Bioimage Informatics

Lecture 9, Spring 2012

Bioimage Data Analysis (II):

Applications of Point Feature Detection Techniques
Outline

• Review: subpixel resolution point feature detection
• Other point feature detection techniques
• Reproducible research in computational science
• Application I: molecule counting
• Application II: single molecule imaging
• Application III: protein colocalization analysis
• Application IV: super-resolution imaging
• Review: subpixel resolution point feature detection

• Other point feature detection techniques

• Reproducible research in computational science

• Application I: molecule counting

• Application II: single molecule imaging

• Application III: protein colocalization analysis

• Application IV: super-resolution imaging
Sub-Pixel Resolution Point Detection

Figure 1
Sub-pixel Detection by Gaussian Fit

- Fit using a Gaussian kernel, which represents the ideal image of a point

\[
K(x, y; x_0, y_0) = K(x - x_0, y - y_0) = A \cdot \exp \left[ -\frac{(x - x_0)^2 + (y - y_0)^2}{B} \right]
\]

- Problem formulation: to minimize the difference between the translated kernel and the image

\[
\min_{(x_0, y_0) \in R^2} \left| I(x, y) - K(x, y; x_0, y_0) \right|
\]
Gaussian Fitting Implementation Details (I)

- Calculation of $|I(x, y) - K(x, y; x_0, y_0)|$

$$E(x_0, y_0) = \sum_{i=1}^{M} \sum_{j=1}^{N} |I(i, j) - K(i - x_0, j - y_0)|$$

May not be integer coordinates

- Intensity interpolation using interp2
Gaussian Fitting Implementation Details (II)

• How to set the Gaussian kernel
  - By fitting the Airy Disk using a Gaussian
    \[ \sigma = \frac{0.61 \cdot \lambda}{NA} / 3 \]
  - By measuring PSF (often using beads) then fitting with a Gaussian

• Spatial sampling: three-times oversampling of Airy disk
  \[ \frac{\text{Airy disk radius}}{\text{pixel size}} \geq 3 \]

• Under high SNR, spatial sampling may be relaxed to 2~2.5.
Gaussian Fitting Implementation Details (III)

• Optimization strategy I: exhaustive search

• Implementation of exhaustive search
  - Oversampling the kernel and images
    → use a small pixel size: e.g. 10nm
  - If multiple minima were identified, use their average position
Gaussian Fitting Implementation Details (IV)

• Initialization:
  - Use detected local maxima to localize the search

• Strengths and weaknesses
  - Good (not necessarily optimal) solution guaranteed
  - Computationally expensive, slow
Gaussian Fitting Implementation Details (V)

- Optimization strategy II: optimization search
  - Many optimization techniques can be applied

- Initialization:
  - Use detected local maxima
  - Use multiple randomly selected initiation points
Gaussian Fitting Implementation Details (VI)

• **Strengths and weaknesses**
  - Not limited by oversampling rate
  - May be trapped in local minimum
Sub-pixel Detection by Correlation

- Detection by maximization correlation

\[ C_{x_0,y_0} = \sum_{i=1}^{M} \sum_{j=1}^{N} I_{i,j}K(i-x_0, j-y_0) \]

- Often the correlation function is normalized

\[ C_{x_0,y_0} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} I_{i,j}K(i-x_0, j-y_0)}{\sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} I_{i,j}^2} \cdot \sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} K^2(i-x_0, j-y_0)}} \]

- Same strategy as in Gaussian fitting as the only difference is the cost function

  - Strategy I: exhaustive search
  - Strategy II: optimization search
Gaussian Fitting vs Correlation

• For point features, Gaussian fitting is the best method overall.

• For larger non-diffraction limited features, correlation gives better resolution.
Limitations of Sub-Pixel Detection

- When the distance between the two point features goes below the Rayleigh limit, they can no longer be resolved reliably unless under very high SNR.
• Review: subpixel resolution point feature detection

• **Other point feature detection techniques**

• Reproducible research in computational science

• Application I: molecule counting

• Application II: single molecule imaging

• Application III: protein colocalization analysis

• Application IV: super-resolution imaging
Other Point Feature Detection Techniques

- Corners rather than individual points are used more often in computer vision.

- A variety of techniques are also available from medical image analysis.

• Review: subpixel resolution point feature detection

• Other point feature detection techniques

• **Reproducible research in computational science**

• Application I: molecule counting

• Application II: single molecule imaging

• Application III: protein colocalization analysis

• Application IV: super-resolution imaging
Open Source & Reproducible Research

An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.
—D. Donoho (http://www-stat.stanford.edu/~donoho/)

• Jon Claerbout is often credited as the first who proposed reproducible computational research.

• There are challenges. But these challenges can be overcome.

• Methods for public-funded biological studies should be open-source.

http://reproducibleresearch.net/index.php/Main_Page
http://sepwww.stanford.edu/data/media/public/sep/jon/
Open Source & Reproducible Research (II)

• Current literatures of image processing and computer vision often are formulated mathematically and do not provide source code.

• Challenges
  - implementation (numerical issues)
  - parameter tuning
  - robustness a major performance issue
• Review: subpixel resolution point feature detection
• Other point feature detection techniques
• Reproducible research in computational science

• **Application I: molecule counting**
• Application II: single molecule imaging
• Application III: protein colocalization analysis
• Application IV: super-resolution imaging
Overview of Cell Cycle

Interphase

- Prophase
- Prometaphase
- Metaphase
- Anaphase
- Telophase

Mitosis

Metaphase-to-anaphase transition

Cytokinesis

DNA replication

Interphase
Dynamic Microtubules in the Mitotic Spindle

Green: microtubule
Red: kinetochore

Alberts et al., MBoC5

25nm

5 μm

microtubule
Molecule Counting by Ratio Imaging

• Review: subpixel resolution point feature detection

• Other point feature detection techniques

• Reproducible research in computational science

• Application I: molecule counting

• **Application II: single molecule imaging**

• Application III: protein colocalization analysis

• Application IV: super-resolution imaging
Confirmation of Poleward Flow of Spindle Microtubules

5 μm

Cameron et al, JCB, 173:173-179, 2006
Fluorescent Speckle Microscopy (FSM)
FSM of Dynamic Spindle Architecture

Fluorescent speckle microscopy
Quantitative Mapping of Spatial-Temporal Spindle Dynamics

Analyzing Dynamics of Individual Microtubules by Single Fluorophore FSM

• Review: subpixel resolution point feature detection
• Other point feature detection techniques
• Reproducible research in computational science
• Application I: molecule counting
• Application II: single molecule imaging

• Application III: protein colocalization analysis
• Application IV: super-resolution imaging
Protein Colocalization Analysis

- To determine whether protein molecules colocalize is essential to many biological studies.

- Protein colocalization analysis can be performed in both fixed and live cells.

Protein Colocalization Analysis (II)

- A commonly used analysis protocol
  - Step 1: Detect point features in a selected channel.
  - Step 2: Identify the intensities at another channel at detected point feature locations.
  - Step 3: Calculate correlation coefficient

- Commonly used definitions of colocalization coefficients

  - Overlap coefficient
    \[ r_o = \frac{\sum A_i \times B_i}{\sqrt{\sum A_i^2 \times \sum B_i^2}} \]

  - Pearson’s coefficient
    \[ r_p = \frac{\sum (A_i - a) \times (B_i - b)}{\sqrt{\sum (A_i - a)^2 \times \sum (B_i - b)^2}} \]
Some General Comments

• It is possible but limiting to consider bioimage analysis as just another application.

• Excellent research opportunities in bioimage informatics.

• Challenges
  - Solid training in image processing and computer vision
  - Interdisciplinary background and thinking
    - For identifying and solving problems
    - For collaboration
• Review: subpixel resolution point feature detection
• Other point feature detection techniques
• Reproducible research in computational science
• Application I: molecule counting
• Application II: single molecule imaging
• Application III: protein colocalization analysis

• **Application IV: super-resolution imaging**
Imaging as a Tool for Studying Cellular Structure

- **Crystallography, NMR, spectroscopy**
  - Resolution: ≤1 nm
  - Live/physiological condition: **NO**
  - Samples must be specially prepared

- **Electron microscopy**
  - Resolution: between 1nm & 100nm
  - Live/physiological condition: **NO**
  - Samples must be fixed

- **Light microscopy**
  - Resolution: ≥ 100nm
  - Live/physiological condition: **YES**
Performance Metrics of a Fluorescence Microscope

- **Resolution:**
  - Rayleigh limit: \[ D = \frac{0.61\lambda}{NA} \]
  - Sparrow limit: \[ D = \frac{0.47\lambda}{NA} \]

- **Numerical aperture (NA)**
  \[ NA = n \cdot \sin \mu \]

  \( n \): refractive index of the medium between the lens and the specimen

  \( \mu \): half of the angular aperture

- Water \( n = 1.33 \)
- Immersion oil \( n = 1.51 \)
Stochastic Optical Reconstruction Microscopy (STORM)


Photoactivation Localization Microscopy (PALM)

Fig. 1. The principle behind PALM. A sparse subset of RFP molecules that are attached to proteins of interest and then fixed within a cell are activated (A and B) with a brief laser pulse at $\lambda_{act} = 405$ nm and then imaged at $\lambda_{em} = 563$ nm until most are bleached (C). This process is repeated many times (C and D) until the population of unbleached, unbinned molecules is depleted. Summing the molecule images across all frames yields a diffraction-limited image (E and F). However, if the location of each molecule is first determined by fitting the expected molecular image given by the PSF of the microscope (G), centered to the actual molecular image (E), left, the molecule can be plotted (G, right) as a Gaussian that has a standard deviation equal to the uncertainty $\sigma_x$ in the fitted position. Repeating with all molecules across all frames (A through D) and summing the results yields a superresolution image (E and F) in which resolution is dictated by the uncertainty $\sigma_x$ as well as by the density of localized molecules. Scale: 1 $\mu$m in (G) and (F), 4 $\times$ 4 $\mu$m elsewhere.
Questions?