Contents

1 ISBI Challenge 2013

1.1 Introduction

The goal of the ISBI Challenge, as announced on their website (Biomedical Imaging Group (2013)), is to give an overview and understanding of availible algorithms for single particle localization microscopy. The focus was on 2d localisation, to give information about the depth of a localisated spot was optional. To benchmark results one needs groundtruth. Therefore the organisers created synthetic datasets of biologically relevant structures such as tubulins. To match realistic conditions the data was transformed to introduce different kinds of noise and background to it.

The participants were given training data sets and the corresponding groundtruth and one month before the deadline of the challenge the test sets. There were two different kind of datasets in principle. One with very dense spots and shorter sequences, the other with longer sequences and fewer spots per frame.

All paricipants were asked to submit their results and also the time it took to run the algorithm and the hardware configuration of the used system.

1.2 Terminology

To be able to compare different algorithms there must be a way to determine the correctly detected spots. To do so for each estimated positon of a flourophor, the nearest correct position of the molecule in the groundtruth data was searched within a lateral tolerance disc. Once a match was found this two spots were taken out of consideration for the matching.

One important parameter for this evaluation is the radius of the lateral tolerance disk, because it has big influence on the number of detections considered to be true positives (TP).

Detections with no associated spot in the groundtruth are called false positives (FP), spots in the groundtruth with no matching detection are called false negatives (FN).

This matching is done frame by frame, it is not possible to match a point from different frames even if the x and y coordinates match perfectly but the frame differs.

The precision (p) of a classification task is defined as the ration between the number of true positives and the sum of true positives and false positives:

$$\text{precision: } p = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1.1}$$

It is a number between 0 at worst and 1 at best, telling how reliable the result is, how likely it is that a labeld sample really belongs to the predicted class. In this context it means how certain a detected spot has its origin in a fluorophore attached to the investigated structure and it's origin is not wrongly detected background noise.

An other important value is the recall r that is defined as the ratio of true positives and the sum of true positives and false negatives:

$$\text{recall: } r = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{1.2}$$

The recall lies also in a range from 0 to 1 and gives an impression on how many relevant spots were found.

1.3 Measures

For the evaluation three different measures were used. The f-score index f, the Jaccard index J and the rot-mean square distance RSME.

$$\text{f-score: } f \quad = \frac{2 \cdot p \cdot r}{p+r} \tag{1.3}$$

1.3.1 Jaccard index

Let A be the set of points of the groundtruth and B be the set of detected points. The Jaccard index J is defined as:

$$\text{Jaccard: } J = \frac{|A \cap B|}{|A \cup B|} \tag{1.4}$$

The intersection is done frame by frame. This means two spots from the groundtruth and the detection set just match if they occure in the same frame.

1.3.2 RSME

The root-mean square distance gives an impression how big the squared distance between a spot in the groundtruth and an associated detection was in average. It can be calculated like this:

$$\text{RSME} = \frac{1}{|A \cap B|} \sum_{i=1}^{|A \cap B|} (p_a(x,y) - p_b(x,y))^2 \tag{1.5}$$
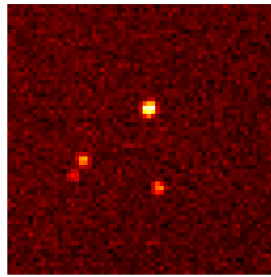
1.4 Trainingsdata

1.4.1 Bundled tubes datasets

There was two kinds of bundled tubes data sets, both created from the same underlying structure, one set with a high spot density and a short sequence of 360 frames, the other
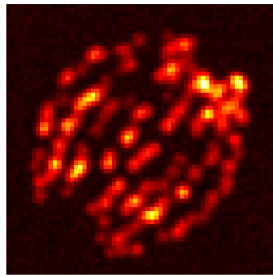
(a) High spot density

(b) Low spot density

Figure 1.1: One frame from bundled tubes training data set
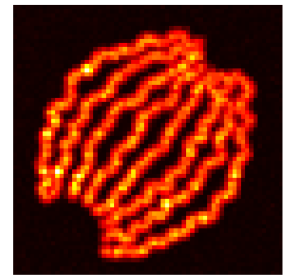

Figure 1.2: Maximum projection of bundled tubes data set

with fewer spots per frame but 12000 frames in total. Picture 1.1 shows one frame of each data set.

The original images very small, just 64 pixels in each dimension. Both sequences had spatial and temporal constant background. Picture 1.2 shows the maximum projection of the bundled tubes data set. The maximum projection is used to reduce the dimensionality of a data set. In this case for each pixel in x- and y-dimension, in the three dimensional dataset, the brightes value from all frames is taken.

1.4.2 Tubulin data sets

The other training data sets models 7 microtubules, a structure that is a long filament up several micrometers long and a diameter of about 25 nanometers. The spot density lies somewhere between the high density and the low density of the bundled tubes data sets. This data sets show strong inhomogeneity in spatial dimensions and moderate inhomogeneity in temporal dimension, see figure 1.3. This is the reason why in the lower left corner of the maximum projection 1.4.2 a brighter area can be seen.
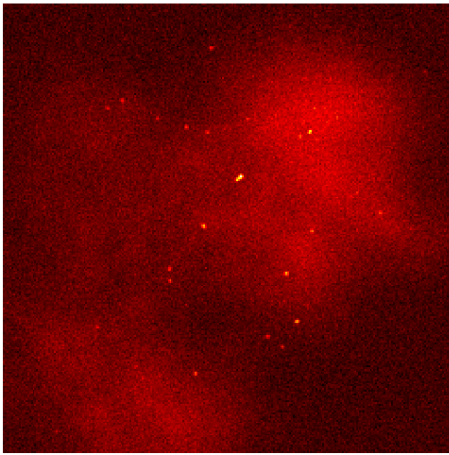
1.5 Submissions
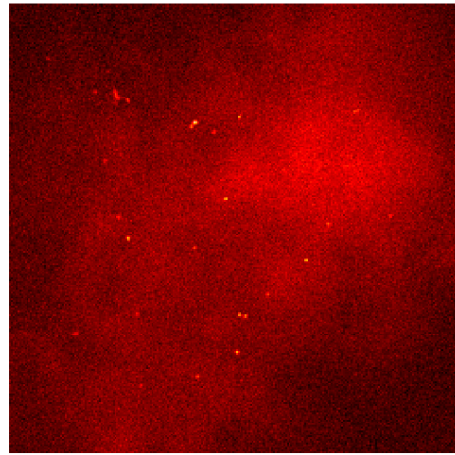
1.5.1 High precision

1.5.2 High score

1.5.3 Highest score via postprocessing

1.6 Results

(a) Tubulin2 frame 10

(b) Tubulin2 frame 1010

Figure 1.3: This pictures show the variability of the background in the spatial and temporal dimensions
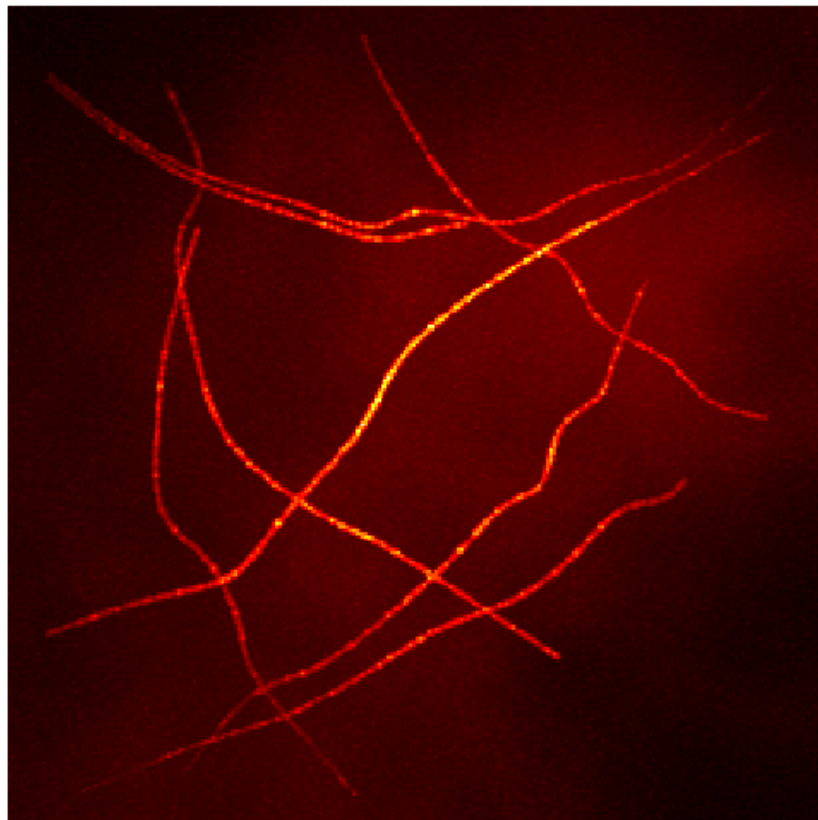
Figure 1.4: Maximum projection of tubulin data set

Bibliography

Biomedical Imaging Group, Switzerland, EPFL (2013), "Localization microscopy isbi 2013 challenge." URL `http://bigwww.epfl.ch/smlm/challenge/`. Accesed on 04.04.2013.