An Overview of Active Contour Using Region Statistics For Image Segmentation

Yunmei Chen

Department of Mathematics
University of Florida

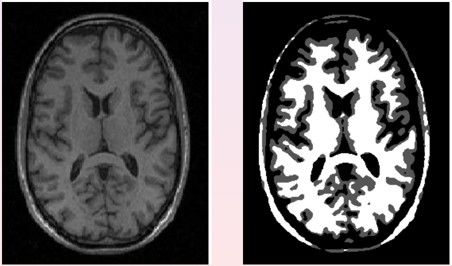
May, 2007

Outline

- Active Contours For Segmentation
- Mumfor-Shah Model
- Active Contour Using Parametric Statistics
- Active Contour Using Non-Parametric Statistics
- Our Recent Results On Active Contour Using Non-parametric Local Statistics

Active Contours in Image Segmentation

- Segmentation is to Partition an image into disjoint, connected components that are homogeneous w.r.t. intensity, texture or certain probabilistic measures.
- Active contour method is evolving contours towards boundaries of interest by designed forces (e.g. edge, region information or prior knowledge).
- Active contour models have a consistent mathematical description, their solutions satisfy certain minimum principles.



Edge and Region Based Active Contours

- Edge based:
  Rely on edge information (high magnitude of image gradient)
  Limitation - sensitive to noise, artifacts, may leak through gaps on boundary
- Region based: Make use of information on regional statistics from image intensities.
  Limitation - high noise level, intensity inhomogeneity, complex intensity distribution
- Combination of edge based and region based

Edge Based Active Contour

Geodesic Active Contour (Caselles, Kimmel, Sapiro, 1997)

$$\min_{C(p)} \int_0^1 g(|\nabla G_\sigma * I|(C(p))|C'(p)|dp$$

$$g(s) = \frac{1}{1+ks^2}, \quad G_\sigma(x) = \frac{1}{\sigma^n} e^{-|x|^2/4\sigma^2}$$

Region Based Active Contours

- Mumford-Shah model in segmentation
  piecewise smooth model / cartoon model
- Likelihood maximization approaches
  - Using parametric statistics
  - Using non-parametric statistics

Mumford-Shah (MS) Model

Mumford-Shah (MS) Model

MS model for simultaneous smoothing and segmentation:

- MS piecewise smooth model (two phases):

$$\min_{C(p),u(x)} \int_{\Omega\setminus C} |\nabla u(x)|^2 dx + \alpha \int_{\Omega} (I(x) - u(x))^2 dx + \beta \int_0^1 |C'(p)| dp. \quad (1)$$

- MS Cartoon model (pw constant):

$$\min_{C(p),m_1,m_2} \sum_i \int_{\Omega_i} (I - m_i)^2 dx + \beta \int_0^1 |C'(p)| dp. \quad (2)$$

Ω_1: region inside C, Ω_2: region outside C, and $\Omega = \cup_{i=1}^2 \Omega_i$.

Mumford and Shah '89, L. Ambrosio et.al. '97, Tsai et.al.'01, T. Chan et. al. '02 and A. Yezzi et.al.'02.

Active Contour Using Parametric Statistics

Likelihood Maximization With Parametric Probability Density Function (pdf)

- Consider $I(x)$ at each $x \in \Omega_i$ $(i = 1, 2)$ as an independent random variable with a parametric pdf $p(I(x)|\theta_i(x), \Omega_i)$,
- joint pdf of I in Ω_i is $\prod_{x \in \Omega_i} p(I(x)|\theta_i(x), \Omega_i)$.
- Likelihood function:

$$L(\theta_i(x)'s, \Omega_i's) = \Pi_{i=1}^{2}(\prod_{x \in \Omega_i} p(I(x)|\theta_i(x), \Omega_i).$$

- Maximizing likelihood estimate (MLE):

$$\min_{C, \theta_i(x)'s} -lnL(\theta_i(x)'s, \Omega_i's) + \beta|C|$$

$$= \sum_{i=1}^{2} \int_{\Omega_i} \ln p(I(x)|\theta_i(x), \Omega_i) + \beta|C|.$$

MLE with Gaussian pdf

- Gaussian pdf $\theta_i(x) = (u_i(x), \sigma_i(x))$

$$p(I(x)|\theta_i(x), \Omega_i) = \frac{1}{(\sqrt{2\pi}\sigma_i(x))} e^{-\frac{|I(x)-u_i(x)|^2}{2\sigma_i^2(x)}}.$$

- MLE approach

$$\min_{C, u_i(x), \sigma_i(x)} \sum_{i=1} \int_{\Omega_i} \{\frac{1}{2\sigma_i^2(x)}|I(x) - u_i(x)|^2 + \ln \sigma_i(x)\} dx + \beta|C|.$$

- Segmentation of homogeneous regions: $u_i(x) = u_i$, $\sigma_i(x) = \sigma_i$ (Zhu. et. al.'96 and Rousson et. al.'02):

$$\min_{C, u_i's, \sigma_i's} \sum_{i=1}^{2} \int_{\Omega_i} \frac{1}{2\sigma_i^2}|I(x) - u_i|^2 dx + |\Omega_i| \ln \sigma_i + \beta \int_0^1 |C'(p)|dp. \tag{3}$$

Connections And Drawbacks of Previous Models

Connections:

- If σ_i's (or $\sigma_i(x)$'s) in Gaussian model are fixed, it reduce to cartoon (or ps. smooth) MS model.
  More references: S.Jehan-besson et al. 01, N.Paragios et al 02, P.Martin et al 04.

Drawbacks:

- Parametric model and MS cartoon model are not good for images with higher level noise or multi-modal intensity distributions.
- Pw. smooth MS Model smooths image by min $\int |\nabla u|^2$, which is not adaptive for edge preserving.

Active Contour Using Non-Parametric Statistics

Nonparametric Density Estimation

Kernel method for nonparametric pdf estimation:

- Estimate pdf $f(X)$ of RV X from data $\{x_1, x_2, ..., x_n\}$:

$$\hat{f}(X = x) = \frac{1}{n}\sum_{i=1}^{n} K(x - x_i),$$

  K: window function: smooth, positive, rapidly decreasing to zero outside the window, whose integral is equal to one.

- Estimate pdf $f(x) = p(I(x)|x \in B)$ of RV $I(x)$

$$\hat{f}(x) = \frac{1}{|B|}\int_B K(I(x) - I(y))dy.$$

- Parzen Window density function:

$$K(Z) = \frac{1}{(\sqrt{2\pi}\sigma)^m} e^{-\frac{Z^2}{2\sigma^2}}$$

Some Related Work

- J.Kim et. al.'05
  Segmentation by maximization of the mutual information between the region labels and image intensities, where the pdf is estimated using Parzen window.
- X.Huang et. al.' 04
  Segmentation by minimizing the energy consisting the shape data term from edge map image, and intensity data term from likelihood function of interior nonparametric statistics estimated by using parzen window.
- G.Aubert et. al. 03
  Segmentation of regions in video sequence by matching histograms of the current frame with the previous one, where the histograms (pdf) are estimated using Parzen window.
- Work in this line: T.Brox et. al. 03 W.Abd-Almageed et. al.'03, A.Hermosillo et. al. 04.

Active Contour Using Non-Parametric Local Statistics (recent joint work with W.Guo)

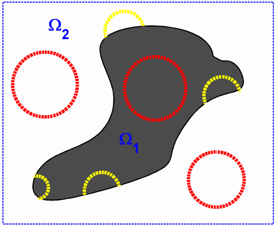
Estimate pdf locally in a neighborhood whose size depends on image gradient to preserve and enhance boundaries of features.

- Consider $I(x)$ at each $x \in \Omega_i \bigcap B(x, r(x))$ as random variables with pdf $f_i(x)$ (Ω_1:inside C,Ω_2:outside C).
- Use data $\{I(y), y \in \Omega_i \bigcap B(x, r(x))\}$ rather than $\{I(y), y \in \Omega_i\}$ to estimate $f_i(x)$.
- Estimation of pdf:

$$\widehat{f}_i(x) = \frac{1}{|B(x, r(x)) \bigcap \Omega_i|} \int_{B(x,r(x)) \bigcap \Omega_i} K(I_i(x) - I(y))dy,$$

$$K(z) = \begin{cases} 3/4(1 - z^2), & |x| < 1 \\ 0, & |x| \geq 1 \end{cases} \tag{4}$$

Radial $r(x)$ Depends On Image Gradient



$$r(x) = \begin{cases} M, & |\nabla \widetilde{I}(x)| \leq s \\ N, & |\nabla \widetilde{I}(x)| > s \end{cases} \tag{5}$$

or more general

$$r(x) = \frac{a}{1 + b|\nabla \widetilde{I}(x)|^2}, \quad \textit{for some } a > 0, \; b > 0.$$

$\widetilde{I}$: a smoothed version of initial image I.

Proposed model

- Find C, $I_1(x)$ and $I_2(x)$ by minimizing

$$-\sum_i \int_{\Omega_i} ln \left(\frac{1}{|B_x^i|} \int_{B_x^i} K(I_i(x) - I(y))dy \right) dx + \beta \int_0^1 |C'(p)|dp \quad (6)$$

- Level set form: Find $\phi(x)$ and $I_i(x)$ by minimizing

$$\begin{aligned} &-\int_\Omega H(\phi(x)) ln \frac{1}{|B_x|} \int_\Omega 1_{B_x} H(\phi(y)) K(I_i(x) - I(y)) dydx \\ &-\int_\Omega (1 - H(\phi(x))) ln \frac{1}{|B_x|} \int_\Omega 1_{B_x} (1 - H(\phi(y))) K(I_i(x) - I(y)) dydx \\ &+\beta \int_\Omega |\nabla H(\phi(x))| dx. \end{aligned} \quad (7)$$

$C = \{x \in \Omega | \phi(x) = 0\}$, inside C $\phi(x) > 0$, $H(x)$:Heaviside function, $B_x = B(x, r(x))$, $B_x^i = B(x, r(x)) \bigcap \Omega_i$.

Features OF The Proposed Model

- Region based active contour using nonparametric local statistics for simultaneous segmentation and adaptive smoothing.
- $u(x) = \sum_i 1_{\Omega_i} u_i(x)$ reduces the oscillation in $I(x)$.
  $u_i(x)$the weighted average of $I(x)$ within the ball $B(x, r(x))$ inside Ω.
  Weighting is determined by the kernel function inside Ω_i.
  smoothing does not cross boundary C.
- Ball size $r(x)$ depends on image gradient, so that the smoothing is adaptive and feature preserving.

Existence of Minimizer

$$\min_{\chi_A \in BV(\Omega)} E(\chi_A) = \beta \int_\Omega |D\chi_A|$$
$$- \int_\Omega \chi_A(x) log \left(\frac{1}{|B_x|} \int_{B_x} \chi_A(y) K(u_1(x) - I(y)) dy \right) dx$$
$$- \int_\Omega (1 - \chi_A(x)) log \frac{1}{|B_x|} \left(\int_{B(x} (1 - \chi_A(y)) K(u_1(x) - I(y)) dy \right) dx \tag{8}$$

where $B_x = B(x, r(x))$,

$$u_1(x) = \frac{\int_{B_x} \chi_A(y) I(y) dy}{\int_{B_x} \chi_A(y) dy}, x \in A$$

$$u_2(x) = \frac{\int_{B_x} (1 - \chi_A(y)) I(y) dy}{\int_{B_x} (1 - \chi_A(y)) dy}, x \in \Omega \setminus A$$

Space of Functions With Bounded Variation (BV)

- **Definition of** $BV(\Omega)$: The space of functions $u \in L^1(\Omega)$ having bounded variation:

$$|Du|(\Omega) = \sup\{\int_\Omega u div\phi | \phi \in C_0^1(\Omega, R^n), |\phi| \leq 1\} < \infty$$

Norm in $BV(\Omega)$:

$$||u||_{BV(\Omega)} = ||f||_{L^1(\Omega)} + |Du|(\Omega).$$

- **Compactness theorem in** BV
Assume Ω is open, bounded with Lipschitz boundary, if $\{u_n\}_{n\geq 1} \in BV(\Omega)$ is bounded, then $\exists$ a subsequence $\{u_{n_j}\}$ and a function $u \in BV(\Omega)$, s.t. $u_j \to u$ strongly in $L^1(\Omega)$.
- **Lower Semi-Continuity in** L^1
Let $u, \{u_j\} \subset BV(\Omega)$ and $u_j \to u$ in $L^1(\Omega)$, then

$$|Du|(\Omega) \leq \liminf_{j\to\infty} |Du_{n_j}|(\Omega).$$

Existence Theorem and Sketch of the Proof

- **Theorem:**
  Let $\Omega \subset R^n$ be open and bounded with Lipschitz boundary, and $I(x) \in L^2$. Then model (8) has a minimizer $\chi_A \in BV(\Omega)$.

- **Sketch of the Proof**
- Energy is bounded below by zero.
- There is a minimizing sequence $\{\chi_{A_n}\}$ s.t. $\lim E(\chi_{A_n}) = \lim\inf_{\chi_A \in BV(\Omega)} E(\chi_A)$
- χ_{A_n} is a bounded sequence in $BV(\Omega)$

Sketch of the Proof (cont.)

- $\{\chi_{A_n}\}$ sub-converges to $u \in BV(\Omega)$ strongly in $L^1(\Omega)$, and a.e. in Ω.
- u takes value 0 or 1 only, $u = \chi_A$ for some $A \subset \Omega$.
- Dominant convergence theorem gives

$$u_{i_{n_j}} \to u_i \; strongly \; in \; L^2(\Omega), \; and \; a.e. \; in \; \Omega, \quad i = 1, 2.$$

$$-\int_\Omega \chi_{A_{n_j}}(x) log\left(\frac{1}{|B_x|}\int_{B_x} \chi_{A_{n_j}}(y) K(u_{1_{n_j}}(x) - I(y)) dy\right) dx$$
$$\to -\int_\Omega \chi_A(x) log\left(\frac{1}{|B_x|}\int_{B_x} \chi_A(y) K(u_1(x) - I(y)) dy\right) dx.$$

- The l.s.c. of BV in L^1 yields

$$E(\chi_A(x)) \leq \liminf E(\chi_{A_{n_j}}) = inf_{\chi_A \in BV(\Omega)} E(\chi_A).$$

Numerical Implementation

- Level Set Method(S. Osher and J. Sethian, T. Chan et. al. etc.)
  Allows for cusps, corners, and automatic topological changes.
  Discretization is on fixed regular grid.
- Additive Operator Splitting(X. Tai, J. Weickert et. al. etc.)
  Solve

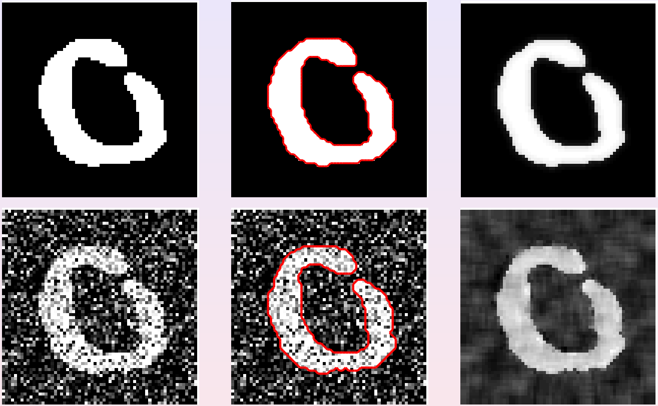
$$\frac{u^{n+1} - u^n}{\Delta t} = \lambda \sum_{l=1}^{n} A_l u^{n+1} + F^n, \quad or$$

$$u^{n+1} = (I - \Delta t \lambda \sum_{l=1}^{3} A_l)^{-1}(u^n + \Delta t F^n).$$

AOS scheme: replace the inverse of the sum by the sum of the inverse:

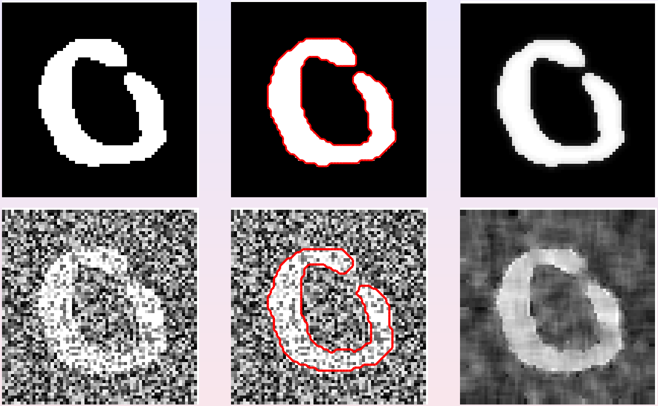
$$u^{n+1} = \frac{1}{3} \sum_{l=1}^{3} (I - 3\Delta t \lambda A_l)^{-1}(u^n + \Delta t F^n).$$

Qualitative Synthetic Results (Gaussian $\sigma = 0.05$)



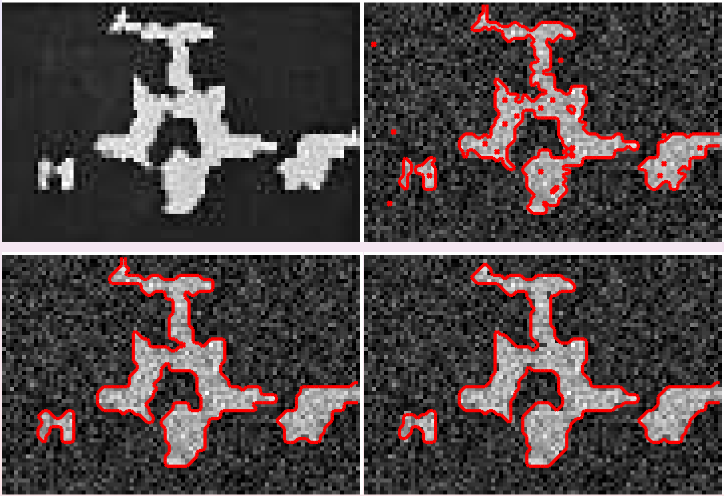
First col.: given image; Sec.: segmentation; Third: smoothing.

Qualitative Synthetic Results (speckle)

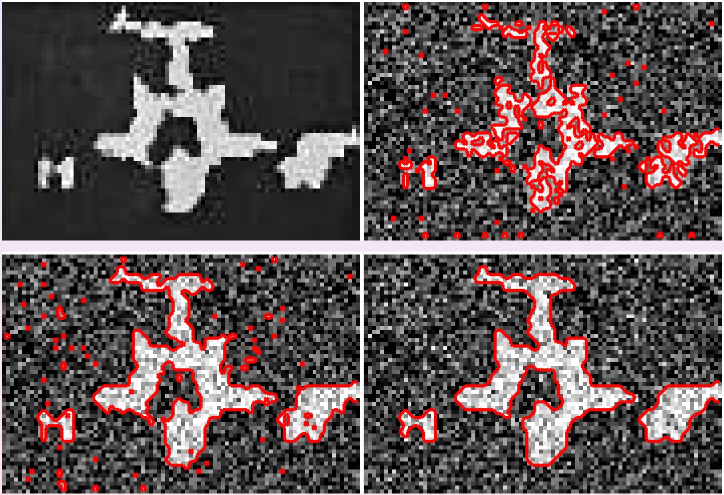


First col.: given image; Second: segmentation; Third: smoothing.

Comparison Between MS, Parametric/Gaussian, Local Nonparametric (With Gaussian $\sigma = 0.01$)



Comparison of 3 Models (With Gaussian $\sigma = 0.05$)

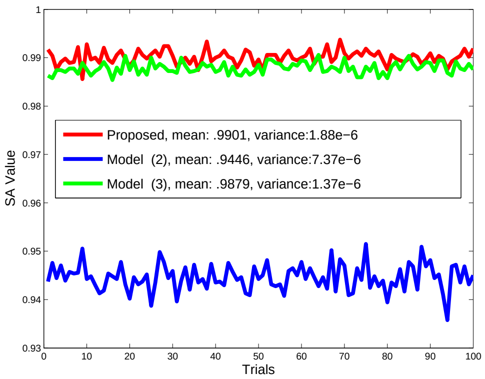


Movie Demo

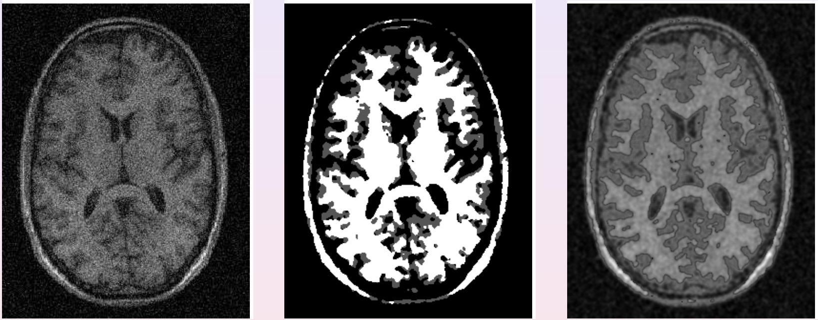
loading movie

Quantitative Comparison In Segmentation Accuracy

SA= number of pixels having same segmentation result with ground truth/total pixel number

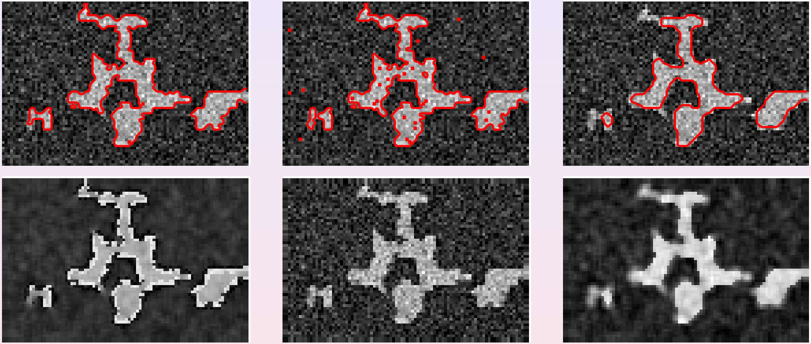


Qualitative Real Data Results



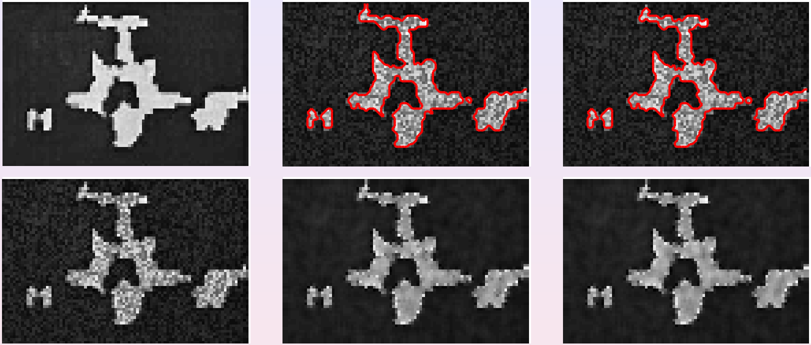
1st col.: noisy image; 2nd: segmentation; 3rd: denoised image.

Compare Choices of $r(x)$ (Gaussian Noise)



1st col: adoptive $r(x)$=2 or 4, 2nd: $r = 2$, 3rd: $r = 4$.

Compare Choices of $r(x)$ (Speckle Noise)



1st col.: given images, 2nd and 3rd: segmentation and smoothing with adaptive radius (M/N), and $r(x)$ depending on image gradient, respectively.