energie discrète s'écrit:

$$E_d(u, B) = \sum_{i,j=1}^{N} (d_{ij} - u_{ij})^2 + \lambda^2 \sum_{(i,j)\notin B} G_{ij}^2 + \alpha \sum_{(i,j)\in B} 1 \tag{3}$$

u est ici une image formée des pixels de valeur u_{ij} et B est l'ensemble des pixels pour lesquels on introduit une discontinuité. G_{ij} est une estimation du gradient de u au pixel (i, j). On peut prendre par exemple: $G_{ij} = \sqrt{(u_{ij} - u_{i-1,j})^2 + (u_{ij} - u_{i,j-1})^2}$.
Posons

$$l_{ij} = \begin{cases} 1 & \text{si } (i,j) \in B \\ 0 & \text{si } (i,j) \notin B \end{cases}$$

L'énergie peut s'écrire alors comme fonction de u et de l:

$$E_d(u, l) = \sum_{i,j=1}^{N} (d_{ij} - u_{ij})^2 + \lambda^2 \sum_{i,j=1}^{N} G_{ij}^2(1 - l_{ij}) + \alpha \sum_{i,j=1}^{N} l_{ij} \tag{4}$$

On élimine maintenant la variable binaire l en remarquant que sa minimisation est simple en chaque pixel. Tout étant fixe par ailleurs, la valeur de l_{ij} qui minimise l'énergie est

$$l_{ij} = \begin{cases} 1 & \text{si } \alpha \leq \lambda^2 G_{ij}^2 \\ 0 & \text{si } \alpha > \lambda^2 G_{ij}^2 \end{cases}$$

C'est à dire qu'on introduit une discontinuité au pixel (i, j) dès que le gradient en ce point est assez grand. On définit donc l'ensemble B à un moment donné comme l'ensemble des points ayant un gradient G supérieur au seuil $\frac{\sqrt{\alpha}}{\lambda}$.

On détermine maintenant u en minimisant l'énergie modifiée:

$$E_g(u) = \sum_{i,j=1}^{N} (d_{ij} - u_{ij})^2 + \sum_{i,j=1}^{N} g_{\alpha,\lambda}(G_{ij}) \tag{5}$$

où

$$g_{\alpha,\lambda}(G) = \begin{cases} \lambda^2 G^2 & \text{si } \lambda^2 G^2 \leq \alpha \\ \alpha & \text{sinon} \end{cases} \tag{6}$$

Ainsi E_g s'exprime comme fonction de u uniquement et on a eliminé les autres variables. On minimise E_g par descente de gradient prenant $u = d$ initialement puis

$$u^{n+1} = u^n + w^n \nabla E(u^n)$$

En réalité, pour simplifier les calculs de $\nabla E(u)$, on sépare le gradient de l'image G_{ij} en ses deux composantes en i et en j:

$$E_g(u) = \sum_{i,j=1}^{N} (d_{ij} - u_{ij})^2 + \sum_{i,j=1}^{N} g_{\alpha,\lambda}(u_{ij} - u_{i-1,j}) + \sum_{i,j=1}^{N} g_{\alpha,\lambda}(u_{ij} - u_{i,j-1}) \quad (7)$$

Cela signifie que l'on a deux ensembles de discontinuités B_i et B_j en i et en j séparément. Cette approximation simplifie beaucoup la formulation du problème et son implémentation. Cependant, elle favorise les contours horizontaux et verticaux par rapport aux autres.

Cependant l'énergie n'étant pas convexe, on peut converger vers un minimum local. Pour éviter cela, la fonction $g_{\alpha,\lambda}$ est approchée par une fonction $g^*_{\alpha,\lambda}$ qui assure la convexité de E_g. Puis on fait varier continuement g entre $g^*_{\alpha,\lambda}$ et $g_{\alpha,\lambda}$ en utilisant une suite de fonctions $g^{(p)}_{\alpha,\lambda}$ pour p dans l'intervalle $[0,1]$. C'est l'idée de base de la "non-convexité graduelle" (*graduated non convexity- GNC*) introduite par Blake et Zisserman (tous les détails se trouvent dans [BZ86, BZ87]).

$$E^{(p)}_g(u) = \sum_{i,j=1}^{N} (d_{ij} - u_{ij})^2 + \sum_{i,j=1}^{N} g^{(p)}_{\alpha,\lambda}(u_{ij} - u_{i-1,j}) + \sum_{i,j=1}^{N} g^{(p)}_{\alpha,\lambda}(u_{ij} - u_{i,j-1}) \quad (8)$$

avec

$$g^{(p)}_{\alpha,\lambda}(G) = \begin{cases} \lambda^2 G^2 & \text{si } |G| < q \\ \alpha - c(|G| - r)^2/2 & \text{si } q \le |G| < r \\ \alpha & \text{si } r \le |G| \end{cases} \quad (9)$$

avec $c = \frac{c^*}{p}$, $r^2 = \alpha(\frac{2}{c} + \frac{1}{\lambda^2})q = \frac{\alpha}{\lambda^2 r}$

Au départ $p = 1$ et $c = c^*$, et $g^{(1)}_{\alpha,\lambda} = g^*_{\alpha,\lambda}$ est convexe. c^* est la plus grande valeur de c qui assure la convexité de E_g.

Puis p décroît de 1 à 0, et $g^{(p)}_{\alpha,\lambda}$ varie de $g^*_{\alpha,\lambda}$ à $g_{\alpha,\lambda}$.

Pour chaque valeur de p, la descente de gradient se fait successivement en chaque pixel utilisant un schéma SOR (Successive Over Relaxation). C'est à dire que pour traiter un pixel (i,j) lors d'une itération, on utilise les nouvelles valeurs obtenues pour les pixels voisins déjà traités:

$$\begin{aligned} u^{n+1}_{ij} = u^n_{ij} - \; & \tfrac{w}{(2+8\lambda^2)}\{2(u^n_{ij} - d_{ij}) + g^{(p)}_{\alpha,\lambda}{}'(u^n_{ij} - u^{n+1}_{i-1,j}) \\ & + g^{(p)}_{\alpha,\lambda}{}'(u^n_{ij} - u^{n+1}_{i,j-1}) + g^{(p)}_{\alpha,\lambda}{}'(u^n_{ij} - u^n_{i+1,j}) \\ & + g^{(p)}_{\alpha,\lambda}{}'(u^n_{ij} - u^n_{i,j+1})\} \end{aligned} \quad (10)$$

La dérivée $g_{\alpha,\lambda}^{(p)}\,'$ est linéaire par morceaux. La détermination du meilleurs coefficient w est donnée dans [BZ87, OR70, PFTV90].
On prend $w = 2(1 - \frac{1}{\sqrt{2}\lambda})$.

Le coefficient λ peut s'interpréter comme un paramètre déchelle. Pour $\lambda = 1$ il y a peu de changement, alors que lorsque λ augmente, les petits détails disparaissent de plus en plus et on obtient les résultats les plus propres mais ils sont peu précis. Le coefficient h s'interprete comme un parametre de contraste définissant la limite pour laquelle une discontinuité est introduite.

References

[BZ86] Andrew Blake and Andrew Zisserman. Some properties of weak continuity constraints and the GNC algorithm. In *Proc. 1986 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 656–661, Miami, 1986.

[BZ87] Andrew Blake and Andrew Zisserman. *Visual Reconstruction*. The MIT Press, 1987.

[Coh91] Laurent D. Cohen. On active contour models and balloons. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 53(2):211–218, March 1991.

[Coh95] Laurent D. Cohen. *Méthodes Variationnelles pour le Traitement d'images*. Université Paris Dauphine, 1995. Mémoire d'Habilitation à diriger des recherches. Accompagné des 10 publications les plus significatives des années 1988-1995.

[Coh96] Laurent D. Cohen. Auxiliary variables and two-step iterative algorithms in computer vision problems. *Journal of Mathematical Imaging and Vision*, 6(1):61–86, January 1996. See also ICCV'95.

[Gri81] W.E.L. Grimson. *From Images to Surfaces: A computational study of the Human Early vision system*. The MIT Press, 1981.

[GY91] D. Geiger and A.L. Yuille. A common framework for image segmentation. *International Journal of Computer Vision*, 6(3):227–234, August 1991.

[GY93] D. Geman and C. Yang. Nonlinear image recovery with half-quadratic regularization. Technical report, Umass, July 1993. to appear in *IEEE Transactions on Image Processing.*

[KWT88] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, January 1988.

[MS85] D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *Proceedings of Computer Vision and Pattern Recognition*, pages 22–26, San Francisco, June 1985.

[MS88] Jean-Michel Morel and Sergio Solimini. Segmentation d'images par méthode variationnelle: une preuve constructive d'existence. *Comptes Rendus de l'Académie des Sciences*, 1988.

[OR70] J. Ortega and W. Rheinboldt. *Iterative Solutions of Nonlinear equations in several variables.* Academic Press, 1970.

[PFTV90] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C. The art of Scientific Programming.* Cambridge University Press, 1990.

[PVY88] T. Poggio, H. Voohrees, and A. Yuille. A regularized solution to edge detection. *Journal of Complexity*, 4:106–123, 1988.

[Ter88] Demetri Terzopoulos. The computation of visible-surface representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-10(4):417–438, July 1988.

[TWK88] Demetri Terzopoulos, Andrew Witkin, and Michael Kass. Constraints on deformable models: recovering 3D shape and nonrigid motion. *AI Journal*, 36:91–123, 1988.

[Wei90] Isaac Weiss. Shape reconstruction on a varying mesh. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-12(4), April 1990.