

Entropy Estimation for Segmentation of Multi-Spectral Chromosome Images

Wade Schwartzkopf

Brian L. Evans

Alan C. Bovik

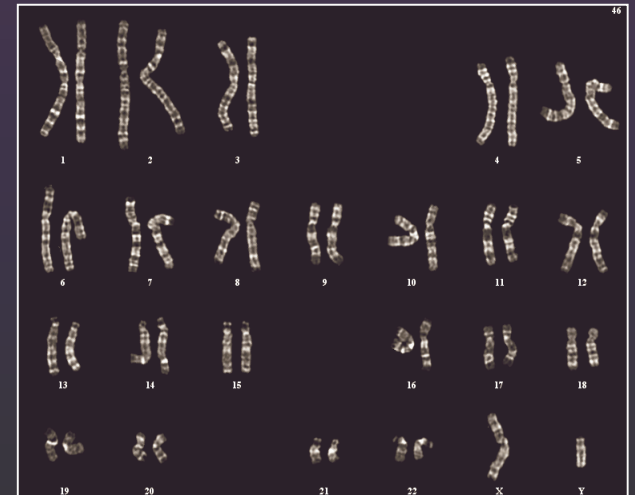
Dept. of ECE, UT Austin

SSIAI

April 9, 2002

Chromosomes

- Store genetic information
- Images reveal state of individual's health
- 46 human chromosomes form 24 classes
 - 22 different pairs
 - 2 sex chromosomes, X and Y
- Goal: Segment and classify each chromosome in image



M-FISH

- Multiplex Fluorescence In-Situ Hybridization [1996]
 - Five color dyes (fluorophores)
 - 32 possible combinations (2^5) to distinguish 24 human chromosome classes
- Each human chromosome class absorbs a unique combination of dyes

M-FISH Images

- Each pixel is a five-dimensional vector
- Each vector element gives contribution of a dye at pixel
- Chromosomal origin distinguishable at single pixel (unless overlapping)



Healthy Male

Multi-spectral Entropy-based Segmentation

- Decouple segmentation from classification
- Classify points in a cluster
- Calculate probabilities for each class
- Choose cut line that
 - Minimizes entropy within objects
 - Based on Shannon's definition of entropy:

$$H = -\sum_{i=1}^N p_i \log_2 p_i$$

p_i : probability of class i
 N : number of classes

Differential Entropy Estimation

- Estimate differential entropy directly from raw pixel values within objects
 - Avoid pixel classification
 - Use nearest neighbor entropy estimation

$$H_n = \frac{1}{n} \sum_{i=1}^n \ln(n\rho_{n,i}) + \ln 2 + C_E$$

$\rho_{n,i}$ is nearest neighbor distance between i and its nearest neighbor

C_E is the Euler constant

Example

- Original M-FISH multi-spectral image



M-FISH Image

Example: Step 1

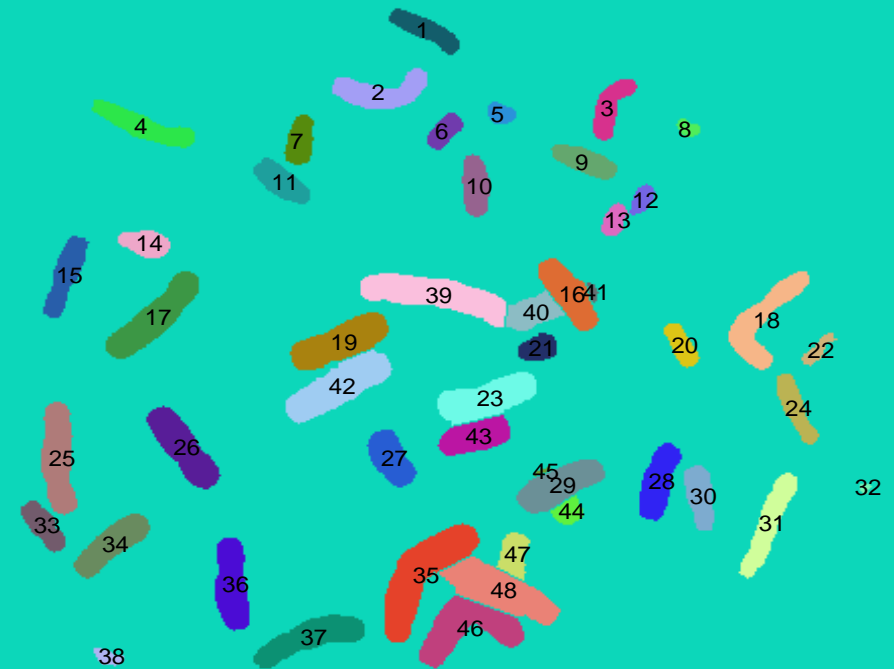
- Separate background and foreground (threshold DAPI plane for M-FISH)
- Label connected components
- Each color represents a connected component



Connected Components

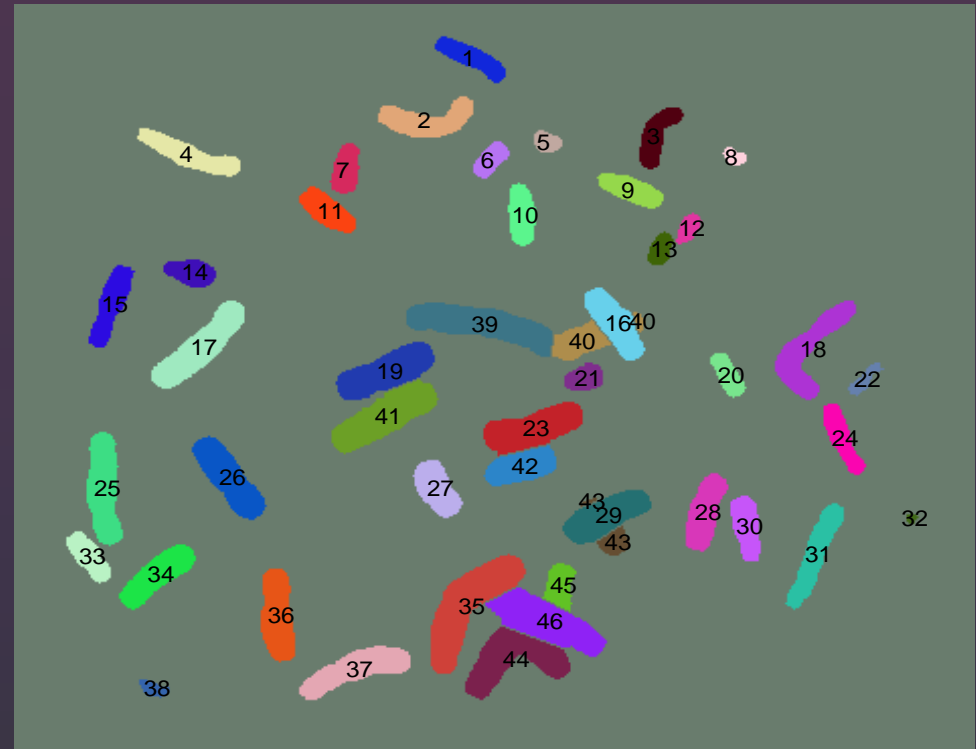
Example: Step 2

- Split touching chromosomes
- For all objects with entropy $>$ threshold,
 - Locate optimal cut line
 - Divide object if resulting entropy decreases
 - Examine each division for further divisions
- Each color represents a segmented object



Example: Step 3

- Merge parts of overlapped chromosomes: merge connected objects when overall entropy decreases
- Each color represents a segmented object



Results

- Adv. Digital Image Res. (ADIR) M-FISH database
 - Publicly available at <http://www.adires.com>
 - Close to 200 hand segmented M-FISH images
- Ten images tested
 - No improvement in performance over discrete entropy
 - Sensitive to algorithm parameters
 - Computationally more expensive than past entropy segmentation techniques

Conclusions

- No improvement in performance
 - Data is noisy but still classified well
 - Noisy data may have high entropy
 - May still be classified as the same class
 - More common than smooth data on classifier decision boundary
- Solution lies in combining chromosome segmentation and classification