# Light Sheet Data Analysis Meeting

Date 02.08.2023 Time: 09:00 – 10:00 Place: ENT conference room Protocol: Elisabeth

Participants: Constantin, Jakob, Bettina, Lennart, Anu, Elisabeth

1. Lennart introduces our light sheet data using Aleyna’s Lab Meeting slides. We identified 2 main problems with the data handling:

* Channel alignment does not yet work (Aleyna tried BigStitcher with Fiji still, but had a mismatch around 1/3 of a cell)
* Data size: When imaging the whole cochlea, we get 600 GB per channel, if we focus on regions of interest in the small mouse cochlea we could go down to 200 GB, still handling the data is difficult

2. Focus on the data size:

* Constantin asks about our current format (BigTIFF) and suggests using zarr/n5 as a file format. These formats save data as 3D blocks instead of slices. They allow easy visualization/ multiscale representation by down sampling the individual volume blocks and only loading what is currently viewed on the fly. For the analysis the down sampling rate can be adapted for the volume blocks.
* Bettina asks about difficulties in converting to zarr/n5 from other formats. Constantin says for now the microscope output and raw data storage should still use tiff, but it can easily be converted to zarr for the analysis (he has the scripts for that).
* Zarr/n5 files can be opened in BigViewer
* There is currently development to establish OME-zarr
* On the long run it may be possible that the microscope directly creates zarr files, but this depends on Jan’s plans

3. Focus on stitching the data:

* BigStitcher-Spark is a newer version of BigStitcher under development that is quite similar but uses another computational backend than Fiji. The program uses zarr/n5 formats, can be run on a laptop or cluster and seems to work similar to BigStitcher. It is still under development, but the developers are happy to give advice if contacted with problems that come up while testing the program.
* Constantin advises to try working in BigStitcher-Spark and use an actual sample to test different registration methods (isotropic/anisotropic scaling or deformation tools). Here it is important to discuss the effects on the data and if the transformation is scientifically sound or may affects the results.

4. Can zarr/n5 be used with Anus analysis tools?

* Using zarr/n5 with python programs (such as StarDist) should not be a problem. You can scan the zarr files and export TIFF files for your areas of interest.
* Matlab has not yet implemented the use those new file formats, so there problems could occur. For now Anus pipeline does not use Matlab.
* The zarr developers focus on the application with deep learning, so the support is very good there.
* Anu describes what has been done on the Flamingo data for now: She ran image crops using the StarDist-Arivis combination which worked okay, running it on the whole cochlea took very long and was not working to well yet (tested both with and without empty annotations)

6. Data handling:

* Currently the data is stored on an SSD drive. Lennart will transfer it to UKON100/archive/imaging/Lightsheet/Huisken\_group
* Jakob gives Constantin access to the server
* We should keep the structure of storing data in folders sorted by year - (month) - experimenter name

5. Next steps:

* Lennart provides Constantin with our latest sample (mouse cochlea, 2 Channels, 200 GB per channel) by Friday
* Constantin tries converting it to zarr, checks if it runs with BigStitcher/BigStitcher-Spark and tries to visualize the data. He will have a quick look next week and report back if it works
* In September/October, when everyone is back from holidays we will meet again (maybe together with someone from Jan’s lab). There we will also talk more about the channel registration once the data handling is established.