Contents

1 Introduction

2 Data processing

2.1 the data

The datasets for STORM microscopy that we recieve from our collaborators from Bioquant are big datasets of several gigabyte in the Andor .sif format. Each file conains a stack of pictures, normaly between 1000 and 10000, taken consecutively. In each picture there are beads and signals. Both result from very small fluorescent molecules attached to the structures that are investigated. The light of this pointlike objects is dissorted to a gaussian shaped signal due to the large magnification. Beads are molecules that emit light at any time contrary to the other signal which blinks that means it is visible in just one frame at an explicit location. The beads are used as landmarks for later alignment of two or more different color channels. The other spots are the structure that the biologist are interested in. Each of the gaussian shaped signals should be recognized and the center will be determined with subpixel accuracy and is stored in the end in a list to be further processed by the colorcomposer application.

2.2 Parameters and options

2.2.1 Necessary

2.3 Import and processing

The STORM data has usually a size of around 3 gigabyte. There are even larger datasets possible, so that it is important to work on smaller parts of the data, instead putting the whole dataset into memory. This is done using chunks of user defined size. The data is processed chunkwise, there is parallelisation for the frames of each chunk. This is possible because the signals in each frame are considered to be independent from each other.

2.4 Workflow

2.4.1 Chose parameters

At the begining the user has the option to set all important parameters, if no parameter is set the default ones are used and will give a good result because all crucial parameters

are either determined from the data or set to reasonable values that work for every data set.

2.4.2 Estimating camera gain and offset

First of all it is checked whether there exists a file containing settings for gain and offset from an earlier run. If this is not the case new parameters are estimated based on the first part of the data, usually 200 frames are sufficiant. The method described by **?** is used to estimate the gain factor. For this methode a Skellam distribution is used. Each dataset is three-dimensional where time is the third dimension. Therefore mean μ and variance σ^2 can be calculated from the data for each pixel individually

$$\mu(i,j) = \frac{\Sigma_t(I_t(i,j) - I_{t+1}(i,j))}{n} \tag{2.1}$$

$$\sigma^2 = \frac{\Sigma_t(\mu - (I_t(i,j) - I_{t+1}(i,j)))^2}{n-1} \tag{2.2}$$

To determine the gain factor the Skellam parameter are plotted over the mean intensities. A straight line can be fitted and its slope is exactly the gain factor.

2.4.3 Recursevly adjusting gain and offset

After the estimation of gain factor and offset, the transformations described in 7.1.1 and 7.1.2 are applied and the background subtracted.
Due to the Anscombe transformation the background pixels of the image should only vary around a mean intensity of zero with a variance of 1. Therefore a histogramm of the pixel intensities is created. The after background subtraction the background pixels should contribute only to the lower intensities in the histogramm. A gaussian function is fitted to the histogramms values. This is done under the assumption that there is much more background in the image than signal or the intensities coming from signals are distributed over a larger range, so the gaussian for the background intensity distribution can be fitted correctly.
If the estimated value for the variance is too far off 1 the originally estimated gain factor is corrected, applied and the fit is done again. This is done until the background variance converges or the maximal number of iterations is reached. In this case the initial gain factor will be used and a warning printed to the screen.

2.4.4 Estimating the width of the point spread function

For a certain number of frames the Fourier transform is calculated and averaged. The result is called mean power spectrum. It can be used to estimate the variance of the point spread function of the signal. A two dimensional gaussian functions Fourier transform is again a gaussian but with inverse variance. This relation is used to determine the variance of the point spread function in spatial domain, using the fit parameter for the variance in frequency domain.

2.4.5 Processing the data

Import Data

Storm data sets can consist of several thousand frames with resolutions up to one mega pixel per frame. This makes it necessary to break the data into smaller parts because otherwise it might be much larger as the RAM of an ordinary machine. Because of the background estimation it is not possible to process every frame completly independent as it was in the older version of this software **??**. It is also faster with some datatypes to load a larger consecutive part of the dataset into memory.
This algorithm uses chunks of user defined size. There are some limitations to the chuncksize that are discussed later. The data set is split into parts of equal size in x- and y-dimensions and independently also in t-dimension. If this partition does not fit at the edge of the data set, the last chunks will be smaller.
The data is transformed to be Poission distributed and after that the Anscombe transform is applied.

Background estimation

For each chunk the median is determined to get a robust estimate of the background value for this chunk. The Bspline interpolation implemented in vigra **??** is used to get interpolated values for the full resolution of the current frames. For this interpolation three chunks in t-dimension have to be available. Therefore the maximal chunksize in t-dimension must not be larger than a third of the total stacksize.
This background is then subtracted from the transformed data to give finally background pixels with zero mean and a variance of one, both in xy- and in t-dimension.

Create mask for background suppression

With the given p-value from the settings a global threshold can be determined, because inhomogeneities of background intensities has been removed. The threshold value is that intensity that is explained from a gaussian distribution with mean zero and variance one with the given p probability. This is possible because the background intensity follows such a distribution after all the applied transformations.
The threshold is applied to the current frame and stored as a mask. Due to the probability that background pixels intensitis might exceed the threshold the connected components of the mask are calculated. Pixels that belong to connected components with too few members are discarded.

Filter data and finding maxima

To imporve the accuracy of the spot detection the transformed signal is convolved with a two dimensional gaussian function with the previously determined or user set width. The convolved image will further be used to find the maxima. Each maxima found is tested to be covered by the mask or discared otherwise. A region of intrest around

the remaining maxima is interpolated to a higher resolution. In the interpolated region it is searched for maxima for the last time. This maxima will be detected with super resolution.
To determin the signal-noise-ratio the unfilterd and uninterpolated pixelintensity is used.

Quality control for detections

Especially in data sets with a high density of spots it can happen, that two spots are near each other and the point spread functions overlap. It may happen that instead of two maxima just one maximum will be detected right between the true ones. This leads to large errors in the localisation. To avoid this a threshold for the asymmetry of the spots can be set.

2.5 Calibration measurement and plausibility

2.6 Accuracy of detection

Unfortunately the position of the flourescent molecules can't be detected perfectly. There are three main contribution to the error in detection.
First, there is the problem of finding the maximum in a noisy signal. Due to noise the pixel next to the true maximum might get some intensity and be therefore brighter.
Second, the choice of the gain factor and the offset might influence the precision.
Third, the position is deteted by upscaling the pixel grid and interpolation. After that the maximums position of the upscaled grid is taken as the resulting position. This gives an error from roughly pixelwidth divided by square root of two.

2.6.1 error from noise

Because there is no ground truth availible for micoscropy data, data must be generated. This was done similar as described by **?**.

2.6.2 error from parameter estimation

2.7 Comparison with older version of the storm algorithm

2.8 Bleaching signal

2.9 Check for slope using calibration

As described by **?** the true slope can be determined.

2.10 New graphical user Interface

3 Multicolor registration

3.1 Background

In microscopy it is often desirable to label different structures in a cell with different colors. To do so our collaborators use different fluoroscent molecules that emit light at different and distinguishable wavelengths. Using different filters it is possible to capture pictures just containing light emited from one fluorophore. To get a mulit-channel picture the different channels must be aligned. Because different flourophores emit different wavelengths, cromatic aberration apears. This means that the light for the same spot but with different wavelenghts is not mapped to the same spot in the image. To align the different channels despite cromatic aberration, beads are used. Beads are flourophores added to the probe, that emit light in all wavelengths the different markers do and therefore are visible in all channels. The beads can be used as landmarks, because their position in the original image is at the same spot. The task is to find a transformation that maps corresponding beads on each other.

3.2 Features of the colercomposer application

3.3 Bead detection

The input for the colorcomposer application is a text file created by the storm algorithm that contains information about the position, intensity, symmetry, framenumber and signal-to-noise ratio of each detection. The beads should ideally be visible in most of the images, this means one must search for detections that appear in almost every frame at the same position. Therefore it is plausible to take every detection of the first 50 frames as initial candidates for beads. After that candidates that are closer than a threshold are merged to get a list of all location where beads might be. Given that list every other detection is tested to belong to one of the bead candidates. If a candidate gets too few members it is no longer considered to be a bead and removed from the list.

3.4 Align Beads

After the beads for each channel are found the next task is to find the same bead in each channel. It can happen that some beads occure in just one channel, if this is the case there will be no corresponding bead in the other channels.
To do so, the minimal number of beads, three to four, that are neccesary to calculate

the transformation are chosen randomly from the first channel. After that, based on a probabilistic approach and a distance matrix containing information about the distances between all beads of the two channels, three to four beads from the second channel are chosen.
Using this pairs of beads linear transformation is found like described by **?**. Using this transformation it is tested how many beads match in total. It is assumed that the correct transformation will match other bead pairs that were not chosen to calculate this transformation. After that the whole procedure is done multiple times. In the end the best transformation is chosen.
In principle shearing should also be allowed for this transformation, but tests indicate that shearing does not occure. There is a problem if there are just three beads in each channel, then every time a perfect transformation is found, but with the constraint of forbidden shearing, the right solution can be identified.

3.5 Accuracy of Registration

3.6 Colocalisation

3.6.1 Global colocalisation

3.6.2 Local colocalisation

3.6.3 Validation of colocalisation approaches

"Image set CBS001RGM-CBS010RGM from the Colocalization Benchmark Source (www.colocaliza
benchmark.com) was used to validate colocalization."

4 Work for the biologists

4.1 Selecting the best camera

4.2 little tool to investigate bleaching

5 Futur work

5.1 3d Storm

5.2 Improved methode to detect maxima

5.3 Colorcomposer implemented in Storm-Gui

6 CCD camera

6.1 Image acquisition

6.1.1 Photon sources with shot noise

The emission of photons is a random process that occures at unpredictible times. Therefore the number of photons passing through a plane is never constant but varies around some average value. The phenomena, that one can never determine exactly how many photons should hit the sensor chip of a CCD camera for example, is called shot noise. It playes a major role if the total number of photons is low, as from dark sources or with short exposure times of the camera.

6.1.2 Quantum efficiency

Quantum efficiency describes the fraction of photons that create a detectable electron in a sensor chip. The quantum efficiency is dependent of the wavelenght of the incoming photon. Photons with energies below the band gap can't produce a free electron that can be detected. The quantum efficiency has a maximum basically caused by two effects. The higher the photons energy the higher the kinetic energy of the freed electron, but it is absorbed earlier and can therefore recombinate with a electron hole more likely.

6.1.3 Gain

There are two different gain factors involved in the capturing process of a camera. First the electric signal for each pixel might be amplified. And there is also a gain factor that describes the proportionality between collected electrons and the digital number that is associated with.

6.1.4 Readout noise

The origin of readout noise is the amplifier. The aplification is never perfect, this means the exact number of electrons at the end of the amplification has some variation around the expected linearly increased value. There might also be some random signals of the electronics that add to the "true" signal. The readout noise is independent of the exposure time.

6.1.5 Dark current noise

Dark current noise is generated by the thermal movement of the atoms in the sensor chip. The movement of molecules and atoms is dependent of the temperature of the material, because of that dark current noise depends strongly on the temperature of the chip and can be reduced by cooling. Dark current noise generates electrons in the bins of each pixel even with closed shutter it is constantly increasing with time and follows Poisson statistics.

6.1.6 Quantisation

The signal must fit into the output color depth. It has to be rounded or truncated to fit in. This process introduces errors that can be seen as additional noise that is dependent on the intensity of the signal. High intensities are disturbed less relative to low intensies.

7 Theoretical background

7.1 Transformations

7.1.1 Transformation to Poisson distributed signal

The images aquired from the camera show not the real intensities I_{true}, which result from the photon emission of the probe, but transformed ones I_{meas}. I consider two main reasons why the taken image differs from the true image, besides noise.
There is dark current which means that even a picture taken with closed shutter would get some intensity, even without any light hitting the sensor chip of the camera. This is a result of thermal movement of the atoms off the sensor chip and can be reduced by cooling. The dark current noise adds an almost constant value o to the output signal. Incoming photons create electrons via inner photoelectric effect. This electrons are collected for each pixel and might be amplified to get the final result. Assuming a linear relation between the number of incoming photons and the number of electrons created and a linear amplifier results in a factor g. This factor is multiplied with the number of photons captured during exposure time for each pixel.
If the gain factor g and the offset o are known the true intensity, the number of photons detected is:

$$I_{\text{true}} = \frac{I_{\text{meas}} - o}{g}. \tag{7.1}$$

7.1.2 Anscombe transformation

The Anscombe transform is used to transform a random variable with a Poisson distribution into one with an approximatly constant standard deviation. The transformation is defined as:

$$A(x) = 2\sqrt{x + \frac{3}{8}}. \tag{7.2}$$

As one can see in figure 7.1 the Anscombe transformations result has for mean intensities greater than 4 a intensity independent standard deviation of one.

7.2 Estimation of camera gain

The

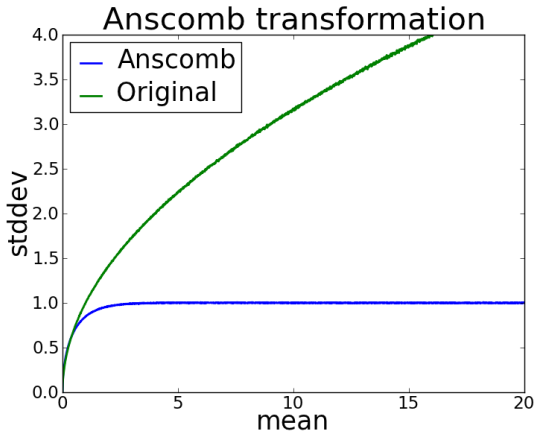


Figure 7.1: Standard deviation over mean intensities of different Poisson distributions

7.3 Distributions

7.3.1 Poisson distribution

One very important probability distribution in physics is the Poisson distribution. It describes the results of "counting experiments" and is therefore very important for image processing as the pictures taken with a camera are in principle counts of photons reaching the camera. Photon counting noise is one important example.
Poisson distributions are just defined for integer values and the variance is the same as the mean value of the distribution. Another important attribute is the skewnes which is the inverse of the squarerot of the mean or variance and describes the assymetry.
The probability mass function is:

$$p(n, \mu) = \frac{\mu^n}{n!} \exp(-\mu) \tag{7.3}$$

7.3.2 Skellam distribution

The probability mass function of a Skellam distribution is a function of the difference between two Poisson random variables

$$p(k; \mu_1, \mu_2) = \exp(-(\mu_1 + \mu_2)) \left(\frac{\mu_1}{\mu_2}\right)^{k/2} I_{|k|}\left(2\sqrt{\mu_1 \mu_2}\right) \tag{7.4}$$

where n_1, n_2 are the Poisson random variables and $k = n_1 - n_2$. $I_{|k|}$ means the modified Bessel function of the first kind.
Mean μ and variance σ of the Skellam distribution are given by

$$\mu = \mu_1 - \mu_2, \qquad \sigma^2 = \mu_1 + \mu_2 \tag{7.5}$$

$$\Rightarrow \qquad \mu_1 = \frac{\mu + \sigma^2}{2}, \qquad \mu_2 = \frac{-\mu + \sigma^2}{2} \tag{7.6}$$

7.3.3 Approach using skewness of poisson distribution

For every pixel there is a set of multiple values in the set. This allows to calculate the different parameters individually for each pixel. One can calculate mean and variance of the measured intensities $I_{\text{meas}}(i,j)$ and gets

$$\text{mean}(I_{\text{meas}}(i,j)) = \text{mean}(I_{\text{true}}(i,j)) + o \tag{7.7}$$

$$\text{var}(I_{\text{meas}}(i,j)) = g \cdot \text{var}(I_{\text{true}}(i,j)) \tag{7.8}$$

$$\tag{7.9}$$

Assuming a Poisson distribution as the true intensity, mean and variance would be the same. Unfortunately the mean true intensities are unknown and it is not possible to determin g and o so far. For large mean Intensities μ the Poisson distribution becomes more and more similar to a Gauss distribution with the same mean. However, for small means, the Poisson distribution is not symmetric. The skewness s_p of a Poission distribution is the inverse of the square root of the mean $(\mu)^{-.5}$. It can also be directly calculated from data

$$s_p = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\sigma}\right)^3 \tag{7.10}$$

The skewness is invariant to shift and multiplication with a constant. This means that the transformation caused by the camera gain and the dark current does not affect the skewness. This gives a third equation to solve for g and o.
This approach has very strict limitation to at least for background not to intense signals. If the mean of the true Poisson distributin is higher than roughly 30 the skewness gives due to noise no stabel results and it is impossible to determin the mean intensity in this way.

8 ISBI Challenge 2013

8.1 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.

8.2 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:

$$\text{precision: } p = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8.1}$$

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:

$$\text{recall: } r = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8.2}$$

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

8.3 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.

$$\text{f-score: } f \quad = \frac{2 \cdot p \cdot r}{p + r} \tag{8.3}$$

8.3.1 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:

$$\text{Jaccard: } J = \frac{|A \cap B|}{|A \cup B|} \tag{8.4}$$

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.

8.3.2 RSME

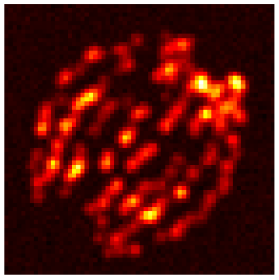
The root-mean square distance gives an impression how big the squared distance between a spot in the groundtruth and an associated detection was in average. It can be calculated like this:

$$\text{RSME} = \frac{1}{|A \cap B|} \sum_{i=1}^{|A \cap B|} \left(p_a(x, y) - p_b(x, y)\right)^2 \tag{8.5}$$

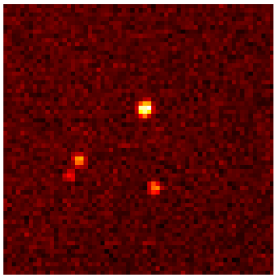
8.4 Trainingsdata

8.4.1 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 8.1: One frame from bundled tubes training data set

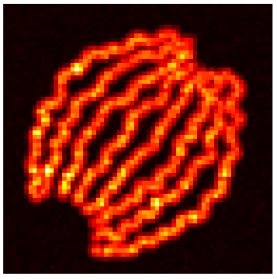


Figure 8.2: Maximum projection of bundled tubes data set

with fewer spots per frame but 12000 frames in total. Picture 8.1 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 8.2 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.

8.4.2 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 8.3. This is the reason why in the lower left corner of the maximum projection 8.4 a brighter area can be seen.

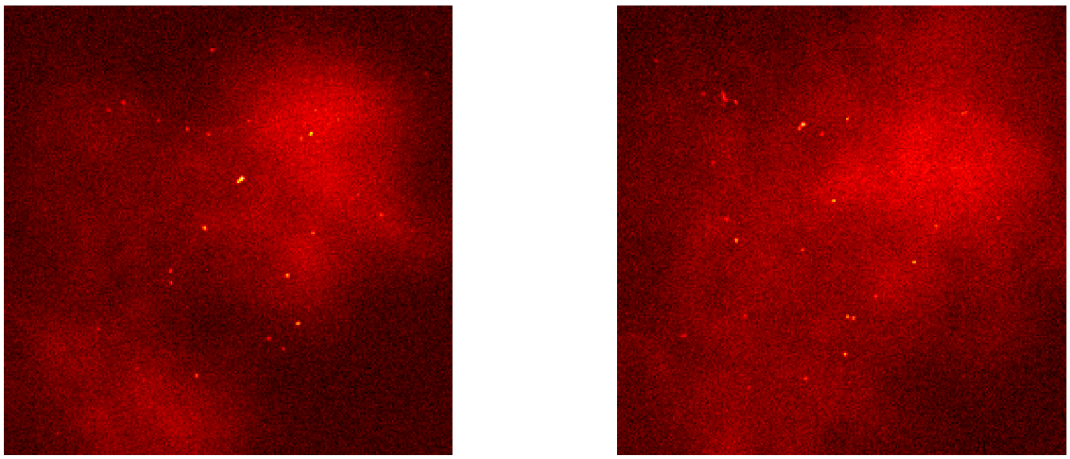
8.5 Submissions

8.5.1 High precision

8.5.2 High score

8.5.3 Highest score via postprocessing

8.6 Results



(a) Tubulin2 frame 10

(b) Tubulin2 frame 1010

Figure 8.3: This pictures show the variability of the background in the spatial and temporal dimensions

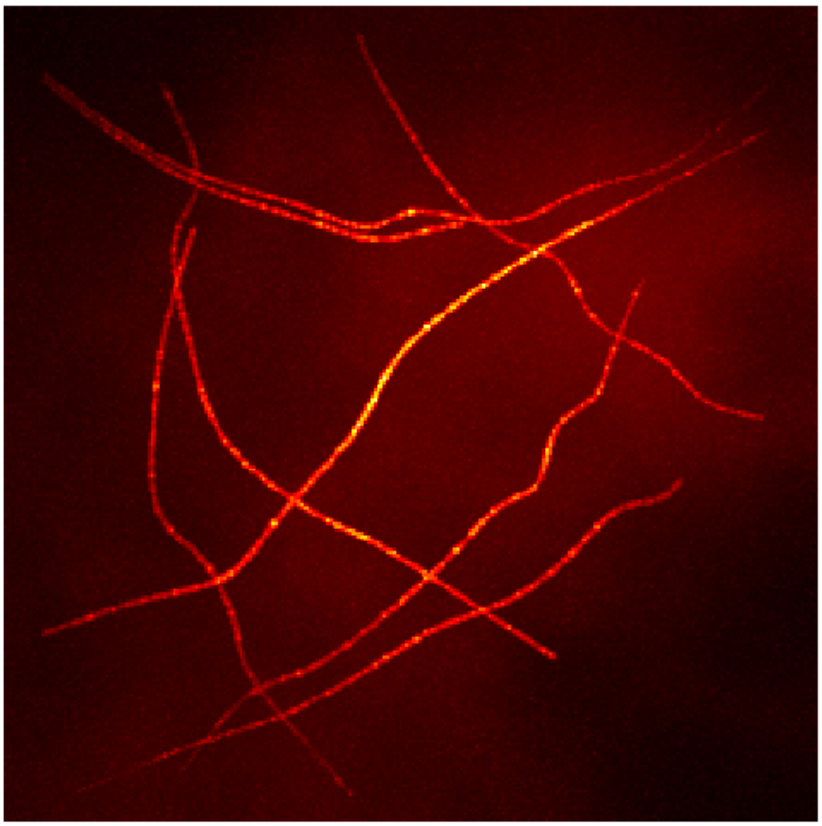


Figure 8.4: Maximum projection of tubulin data set

Bibliography

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