A COMPARISON OF SINGLE PARTICLE TRACKING ALGORITHMS[1]

PRATEEK TANDON (Ptandon1)

Reading Assignment 6

SUMMARY

To develop an in-depth understanding of the processes taking place at the particle (molecular) level, we need to be able to accurately analyze the time series data generated by microscopy. Single particle tracking is one of the fundamental challenges that the whole area of computer vision and bioimage informatics faces. To be able to accurately identify each particle and track its position across the large amount of frames captured. In the paper under consideration[1], the authors summarize the results of an open 3-week challenge having a common dataset and compare the algorithms that the participants followed.

INTRODUCTION

With hundreds of frames(images) generated per bioimaging experiment, and hundreds of particles in each frame, it is practically impossible to do the task of tracking the particles manually. This drives the need for automation of the process using computational methods of machine learning. As apparent in the field of research, often the ground truth pertaining to an imaging based experiment is unavailable, thus the choice of the machine learning methods becomes limited to the category of unsupervised learning methods. The main objective is thus to efficiently mine the parameters pertaining to the unsupervised algorithm under consideration and optimize the desired level of accuracy at least up to a threshold.

DETAILED SUMMARY

The generic approach of analyzing imaging data can be broken down into two distinct steps. First, accurate identification of particles in an image. Secondly, once the particles are identified, link the subsequent image to the current frame by locating the previously identified particles in the new image. This way, the motion, directionality and other aspects pertaining to the molecular dynamics can be assessed.

The particle identification process begins by denoising the image so as to enhance the actual difference between the background and the object intensity. This is also called object enhancement and is done using median, wavelet based, Gaussian, Laplacian of Gaussian filters. Having denoised the image, we now begin search for local maxima as potential positions of particles. We then localize the position using centroid based algorithms or Gaussian fitting. Performance optimization measures that are taken in this localization include non-linear interpolation and iterative centroid calculation.

Having identified the particle positions, now the job is to link two frames together by means of mapping particle positions. While the most common approach to this problem would be identification of k-nearest neighbors of each particle, in practice, however, it is bound to perform poorly due to a wide range of possible positioning of particles (example, directional movement consisting of long range jumps). To overcome such problems, multiframe and multi-track optimization are found to perform better. Particularly, Kalman filtering, multiple hypothesis evaluation schemes and other time-series association optimization techniques score high in terms of accuracy.

Since directional motion of particles is a specific case, and the techniques are based on much more generic (diffusive motion) case, overfitting to the training data upto a certain level tends to improve the accuracy. This kind of overfitting is somewhat acceptable in such cases as most of the times, the ground truth is not available to the researchers and thus using domain knowledge to generating synthetic data with a ground truth, fitting the model to this synthetic data and applying this to the real data is a means to reduce the search space and is the best choice for the researchers.

It should be noted that none of the methods alone are capable of solving the problem but a combination of them incorporating tracking information from multiple frames rather than a pair of frames is necessary to yield satisfactory results.

For performance measurement of these particle tracking methods, two tracks were considered. X: ground truth tracks and Y: estimated tracks. The distance between tracks was calculated using the Munkres algorithm with Z (dummy, extended version of Y) as follows: $d(\theta_k^X, \theta_k^Z) = \sum_t |\theta_k^X(t) - \theta_k^Z(t)|_{2,\epsilon}$. Various metrics used for the comparison of algorithms using this computation were: 1) $\alpha(X,Y) = 1 - \frac{d(X,Y)}{d(X,\phi)}$, 2) $\beta(X,Y) = \frac{d(X,\phi) - d(X,Y)}{d(X,\phi) + d(\bar{Y},\phi)}$ where $\bar{Y}$ is the set of spurious tracks; 3) Jaccard Similarity Coefficient $JSC = \frac{TP}{TP+FN+FP}$ 4) $JSC_\theta = \frac{TP_\theta}{TP_\theta + FN_\theta + FP_\theta}$ and RMS error.

Using these performance metrics under the given scenario, it was observed that the efficiency of the tracking algorithms significantly depends on the particle density and SNR. The methods were found to possess some inherent robustness with respect to particle density. It was found that Increase in particle density by a fixed amount does not drop the performance by the same amount. In case of SNR however, a steep performance drop was observed below SNR = 4 and with rise in complexity of particle shapes.

DISCUSSIONS

Since the algorithms used by the different participating groups were conceptually pretty similar, one main conclusion from the study is that the key to generating meaningful conclusions from the data is finding the optimal set of parameter values for the algorithms. Furthermore, research can be carried out from the perspective of performance evaluation of the algorithms considered and the quality of images in the dataset.

REFERENCES

1. Chenouard, N. *et al.* Objective comparison of particle tracking methods. *Nature methods* **11,** 281–9 (2014).