

# Probabilistic segmentation and analysis of horizontal cells

Vebjorn Ljosa and Ambuj K. Singh  
 Center for Biimage Informatics and Dept. of Computer Science  
 University of California, Santa Barbara  
 ICDM 2006 proceedings p. 980–985

[www.liv.ac.uk](http://www.liv.ac.uk), Geoff Lewis, Mark Verardo

## Horizontal cells: Hard to segment

Rat's nest of neurites, many of which do not even belong to any of the cells in the image

Laser-scanning confocal (calretinin, neurofilament, calbindin)

[Geoff Lewis]

## Horizontal cells

- Patterns
  - Number of neurites
  - Neurite thickness
    - Do they taper?
  - Branching
  - Connection
- Response to injury
  - Grow more?
  - Branch more?
  - Different distribution of neurofilaments?
  - Connections

[webvision.med.utah.edu]

## Top-down approach: Visual vocabulary [ICDM 2005]

Step 1: Tile image

Step 2: Extract tile features {CS, CL, HT} MPEG-7

Step 3: ViVo generation ICA, PCA

Visual vocabulary

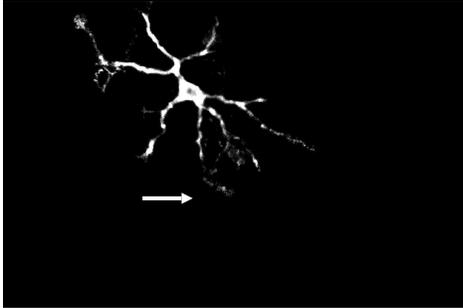
Dim<sub>2</sub>

Dim<sub>1</sub>

## Binary segmentation leads to large segmentation errors

Watershed algorithm

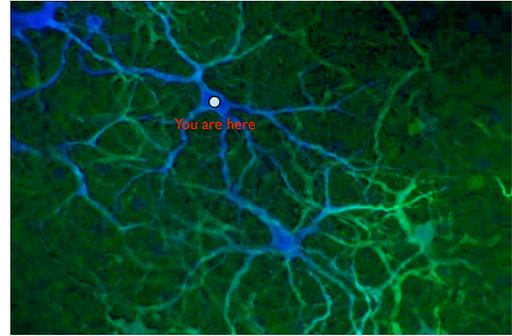
## Can we find a probabilistic segmentation?



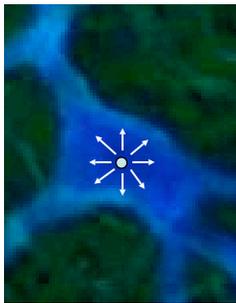
For each pixel, a probability that the pixel belongs to the cell

7

## Segmentation by repeated random walk

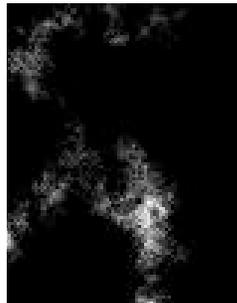


8



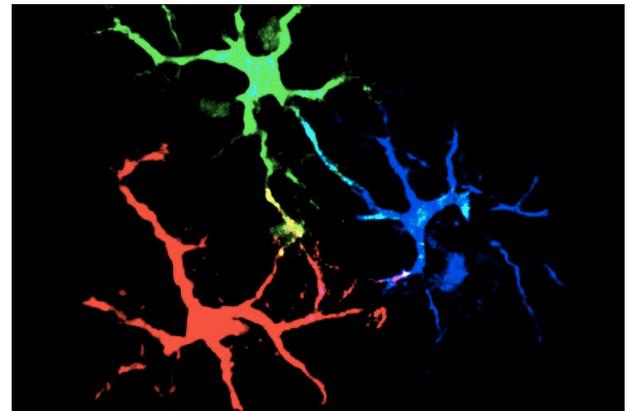
Random step in one of eight directions, biased by the intensity.

Return to start with probability  $c = 0.0001$



Result after 1000 steps.  
Repeat until convergence.

9



Result at convergence (3 cells)

10

## Eigenvector problem

$$\mathbf{x} := (1 - c)\mathbf{P}\mathbf{x} + c\mathbf{s}$$

Probabilistic mask

Transition matrix (very sparse)

Restart probability

Initial location

Define new transition matrix  $\mathbf{Q}$ :  
[Camoglu et al., Bioinformatics, 2006]

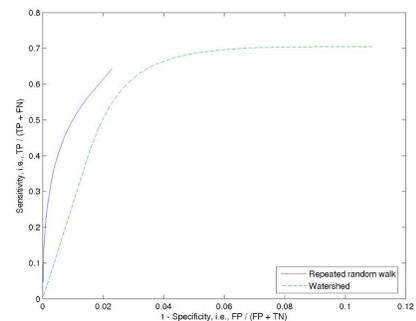
$$Q_{ij} = \begin{cases} (1 - c)P_{ij} & \text{if } s_i \neq 1 \\ (1 - c)P_{ij} + c & \text{if } s_i = 1. \end{cases}$$

At convergence,  $\mathbf{x} = \mathbf{Q}\mathbf{x}$ .  
Largest eigenvalue is 1.

Corresponding eigenvector is the probabilistic mask (stationary probability of the Markov chain).

11

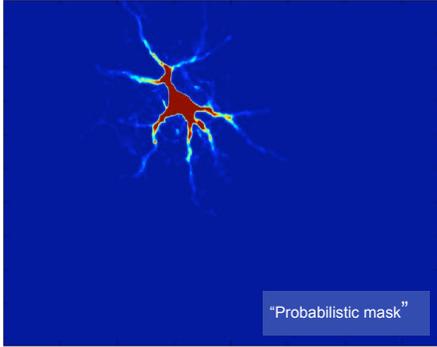
## Better sensitivity and specificity than watershed



... but that is not the main point

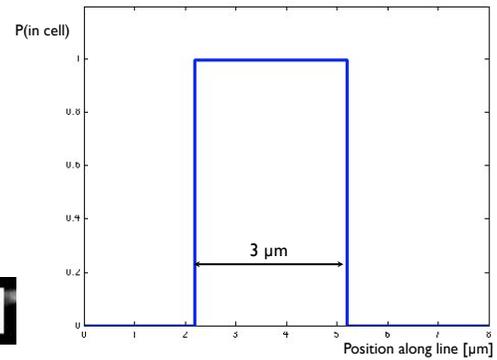
12

## Cell segmentation as a probabilistic value



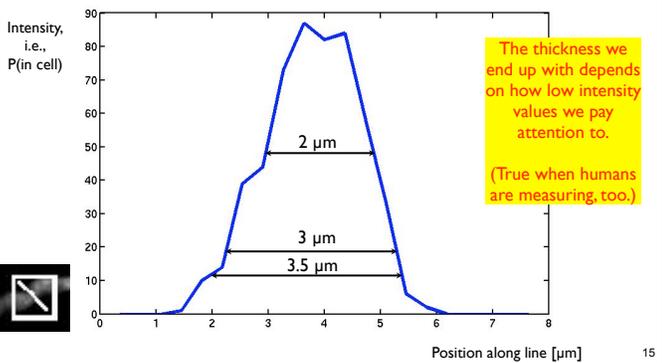
13

## How does this affect analysis and mining? A simple example: How to measure neurite thickness?



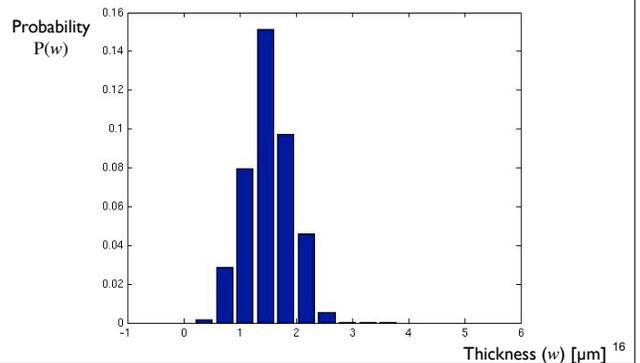
14

## How to measure neurite thickness?



15

## Probabilistic thickness measurement



16

## Summary

- Cannot always have perfect data
  - The curse of research biology
  - Must make use of images we cannot segment perfectly
- Probabilistic segmentation
  - Better segmentation result, more information
  - Must revamp analysis, mining, and database techniques
- How does this affect higher-level analysis and mining?
- One step closer to bringing data-driven research to microscopy-based fields