

Automatic segmentation of DNA molecules

The Laboratoire d'Optique Biomédicale (LOB) at EPFL is offering a semester (master) project related to the image processing of super-resolved stretched DNA.

DNA analysis methods have evolved tremendously over the past decade. One of the goal of such techniques is to be able to recognize the species the DNA strand is coming from. As an alternative to DNA sequencing (i.e. reading the whole DNA sequence), we have developed in our lab a way to map the DNA to its corresponding species while avoiding complicated PCR reactions and DNA sequencing.

The method is based on sequence specific labelling of DNA and subsequent stretching on a glass surface. The stretched DNA is then imaged with a super-resolution microscope resulting in a sort of bar-code image (fig. 1 Left). The intensity profile (fig. 1 Right) of each DNA molecules is extracted and matched against a database of species.^{1, 2}

In our lab we intend to use this method for the analysis of microbiome samples in mice that develop Alzheimer disease³. In order to study the entire microbiome, we need to analyse thousands of images similar to fig.1, extract all the individual DNA molecules and match them to their sequences. Currently we mainly detect the DNA strands by hand, if we want to process thousands of images it is critical that this process is automated.

The task of the student will be to develop and optimize a DNA segmentation algorithm, based on Hough transform, steerable filters^{4, 5}, morphological operators, etc... We will provide the student with several experimental images of various density and quality allowing him to evaluate the performances of the algorithm.

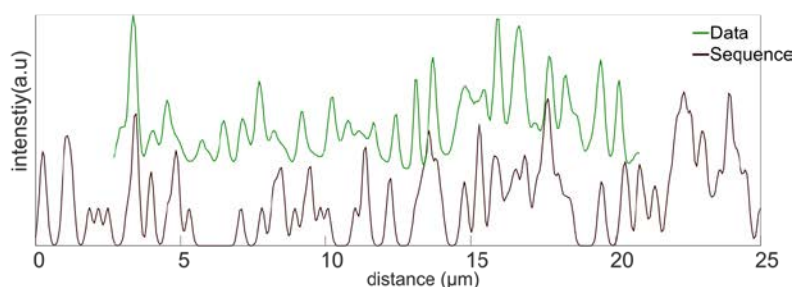
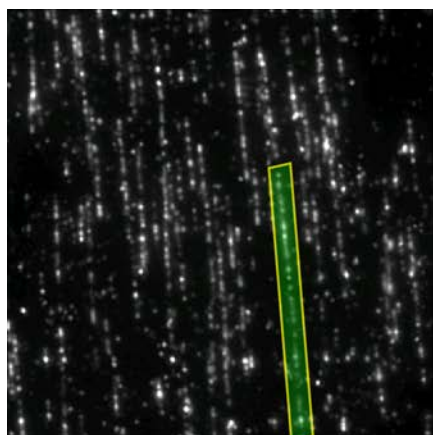


Figure 1 Left. Typical high resolution image of labeled DNA and extraction of a single line (in green). Right. matching of the labeled data to the sequence data.

Skills: Knowledge in optical imaging and coherence, signal analysis and programming (MATLAB).
Assistant: Dr. J. Deen (jochem.deen@epfl.ch), A. Descloux (adrien.descloux@epfl.ch)
Supervisor: Prof. Theo Lasser (Theo.Lasser@epfl.ch)

References:

1. Deen, J.; Vranken, C.; Leen, V.; Neely, R. K.; Janssen, K. P. F.; Hofkens, J. *Angew. Chem. Int. Ed.* **2016**, doi:10.1002/anie.201608625
2. Neely, R. K.; Deen, J.; Hofkens, J. *Biopolymers* **2011**, 95, (5), 298-311.
3. AD-gut. European commission Horizon 2020: 2016; project ID 686271.
4. Berlemont, S.; Bensimon, A.; Olivo-Marin, J. C., IEEE ICASSP '07, 15-20 April 2007, 2007; pp I-1225-I-1228.
5. Qiguang, M.; Mei, Y.; Jingjing, L., IEEE ICCASM '10, 22-24 Oct. 2010, 2010; pp V9-120-V9-124.