Detection of Sub-resolution Dots in Microscopy Images

Karel Štěpka, 2011
Centre for Biomedical Image Analysis, FI MU
supervisor: doc. RNDr. Michal Kozubek, Ph.D.

Outline

- Fluorescence microscopy
- Image degradations
- Evaluation of analysis
- Existing approaches to dot detection
- Further Work

Fluorescence Microscopy

Fluorescence Microscope

- Light source
- Excitation filter
  - Allows only the excitation part of the spectrum to pass through
- Sample
  - Absorbs incoming light
  - Emits light with a lower frequency (fluorescence)
- Emission filter
  - Allows only the emission part of the spectrum to pass through
- Sensor

Fluorescence In-Situ Hybridization

- Allows to stain individual chromosomes or their parts
- Probes appear as small dots in the result

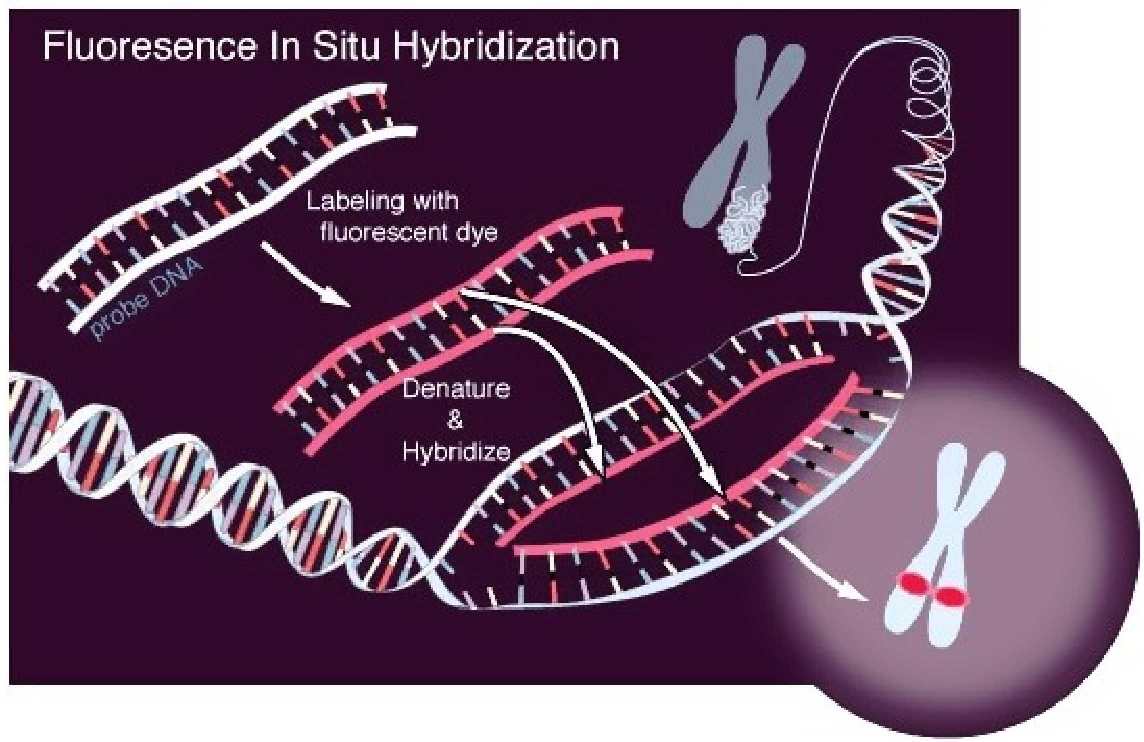


Image courtesy of Wikimedia Commons

Observable Parts of a Cell

- Cytoplasm
- Cytoskeleton
- Nucleus
- Whole chromosomes
  - Conditions related to the number of chromosomes (e.g. Down syndrome)
- Telomeres
- Kinetochores
- Individual genes
  - Translocations (e.g. BCL/ABR genes and their relation to certain kinds of leukemia)

Observable Parts of a Cell – Dots

- Cytoplasm
- Cytoskeleton
- Nucleus
- Whole chromosomes
  - Conditions related to the number of chromosomes (e.g. Down syndrome)
- Telomeres
- Kinetochores
- Individual genes
  - Translocations (e.g. BCL/ABR genes and their relation to certain kinds of leukemia)

Fluorescence Dots

- Real size on the order of 10 nm
- In the resulting image, often 1 pixel > 60 nm
- Because of the diffraction limit of visible light, the magnification cannot be easily improved

- Due to image degradations, the sensor detects a blurred image of the dot
- Image of a dot has a few pixels across

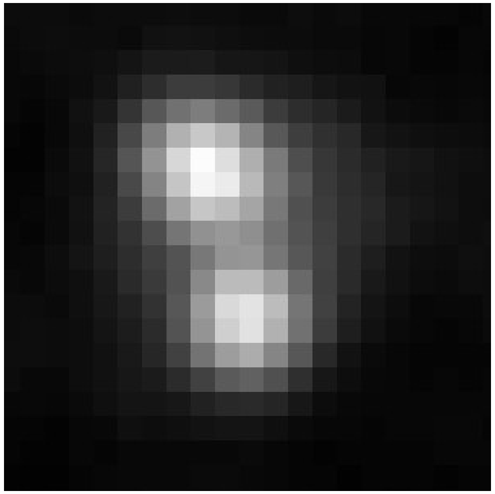


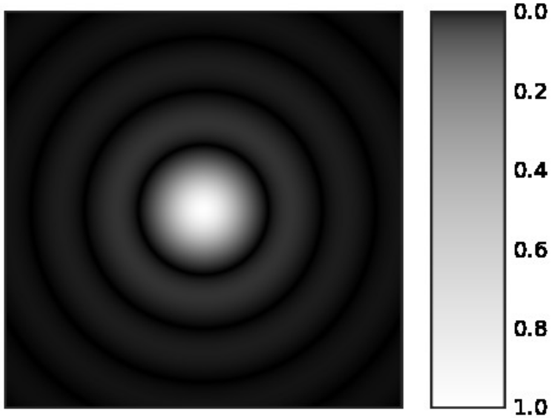
Image Degradations

Types of Image Degradation

- Noise
  - Many kinds, with different causes and statistical distributions:
    - Photon shot noise (Poisson)
    - Impulse noise (often fixed pattern)
    - Readout noise (Gaussian)
    - Dark current noise
    - Laser speckle noise
  - Can be suppressed using various methods
    - Dark frame subtraction
    - Gaussian blurring
    - Non-linear filters (median, non-linear diffusion)

Types of Image Degradation

- Degradation by point spread function (PSF)
  - Every optical system has a characteristic PSF
  - Describes scattering of photons travelling through individual components of the system
  - Even in an ideal optical system, a point light source produces signal equivalent to the Airy disk



  - PSF can be experimentally measured
  - Degradation can be suppressed using deconvolution

Types of Image Degradation

- Chromatic aberration
  - Different wavelengths have different refractive index
- Field curvature
  - Sensor is planar, but the focal area is curved
- Spherical aberration
  - Related to the shape of the lens
- Degradations related to sensor technology
  - Smear in CCD chips

Evaluation of Analysis

Measures to Consider

- Detection
  - $\text{precision} = \dfrac{TP}{TP + FP}$
  - $\text{recall} = \dfrac{TP}{TP + FN}$
  - $\text{F1 score} = \dfrac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$

	present	not present
found	TP	FP
not found	FN	TN

- Distinguishing between large dots and double-dots
  - To identify chromosomal conditions such as polysomy

Measures to Consider

- Localization
  - Absolute position
    - To determine the number of dots inside/outside the nucleus
  - Relative position of individual signals
    - To identify chromosomal translocations
  - Mean squared error

- Overall intensity
  - To determine the amount of fluorescent dye or protein
  - Mean squared error

Evaluation of Analysis

- Comparison of the results with the ground truth (GT)
  - We can obtain GT by manually annotating real images
  - We can generate synthetic (simulated) images together with their GT
- Real testing data, manual GT
  - Different people, or the same person over multiple attempts, generally annotate images differently
  - Time consuming, expensive
- Synthetic testing data, generated GT
  - GT is accurate and undebatable (created before the images)
  - The synthetic data must correspond to the real images

Existing Approaches to Dot Detection

“Classical” Detection Methods

- Thresholding
  - Fixed
  - Otsu
  - Unimodal
  - Adaptive
- Mathematical morphology
  - Top-hat transform

Recent "Classical-Based" Methods

- EMax
  - Extended maxima transform, size-based filtering
- Gué
  - Top-hat, thresholding, region growing, morphological closing and opening
- HDome
  - HDome transformation, mean shift clustering, cluster filtering

Recent “Classical-Based” Methods

- Kozubek
  - Gradual thresholding, size-based filtering
- Netten
  - Top-hat, dot label (“sweep” through all intensity levels)
- Raimondo
  - Top-hat, modified unimodal thresholding, pattern matching (using a model of a dot)

Machine Learning Approach

- Examine all potential dot locations and classify them as positive/negative
  - Usually using a sliding sub-window
- Training is required, overtraining is undesirable
  - Training set contains image patches from which the classifier learns
    - Positive examples
    - Negative examples
  - Test set is used to determine the quality of the classifier
  - Ideally, $\mathit{training\_set} \cap \mathit{test\_set} = \emptyset$
  - We train on the training set, until the results on the test set are satisfactory

Machine Learning Approach

- Neural networks
  - Multilayer perceptron
  - Each input neuron corresponds to one pixel
- AdaBoost
  - Haar-like features used for weak classifiers
  - Combines several weak classifiers into one strong
  - Computationally intensive in 3D
- Fischer discriminant analysis
  - Computationally intensive in 3D

Recent Survey by I. Smal et al.

- Compared performance of several methods (including machine learning)
- 2D data
  - Real images
  - Simplified synthetic images
    - Dots represented by Gaussian profiles
- Did not evaluate the influence of method parameters
- Good starting point

Ihor Smal et al.: Quantitative Comparison of Spot Detection Methods in Fluorescence Microscopy. *IEEE Transactions on Medical Imaging* 29(2): 282–301 (2010)

Parametrization – No Size Fits All

- No method can be used on all types of images without any adjustments
- On the data/pixel level, images can be very different, even when displaying the same class of objects
  - Noise level
  - Base intensity
  - Dynamic range
  - Contrast
  - Background (non-)uniformity
  - Illumination artifacts
  - Amount of objects of interest

Parametrization – Usability

- Usability of a method depends on:
  - Number of its parameters
  - Sensitivity to parameter changes
  - Intuitiveness of its parameters for the end user

- A thorough parametric study is required
- Curse of dimensionality
  - Some of the methods have 4–6 free parameters

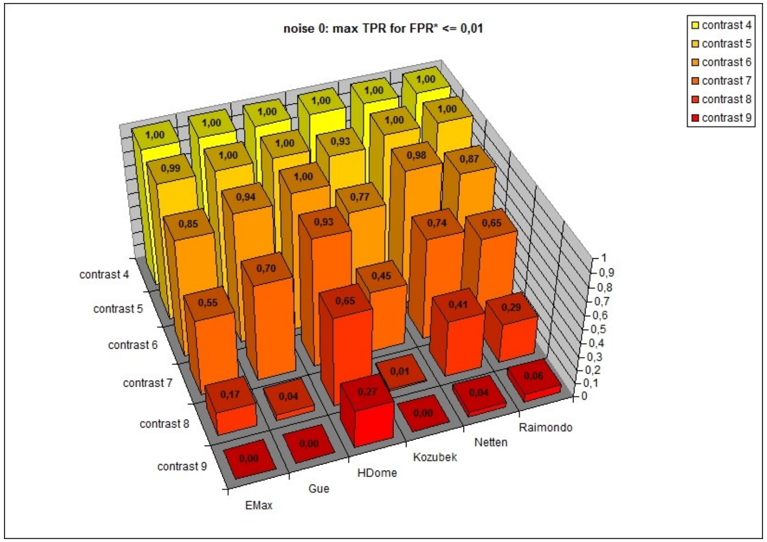
Further Work

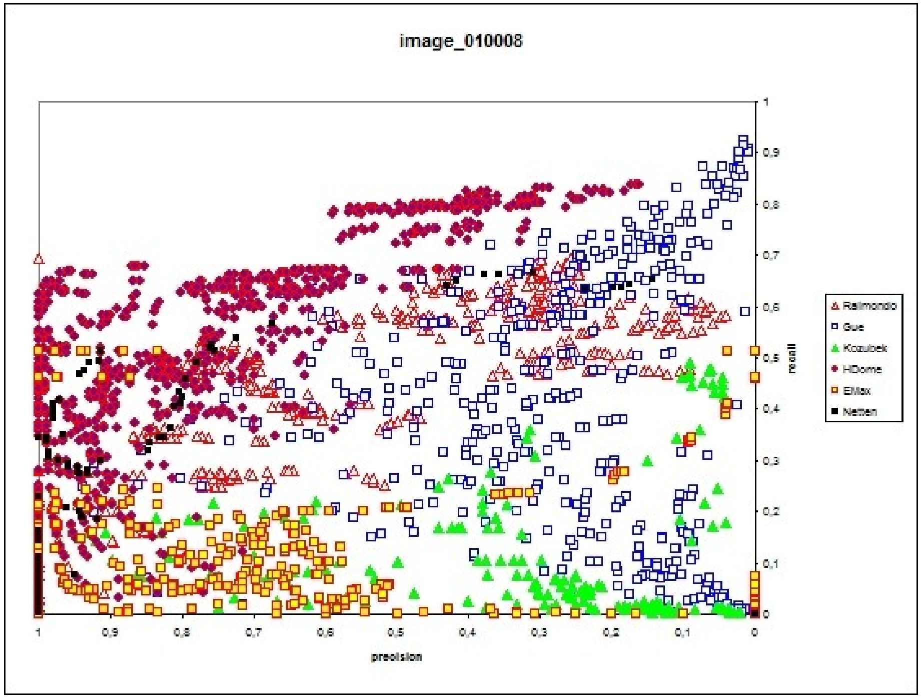
Further Work

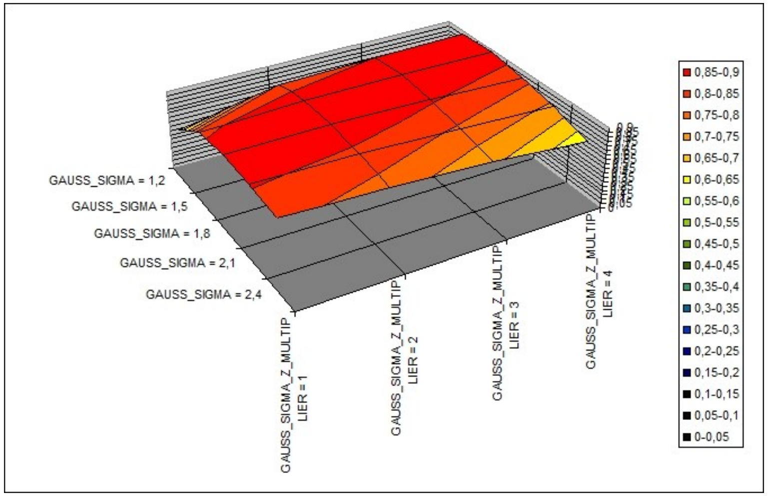
- Prepare a set of benchmark data
  - Cover testing of all important measurements
    - Detection, localization, intensity
  - Possibly make the set publicly available through CBIA web-site
- Perform a thorough evaluation of existing methods
  - Test the methods on various images
    - Real, manually annotated data
    - Simulated data with known GT
  - Investigate their behavior when used on 3D data
  - Parametric study
  - Publish the results

Further Work

- Intermediate results







Further Work

- Investigate the conceptual difference between 2D and 3D fluorescence images
  - Dots do not lie in the same focal plane
  - 2D images are usually obtained via max. intensity projection
  - Microscopy images exhibit strong anisotropy
  - Per-slice processing or direct extension to 3D do not take any of this into account
- Design a method natively working with 3D images
  - Most of the existing methods are natively 2D (or nD), and use no special approach for 3D data
  - Investigate localization using model fitting
  - Include the new method in the comparison