Working title: the most amaz-zing SPIM and how we built it.

M. Caroline Müllenbroich,[a,b] Ludovico Silvestri, [a,c] Leonardo Onofri,[e] Irene Costantini,[a] Marcel Van t'Hoff[f] Leonardo Sacconi,[a,c] Giulio Iannello,[e] Francesco S. Pavone,[a,b,c,d]

[a]European Laboratory for Non-linear Spectroscopy (LENS), University of Florence, Italy
[b]Department of Physics and Astronomy, University of Florence, Italy
[c]National Institute of Optics, National Research Council, Italy
[d]International Center for Computational Neurophotonics (ICON Foundation), Italy
[e]Integrated Research Centre, University Campus Bio-Medico of Rome, Italy
[f]Distrio, Murmex

Abstract. 200 words limit. no numerical references presenting concisely the objectives, methodology used, results obtained, and their significance.

Keywords: Light sheet microscopy, Big data, Optical clearing, Data management, whole brain imaging, rolling shutter, 7,8..

Address all correspondence to: First author, University Name, Faculty Group, Department, Street Address, City, Country, Postal Code; Tel: +1 555-555-5555; Fax: +1 555-555-5556; E-mail: myemail@university.edu

1 Introduction

The highly ambitious project of mapping and understanding each and every neuronal connection in the whole brain has been moved from the realm of wishful longing to feasible reality by the recent advent of light sheet fluorescent microscopy (LSFM). With this technique 3D data sets can be acquired with a resolution that is high enough to identify neurons and their dendritic, axonal and spine features in time scales which are no longer the bottle neck of high-throughput acquisition. In LSFM, the sample is illuminated with a thin sheet of light confined into the focal plane of the detection objective, which collects the fluorescence emission along an axis perpendicular to the illumination plane.[1] This technique drastically reduces the imaging acquisition time by recording millions of pixels in parallel and affords optical sectioning by operating fluorescence excitation and detection on separate, perpendicularly oriented paths where the excitation light sheet and the detection focal plane overlap (Figure1,A). Consequently fluorophores outside the light sheet are neither bleached nor contribute blurring out-of-focus noise. Consequently LSFM reduces phototoxicity and photobleaching while achieving excellent resolution at high penetration depths; however, it requires the sample to be transparent.

Several challenges remain to be overcome, however, to allow fast and, most of all systematic, production of reliable datasets and their meaningful interpretation to further our understanding of neuronal networks. Those challenges include fast, cheap and reproducible sample preparation, automated image acquisition that does not require constant attention by an expert and, most crucially, the storage, interpretation and analysis of the unprecedented huge data sets light sheet microscopy routinely produces. The mapping and understanding of this "big data" is a colossal task that requires the expertise of computer scientists to employ fully automated post-processing, for example, to do cell counting or blood vessel segmentation. On the other hand, the imaging of large, intrinsically opaque samples in light sheet microscopy necessitates clearing protocols based

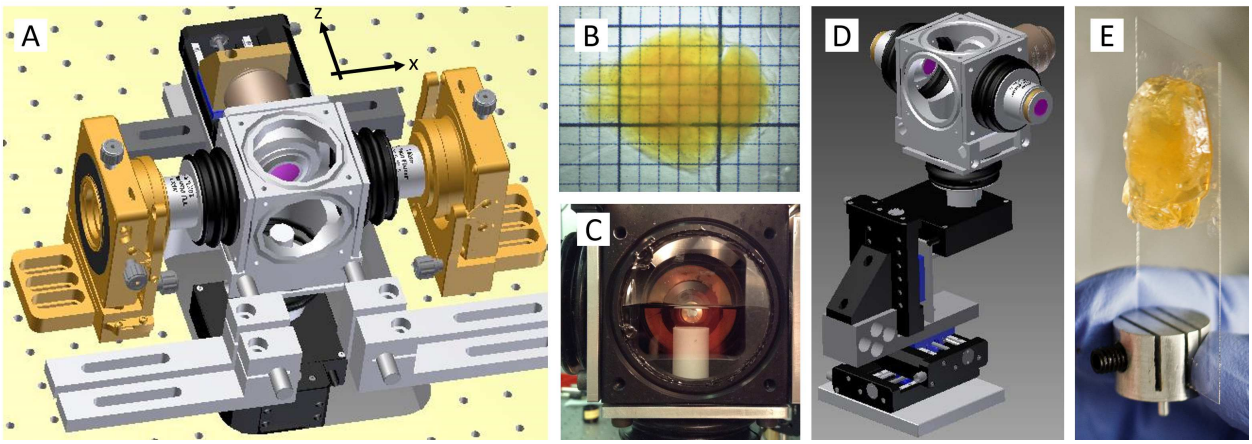

Fig 1: Schematic of light sheet microscopy. (A) Fluorescence excitation (along x axis) and detection (along z axis) are operated on independent, perpendicular light paths where the excitation light sheet and the detection focal plane overlap. The clarified, fluorescently-labelled brain (B) is mounted on a Teflon cylinder inside the watertight chamber (D) and can be translated and rotated freely with piezo motors (D). The sample is glued onto a cover slip which is slid into a custom made adapter mount that is inserted into the Teflon cylinder (E).

on refractive index matching which render the tissue transparent. This makes LSFM a truly interdisciplinary field in which the technological advances by optical developers need to be matched by novel development in the area of information and biotechnology.

Here we will review a state-of-the-art LSFM, as it is implemented in our lab, which can be used to acquire 3D data sets of clarified mouse brains. The LSFM features double-sided illumination with a digitally scanned light sheet and a sample chamber which has been specifically designed for the the imaging of large ($> 1cm^3$), immersed and freely movable samples. After briefly presenting the optical clearing protocol we employ to render our samples transparent, we explain how to prepare and mount the samples for stable, 3D imaging for several days. In addition to a detailed description of all opto-mechanical components, we further present a practical guide for the alignment of a LSFM, a non-trivial task which in previous publication has too often been summarised in a single sentence. A systematic approach to handle, store and analyse the data is also given. We further explain about the data volume we produce how it can be reduced and handled. We finish with some nice data and an outlook. The final sentence of the introduction needs to be amazing.

2 Optical clearing and sample mounting

Biological tissue is, bar of very few exemptions like the cornea or zebrafish larvae, highly scattering. Scattering is the deflection of light rays due to the heterogeneous mixture of cells and sub-cellular structures with varying refractive indices. Rayleigh scattering, the nearly isotropic and strongly wavelength dependent scattering of light by particles much smaller than the wavelength of light, and Mie scattering, which occurs on cellular structures such as organelles and other refractive structures which are larger than the wavelength of light, far outweigh absorption and cause most biological tissues to be opaque. In the visible spectrum, the scattering of fluorescence photons is dominant over ballistic fluorescence such that for the imaging of large samples optical

clearing methods based on refractive index matching have been employed. These methods involve several chemical sample preparation steps to generate transparent yet structurally and anatomically intact tissue.

A successful optical clearing method not only renders the tissue transparent but also causes neither quenching of protein fluorescence nor tissue distortion and is compatible with repetitive immunostaining. A very promising method called CLARITY[2,3] transforms tissue into a nanoporous, hydrogel-hybridized, lipid-free form. Crucially for high through-put clearing and imaging of large samples, the protocol, however, is ideally also fast, easy and cheap. Our lab therefore development a versatile, simple, rapid and inexpensive clearing method based on a water-soluble agent, 2,2'-thiodiethanol (TDE) which in combination with CLARITY allows for the optical clearing of entire mouse brains or human biopsies. With this protocol (see [cite Irene] for a detailed description) a mouse brain (Figure1,B) takes 1 month to clear at a costs of $XY.

To elucidate neuronal projections and functional connections in structurally intact tissue it is paramount to be able to image centimeter-sized, clarified samples, such as mouse brains, with high resolution in whole mount preparation. The question of sample preparation and mounting in light sheet microscopy is not trivial but requires novel approaches to such extent that it is becoming a separate field of research and very diverse strategies such as FEP tubes, 3D printed chambers and Quarz cuvettes have been reported (cite), however these approaches are limited to much smaller sample volumes.

We designed a cubic, water-tight sample chamber (Figure1,C) that allows access from all six sides while maintaining the 3D integrity of large, clarified and fluorescently-labelled mouse brains. A motorized x-, y-, z-, θ-stage (M-122.2DD and M-116.DG, Physik Instrumente) allows free 3D motion and rotation of a Teflon cylinder which reaches into the centre of the chamber (Figure1,D). The clarified brains are fixed with super glue to a coverslip which is slid into a custom made adapter and tightened with a plastic-capped screw. The adaptor is then inserted into the Teflon cylinder (Figure1,E) with the coverslip being positioned on the far side of the detection objective. Three different slits in the adapter correspond to varying distances to the detection objective and give variability in sample thickness for example to allow also for the mounting of rat brain hemispheres. The clarified brains are imaged immersed in clearing solution composed of 63% TDE in Phosphate buffered saline (PBS) and a refractive index of 1.45. Stable mounting is a key concern as a whole brain tomography can require image acquisition in excess of 3 days. We found that the effect of gravity, tissue shrinking/expansion and evaporation of PBS over such time spans are negligible, however, after that point the super glue is starting to dissolve.

3 Optical path

3.1 Illumination

The custom-made, confocal light sheet microscope is equipped with 5 linearly polarised, cw lasers for fluorescence excitation (Figure2,A). The wavelengths were chosen to excite the most common fluorophores (see Table2 for the manufacturer and specifications of opto-mechanical components) and have each an output power of 50mW. The laser light from each laser is first collimated and expanded with a telescope (f140, f200) and then combined into a common path with a beam steering mirror and a long-pass filter (Semrock, LaserMUX^{TM} series). An acousto-optical tunable filter (AOTF) acts as fast (μs) electronically tunable filter which uses the acousto-optical interaction inside an anisotropic medium to select and transmit any combination of up to four of the laser lines.

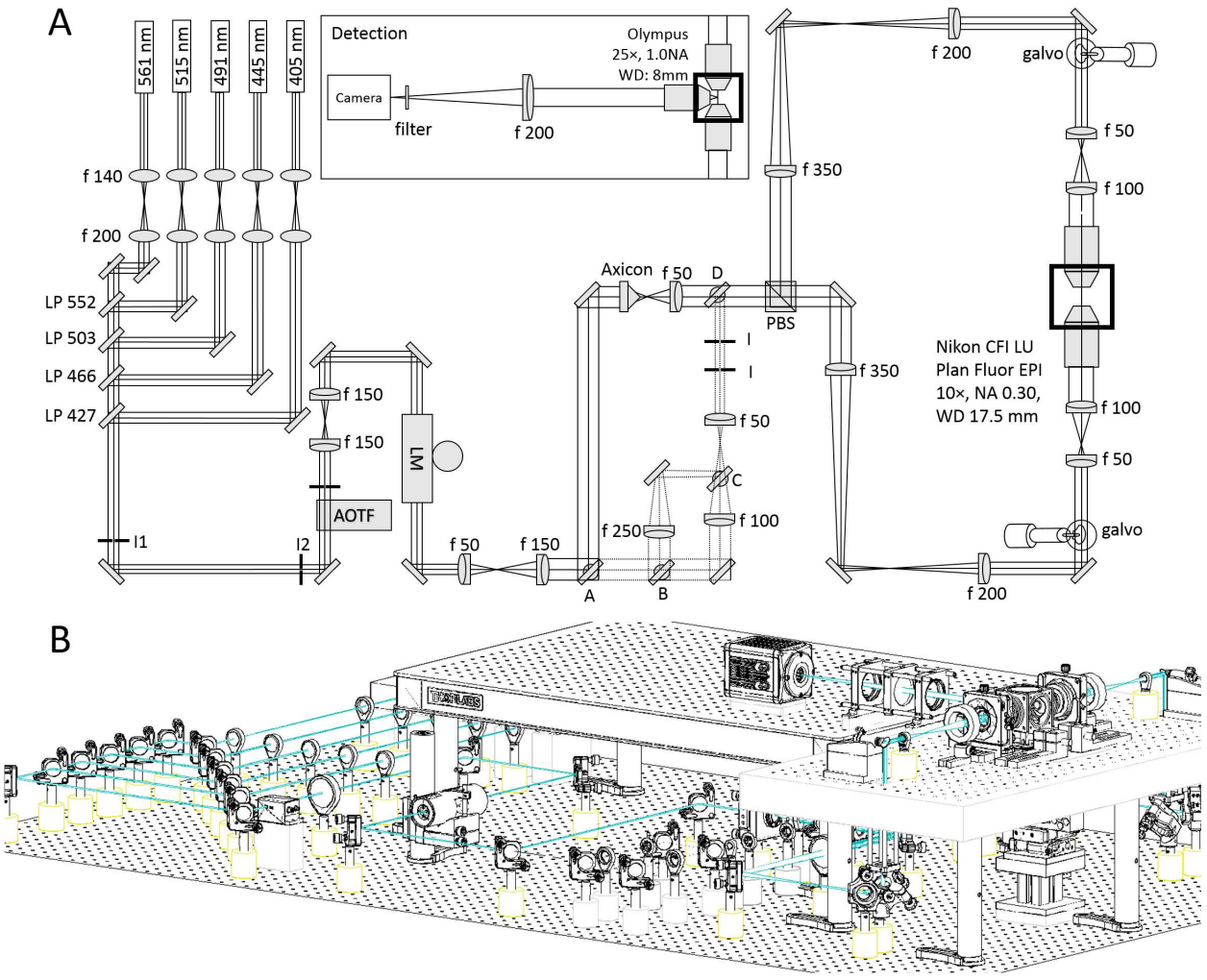

Fig 2: (A) Topview of the excitation path. The galvo scanners are mounted above periscopes. LP: long-pass filter, I: iris, AOTF: acousto-optical tunable filter, LM: laser modulator, PBS: polarisation beam splitter, ABCD: flip mirrors. Inset: detection. (B) Oblique view of the microscope. A custom-made breadboard serves to mount the sample chamber and objectives at an elevated height and features two circular holes at the edges for the periscopes and a large central cut out for the translation stages. A second breadboard is used for the camera.

The radio-frequency applied on the AOTF transducer controls the wavelength being transmitted into the first order and the radio-frequency amplitude allows to adjust the transmitted light intensity. Due to the nonlinear response of the AOTF we measured the AOTF light transmittance for each wavelength as a function of radio frequency amplitude and determined a linear look-up table through interpolation. The zero order light is blocked by an iris. An electro-optical laser modulator acts like a wave plate with electrically controlled retardation and rapidly rotates (hundreds ns/ms?) the input polarisation of the excitation light by $90°$. The wavelength-dependent, high voltage that needs to be applied to the birefringent crystal inside the modulator to change the optical path length is provided by a two-step pre-amplification system. First, a low voltage, analog signal from the DAQ board is pre-amplified with custom made electronics and then fed into a commercial high voltage amplifier. A digital line is used to electronically control the frequency of the polarisation shift. After the laser modulator the excitation beam is further expanded by a factor of 2 (f100, f200) with two achromatic doublets. From here on three different light paths can be chosen through flip mirrors. One light path option guides the light through an axicon (apex angle of $160°$) and an achromatic doublet (f50) to create a Bessel beam. The other two options create Gaussian beams of different beam diameter. This diameter variation translates to different fields of view later on in the detection path. The three options are recombined before a polarising beam splitter cube which splits the excitation light depending on its polarisation into one of the two excitation arms which are built to be identical. The beam is reduced with a telescope (f350, f200) whose telecentric plane coincides with the mirrored surface of a galvanometric scanner (galvo). The galvos are mounted on a custom made optical breadboard which features two circular holes to pass the periscopic beams and a large central cut-out for the sample chamber and motor stages. The light sheet is generated digitally by scanning the excitation beam with a sawtooth signal applied to the galvo. This generates perfectly incoherent illumination resulting in fewer artefacts.The scan mirror surface is reimaged with a telescope (f50, f100) onto the back aperture a long working distance, low magnification objective (Nikon, 10x 0.3NA WD 17.5mm). The two excitation objectives are designed for air immersion. A coverslip glued to the front housing edge serves the dual purpose of maintaining the first diffractive surface between the front optical element and the air and to protect the glue/mounting of the lens elements which is compromised by the organic solvent. The sample chamber is tightly bolted to the optical breadboard while soft connections through silicone bellows allow for adjustive movements of the objectives and free 3D motion of the motor stages. All connections are sealed with rubber rings and silicone caulk and additionally tightened with cable binders. In this way the objectives can be refocused and realigned without compromising the watertight seal of the chamber.

3.2 Detection

Fluorescence is collected on a axis perpendicular to the excitation with an objective that was specifically designed for immersion in clearing solutions of various refractive indices. The objective is equipped with a correction collar allowing for immersion in media with refractive indices ranging from 1.41 to 1.52 (Olympus, XLSLPLN25XGMP). To image the whole volume of interest the objective has a relatively low magnification of 25x and a large working distance of 8mm yet a high numerical aperture (1.0NA) affords high resolution imaging. HERE WE'LL NEED SOME CHARACTERISATION DATA. An overview of imaging properties and derived quantities is given in Table3. We use a tube lens of 200mm to create an image on a sCMOS camera (Orca Flash4.0,

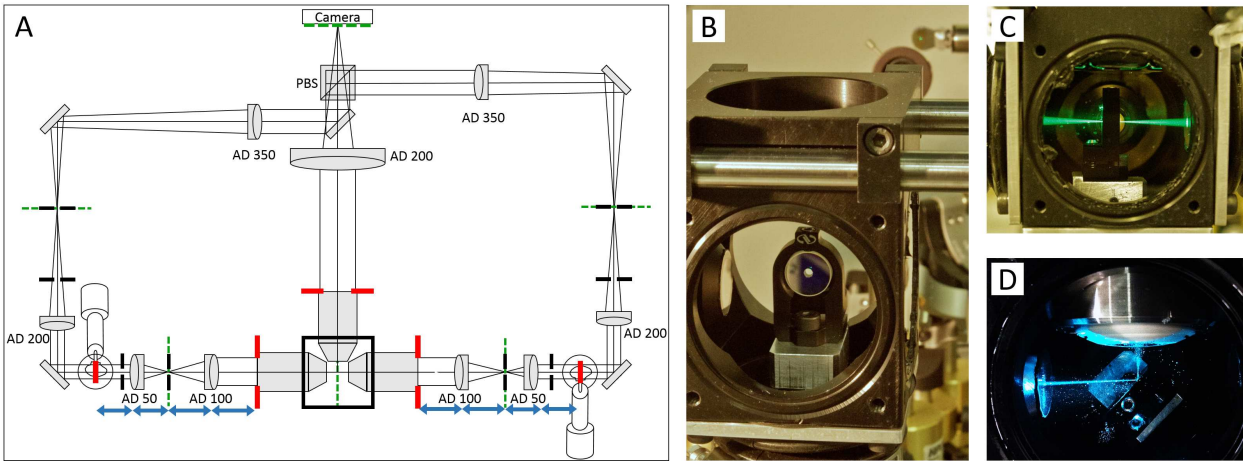

Fig 3: (A) Basic geometrical optics of a double-sided illumination light sheet microscope. AD: achromatic doublet, PBS: polarising beam splitter, red: objective back focal planes and conjugated telecentric planes, green: image planes, blue: 4f telecentric lens system. (B) The alignment mirror can be rotated to precisely reflect at $45°$. (C) With lateral movement the alignment mirror can be placed such that light is transmitted into the opposing excitation arm. (D) Alignment mirror with drilled hole for light transmission, mount for tip adjustment and adaptor for the Teflon tube.

Hamamatsu). A fluorescent filter is used to block out any excitation light. The camera chip of over 4 megapixels is read out in one sweep from top to bottom in the so-called light sheet mode in which only a subset of horizontal lines is exposed and consequently read out and any time. This sweeping line exposure creates a virtual confocal slit. Confocality is achieved by synchronising a single line which generates the light sheet with the exposure of a few lines on the CMOS chip.

For image acquisition the sample is slowly translated at a fixed speed along the detection axis through the light sheet. Images are recorded at a frame rate of 44Hz in different planes along the optical axis of the detection lens (here z) to acquire a stack. After successful stack acquisition the sample is moved to an adjacent vertical position (here y) to acquire the next image stack. Once the entire vertical slice composed of z stacks is acquired the sample is moved to an incremented x position. In this way the entire cubic volume is imaged.

3.3 Alignment

For light sheet microscopy it is crucial that excitation and detection occur on perpendicular axes because any deviation from this geometry results in obscured images of reduced resolution and contrast. The objectives need to be perfectly confocal so that the fluorescence that is generated in the swept excitation beam also falls within the detection objective's depth of field. Telecentricity, that is the imaging with two lenses which are the sum of their focal lengths apart, has to be strictly observed in all three arms (Figure3,A), which means the galvo scanners need to be precisely placed in telecentric planes of the excitation objectives while in the detection path, the camera chip needs to be positioned in an image plane of the detection objective. Additionally, homogeneous illumination from both sides impose strict symmetry considerations on both illumination arms that have not only to be sufficiently aligned within themselves respectively but function as pair with recursive dependence.

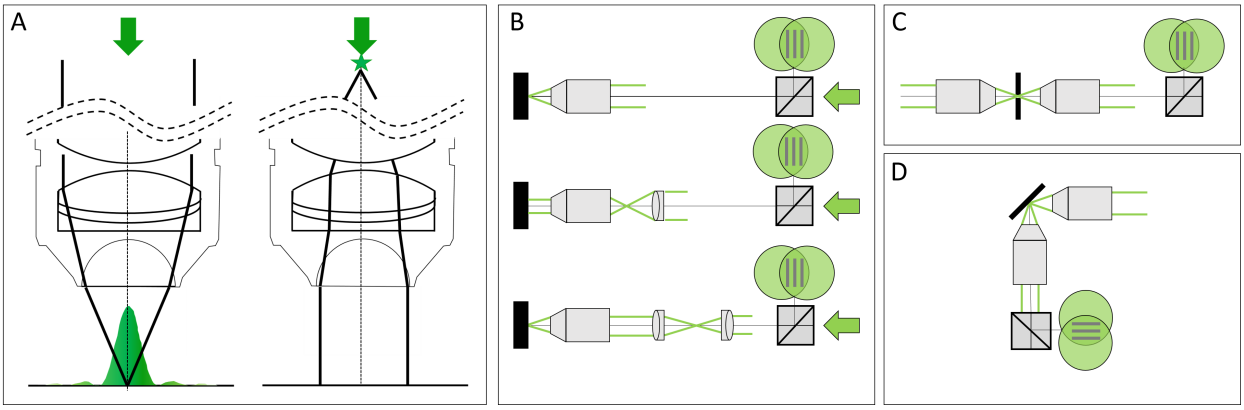

Fig 4: (A) Working principle of a microscope objective (left): a collimated input beam in converted into a spherical wavefront. Right: inverted light path where a focal spot in the back aperture creates a collimated output. (B) Recursive placement of lenses using the sample mirror in back-reflection mode, (C) in transmission mode and (D) in reflection mode.

Our microscope was aligned with two tools, a shear plate to qualitatively assess collimation and a small mirror that can be mounted inside the sample chamber where the sample would usually be. The 0.5in mirror was first pierced with a drill using a ceramic drill bit to produce a hole roughly in its centre of the approximately the same diameter as the excitation beam inside the sample chamber. Using a compact single axis adjustable mirror mount (V50-AX, Newport) attached to the Teflon cylinder inside the chamber allows to to adjust the pitch of the reflection with the mount and the yaw with the rotation stage while at the same time providing a very space efficient mounting for the pierced 0.5in mirror. With this "sample mirror" three different position can now be easily implemented, firstly, back reflection by hitting the reflective surface at $0°$, secondly, transmission by laterally displacing the mirror until the light passes through the drilled hole, and thirdly, reflection of the light by $90°$ by precisely turning the mirror mount by $45°$ using the rotation stage.

The excitation objectives are fixed on mounts which allow three dimensional translation plus pitch and yaw adjustment (LP-1A, Newport). In a first step the beam paths of both excitation arms where brought to overlap through irises placed on the breadboard and the optical bench without any microscope objectives. We found it useful to use two mirrors in each periscope, one vertically mounted and one mounted at $45°$ for vertical deflection of the incoming beam. In this way full beam steering can be achieved to realign the periscope before hitting the galvo scanner. After this initial alignment the first excitation objective was then placed into the beam path and brought to focus onto the sample mirror, which was adjusted to reflect the light back into the same objective. Using a shear plate and a polarising beam splitter cube the back reflected light was qualitatively checked for collimation to ensure the sample mirror was correctly placed in the focal plane of the objective (Figure4,B). The sample mirror was then moved laterally to allow the excitation beam to pass through the drilled hole and the second excitation objective was placed (Figure4,C). This time collimation was checked with the light going through both objectives and so their confocal placement was insured. The microscope was then aligned recursively by starting at the confocal point between the objectives and placing successively lens after lens in the up-stream direction. Finally, the sample mirror was position in reflection mode and the detection objective was aligned (Figure4,D).

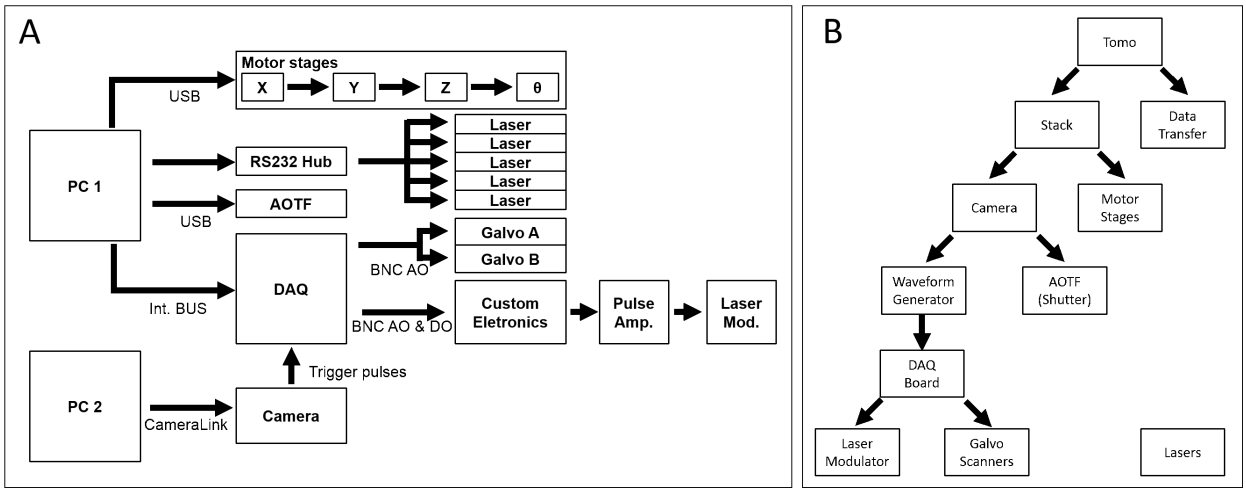

Fig 5: Schematic views of the hardware (A) and software (B) architectures of the microscope, showing physical and logical connections, respectively. The software architecture scheme clarifies the hierachic structure of the Murmex components. AO analog output, DO digital output, DI digital input.

4 Optical characterisation

5 Microscope management software

5.1 General design principles

The many components of the microscope (camera, lasers, AOTF, stages, etc.) are orchestrated together and operated through a custom-made software. Since the microscope should work mainly without human intervention for many hours (a whole-brain tomography can last a couple of days) without losing proper synchronization, software design is not a secondary issue when building a light sheet microscope for whole-brain imaging. The overall architecture of the software and its practical implementation were chosen to satisfy several principles:

- *User-friendliness*. The light sheet microscope here presented was designed and developed by optical technology specialists yet needs to be operational in a multi-user environment. The greatest challenge therefore consists in making the use of the microscope easy and intuitive such that no formidable expertise is needed when it is operated by researchers with no expertise in optics and/or software development.

- *Working in a distributed environment*. The apparatus should deal with many independent hardware components, and with several data streams (see below). Individual processing steps and/or data transfer can be computationally intensive, and may need to be distributed over several computers to perform properly at run-time.

- *Robustness and ease of maintenance*. Any kind of hardware problem, from timing delay to real failure, must be automatically handled to avoid loss of precious data. Furthermore, developers must be able to fix existing code or to develop new one without having to change the whole software.

- *Reusability and scalability*. When a second or third microscope with basic similarity yet slightly specialised equipment is being designed in the same or in other laboratories, re-writing all the code from scratch is pure loss of time. Writing well structured, scalable and easily maintainable scientific software can help avoiding redundancy and boosting up productivity.

Our microscope management software has been written in LabVIEW (National Instruments, Austin, US), as this language is particularly well suited for hardware control and data acquisition through pre-developed methods. An object-oriented programming framework[7] has been used, dividing the entire project into well-defined self-contained modules. This framework, besides helping defining the proper scope of each software component, guarantees robustness, ease of maintenance, scalability and reusability. Code is stored on GitHub (www.github.org) to help multiple users developing software simultaneously. The graphical user interface (GUI) provides intuitive control of the apparatus as well as deep personalization. In general, the intrument parameters are already set to optimal values stored on disk, and on-line tips on how to set or change the various settings are provided. Both GUIs and code are developed according to the LabVIEW Style Guide.[10]

5.2 Murmex

To let the software operate efficiently and synchronously on different machines, the Murmex software development kit was used. This kit integrates the object-oriented programming paradigm with a messaging system, and can be used to create communicating yet independent software modules within a standardized programming scheme.

Basically, Murmex creates distributed finite state machines. These are design patterns in which one component proceed from one particular state to another based on the message it receives either from itself or from other components. For sending and receiving messages, Murmex uses the library LabbitMQ, which is a wrapper of RabbitMQ, a message-oriented middleware facilitating message shipment and receival. These layers of abstraction are mostly hidden for the developers. The RabbitMQ server (message broker) needs to be installed either locally or somewhere in the network. The message-based communication allows for buffered, asynchronous and reliable message communication between several software components. Messages are routed by the broker to the specific software component by its ID. The LabbitMQ wrapper allows use of RabbitMQ inside the LabVIEW programming environment.

Murmex design pattern has twelve pre-defined common states (Initialise, Configure, Start, Stop, Proceed, Update, Reset, Inspect, Inform, Acknowledge, Fetch, Terminate) and a generic state for implementing custom states. Usually, just a few of this states are actually used in standard application. An example of a standard implementation of the Murmex state machines for a scanning system is shown in Fig. BLABLA. Each Murmex component streams information about its current status to its 'observers', other components specified in its configuration file. Using this oberserver/observee hierarchy, it is possible to design large and complex network of different components, running on whatsoever number of computers connected through a local connection or the Internet. A special component, named ServiceManager (also based on Murmex), which must be running on each computer, can start executables remotely. This allows the user to start up a sophisticated configuration with multiple components executed on different computers, but steered from one computer.

5.3 System timing and software organization

A schematic overview of the hardware and software architecture of our light sheet microscope is shown in Fig. BLABLA, respectively. One personal computer (Precision T5600, Dell, Round Rock, USA) controls all the remote hardware components with the exception of the camera, which is controlled by a second, dedicated computer (Precision T7500, Dell, Round Rock, USA) to handle and manage the enormous streams of data produced (see next section). Various instrumentations are addressed either through computer ports (as USB, RS232 and CameraLink) or via analog and/or digital signals generated by a DAQ board (NI PCIe-6353, National Instruments, Austin, USA).

Synchronization of the camera rolling shutter to the scanning of the light sheet is achieved by letting the camera acquiring in free-run mode, and using a trigger output from the camera itself to time the generation of the waveforms sent to the galvo drivers. In particular, the HSYNC signal from the camera is used, providing a true-type-logic (TTL) pulse everytime the rolling shutter steps from one line to the next one. This trigger signal is used as a clock by the DAQ board, so that the line producing the light sheet moves step by step with the rolling shutter. The same HSYNC signal is used to trigger the DAQ digital signal sent to the electro-optic laser modulator. In this case, since at each position of the scanning line the modulator should go trough at least two states (vertical polarization and horizontal polarization), HSYNC is used as a trigger while the clock used is the (much faster) internal clock of the DAQ board.

Each hardware component (AOTF, galvo scanners, electro-optical modulator, sample stages, camera, lasers) is managed through its own Murmex component. Since the galvo scanners and the electro-optical modulator are not addressed directly, but via analog and digital signals generated by the DAQ board, their software component just send analog and digital waveforms to an additional intermediate component responsible for DAQ board management. Beyond hardware-related software, two additional Murmex components (named Stack and Tomo) orchestrate the operation of the entire system.

The stack component is used to acquire single image stacks along the z direction, and to this aim send specific messages to coordinate the Camera and the Stages components. Stacks are collected by moving continuously the sample along z while keeping the camera acquiring in free run. The translation speed is such that in the time needed to collect a single frame the specimen has moved by a z step specified in the GUI. The Camera component uses low-level libraries from Hamamatsu to save the data, producing one single *.cxd* file (which is a proprietary tiff-like format from Hamamatsu) for each image stack. The name of this file and the path to it is specified by the Stack component.

The tomo component is used to image the entire sample by collecting many adjacent, parallel, partially-overlapping stacks. To this aim it exchanges messages with the Stack and the Stages components. The user can specify the volume to be imaged either by inserting the maximum x, y and z coordinates of a parallelepiped, or by providing a text file with the list of stacks to be collected (x and y coordinates, starting and ending zs). The Tomo component also inform the experimenter about the state of the acquisition by sending an email when it's finished or if any trouble is experienced.

All the software described here can be freely downloaded from GitHub. Beacuse of its modular structure and of the extensive documentation, it can be easily integrated and adapted in other microscopy systems.

6 Data Management

The volume of a mouse brain have been estimated from MRI measurements[9] as $\simeq 0.5$ cm^3; the same Authors find out that a parallelepiped encompassing the whole brain would have a volume of $\simeq 0.9$ cm^3. Given the pixel size of the camera and the magnification of the system, and assuming a z step of 2 μm, the voxel size results $0.234 \times 0.234 \times 2$ μm^3. This means that a mouse brain will be represented with about 4.6 TeraVoxels, while a parallelepiped encompassing the brain will result in 8.2 TeraVoxels. Since the camera produces 16 bit images with no compression, each raw voxel occupies 2 bytes of disk space, resulting in 8.3 TB for the whole brain "alone" and 14.8 TB for the parallelepiped (see Table XXXX). However, image stacks are acquired with a 10% of linear superposition, which is used later for image stitching. This redundancy expands the size of raw data to 10.2 TB and 18.3 TB, respectively.

This unprecedented amount of data challenges to the traditional way we interact with our data. Having far passed the point of taking your data home with you on a hard drive at the end of the experiment, data management in light sheet microscopy requires the development of a robust data management infrastructure, and of novel software tools to process the images in order to extract meaningful information.

6.1 Data handling and storage

The camera used in our light sheet microscope (Hamamatsu Orca Flash 4.0 v2) flushes the datastream directly to a dedicated workstation with 12 solid-state drives (SSD) operating in RAID 0 configuration, resulting in a virtual drive of 10.9 TB. Although this is already a large and expensive system for the current SSD technology, it is way smaller than the size of the raw datasets produced in each tomography. We thus connected the workstation to a larger network area storage (NAS) via a 10 Gbit/s network copper cable. The NAS operates in a safer data modality (RAID 5) with 32.4 TB of useful disk space. The NAS is then connected by a dedicated 10 Gbit/s link provided by the Consortium GARR (the Public Institution managing research networks in Italy) to CINECA (the main italian supercomputing center). The data flow scheme is summarized in Figure 6.

While transferring the data from the SSD to the NAS, we also perform a file conversion from the proprietary Hamamatsu format to a more manageable one. We chose the TIFF format among the others because it permits to store 3D image with lossless compression in a single file. Furthermore, it is a flexible and platform-independent format supported by numerous image processing applications. We thus developed a conversion module that leverages the functionalities offered by the interface of the DCIMGAPI library, supplied by Hamamatsu, and the freely available LibTIFF library. The module reads a single .cxd 3D image and writes it in a variable number of multi-page TIFF files, depending on the user-defined limit of pages per file. Since the output of the Hamamatsu camera may exceed tens of Gbytes, the image in proprietary format is read slice by slice to minimize memory requirements. The module has also an option to downsample the image on save the input image at full or reduced resolution, in order to adapt to storage availability. Usually we downsample the image in the XY plane by a factor 2 and leave the Z sampling as usual. This kind of downsampling, in combination with the lossless compression offered by the TIFF format, reduces data size by about one order of magnitude. Thus the parallelepid volume cointaining the whole mouse brain can be represented using about 2-3 TB of disk space.

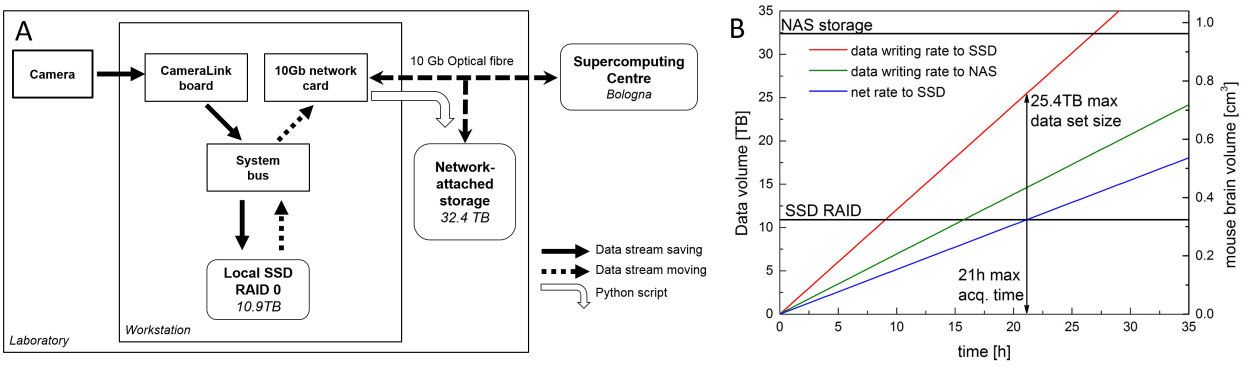

Fig 6: Scheme of the data flow. (A) Data streaming on the microscope. SSD: solid-state disk, RAID: redundant array of independent disks, NAS: network-attached storage. (B) Data production rate versus data conversion rate showing the maximum acquisition time and tomography size feasible before the SSD are filled up.

Although file conversion drastically reduces data size, in the current implementation it is quite slower than data production. The Camera component flushes data on the SSD with a rate R_{data} dependent on frame rate R_{frame} and on the image size:

$$R_{data}[\text{Bytes}] = 2 \times R_{frame} \times \#_{pixels}\,, \tag{1}$$

where the factor takes into account the 16 bit grey depth of images. The volumetric imaging rate is given by $R_{volume} = 0.5 R_{data} \times$ Voxel size.

With a standard image size of 2048×2048 pixels, and $R_{frame} = 44$ Hz, data are produced at a rate $R_{data} = 352$ MB/s ≈ 1.2 TB/h, which results in a volumetric rate $R_{volume} \approx 2 \times 10^7\ \mu\text{m}^3/s$ ≈ 0.07 cm^3/h. The conversion module can empty the SSD wih an average rate of about 0.7 TB/h, meaning that the SSD is actually filled up during imaging at a rate of about 0.5 TB/h. Imaging sessions should thus be shorter than 21.8 hours: this limits the maximum amount of raw data being acquired continuously to approximately 25.4 TB. This is well above the disk space needed to image the parallelepieped volume encompassing a whole mouse brain. Larger specimens, as rat brain or portion of human brain, must be imaged in consecutive session until faster strategies for image conversion are devised.

6.2 Stitching and visualisation

Given the limited field of view of the microscope, the acquisition of macroscopic specimens required many parallel image stacks to cover all the volume. Thus, in order to achieve a 3D image of the whole specimen from raw data, the Terastitcher[6] has been recently proposed, i.e. a stitching tool capable to deal with teravoxel-sized images. However, the Terastitcher does not support input data acquired through the serial sectioning procedure, which leads to a specimen partitioned in different layers. Furthermore, only single channel images can be processed. For these reasons, we extended the Terastitcher functionalities introducing the two following additional features: i) stitching of a specimen partitioned in a number of overlapping layers for the hippocampus reconstruction and, ii) coping with images containing more than one channel for the human brain tomography. With respect to the first requirement, that is allowing a complete reconstruction of a multi-layered raw

Table 1: Data production rates with two python scripts running in parallel.

frame rate	44	Hz
size of 1 frame	8	MB
frames per stack	3900	
size of original stack	30.4	GB
size of converted stack	3	GB
stack acquisition time	88.6	s
stack conversion time	140	s
time ratio conversion/ acquisition	1.6	
data production rate	+1.2	TB/h
data conversion rate	-0.7	TB/h
net rate	+0.5	TB/h
SSD RAID	10.9	TB
max acquisition time to fill up SSDs	21	h
max tomo size to fill up SSDs	25.4	TB

data, we schematically depicts in Fig.1 the adopted strategy. First of all, the various input layers, each of which is composed of several parallel overlapping stacks, are separately stitched by using the existing Terastitcher tool. After this preliminary step, leveraging the layer coordinates provided by the instrument, we import the processed layers as a new volume where each layer has a partial overlap with adjacent layers. Furthermore, each layer is organized in a non-overlapping tiled format, enabling the application of a multi MIP-NCC approach,[6] which computes the displacement between two adjacent layers along all of the three directions (Fig. 1). When a displacement computation for each pair of adjacent layers has been computed, the overlapping regions are merged through a blending procedure which smooths the transition between layers, resembling once more the procedure used by Terastitcher for combining adjacent stacks. Finally, we note that due to the repositioning of each layer, the volume containing the reconstructed specimen might contain empty regions, which are thereby filled with black voxels. Turning our attention to the second additional feature, that is handling multi-channel images, the Terastitcher has been extended so that the MIP-NCC algorithm used for displacement computation can work on an image, which can be either the fusion of the input channels or one of the input channels. This permits to privilege the channel with more information content over the other channels as well as discard noisy channels. Note that the reconstructed 3D image can be produced with the same channel composition of raw data.

6.3 Example for postprocessing: Automated cell counting

See Reference[5] for details.

7 Data

Everybody take a look at Figure7 because it is the best we have.

8 Conclusion

summary Light sheet microscopy has already been a game changer for large-scale imaging by yielding data with a combination of unprecedented spatio-temporal scale. The impact of this

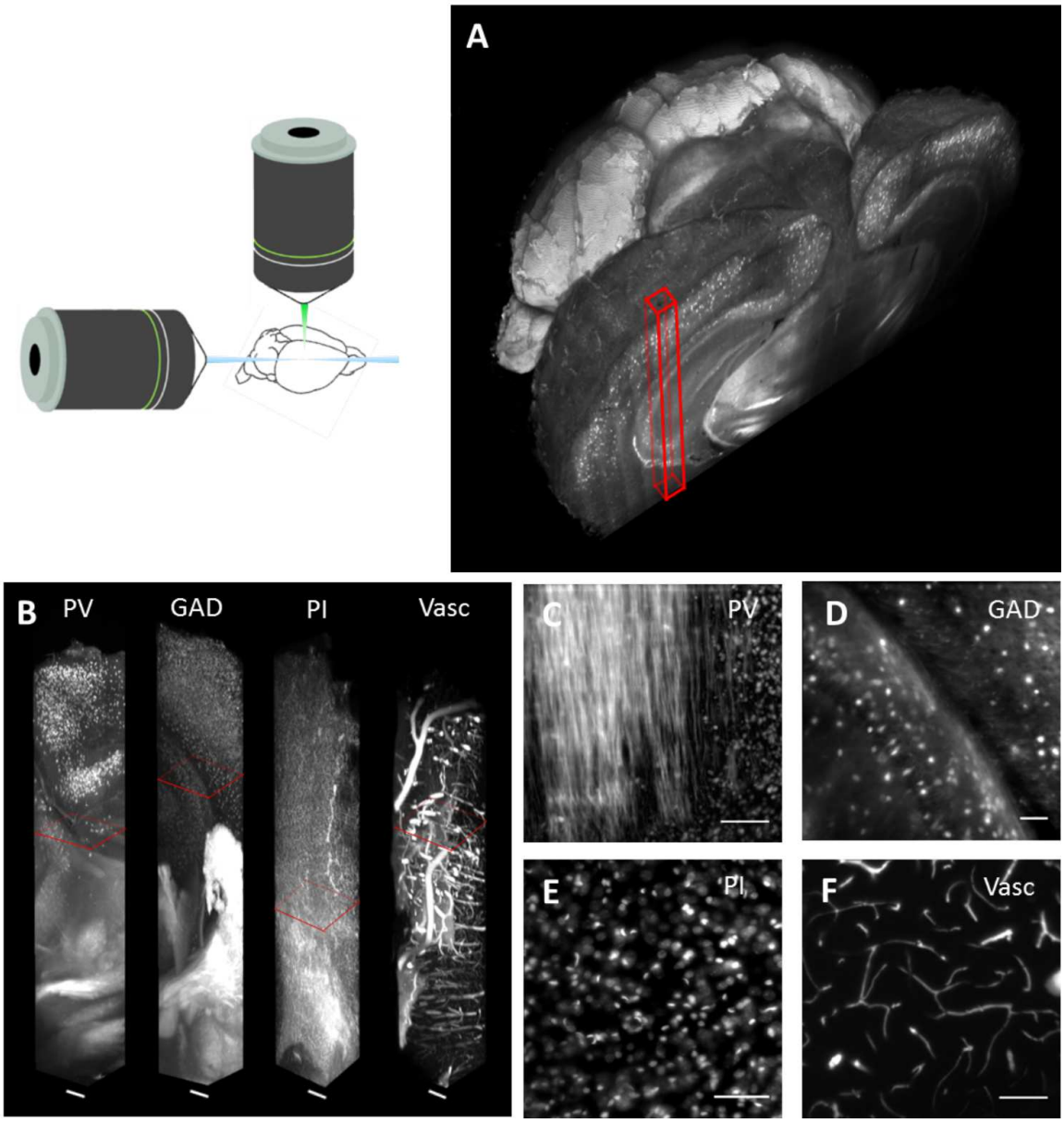

Fig 7: Whole mouse brain tomography. Imaging of whole transgenic mouse brains treated with CLARITY and cleared with TDE 63% imaged with LSM (Olympus, 25X objective). (A) 3D rendering of a parvalbumin-dTomato brain. (B) 3D rendering of stacks from PV-dTomato mouse brain, GAD-dTomato mouse brain, PI stained mouse brain, FITC-albumin labeled mouse brain, scale bar = 400 m. (C,D,E,F) High resolution insert, of the stack corresponding to red boxes in C. Scale bar = 100 m.

unique measurement technique will continue to revolutionise the field of whole brain connectomics due to its ability to record millions of images over the course of days or even weeks. The data produced in this fashion easily amounts several TB per data set and needs to be stored, transferred, retrieved, processed and visualised necessitating the concurrent development of novel computational interface and analysis methods. The latter need to be robust, standarised and fully automated processes yet allow for flexible, exploratory and tailored analysis.

9 Outlook

where are we going next with this? ask Leo Onofri and Giulio to contribute to this, issues data mangament, aberration correction mapping of complementary data sets obtained in the same system, eg structural connectivity, morphology and gene expression.

Table 2: Overview of optomechanical components.

Component	Manufacturer	Part#	Specifications
Lasers	Cobolt AB, Sweden	MLD	405nm (for DAPI), 50mW, s-polarised
		MLD	445nm (for CFP), 50mW, s-polarised
		Calypso	491nm (for GFP, FITC), 50mW, s-polarised
		Fandango	515nm (for Venus, YFP), 50mW, s-polarised
		Jive	561nm (for dTomato, PI, RFP), 50mW, s-polarised
AOFT	AA Opto-Electronic, France	AOTFnC-400.650-TN	$> 90\%$ diffraction efficiency, 3nm resolution, low cross talk between laser lines, high separation angle
Laser modulator	Qioptiq GmbH, Germany	LM 0202 VIS ADP	400-650nm , $\lambda/2$-voltage (633nm): 210V
Pulse amplifier	Falco Systems, The Netherlands	WMA-300	50x amplification up to $\pm$ 150V, DC to 5MHz signal bandwidth
Galvo scanner	Cambridge Technology, USA	6220H	small angle step response $200\mu s$
Objectives	Nikon; Japan	Plan Fluor EPI	10x0.3NA, WD 17.5mm, EFL 20mm (excitation)
	Olympus, Japan	LMPLFLN20X	20x0.4NA, WD 12mm, EFL 9mm (detection)
Motor stages	Physik Instrumente, Germany	C-863.11	DC servo-motor controller
		M-122	Travel range 25mm, $0.1\mu m$ resolution, max. velocity 20mm/s
		M-116	Precision Rotation Stage, $2.5\mu rad$, max. velocity $20°/s$
Camera	Hamamatsu, Japan	Orca Flash4.0	sCMOS sensor, 2048(H) x 2048(V), cell dim.: $6.5\mu m$, active area: 13.3mm x 13.3mm, 16bit images
DAQ board	National Instruments, USA	NI PCIe-6353	AI: 1 MS/s multichannel; 16-bit resolution, 10 V; AO: 2.86 MS/s, 16-bit resolution, 10 V; digital I/O lines (hardware-timed up to 10 MHz), 100MHz max counter frequency
Workstation	Dell, USA	T7500	12GB RAM, Intel Xeon Processor X5647 @ 2.93 GHz, 64bit OS, Win7

Table 3: Imaging properties and derived quantities. The thickness of the light sheet is approximately of the same size as the depth of field of the high NA detection lens.

Detection				
wavelength	λ		0.5	μm
refractive index	n		1.45	
numerical aperture	NA_d		1	
angular aperture	α	$= sin^{-1}(\text{NA}_d/n)$	43.6	deg
magnification	M		25	
tube lens	f_{TL}		180	mm
effective focal length objective	EFL	$= f_{\text{TL}}/M$	7.2	mm
diameter back focal plane	BFP	$= 2\text{EFLNA}_d$	14.4	mm
field number	FN		18	mm
field size in specimen	S	$= \text{FN}/M$	0.72	mm
depth of field	Δ	$= \lambda * n/(\text{NA}_d)^2 + n * e/M\text{NA}_d$	1.10	μm
Airy radius lat.	r_A	$= 0.61\lambda/\text{NA}_d$	0.31	μm
Excitation				
numerical aperture	NA_e		0.3	
magnification, nominal	m		10	
tube lens	f_{tl}		100	mm
effective focal length, objective	efl	$= f_{\text{tl}}/m$	20	mm
magnification, effective	m_e	$= f_{\text{tl}}/\text{efl}$	5	
refractive index	η		1	
beam radius	ω		4.5	mm
min light sheet waist	ω_0	$= \lambda\text{efl}/\pi\omega$	1.41	μm
confocal parameter	b	$= 2\pi\omega_0^2/\lambda$	25.15	μm
Camera				
cell size	e		6.5	μm
effective area	r^2		13.3^2	mm^2

Acknowledgments

Human Brain Project tutta la vita. GARR gave us a beautiful 10gb fiber to cineca CINECA for hosting us on pico

References

1 J. Huisken and D. Y. Stainier, "Selective plane illumination microscopy techniques in developmental biology," *Development* **136**(), 19631975 (2009). 1

2 Chung, K., Wallace, J., Kim, S. Y., Kalyanasundaram, S., Andalman, A. S., Davidson, T. J. & Deisseroth, K. Structural and molecular interrogation of intact biological systems. *Nature* **497**, 33233 (2013). 3

3 Tomer, R., Ye, L., Hsueh, B. & Deisseroth, K. Advanced CLARITY for rapid and high-resolution imaging of intact tissues. *nature protocols* **9**, 1682-1697 (2014). 3

4 L. Silvestri, A. Bria, L. Sacconi, G. Iannello, and F. S. Pavone, "" *Opt Express* **20**(), 20582-20598 (2012).

5 P. Frasconi, L. Silvestri, P. Soda, R. L. Cortini, F. S. Pavone and G. Iannello, "" *Bioinformatics* **30**(), i587-i593 (2014). 13

6 A. Bria and G. Iannello, "TeraStitcher-A tool for fast automatic 3D-stitching of teravoxel-sized microscopy images," *BMC bioinformatics* **13**(), 316 (2012). 12, 13

7 G. Castagna, *"Object-Oriented Programming A Unified Foundation,"* Birkhäuser, Boston (1997). 9

8 I. Costantini, J.-P. Ghobril, A. P. Di Giovanna, A. L. Allegra Mascaro, L. Silvestri, M. C. Muellenbroich, L. Onofri, V. Conti, F. Vanzi, L. Sacconi, R. Guerrini, H. Markram, G. Iannello, and F. S. Pavone, *submitted*.

9 N. Kovacevic, J. T. Henderson, E. Chan, N. Lifshitz, J. Bishop, A. C. Evans, R. M. Henkelman, and X. J. Chen, "A Three-dimensional MRI Atlas of the Mouse Brain with Estimates of the Average and Variability", *Cereb. Cortex* **15**(5), 639-45 (2005). 11

10 P. A. Blume, *"The LabVIEW Style Book (National Instruments Virtual Instrumentation Series),"* Prentice Hall PTR (2007). 9

List of Figures

List of Tables