Automatic Microarray Spot Segmentation Using a Snake-Fisher Model

Jinn Ho and Wen-Liang Hwang

Institute of Information Science, Academia Sinica

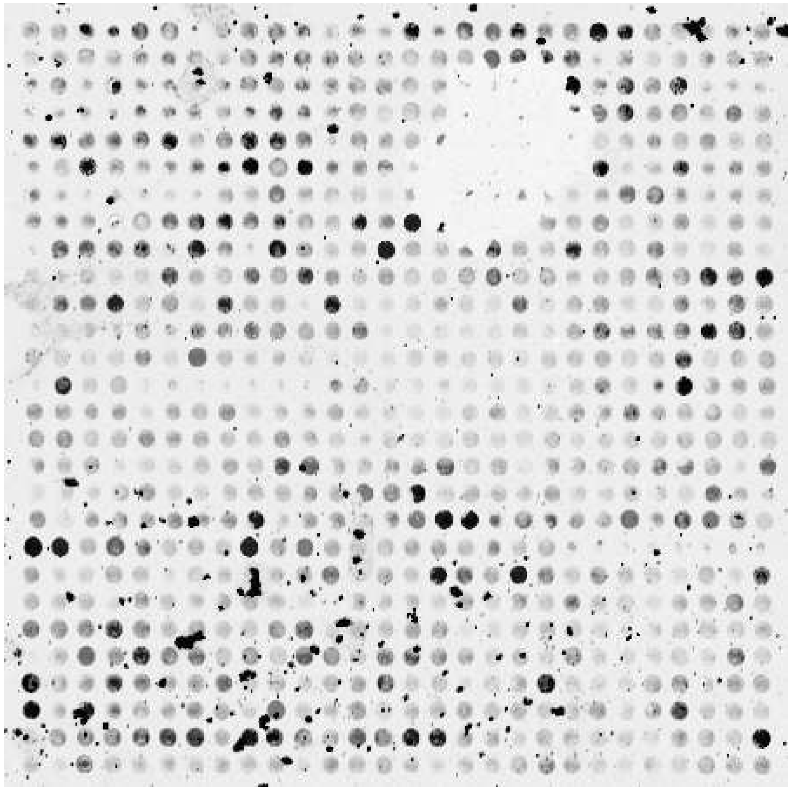
Microarray Image Processing

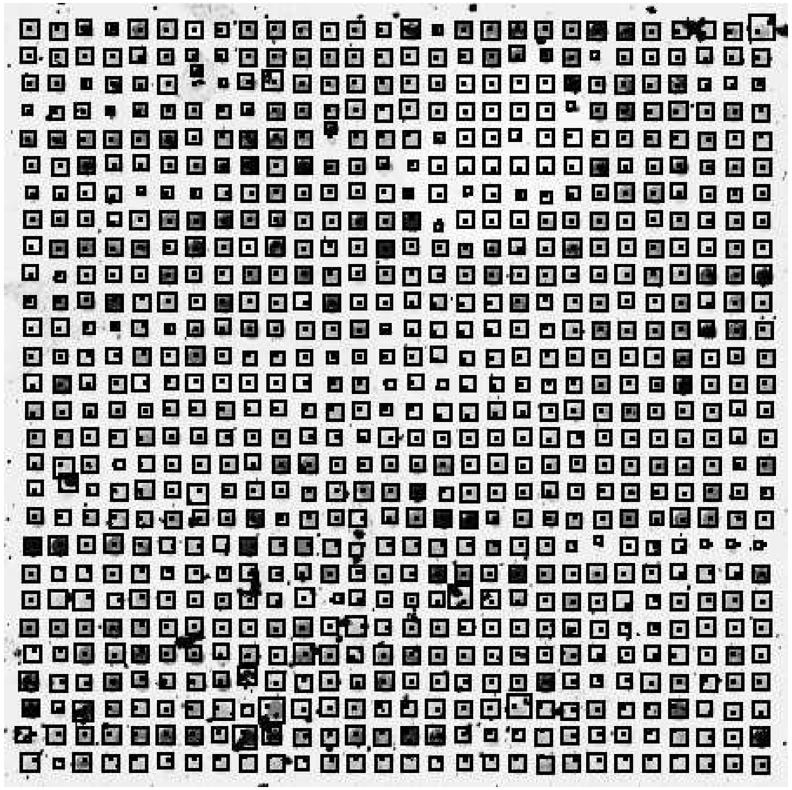
- The processing of microarray images involves three steps:
  - Detect the positions of the spot centers and identifies their coordinates (spot gridding).
  - Segment a spot
  - Extract intensity from a spot

Motivation

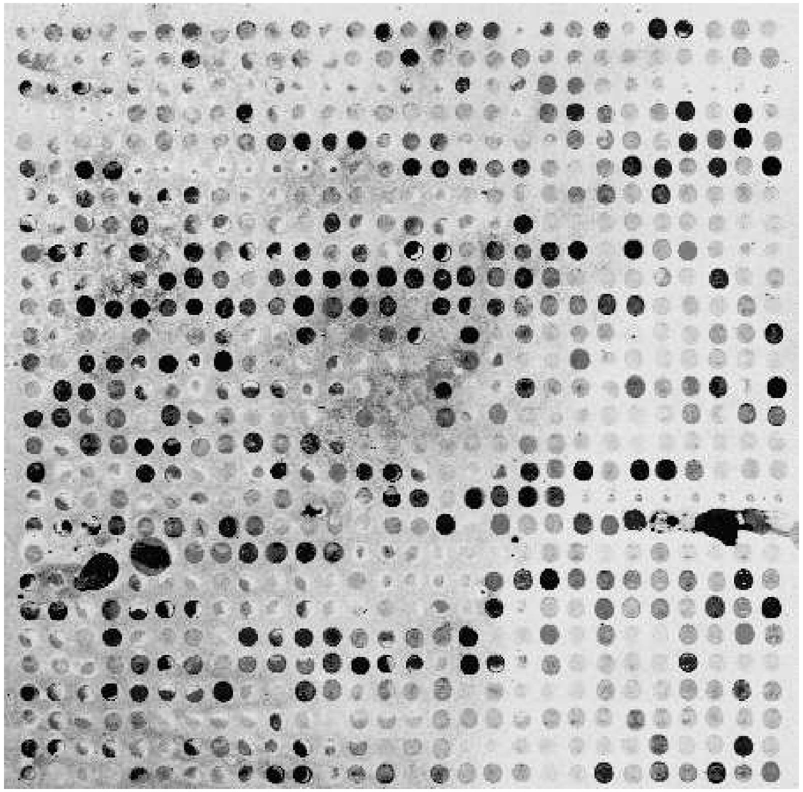
- Importance to have an automatic and accurate algorithm to perform image analysis tasks of an microarray image.
- Our previous work shows that **spot gridding** can be accurately and automatically solved.
- Inspired by Paragious and Deriche's work, which unifies boundary-based and region-based image partition approaches, we integrate the snake model and the Fisher criterion to **segment** a microarray image.

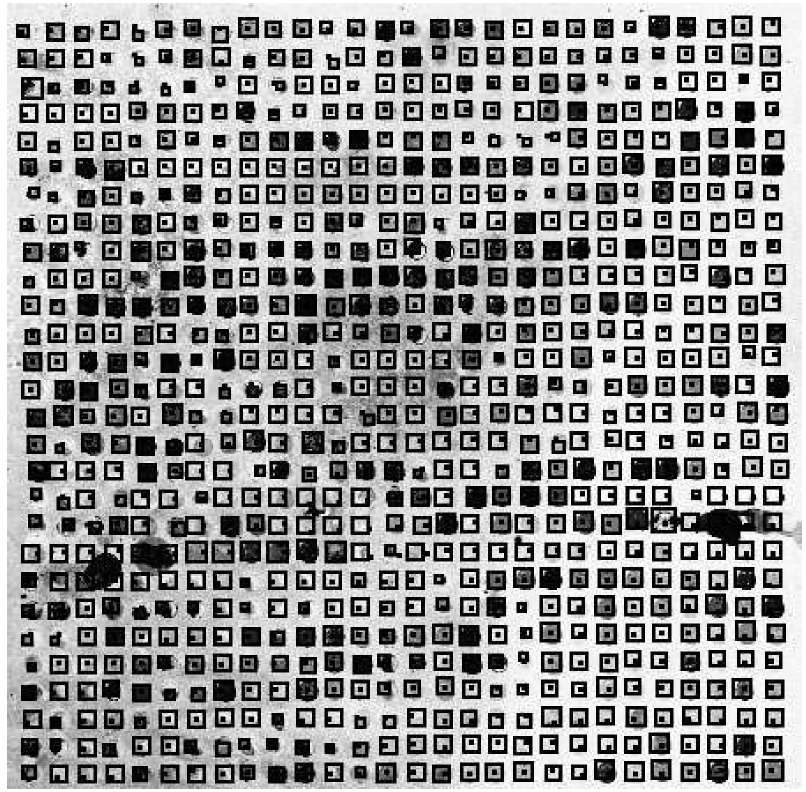
Block (3,4) of lc30n008





Block (1,1) of lc30n010





Contributions

- Our algorithm is automatic because the parameters and the contours are adaptively estimated from the data without human intervention.
- Our results outperforms those of *GenePix Pro* 5.0 and *Spot* 2.0.

Segmentation

- Region-based approach - Segment regions based on regions’ statistics.
- Boundary-based approach - Segment regions based on boundaries’ properties.
- Several approaches have been proposed that combines the two approaches for segmentation.

Proposed Spot Segmentation Approach

- We propose using the snake model to capture boundary information and the Fisher criterion to capture region information.

$$E_{snake}(\Gamma) = \int_{\Gamma} (\frac{\alpha}{2}|\Gamma_s|^2 + \frac{\beta}{2}|\Gamma_{ss}|^2 - \| \bigtriangledown I\|^2)ds,$$

$$E_{region}(\Gamma) = [\iint_{R_1} (I - M_1)^2 dxdy + \iint_{R_2} (I - M_2)^2 dxdy]/(M_1 - M_2)^2.$$

$$E_{total}(\Gamma) = E_{snake}(\Gamma) + \tilde{\gamma} E_{region}(\Gamma).$$

- Setting $\beta = 0$ because it gives fourth derivative in Euler solution.

Euler Equation

- Given the parameters (α, γ) in $E_{total}(\Gamma)$, $[x \;\; y]$ on the boundary curve Γ^* that minimizes $E_{total}(\Gamma)$ should satisfy:
-

$$-\frac{\partial \| \bigtriangledown I \|^2}{\partial x} - \alpha x_{ss} + \gamma[(I - M_1)^2 - (I - M_2)^2] y_s = 0, \quad (1)$$

$$-\frac{\partial \| \bigtriangledown I \|^2}{\partial y} - \alpha y_{ss} - \gamma[(I - M_1)^2 - (I - M_2)^2] x_s = 0. \quad (2)$$

- $\Gamma^*_{ss} = [x_{ss} \;\; y_{ss}] = \kappa \vec{n}$, where κ is the curvature, and $\vec{n}$ is parallel to $[y_s \;\; -x_s]$. We have

$$-\nabla \| \bigtriangledown I \|^2 - \alpha \kappa \vec{n} - \gamma' [\frac{(I - M_1)^2 - (I - M_2)^2}{(M_1 - M_2)^2}] \vec{n} = 0. \quad (3)$$

- A point on the optimal contour must satisfy (4) in the tangent direction $(\vec{t})$, and (5) in the normal direction $(\vec{n})$:

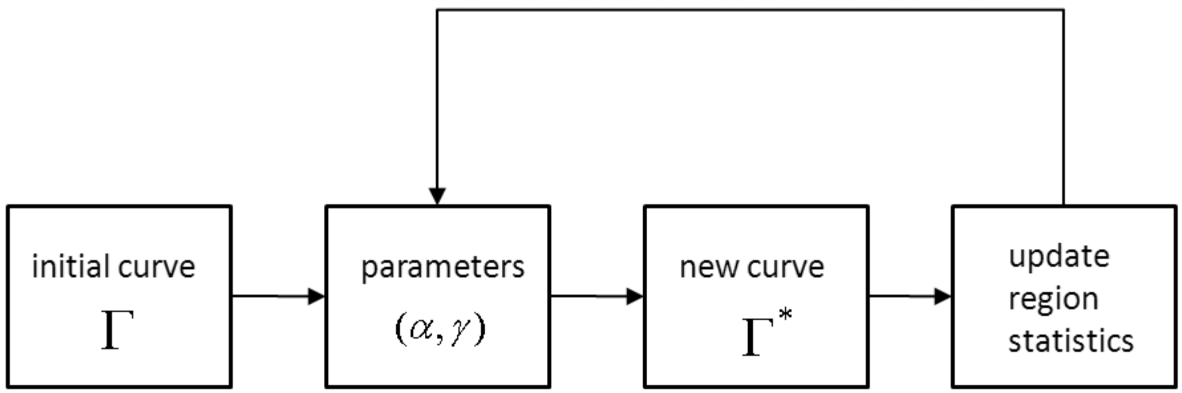
$$\nabla \| \bigtriangledown I \|^2 \cdot \vec{t} = 0, \tag{4}$$

$$-\nabla \| \bigtriangledown I \|^2 \cdot \vec{n} = \alpha\kappa + \gamma' [\frac{(I - M_1)^2 - (I - M_2)^2}{(M_1 - M_2)^2}]. \tag{5}$$

Equation (5) balances three terms: the first term is provided by the normal component of the gradients of the image, the second term is proportional to the curvature, while the last term measures the class separation.

Automatically Determining Parameters and Boundary Curve

- Estimate parameters by using Euler equation from a contour.
- Refine the contour from the derived parameters.
- Modify the region statistics (Fisher criterion).
- Repeat the above steps.



Determining Initial Contour - Climber Algorithm

- Climber movement: each climber moves freely in the tangent direction, while moves restrictively in the gradient direction.
- A climber climbs to the peak of the magnitude of the gradient function by a Hastings-Metropolis penalization and a temperature schedule similar to that in the simulated annealing algorithm.

Climber Movement

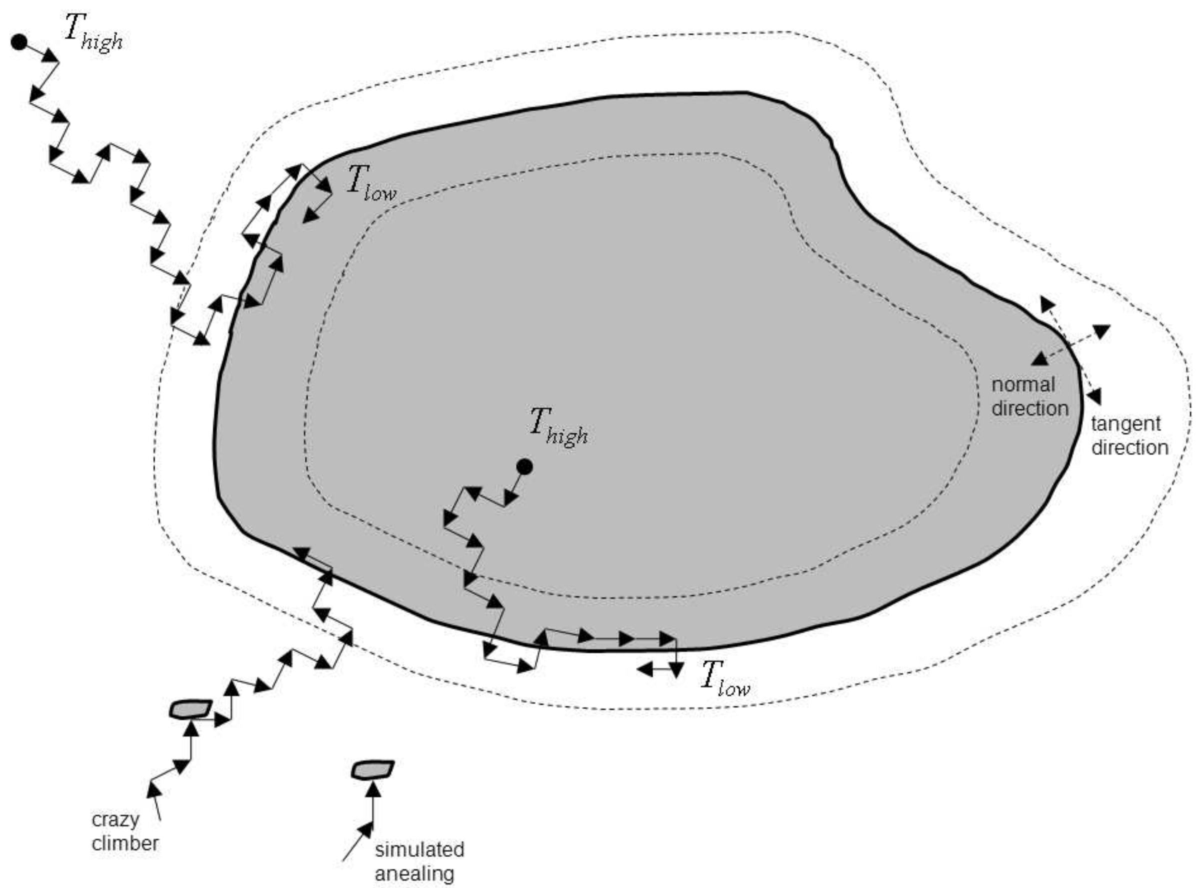
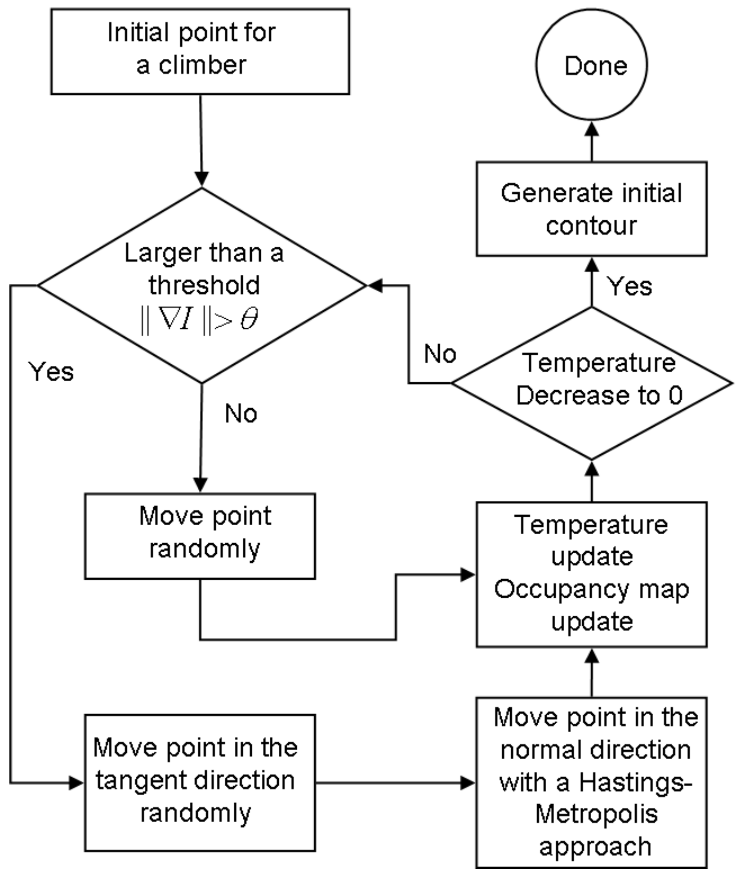


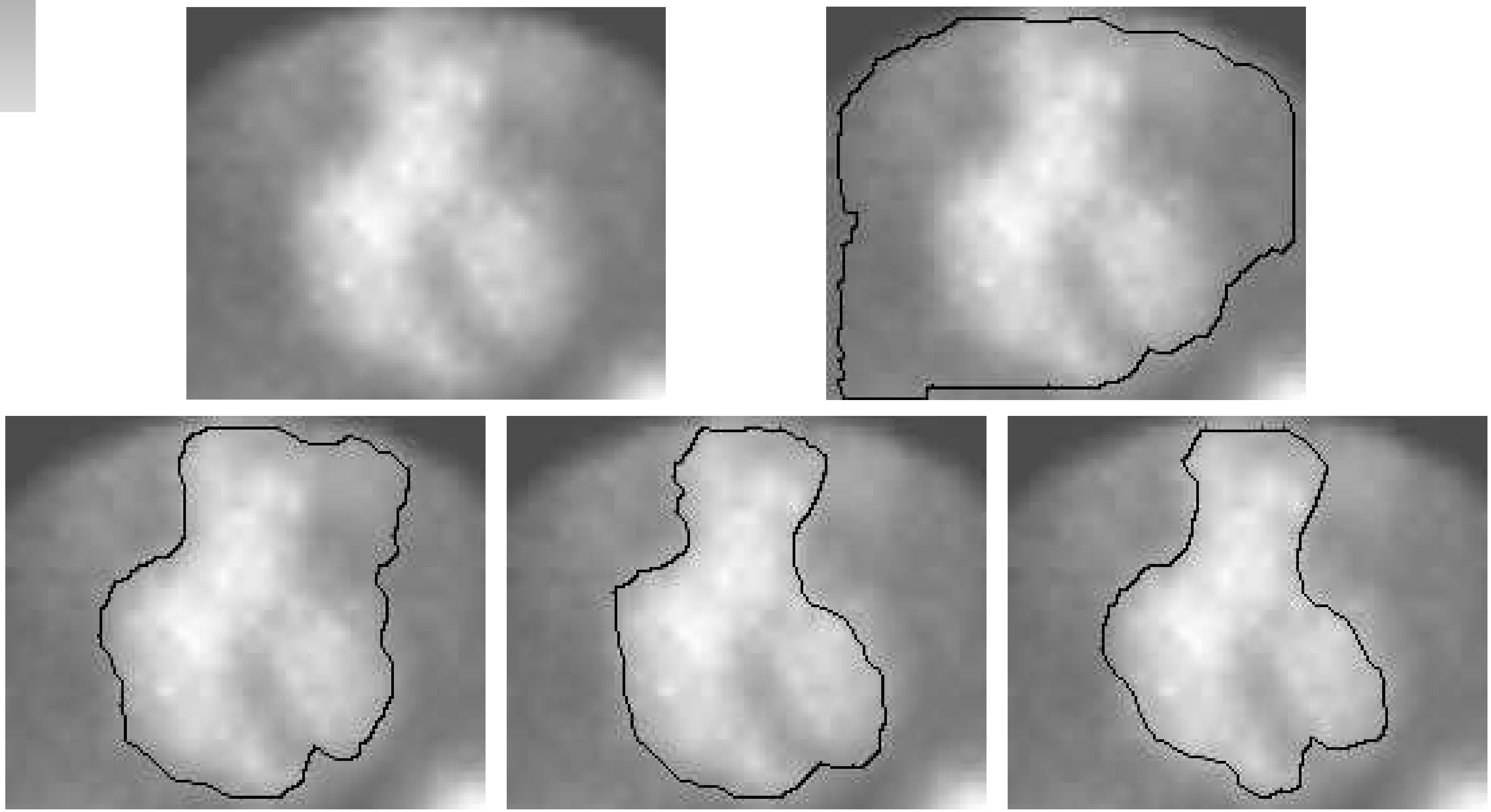
Diagram of the Climber Algorithm



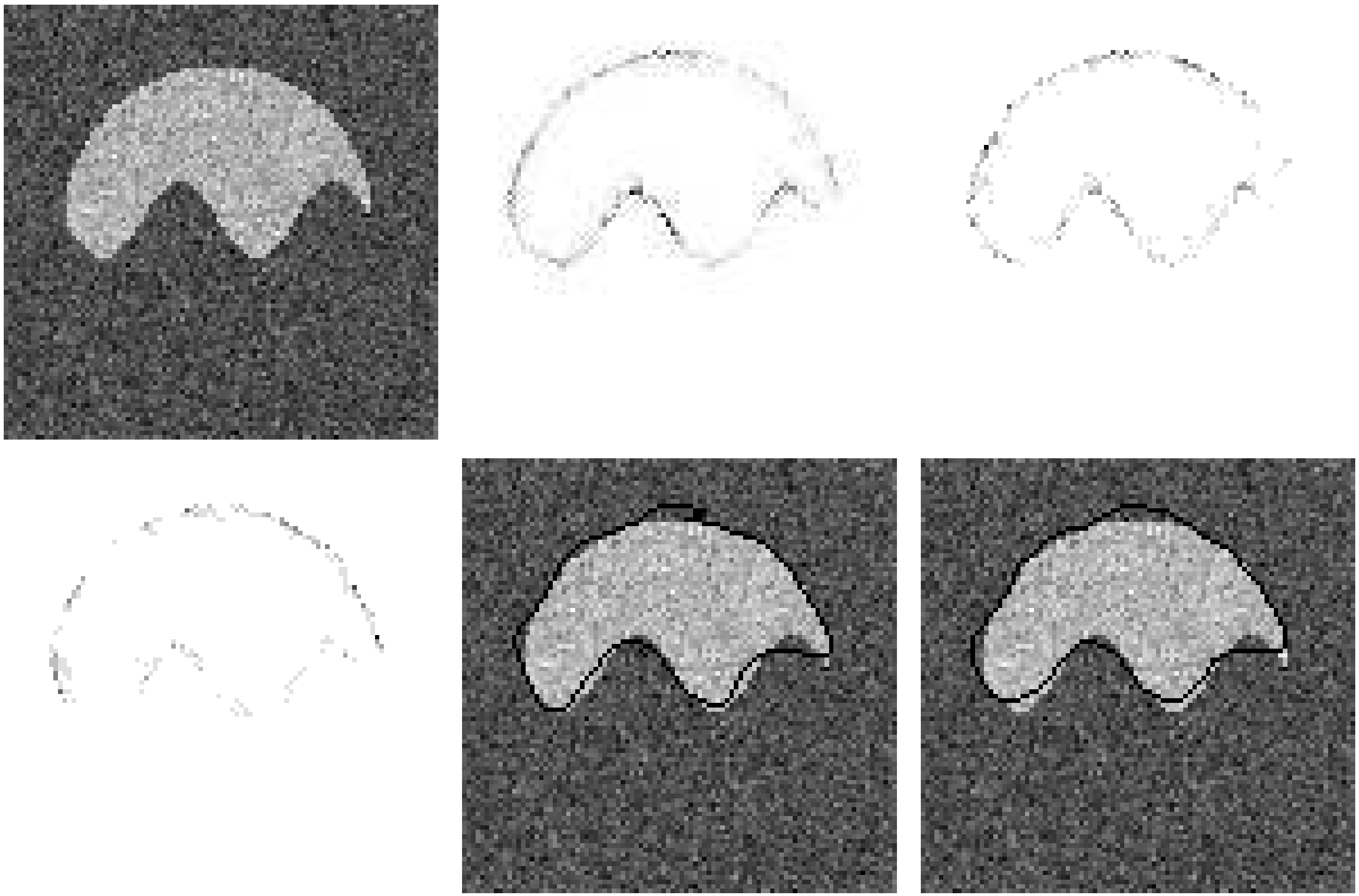
Occupation Measurement

- The occupation of a point is the number of times that the point is visited by all climbers.
- The occupation measurement can be normalized to be a probability.
- We only retain those points in the occupation measurement having large enough probability. They are likely to be the points of large gradient magnitude.
- These points are then linked into closed contours.

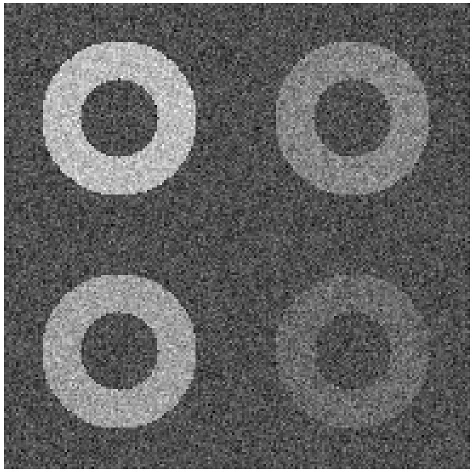
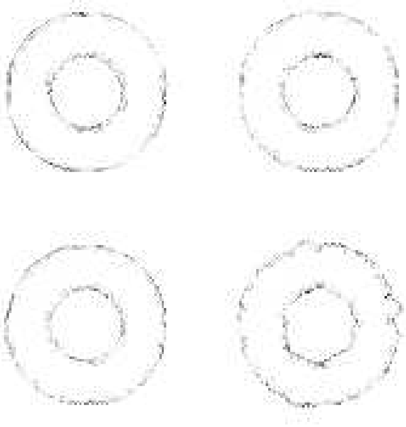
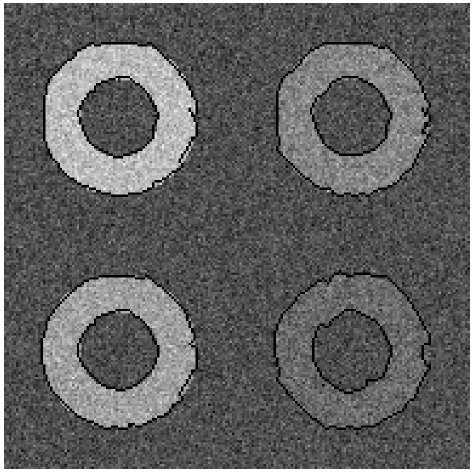
The Evolution of a Climber’s Contour



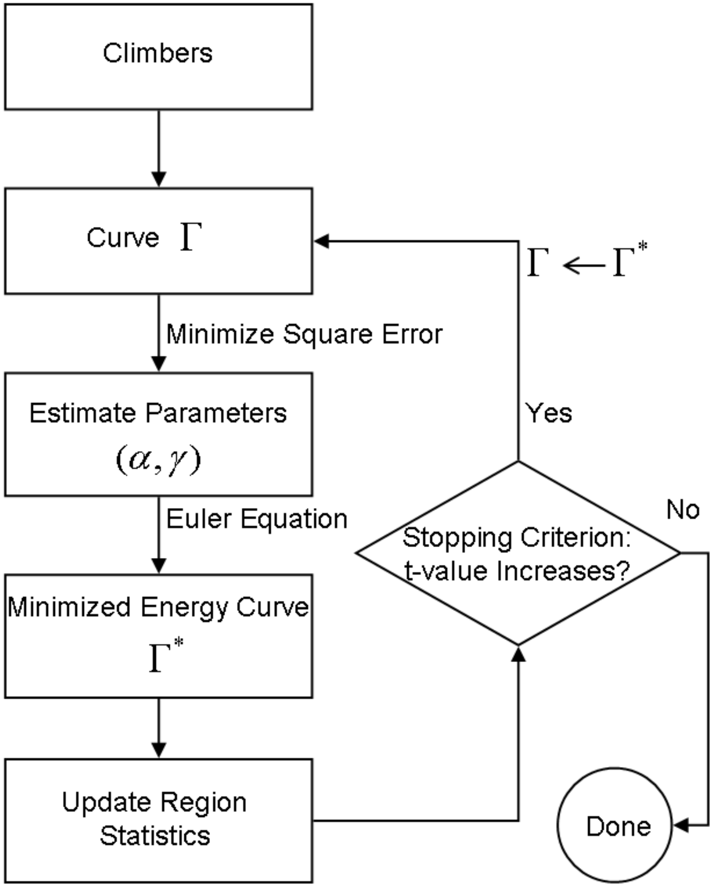
Number of Climbers vs Initial Contour



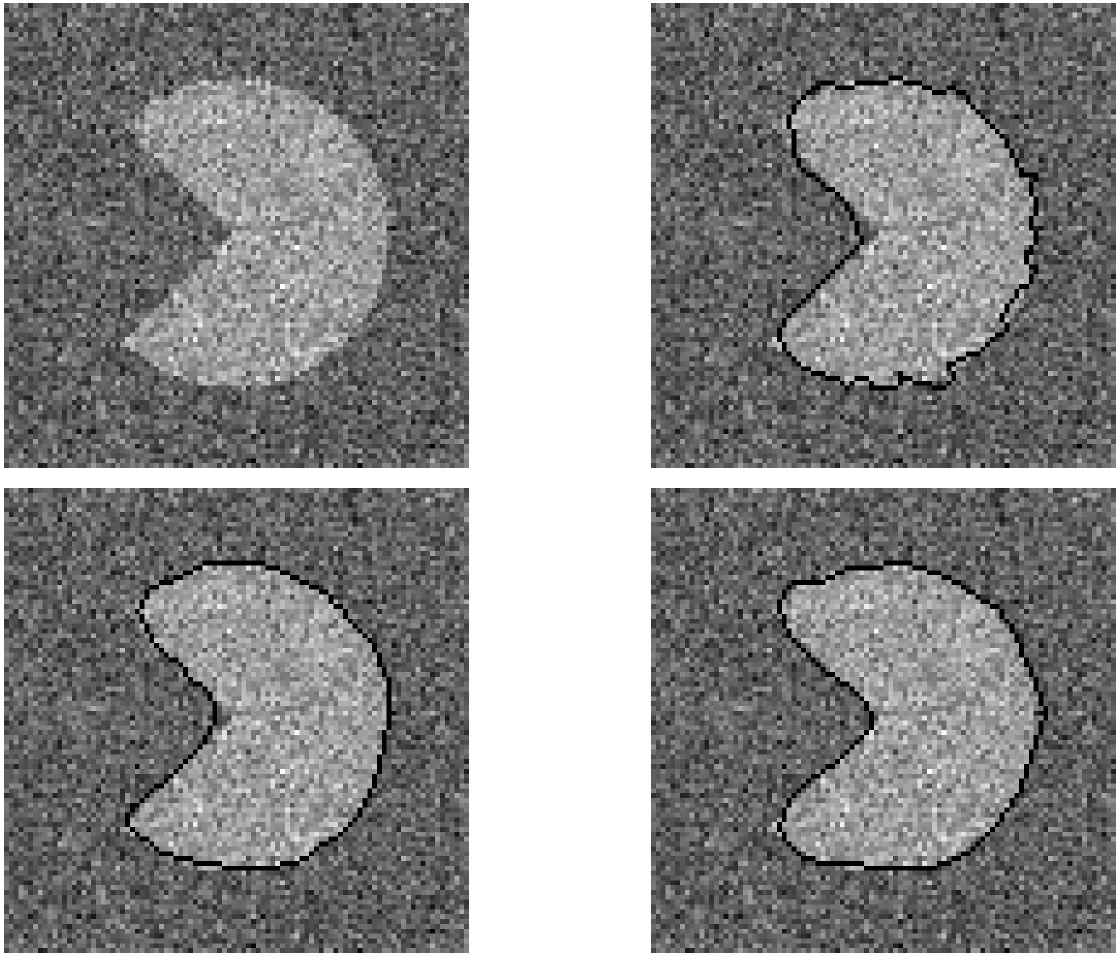
Detecting Multiple Contours

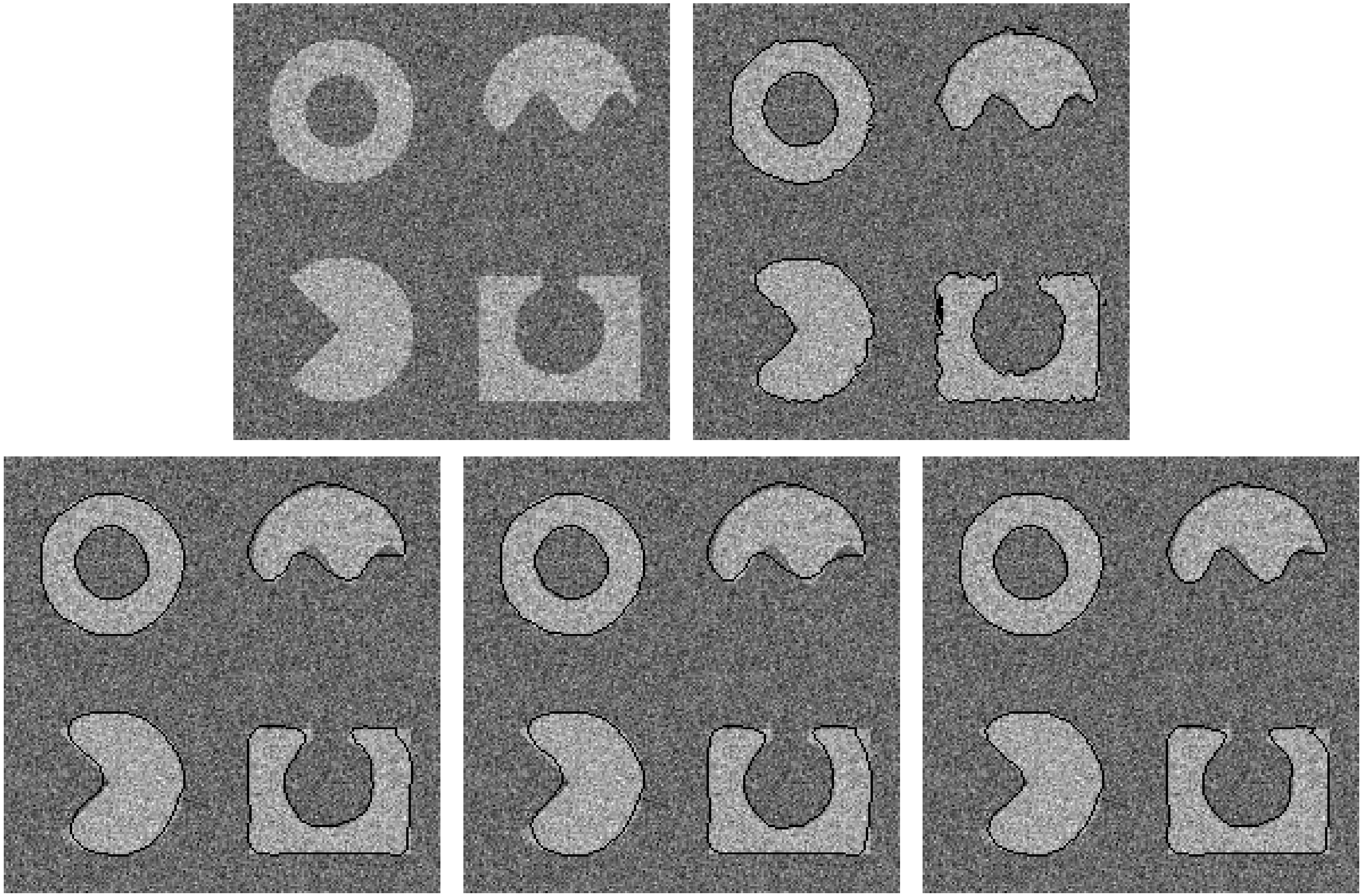
Spot Segmentation Algorithm



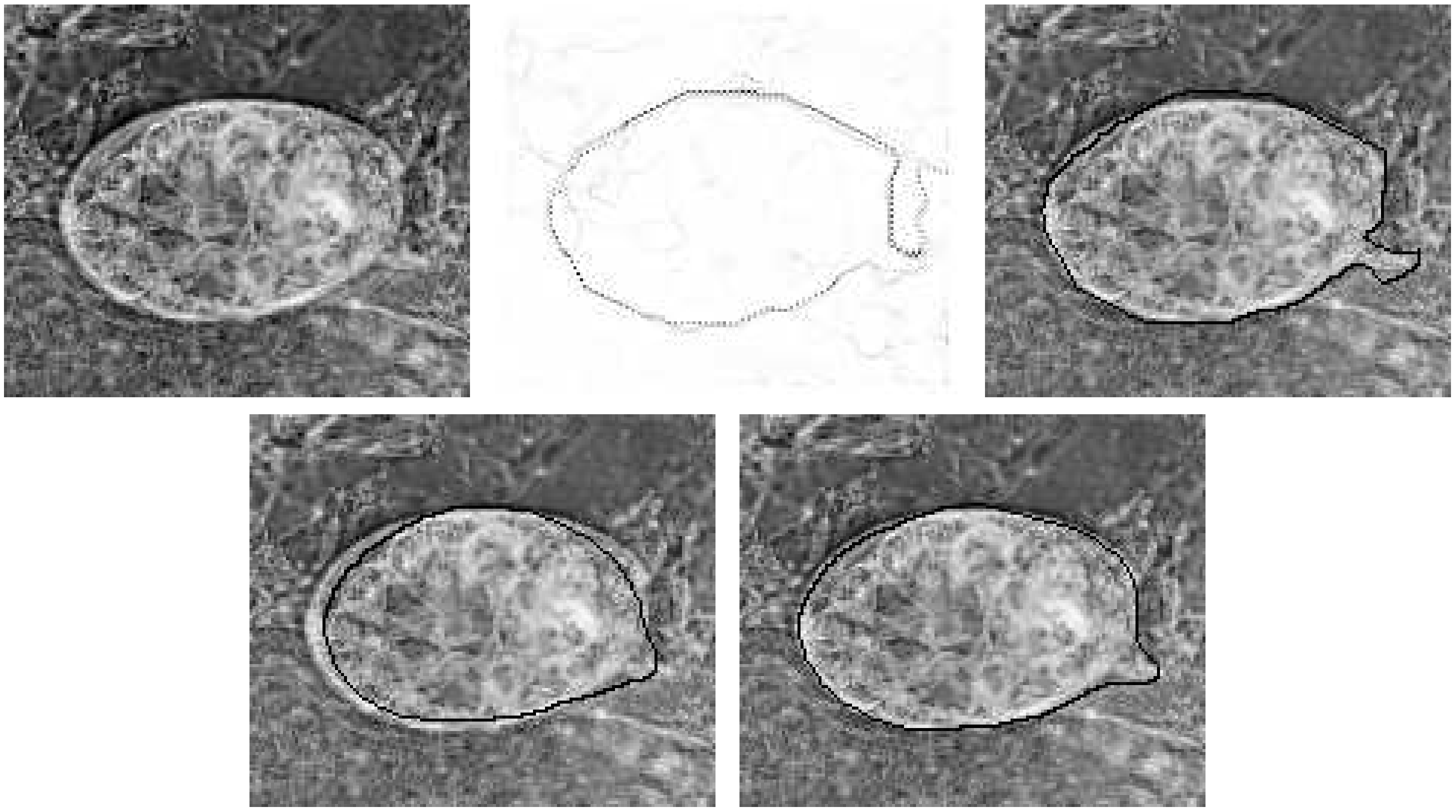
Two Iterations on a Noisy Image



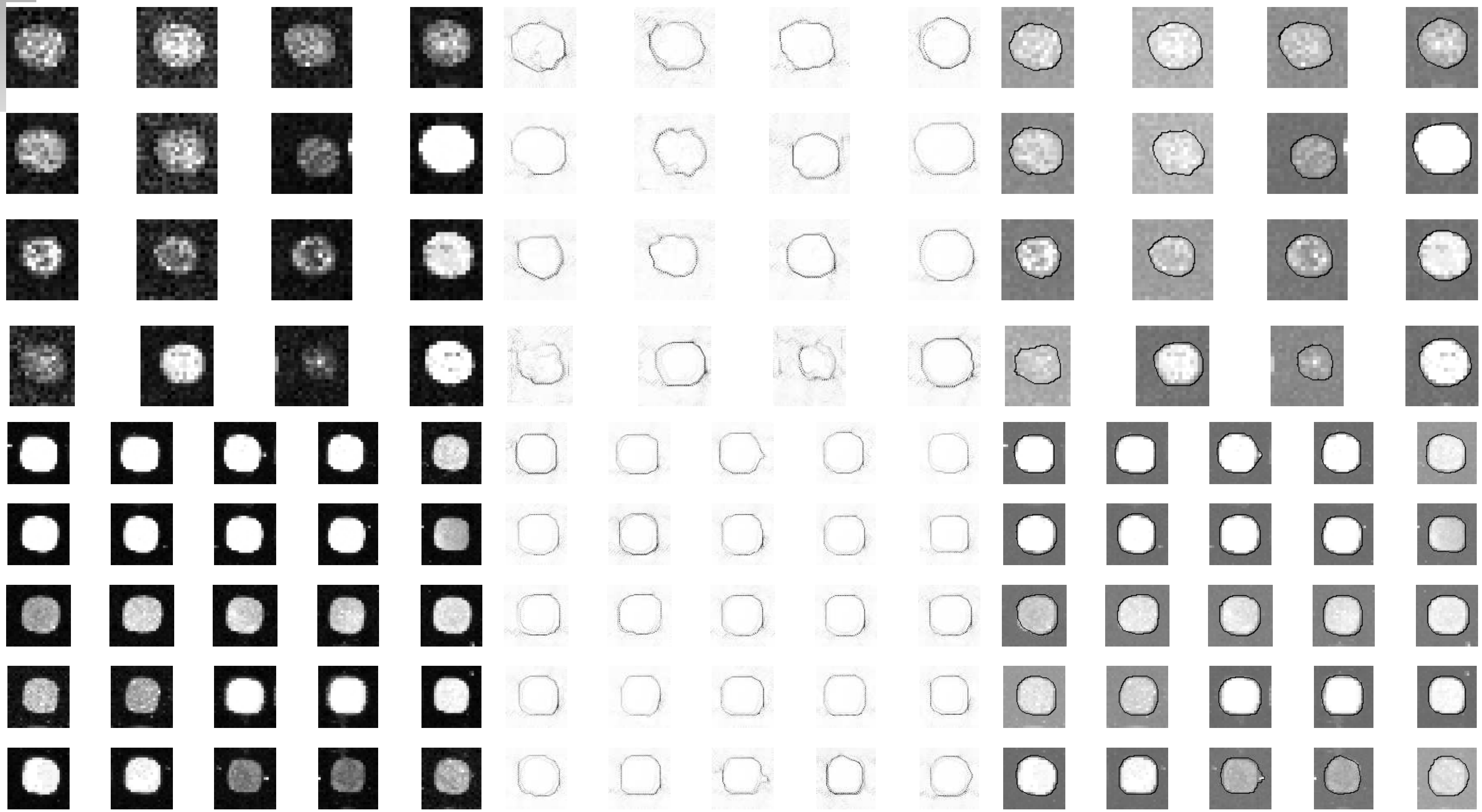
Compare with Snake with the same Initial Contour



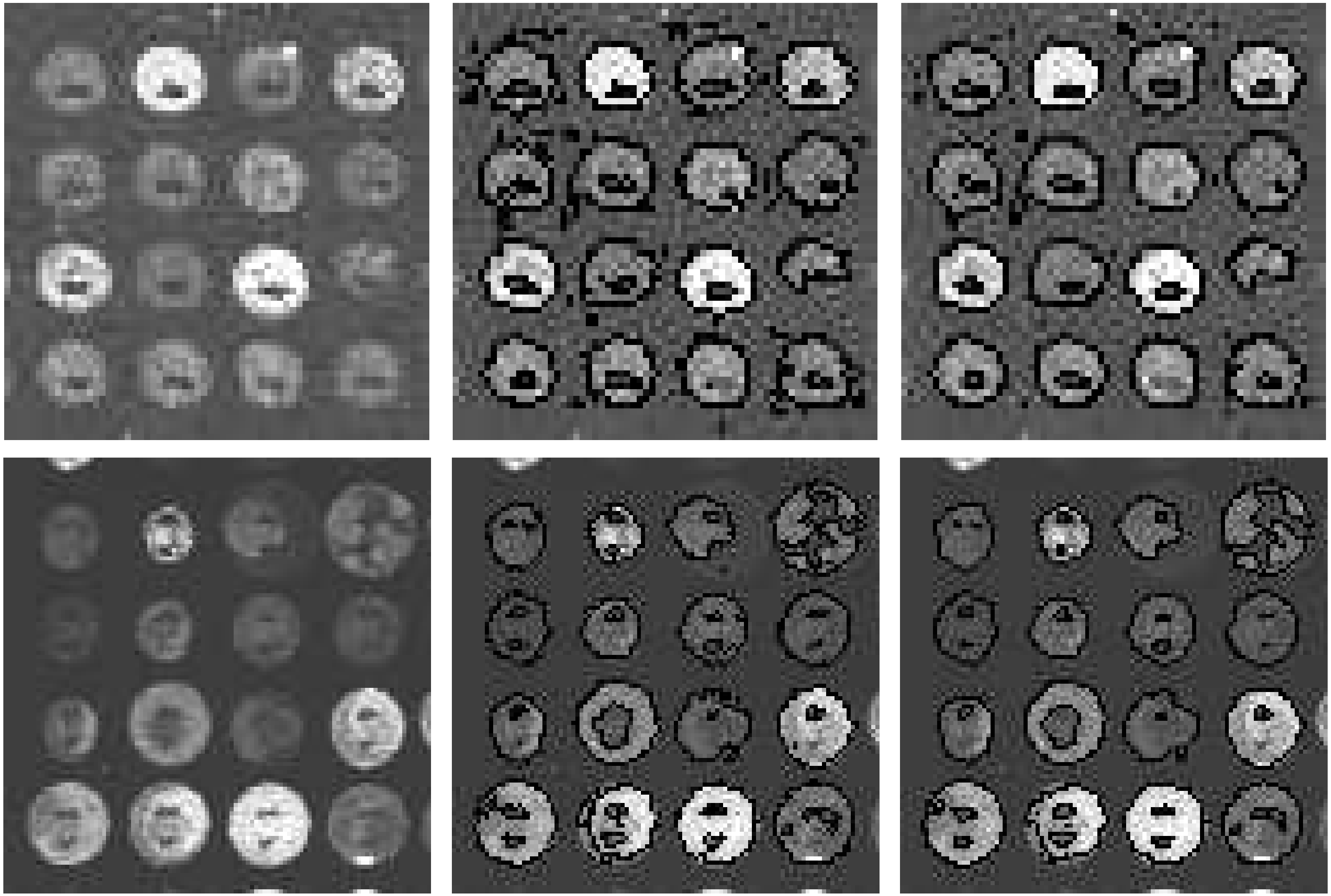
Egg Image



Top: cDNA($LC23N085$ in the SMD) and Bottom: Oligonucleotide

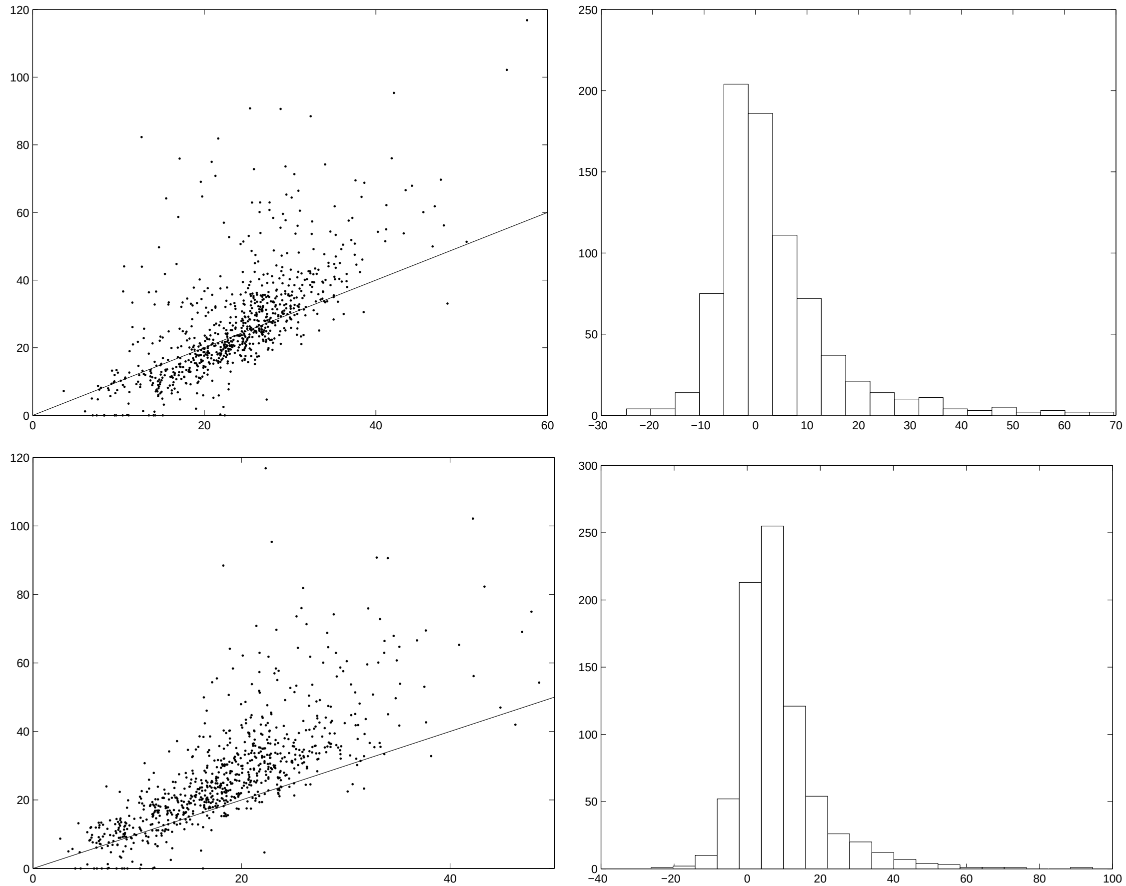


$LC23N085$ **(Top) and** $hp7004b$ **in the SMD**



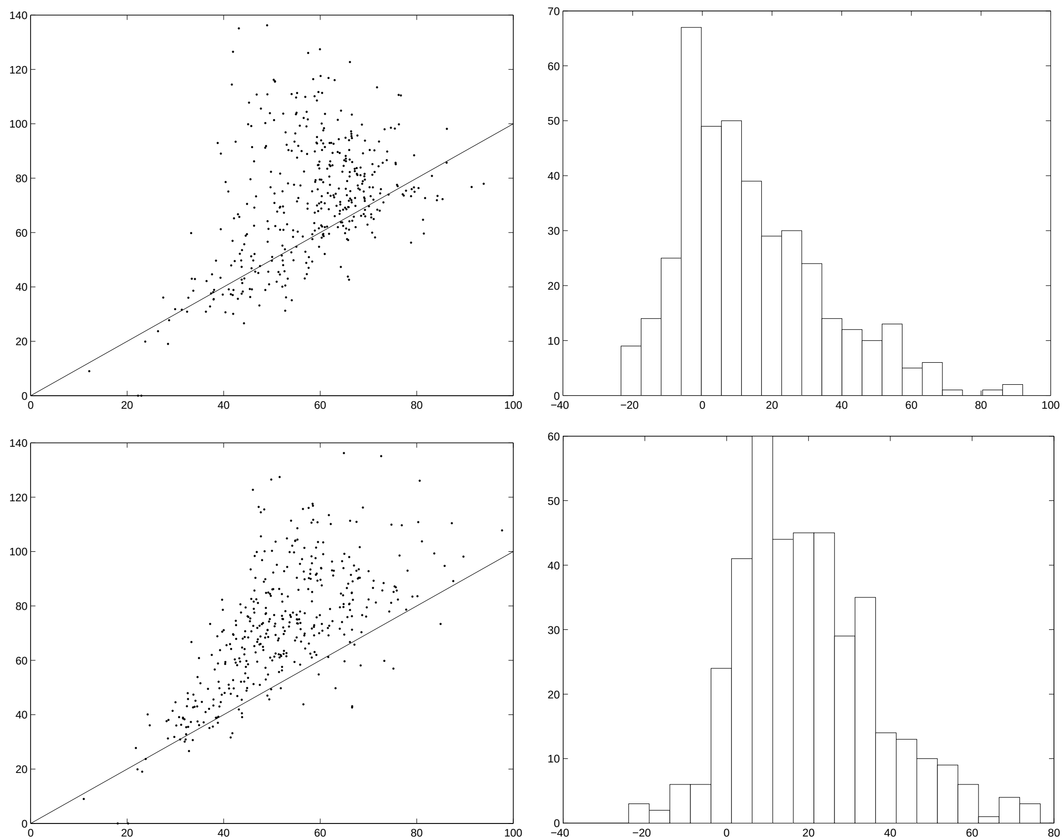
Spot *2.0 (Top) and with* GenePix Pro *5.0*

- Subblock $(2,1)$ of the $LC23N085$



Spot *2.0 (Top) and with* GenePix Pro *5.0*

- 400 spots of a subblock of an oligonucleotide microarray image.



Conclusion

- Integrate snake and Fisher criterion as an objective function to segment regions.
- Parameters and the contours of the objective function are determined without human intervention.
- The initial contour is estimated by climber algorithm. The climber algorithm is robust.
- Our method outperforms on spot segmentation task over commercial software, *Spot* 2.0 and *GenePix Pro* 5.0.