Towards an Automated Processing and Analysis System for multi-dimensional light-sheet microscopy big data using ImageJ and OMERO.

Whitehead, L.[1,2], Wimmer, V.[1,2], Lafouresse, F.[1,2], Ratnayake, D.[3,4], Currie, P.[3,4], Groom, J.[1,2], Rogers, K.[1,2] and Boudier, T.[1]

[1] Walter&Eliza Hall Institute of Medical Research, Australia, [2] University of Melbourne, Australia, [3] Australian Regenerative Medicine Institute (ARMI), Australia, [4] Monash University, Australia

Recent years have witnessed an explosion of multi-dimensional data thanks to the development of new fast volumetric Light-Sheet Microscopy technologies(Huisken et al., 2004). These systems can acquire volumetric data for a very long period of time greatly without harming the sample, with a typical experiment yielding up to 100 GB of data per hour. Storage and automated processing and analysis of such data is still a challenge(Meijering et al., 2016).

We have developed TAPAS (Towards an Automated Processing and Analysis System) for dealing with the issue of both storage and automated processing. Raw data are stored within an OMERO database (Allan et al., 2012), a simple text file defines a series of basic processing tasks to be undertaken (such as "crop", "scale", "filter","measure", "), then our system will apply the processing pipeline to the specified datasets within OMERO.

We applied our system to two different datasets, the first is a series of 3D light-sheet data of labelled T-cells inside lymph nodes, totalling almost 1TB of data. This data required downscaling, filtering, thresholding and cell detection. This was preformed in an semi-automated fashion based on our library for 3D processing and analysis (Boudier, 2012; Ollion et al., 2013). Initially, in a completely automated process, the data is pulled from OMERO, scaled down in X and Y by a factor of 2 and a 3D median filter is applied to detect the main structure of the lymph node. The cells are detected by applying a 3D TopHat filter to remove the underlying background. At this point a manual thresholding step is required, the pre-processed data are downloaded, thresholded by the researcher, binarised and sent back to omero. TAPAS then performs a layer analysis to study the localisation of the cells within the lymph node. The results are automatically saved as an attachment to the raw image file within OMERO. The researcher can easily retrieve the results and perform additional statistical analysis if required.

The second dataset we apply TAPAS to is a 3D light-sheet time-lapse of a zebrafish embryo, acquired for 25 hours every 90 secs and comprising of more than 1TB of data. This dataset contains labelled macrophages, which we want to track over time. First we wanted to generate a preview of the data as a maximum projection video. The TAPAS pipeline for this dataset involves a downscaling of the data by a factor of 2, a 3D median filtering and a maximum intensity projection along the Z-axis. Thus appropriately downscaled, the data is automatically saved to OMERO to allow easy visualization and interactivity to the researcher.

We have designed a framework for an automated processing and analysis system using open-source software ImageJ/Fifi and OMERO (Arena et al., 2016). This system is based on a simple text based description of the processing pipeline and does not involve any programming. We hope to make this system available soon to larger microscopy communities.

Allan, C., Burel, J.M., Moore, J., Blackburn, C., Linkert, M., Loynton, S., MacDonald, D., Moore, W.J., Neves, C., Patterson, A., 2012. OMERO: flexible, model-driven data management for experimental biology. *Nat. Methods* 245.

Arena, E.T., Rueden, C.T., Hiner, M.C., Wang, S., Yuan, M., Eliceiri, K.W., 2016. Quantitating the cell: turning images into numbers with ImageJ. Wiley *Interdiscip. Rev. Dev. Biol.*

Boudier, T., 2012. 3D ImageJ Suite. http://imagejdocu.tudor.lu/doku.php?id=plugin:stacks:3d_ij_suite:start.

Huisken, J., Swoger, J., Bene, F.D., Wittbrodt, J., Stelzer, E.H.K., 2004. Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy. *Science* 305, 1007 - 1009. https://doi.org/10.1126/science.1100035

Meijering, E., Carpenter, A.E., Peng, H., Hamprecht, F.A., Olivo-Marin, J.-C., 2016. Imagining the future of bioimage analysis. *Nat. Biotechnol.* 34, 1250 - 1255. https://doi.org/10.1038/nbt.3722

Ollion, J., Cochennec, J., Loll, F., Escudé, C., Boudier, T., 2013. TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinformatics* 29, 1840 - 1841.