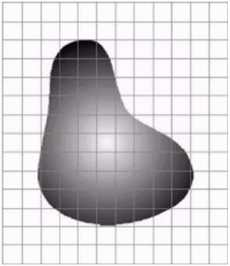
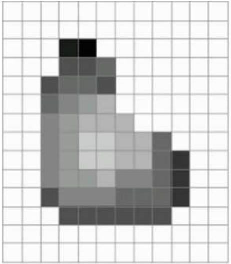
Modèles mathématiques pour le traitement et l'analyse d'images

Marie-Odile Berger, INRIA Nancy Grand Est

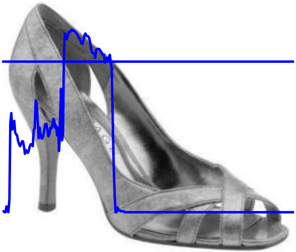
Colloque *Les mathématiques dans la société*

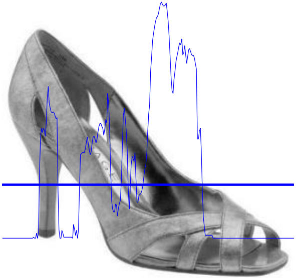
- Une image est un tableau à deux dimensions à valeurs discrètes
- L'image est issue d'un monde continu par pixelisation et quantification des niveaux de couleur
- En pratique, on considérera une image comme une fonction $R \rightarrow R$ mais elle n'est connue qu'en un ensemble discret de valeurs.



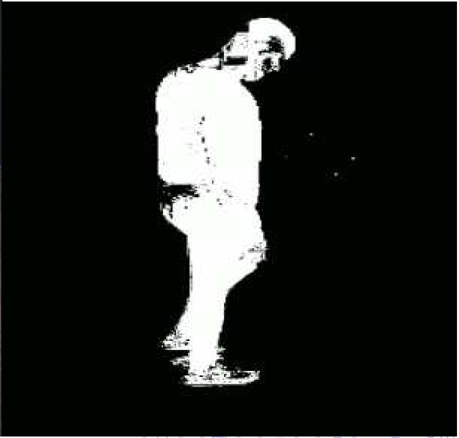


La réalité d'une image et ses *imperfections*



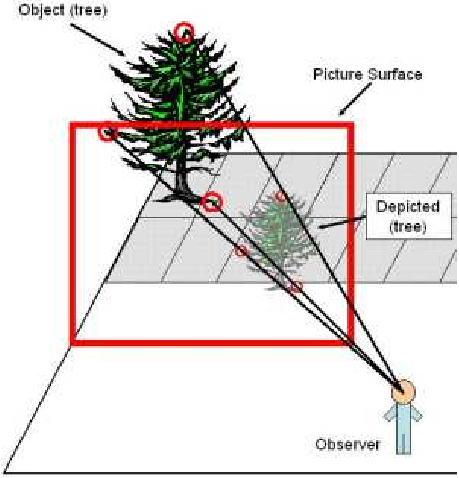


Un bas niveau bruité

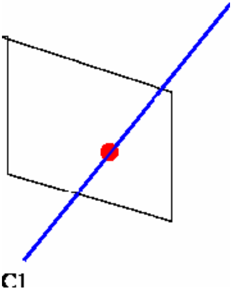


Objectif du traitement et de l'analyse d'images

- Objectifs purement 2D: on s'intéresse aux objets présents dans l'image sans chercher à les interpréter en 3D
- Objectifs 3D: interpréter l'image en terme de contenu 3D



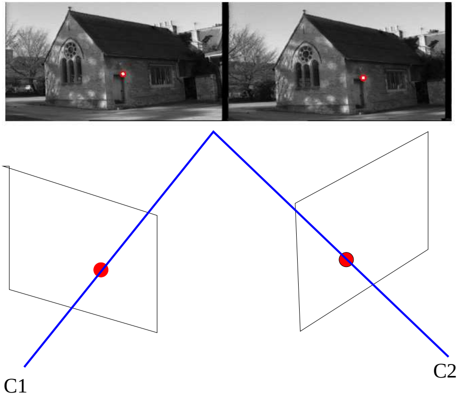
Difficultés de l’analyse en 3D



- À un point dans l’image est associée une ligne de vue

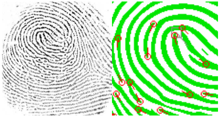
Difficultés de l'analyse en 3D

- À un point dans l'image est associée une ligne de vue
- On peut inférer de l'information 3D:
  - avec plusieurs caméras (stéréo)
  - avec des connaissances sur la forme à identifier
  - ...
- La géométrie est très présente en vision par ordinateur mais ce n'est pas la seule discipline mathématique!



Exemples d'applications (2D)

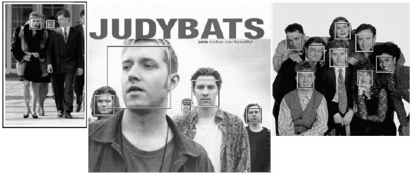
- Analyse 2D



- Construction de panoramiques [BL07]

  $\Rightarrow$ 

- Reconnaissance [VJ01]

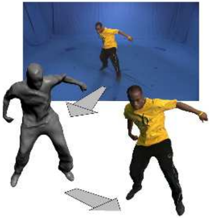


- Reconstruction stereoscopique [IGN]

- Structure from motion [ASS+09]

Exemples d'applications: acquisition de modèles structurés

- Capture de mouvement [SH07]



Où trouve-t-on des mathématiques en image?

- A de très nombreux endroits, avec un caractère plus ou moins pointu selon le domaine (de l'utilisation des mathématiques à la recherche en mathématiques)
- Quelles sortes de mathematiques?
  - traitement du signal, filtrage 2D
  - géométrie: formation de l'image par projection perspective, géométrie inter-images
  - extraction et modélisation de structures 2D ou 3D, tenant compte du fait que le bas niveau est imparfait (probabilites, régularisation, EDP)
- De façon transverse:
  - créer des schémas numériques adaptés permettant de calculer les solutions
  - créer des schémas numériques robustes, c'est à dire peu sensibles aux imperfections du bas niveau

Segmentation/ EDP/ level sets

Objectifs:

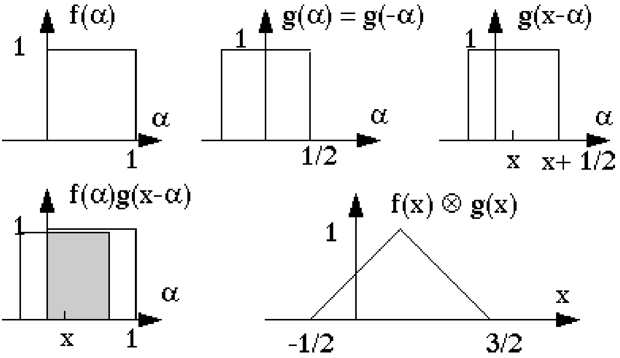
- Présenter les méthodes historiques, assez ad hoc
- Montrer les méthodologies mathématiques ayant conduit à des améliorations fortes en qualité de segmentation
- Introduire de la connaissance en segmentation
- Exemples de travaux récents

Historique de la segmentation

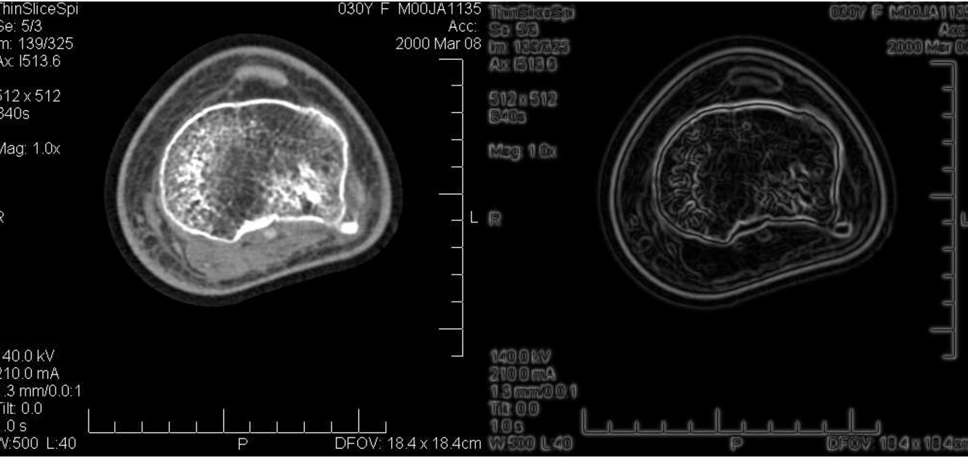
- Contour: changement brutal de l'intensité lumineuse.
- $\rightarrow$ extrema de la dérivée première
- mais l'image est numérisée: calcul approximatif de la dérivée $f'x) = lim_{h \rightarrow 0} \frac{f(x+h)-f(x)}{h}$ en prenant h=1...
- Le calcul de dérivée nécessite un pré-filtrage des images (souvent filtre gaussien) $G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * I$; $G_y = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I$
- Point de contour détecté si $|\nabla(I)|$ est assez grand. Reste ensuite à chainer ces points

Convolution

$$f * g(x) = \int f(\alpha)g(x-\alpha)d\alpha$$

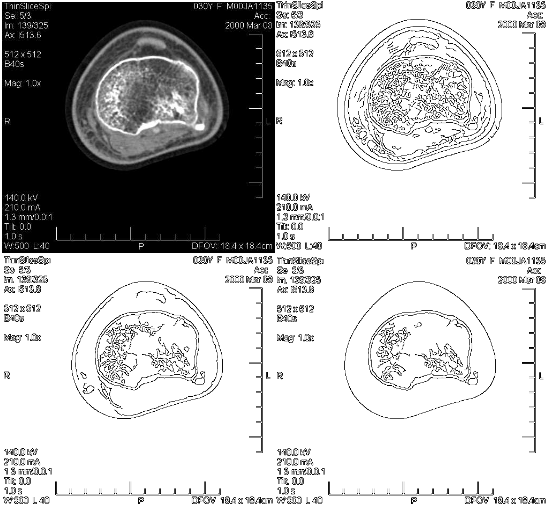


Une image du module du gradient



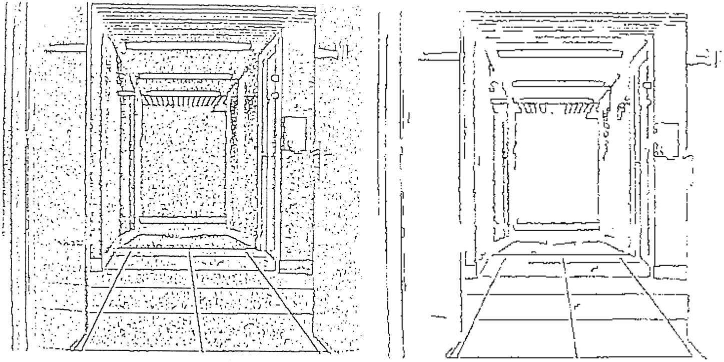
Exemples d'extraction de points de contours

une forte dépendance au seuil...



Exemples d'extraction de points de contours

reste à chainer et à éliminer les petites structures pour obtenir des courbes:

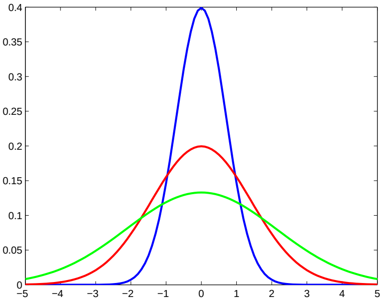


Difficulté: le contour n'est pas extrait en tant que courbe mais comme un ensemble de points isolés.

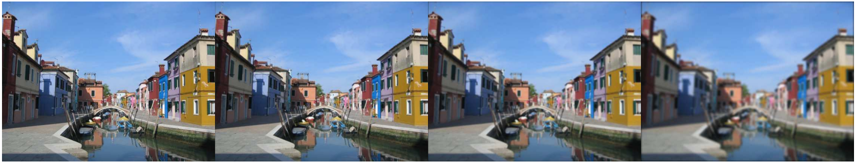
Lissage: le filtre Gaussien

- L'opérateur gaussien de moyenne nulle est d'écart type σ est donné en dimension 1 par

$$G_\sigma = \frac{1}{\sqrt{2\pi\sigma}} exp^{-x^2/2\sigma^2}$$

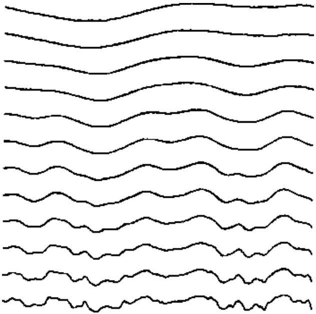


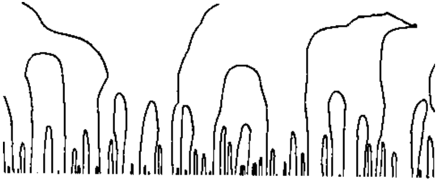
- la convolution $u_\sigma = G_\sigma * u$ d'une image u par une gaussienne G_σ permet de lisser l'image à différentes échelles (initiale, 3,5,10).

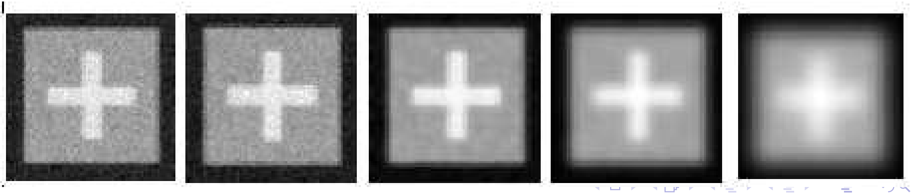


Lissage et extraction de contours

Lisser une image (ici filtre gaussien), entraine une délocalisation des points de contours...







Vers des méthodes d'extraction bien fondées

- A partir des années 1990, volonté d'avoir des processus d'extraction
  - moins ad hoc!
  - respectant des propriétés d'invariance (pas de délocalisation par filtrage)
  - considérant les contours en tant que courbes (et non comme un ensemble de points)
- Les méthodes:
  - filtrage anisotropique: lisser l'image en préservant les strutures
  - formuler les problèmes en terme d'optimisation: l'indice cherchée est un object complexe (courbe, surface) minimisant une fonctionnelle traduisant l'adéquation à l'image et les propriétés souhaitables de cet objet.
  - les équations aux dérivées partielles permettent de résoudre ces problèmes d'optimisation.

Equation de la chaleur et filtre gaussien

- Rappels:

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$$
$$div(u) = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y}$$
$$\Delta u = div(\nabla u)$$

- L'équation de propagation de la chaleur:

$$\frac{\partial u(t,x)}{\partial t} = \Delta u(t,x),$$
$$u(0,x) = u_0(x)$$

- Propriété fondamentale:
  $I * G_\sigma$ est solution de l'équation de la chaleur à $t = \sigma^2/2$.

Schéma numérique

- en dimension 1, $\frac{\partial u(t,x)}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ avec la condition aux limites $u(0,x) = I$
- soit δt le pas de discrétisation. $\frac{\partial u(t,x)}{\partial t} \approx (u(t+\delta t, x) - u(t,x))/\delta t$
- Discrétisation spatiale avec un pas de 1 (pixel):
  $\frac{\partial u}{\partial x} \approx u(x+1) - u(x) \approx u(x) - u(x-1) \approx (u(x+1) - u(x-1))/2$
- Premier pas de temps:

$$
\begin{aligned}
&(u(\delta t, x) - u(0,x))/\delta t = (u(x+1) - u(x-1))/2 \\
&\rightarrow u(\delta t, x) = u(0,x) + \delta t * (u(x+1) - u(x-1))/2
\end{aligned}
$$

- itérations suivantes $u(t+\delta t, x) = u(t,x) + \delta t(u(x+1) - u(x-1))/2$

Filtrage anistropique: le filtre de Perona-Malik [PM90]

- Modifier l'équation de la chaleur en n'autorisant la diffusion que lorsque le gradient est faible (pas de contour)
- $$\frac{\partial u}{\partial t} = \Delta u = div(\nabla u)$$

  devient
- $$\frac{\partial u}{\partial t} = div(c(|\nabla u|)\nabla u)$$

  ou c est une fonction décroissante avec $c(0) = 1$ et $lim_{x\to\infty} c(x) = 0$, par exemple $c(x) = \frac{1}{\sqrt{1+x^2}}$
- si $c = 1$, on retrouve l' équation de la chaleur $\rightarrow$ diffusion,
- si c est faible, la diffusion est stoppée ce qui préserve les bords.

Filtrage anistropique: résultats avec Perona-Malik


Initiale


50 it


100 it

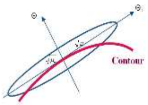

Zoom


50 it


100 it

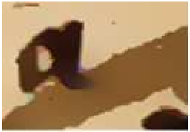
Filtrage anisotropique: améliorations ultérieures

- Le filtre de Perona réduit la diffusion dans les zones à fort gradient mais il ne tient pas compte de la structure locale
- Amélioration ulterieures permettant des lissages directionels: on lisse tangentiellement au contour mais pas au travers du contour (ex Alvares, Weickert [Wei97]) Toujours avec un schéma de type EDP.
- utilisation d'un tenseur de diffusion au lieu d'une diffusion scalaire



Tenseur

I originale



filtre Perona

filtre weickert

Optimisation, EDP et segmentation

- De nombreux problèmes de segmentation sont formulés comme des problèmes d'optimisation d'une énergie
- Avec un (des) termes d'adéquation à l'image et un terme assurant la régularité de la solution
- Exemple remarquable:
  - Contours actifs [KWT88]]: un contour est une courbe suffisament régulière dont les points ont en moyenne un gradient fort.

$$min_{v(s)} \underbrace{\int \alpha|v'(s)|^2 + \beta|v''(s)|^2}_{regularite} - \underbrace{|\nabla I(v(s))|^2 ds}_{Contour}$$

Les modèles de contours actifs

$$\int \alpha|v'(s)|^2 + \beta|v''(s)|^2 - |\nabla I(v(s))|^2 ds$$

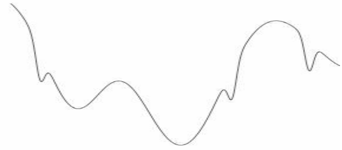
- Pas de solution explicite: v est solution de l'EDP

$$\alpha v''(s) + \beta v''''(s) - \frac{\partial}{\partial v}|\nabla I(v(s))| = 0$$

- Solution itérative à partir d'une courbe initiale: la solution initiale doit être assez proche de l'objet à segmenter
- Il est impossible de changer de topologie: on ne peut détecter qu'une seule courbe, pas plusieurs.

Evolution des modèles

- Le type de l'EDP à résoudre dépend fortement de l'énergie utilisée... (forme de l'energie et non convexité, calcul de la dérivée difficile, choix et convergence du schéma numérique)



- Est il possible d'envisager d'autres energies, véhiculant la même idée (régularité des frontières) qui soient plus faciles à optimiser? avec des schémas numériques bien établis?
- Peut-on autoriser les changements de topologie?

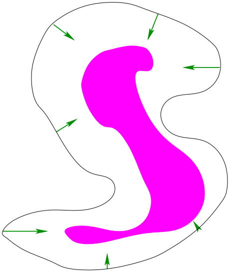
Evolution

- De nouvelles énergies, de nouvelles EDP avec de meilleures propriétés:

$$\int g(|\nabla I(v(s))|)|v'(s)|ds$$

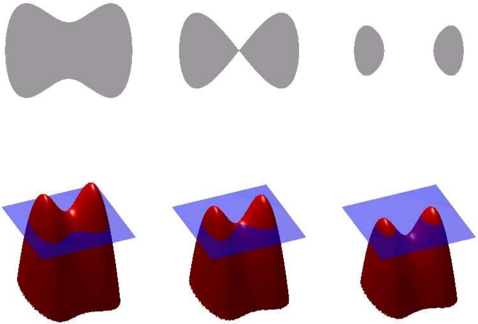
ou g est une fonction décroissante du gradient, comme celle de Perona Malik ($g = \frac{1}{1+|\nabla I|^2}$)

- Même comportement intuitif: minimum atteint pour des gradient forts et une bonne régularité.
- La courbe suit une équation d'évolution $\frac{\partial c}{\partial t} = \alpha N$ avec α dépendant de la courbure et du gradient.



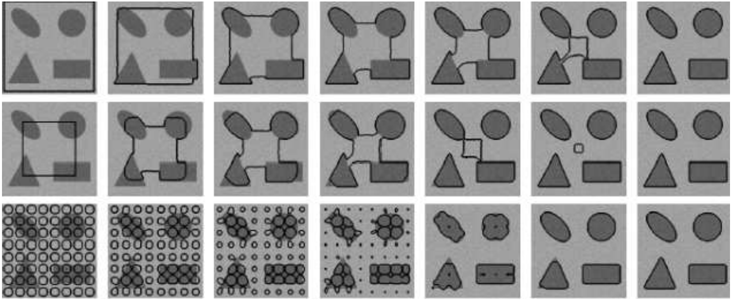
Level-sets et changement de topologie

- Idée des level-sets: modéliser les objets comme courbe de niveau d'une surface [Set96]



- l'équation d'évolution sur les droites est remplacée par l'équation d'évolution sur la surface.
- Les changements de topologie deviennent possible!

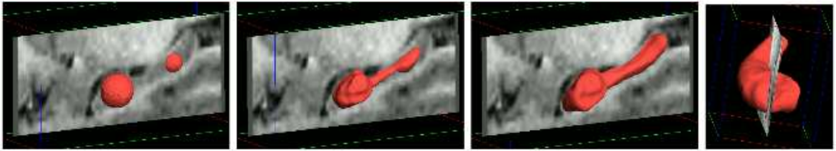
Levels sets: exemples [PD00]



Levels sets

Inombrables améliorations

- en 3D [HCG03]



(a) (b) (c) (d)

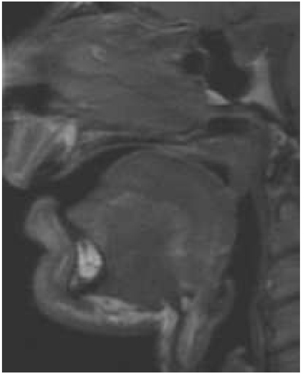
Fig. 1. Segmentation of left hippocampus from MRI: (a) initialization by two bubbles, (b) 6 iterations, (c) 18 iterations (final), and (d) final segmentation with rotated view ($c_{\mathrm{MCF}} = 1, r_{\mathrm{MCF}} = -1, r_{\nabla g} = 1, r_c = 2, \alpha = 1.4, r_s = 0$, and $c_s = 0.8$).

- Avec des indices autres que les contours: régions, texture
- Application à la reconstruction multi-vues

Introduire des a priori de formes dans la segmentation

- Les approches précédentes introduisent des directives assez floues sur les formes à extraire (frontières assez régulières, ...)
- Introduire des modèles de forme dans ce processus est indispensable quand les indices images ne suffisent pas à inférer la forme
  - Construire un modèle de forme paramétrique (formes acceptables)
  - Optimiser dans cet espace de formes

- Cadre de l'étude: segmenter automatiquement le conduit vocal pour synthétiser (visuel+audio) une tete parlante réaliste
- L'information photométrique ne suffit pas à extraire les contours de langue: impossibilité de séparer langue, palais et dents selon les sons.



Segmentation guidée par la connaissance

- Une méthode de segmentation guidée par un modèle de référence obtenu par détourage manuel sur un locuteur.
- Le (un) modèle de référence est construit par analyse en composantes principales sur les courbes sagittales manuellement détourées sur 39 sons acquis par IRM.

$$C(w) = \bar{C} + \sum_{i=1}^{p} w_i \delta C_i,$$

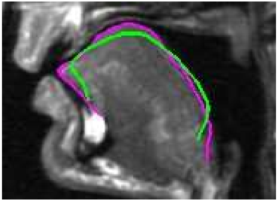
  les δC_i sont les composantes principales,
  p est assez grand, $w_i \in [-5\sqrt{\lambda_i}, 5\sqrt{\lambda_i}]$ pour augmenter les possibilités du modèle au dela de la référence
- Segmentation: On cherche une courbe dans cet espace de formes admissibles délimitant deux zones homogènes en intensité

Une méthode variationnelle de segmentation intégrant un modèle de formes

- On cherche la courbe C minimisant une energie E traduisant l'adéquation de la forme à l'image.
- E traduit le fait que les variances à l'intérieur et à l'extérieur du contour sont minimales. Idéalement, deux valeurs constantes de part et d'autre de la courbe.

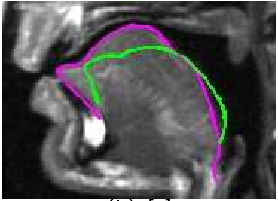
$$E_G(C) = \int_{C_{\text{in}}} (I(\mathbf{x}) - \mu)^2 d\mathbf{x} + \beta \int_{C_{\text{out}}} (I(\mathbf{x}) - \nu)^2 d\mathbf{x}$$

Résultats

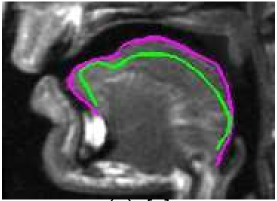


.ps

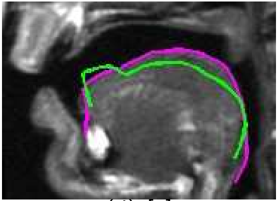
(a) [i]



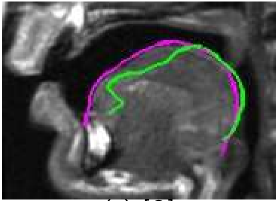
(b) [e]



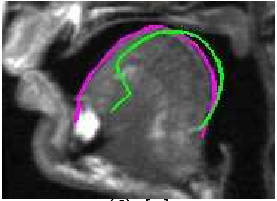
(c) [ε]



(d) [a]



(e) [O]



(f) [u]

Figure: vert: courbe intiale. magenta: resultat

Conclusion

- Illustration de quelques contributions des mathématiques à la vision. Il y en a bien d'autres, notamment en apprentissage...
- Les collaborations avec les mathématiciens ont permis des avancées importantes
    - formalisation des problèmes (propriétés souhaitables, invariance)
    - méthodes de modélisation/segmentation élaborées
    - des schémas numériques de plus en plus efficaces
- Des outils sophistiqués souvent diffusés librement,
- Innombrables applications

S. Agarwal, N. Snavely, I. Simon, S. Seitz, and R. Szeliski.
Building Rome in a Day.
In *Proceedings of 9th International Conference on Computer Vision, Kyoto, Japan*, October 2009.

Matthew Brown and David G. Lowe.
Automatic panoramic image stitching using invariant features.
International Journal of Computer Vision, 1(74):59–73, 2007.

Sean Ho, Heather Cody, and Guido Gerig.
Snap: A software package for user-guided geodesic snake segmentation.
In *Miccai2003*, 2003.

M. Kass, A. Witkin, and D. Terzopoulos.
Snakes: Active Contour Models.
International Journal of Computer Vision, 1:321–331, 1988.

N. Paragios and R. Deriche.
Geodesic active contours and level sets for the detection and tracking of moving objects.

IEEE Transactions on PAMI, 22(4):415, April 2000.

T. Peng, E. Kerrien, and M.O. Berger.
A shape base framework to segmentation of tongue contours from mri data.
In *35th IEEE International Conference on Acoustics, Speech, and Sign al Processing - ICASSP 2010*, pages 662 – 665, Mar 2010.

P. Perona and J. Malik.
Scale Space and Edge Detection Using Anisotropic Diffusion.
IEEE Transactions on PAMI, 12(7):629–639, July 1990.

J. A. Sethian.
Level Set Methods.
Cambridge University Press, 1996.

J. Starck and A. Hilton.
Surface capture for performance based animation.
Computer Graphics and Applications, 3(27), 2007.

Paul Viola and Michael Jones.
Robust real-time object detection.

International Journal of Computer Vision, 2001.

J. Weickert.
A review of nonlinear diffusion filtering.
In *Scale Space 97, Utretch (The Netherlands), Lecture Notes in Computer Science*, pages 3–28, 1997.