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

\[
E_{\text{snake}}(\Gamma) = \int_{\Gamma} \left( \frac{\alpha}{2} |\Gamma_s|^2 + \frac{\beta}{2} |\Gamma_{ss}|^2 - \| \nabla I \|^2 \right) ds,
\]

\[
E_{\text{region}}(\Gamma) = \left[ \iint_{R_1} (I - M_1)^2 dx dy + \iint_{R_2} (I - M_2)^2 dx dy \right] / (M_1 - M_2)^2.
\]

\[
E_{\text{total}}(\Gamma) = E_{\text{snake}}(\Gamma) + \tilde{\gamma} E_{\text{region}}(\Gamma).
\]

Setting \( \beta = 0 \) because it gives fourth derivative in Euler solution.
Euler Equation

• Given the parameters \( (\alpha, \gamma) \) in \( E_{\text{total}}(\Gamma) \), \([x \ y]\) on the boundary curve \( \Gamma^* \) that minimizes \( E_{\text{total}}(\Gamma) \) should satisfy:

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

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

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

\[
- \nabla \| \nabla I \|^2 - \alpha \kappa \vec{n} - \gamma' \left( \frac{(I - M_1)^2 - (I - M_2)^2}{(M_1 - M_2)^2} \right) \vec{n} = 0. \tag{3}
\]
A point on the optimal contour must satisfy (4) in the tangent direction ($\vec{t}$), and (5) in the normal direction ($\vec{n}$):

\[
\nabla \parallel \nabla I \parallel^2 \cdot \vec{t} = 0,
\]

\[
-\nabla \parallel \nabla I \parallel^2 \cdot \vec{n} = \alpha \kappa + \gamma' \left[ \frac{(I - M_1)^2 - (I - M_2)^2}{(M_1 - M_2)^2} \right].
\]

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.
• 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

- Initial point for a climber
  - Larger than a threshold $\|\nabla f\| > \theta$
    - Yes: Move point randomly
      - No: Move point in the tangent direction randomly
    - No: Temperature update Occupancy map update
  - No: Temperature Decrease to 0
    - Yes: Generate initial contour
      - No: Done

• 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
Two Iterations on a Noisy Image
Compare with Snake with the same Initial Contour
Egg Image
*Top:* cDNA($LC_{23}N_{085}$ in the SMD) and *Bottom:* Oligonucleotide
LC23N085 (Top) and hp7004b in the SMD
Compare our t-value (y-axis) with Spot 2.0 (Top) and with GenePix Pro 5.0

- Subblock (2, 1) of the LC23N085
• 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.