

Image-processing Pipelines: Applications in Magnetic Resonance Histology

G. Allan Johnson*^{a,b}, Robert J. Anderson^a, James J. Cook^a, Christopher Long^{a,b}, Alexandra Badea^{a,b}
^aDuke Center for In Vivo Microscopy, Radiology, 311 Research Drive, Durham, NC, USA 27710;
^bDepartment of Radiology, Duke University Medical Center, Durham, NC, USA 27710

ABSTRACT

Image processing has become ubiquitous in imaging research—so ubiquitous that it is easy to lose track of how diverse this processing has become. The Duke Center for In Vivo Microscopy has pioneered the development of Magnetic Resonance Histology (MRH), which generates large multidimensional data sets that can easily reach into the tens of gigabytes. A series of dedicated image-processing workstations and associated software have been assembled to optimize each step of acquisition, reconstruction, post-processing, registration, visualization, and dissemination. This talk will describe the image-processing pipelines from acquisition to dissemination that have become critical to our everyday work.

Keywords: magnetic resonance microscopy, image processing

*gjohnson@duke.edu

phone 1-919-684-7754

fax 1-919-684-7158

www.civm.duhs.duke.edu

1. INTRODUCTION – MAGNETIC RESONANCE HISTOLOGY

Paul Lauterbur closed his seminal article ¹ on magnetic resonance imaging with the statement “Zeugmatographic (imaging) techniques should find many useful applications in studies of the internal structures, states and composition of microscopic objects.” Magnetic resonance imaging (MRI) has certainly revolutionized modern clinical medicine. But, Professor Lauterbur’s observation that MRI could be applied at microscopic resolution has also been transformative in the basic sciences and has been the basis of our research at the Duke Center for In Vivo Microscopy (CIVM) for the last 30 years. The interplay of engineering, physics, and computer science with applications across a broad field of biomedical science has provided fertile ground for innovation and discovery.

One of the most fascinating applications of magnetic resonance microscopy has been its use in histology—the study of tissue architecture using magnetic resonance microscopy. The concept was first articulated in 1993 ². While the spatial resolution of magnetic resonance histology (MRH) does not approach that of conventional optical methods, MRH provides three benefits that differentiate it from more traditional optical techniques: 1) MRH is non-destructive and provides wide field-of-view coverage (e.g. whole brain); 2) MRH provides a unique source of tissue contrast based on water and how that water is bound in the tissue; 3) MRH is inherently three- (and many more) dimensional.

We, and others have chased the resolution limits in MRH through a variety of approaches, including more sensitive radio frequency coils, novel image encoding strategies, and sophisticated image processing. As resolution increased, the size of the imaging arrays increased. The introduction of isotropic imaging in 1991 highlighted the need for larger three-dimensional arrays ³. The use of multi-echo sequences to enhance contrast and signal-to-noise extended the arrays to four dimensions ⁴. Diffusion tensor imaging has pushed the arrays to higher dimensions ⁵. And, the development of population atlases has added yet one more dimension ⁶. The key to all of these innovations has been the development of image-processing pipelines to automate and standardize the process, so we can routinely handle very large, multidimensional image arrays.

2. METHODS

Figure 1 shows an overview of the process for handling large multidimensional data sets. We have taken advantage of the extraordinary work of many other investigators who have developed sophisticated packages—mostly for clinical MRI. We have extracted and adapted modules for our specific interest in MRH. Data flow is compartmentalized into

five separate pipelines: acquisition, reconstruction, post-processing, archive, and dissemination. Dedicated hardware and software are optimized for each pipeline.

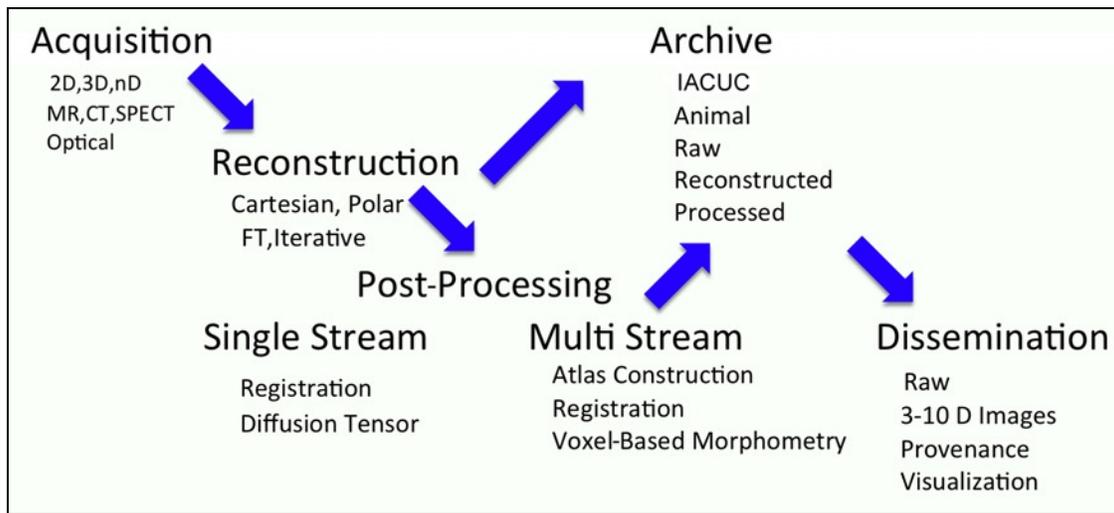


Figure 1. By breaking the process into modular, specific tasks, hardware and software for each task can be optimized to allow routine processing of very large multidimensional images.

The CIVM has 14 different imaging systems. Separating acquisition and reconstruction allows us to reuse reconstruction algorithms initially developed for one system or modality as a starting point for another system. Two MR systems are dedicated to MR histology: a 7.0 T 210-mm bore magnet and a 9.4 T 89-mm bore magnet, both with Agilent Direct Drive consoles. Scripts on each system automate the acquisition of data for multidimensional arrays that can run unattended for several weeks. Web-based sensors on the system allow us to remotely monitor temperature and data acquisition.

Six dedicated reconstruction engines serve five MRI systems (including the two dedicated to MRH). A typical system consists of a multi-(12) processor Mac Pro® with at least 64 GB of memory, and at least 1 TB of RAID that is used as a local scratch disk. Reconstruction is done with a combination of MATLAB, C, and Perl scripts.

The same hardware can be repurposed for post-processing single stream pipelines. A “single stream” pipeline works with one data set at a time. The most heavily used pipeline is dedicated to diffusion tensor imaging (DTI), and it employs MATLAB, FSL <<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>>, Advanced Normalization Tools (ANTs) <<http://picsl.upenn.edu/software/ants/>>, and the Diffusion Toolkit/TrackVis <<http://trackvis.org/dtk/>>. Avizo (FEI Visualization Sciences Group) is resident on every system to inspect the data through the course of processing.

A system designed to handle multiple streams simultaneously is used for generating atlases, registration of groups of images, and voxel-based morphometry. It employs a compute cluster with 11 nodes (32 logical cores/node), with 256 GB memory/node (2.8 TB total). That system includes a 32 TB scratch array linked of via 10 Gb/sec NAS.

A dedicated Oracle database coupled via a high-speed (10 Gb/sec) switch to 60 TB NAS provides a central store for all curated data. Software robots distributed on all of the pipelines automate the archive of large arrays and ingestion of metadata into the Oracle database. A second RAID provides local backup. Daily access is directed through the primary archive and permissions on this backup are limited to preclude potential corruption.

All the pipeline’s modules communicate through the use of an image header. The header, generated at the time of acquisition, documents acquisition, processing parameters, job status data on the project, as well as the project owner, pulse sequence used, and acquisition parameters. These data, in turn, inform the reconstruction engine how the data is to be reconstructed. Details of the reconstruction, e.g. the use of a Fermi filter, zero padding etc., are added to the header providing provenance for the processing of the data set as it passes through the pipelines. A diffusion tensor image for

example, includes the gradient values and gradient angles for each image. These data are accessed via the diffusion tensor pipeline to fit the tensor data and derive the scalar images, described in the following sections.

As data volumes continue to grow in both absolute volume and number of dimensions, dissemination has become more challenging. Data dissemination is accomplished by two systems. The first is a dedicated MySQL database with 10 TB of RAID. The majority of our publications include a link to this MySQL database that allows readers to download any of the image sets discussed in a given publication (see <http://www.civm.duhs.duke.edu/SharedData/DataSupplements.htm> for examples). A second, newer system is being developed to address even larger data sets—libraries of images with total volume exceeding 100 GB. Simple download of these image libraries, while possible, is problematic for two reasons: 1) The large volume makes download tedious even with a high bandwidth connection; and 2) once downloaded, the user must have specialized hardware and software to effectively interact with the data. We have addressed this through the addition of two NVIDIA GRID servers, each with two K2 graphics cards providing more than 6000 cores for interactive visualization. Citrix and a middle level of software generate virtual GPUs for every user. Software derived from 3D Slicer <<https://www.slicer.org/>> provides the users interactive visualization protocols designed to help navigate the complex multidimensional arrays. Since the raw data remains at the server, users with very modest hardware can interact with very large image arrays.

3. RESULTS

Figure 2 shows the steps in the reconstruction pipeline. Modules for remote copy automate the secure transfer of data from any scanner. This may include multiple individual runs that will be aggregated into a single image. Or, it may include a single file, which contains Fourier data for multiple images that must be separated and sorted before reconstruction.

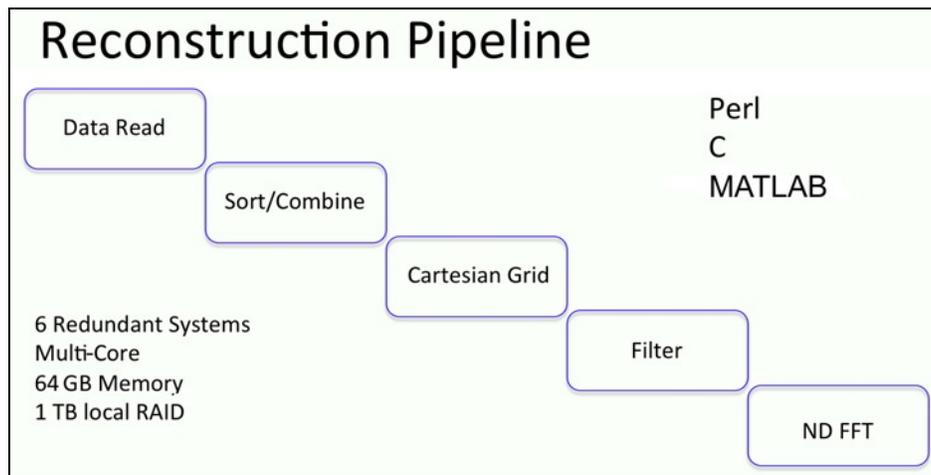


Figure 2. The reconstruction pipeline provides an automated way to separate and combine raw data from multiple acquisitions, resample data from unique trajectories in Fourier space, and include 3D filters, all prior to Fourier transform and magnitude or phase calculation.

Use of this pipeline is best demonstrated by the example in Figure 3. We introduced the concept of MR histology with isotropic spatial resolution using “large arrays” in 1991³. At that time, a 256^3 array was considered large. Raw (Fourier) data for such an array is ~ 134 MB. By contrast, a recent scan of the macaque brain at 75- μm isotropic resolution yielded a $1500 \times 1200 \times 1000$ array, which is more than 100-times larger⁷. The problem at acquisition is that the dynamic range of the digitizer may not be sufficient. Figure 3 explains how the reconstruction engine can be used to extend the spatial resolution.

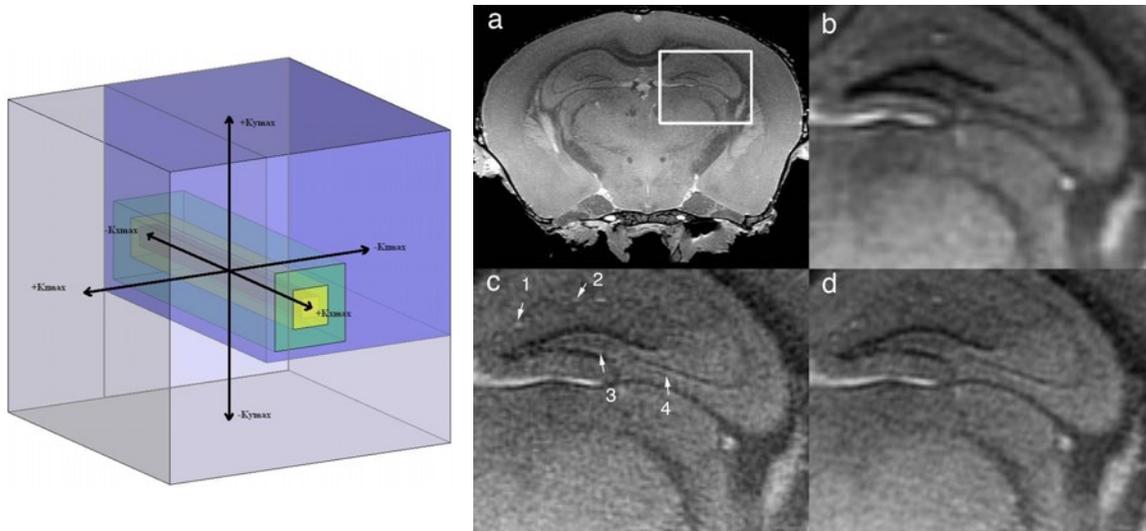


Figure 3. In 3D Fourier encoding, the signal at the center of Fourier space can easily be 100,000-times that of the signal at the periphery, exceeding the dynamic range of a 16-bit digitizer. The graphic on the left describes how we acquire three separate data sets (center of k space - yellow; intermediate k space - green, and peripheral segment of k space - blue). The spectrometer gain is reduced for the first set (center) to limit over-ranging at the center of Fourier space. The gain is successively increased moving to the periphery, enabling one to digitize the lower signal in the periphery of Fourier space with sufficient bit depth to facilitate the averaging inherent in the 3D encoding. By using the conjugate properties of Fourier space, we can limit the total volume of Fourier space sampled to a subset of the whole volume. The part of the figure on the right (reprinted from GA Johnson, et al., *High-throughput morphologic phenotyping of the mouse brain with magnetic resonance histology*, *Neuroimage* 37(10): 82-89, 2007, with permission from Elsevier), and the text below outline a typical use of the reconstruction pipeline.

The right panel in Figure 3 demonstrates two advantages of the independent reconstruction pipeline. (a) Shows a single slice from the 3D reconstruction at isotropic resolution (21.5- μm , $512^2 \times 1025$ array) using the *fully sampled* data. The box highlights an area of the hippocampus for comparison in the three remaining panels. In (b), the central $384 \times 384 \times 1024$ volume of Fourier space is reconstructed with a single acquisition, i.e. without the benefit of extending the dynamic range. (d) Shows the same slice reconstructed from a fully sampled $512^2 \times 1024$ array acquired with extended dynamic range. The total acquisition time was 7.3 hours. (c) Shows the intermediate case demonstrated in the graphic in the left pane. An asymmetric $384 \times 384 \times 1024$ array is used with extended dynamic range (acquisition in 4.1 hours). The arrows demonstrate small vessels (1 and 2), a thin layer between the granular and polymorphic layers of the dentate gyrus (3), and the stratum lucidum (4). Newer spectrometers incorporate digitizers with more dynamic range. But, larger arrays required for MRH and the use of multi-view sequences (e.g. Car-Purcell-Meiboom-Gill [CPMG]) will continue to push the need for such methods. The reconstruction pipeline also supports a much wider range of k space sampling strategies and combinations from multiple scans.

The generation of “calculated” images acquired with variants of the sequence parameters has been common for years. Early work focused on calculated spin-lattice or spin-spin relaxation times from varied TR or TE, which require relatively straightforward voxel-by-voxel curve fitting. The advent of diffusion tensor imaging (DTI) adds a new level of complexity, particularly for volumetric (3D) DTI. Figure 4 shows the post-processing pipeline for volume DTI.

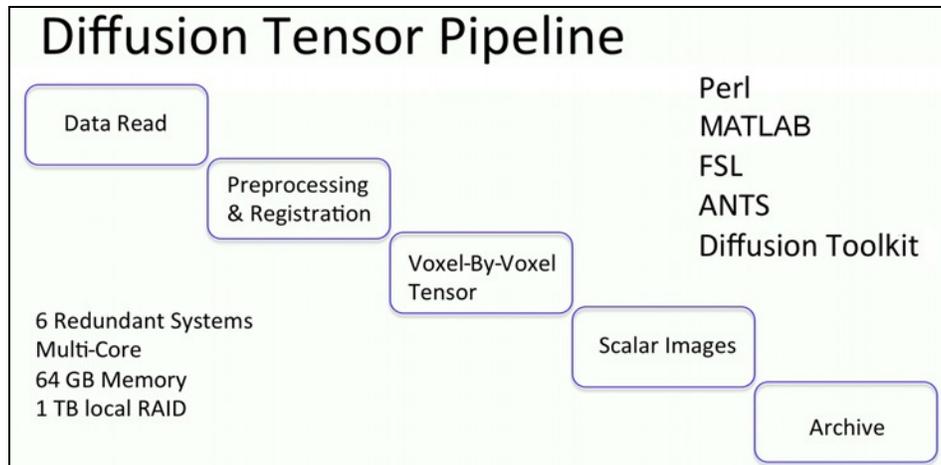
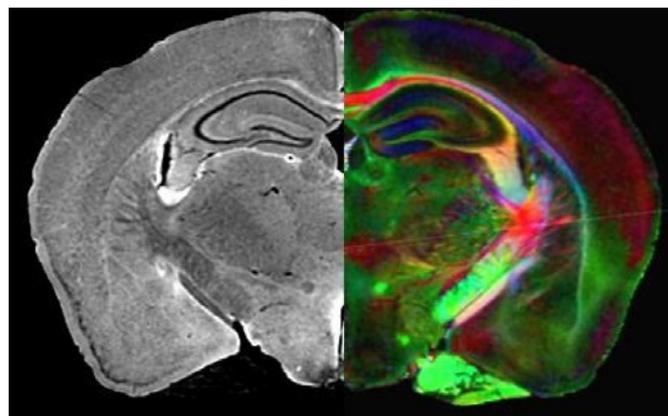


Figure 4. The diffusion tensor pipeline reads multiple 3D volumes, ingests the metadata (e.g. gradient amplitude and gradient angle), performs a skull-stripping algorithm, registers data, and performs a voxel-by-voxel tensor calculation. As an option, fibers can also be reconstructed.

The DTI pipeline takes advantage of several open source packages. A locally written routine strips the skull to facilitate the registration of multiple data sets to a common space. Residual eddy currents can cause registration errors that compromise the hard-fought resolution. The user has a choice of modules from FSL or Diffusion Toolkit for the tensor calculation. A final step generates the output “package” of images—a 4D Nifti file containing the tensor and a series of 3D scalar images, e.g. axial diffusivity (AD), radial diffusivity (RD), fractional anisotropy (FA), color fractional anisotropy (ClrFA), and optional fiber tracks. Multiple open source packages are incorporated to provide the range of options for diffusion calculation and display. Video 1 shows a side-by-side comparison of the diffusion-weighted image and color fractional anisotropy volume from a diffusion tensor acquisition of a mouse brain at 25- μ m resolution.



Video 1. Scalar images output from the diffusion tensor pipeline. Diffusion-weighted (left) and color fractional anisotropy images (right) have been derived from 132 3D acquisitions, a 60 GB 4D array. The complementary nature of the two contrasts provides improved anatomic definition. <http://dx.doi.org/10.1117/12.2203525.1>

Figure 5 shows a multi-stream pipeline running on a cluster that integrates images from multiple specimens into population atlases. Population atlases have been widely used in clinical neuroscience. By registering a collection of data from multiple patients (or specimens in our case), one can derive statistical insight into the variability of that population, and differences between distinct populations/treatments/strains. The multi-stream pipeline registers images from many different individuals into a common space, using multi-threaded and parallel processing. This work relies on iterative

deformation among a number of (3D) images to form a minimum deformation template (MDT). We have delineated more than 200 subfields for the mouse brain and more than 600 subfields for the rat, which can be mapped to the MDT. Since transformations are invertible, one can derive morphologic statistics from all of these subfields. Voxel-based analysis for such things as morphometry or DTI parameters extends the concept to a higher domain for statistical analysis on a voxel-by-voxel basis.

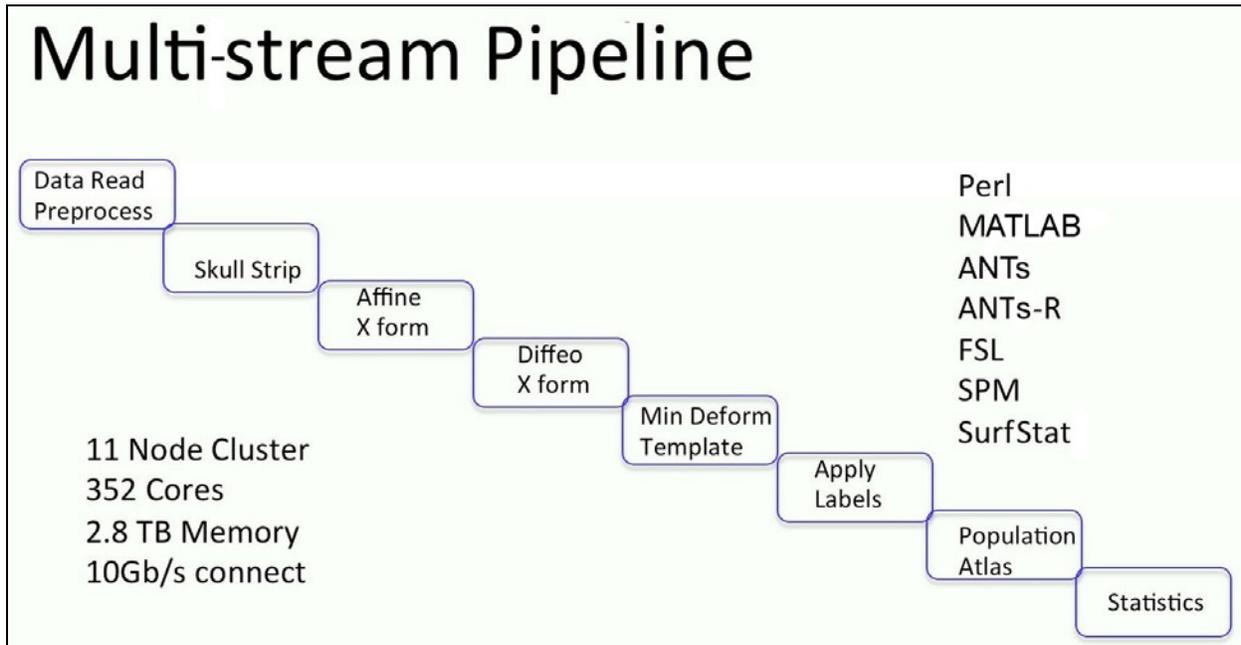


Figure 5. Multi-stream pipelines that aggregate data from multiple specimens are implemented on a high-performance cluster, which enables parallelization that drastically reduces the time required for formation of the minimum deformation template.

Figure 6 shows an example in which five 3D gradient recalled MRH images of an adult Wistar rat have been registered together to produce an average atlas⁸⁻¹⁰.

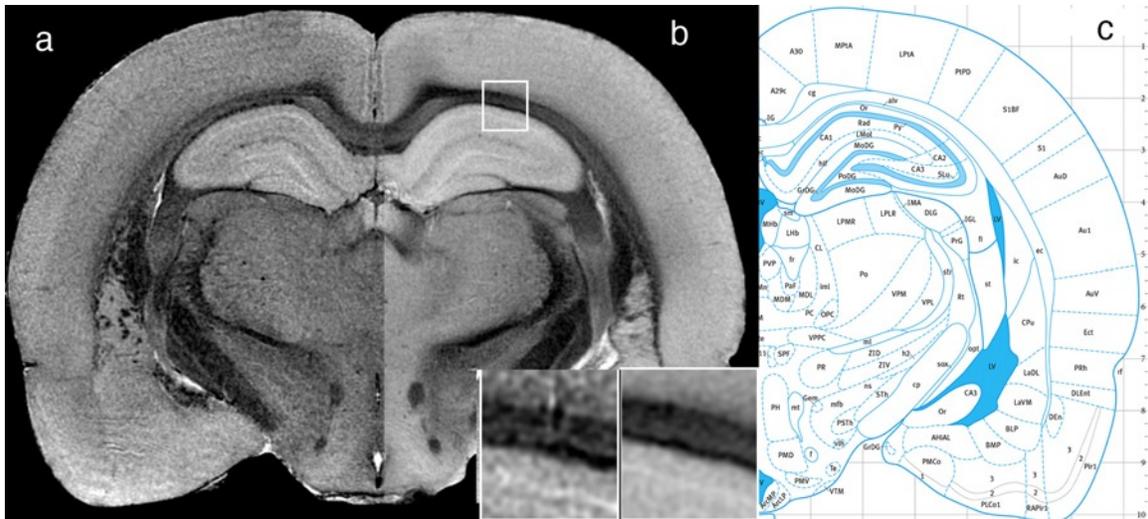


Figure 6. (a) GRE (at 50 μm) image from an adult Wistar rat brain. (b) The result of registering 5 different specimens. The inset shows that the registration is sufficient to maintain the boundary of the corpus callosum, as well as some of the layering. (c) A comprehensive set of labels for this average atlas assists in calculating volumes from any of the many subfields¹⁰. Components of this figure are extracted from Figure 3⁸ and Figure 45 (interaural 5.14 mm)¹⁰.

Figure 7 demonstrates an example that exploits all of the pipelines discussed thus far. The study was designed to understand the mechanism of blast-induced traumatic brain injury (bTBI) in a rat model using DTI magnetic resonance histology¹¹. The study tested the hypothesis that an initial blast injury might sensitize the animal to a second injury. Three groups of animals were examined: Group A - Control (n=9); Group B (n=9) subjected to a single blast; Group C (n=9) subjected to an initial blast followed by a second blast 1 minute later. MR histology was performed on all 27 animals with a gradient echo image and DTI acquisition using a baseline and 6 different diffusion-weighted images. For each specimen, there were 8 raw and reconstructed images, a total of ~ 3 GB/specimen. Magnitude images were completed on the reconstruction pipeline. Processing through the DTI pipeline added an additional 3 GB/specimen. Atlases were generated on the multi-stream cluster pipeline for the control group for the mean diffusivity (MD), axial diffusivity (AD), radial diffusivity (RD), and fractional anisotropy (FA) using pairwise, nonlinear image registration to create the minimum deformation template (MDT). A total of 36 pairwise registrations, or $n(n-1)/2$ were executed to produce the MDT for n=9 animals. The two treatment groups were spatially registered to the control group MDT, and for each animal, the same transforms were applied for each of the different contrasts. Voxel-based statistics were computed for each of these four imaging contrasts using SurfStat <<http://www.math.mcgill.ca/keith/surfstat/>>. Figure 7 shows the power of the collective set of pipelines in reducing this enormous data set (~ 200 GB) into a meaningful and easily understood result.

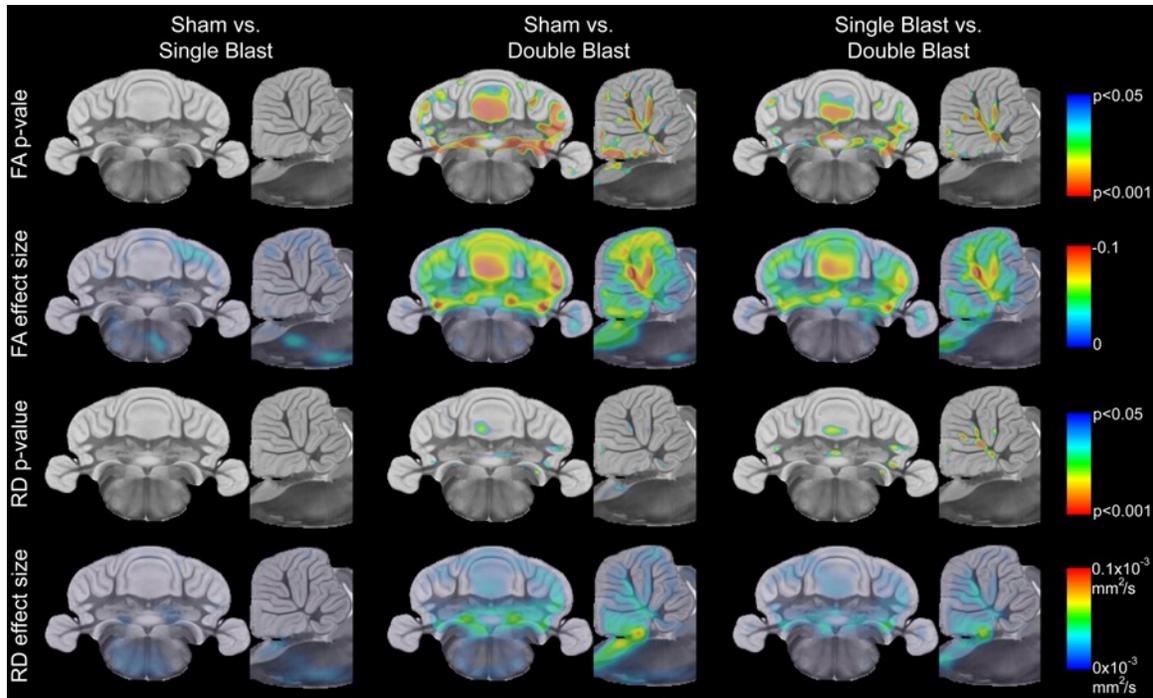


Figure 7. Voxel-based analysis enabled by efficient multi-stream pipelines allows quantitative statistical comparisons between the sham and single blast group, sham and double blast group, and single vs. double blast group. Results show significant injury in the cerebellum. Fractional anisotropy is the most sensitive biomarker. But, radial diffusivity is also altered, consistent with myelin degeneration. (Copyright permission was granted for use of Figure 5 from Calabrese et al., *Journal of Neurotrauma* 2014, published by Mary Ann Liebert, Inc., New Rochelle, NY.)¹¹

Finally, Figure 8 shows a screen shot of the Citrix window connecting external users to our Mammalian Brain Library, which is served from the NVIDIA GRID server. This library consists of GRE and DTI images from one individual and a population atlas of at least five specimens for each of these species: mouse, rat, and monkey. The aggregate data of ~ 30 GB has been carefully curated and organized to match existing paper atlases. In the screen shot, a registration protocol has been chosen using the single mouse brain atlas. The default display puts a sagittal diffusion-weighted image in the upper-left panel along with cursors that allow one to interactively position the location and angle of the slices seen in the lower-left and lower-right panels. In each of those two bottom panels, a drop-down menu provides choice of seven different contrasts (AD, DWI, FA, GRE, Susceptibility, ClrFA, and RD). All of the images are registered to the same space. In this example, the DWI (lower left) and ClrFA (lower right) provide decidedly different perspectives of the anatomy, but it is the very same tissue viewed through a different contrast mechanism. For this interactive display protocol, the user has the option of uploading (upper-right window) one of their data sets (in this example, a Nissl-stained section). Since both the plane and angle can be interactively manipulated, it is straightforward for the user to visually align the plane of the atlas to whatever plane was used in the acquisition of the test Nissl. At this point, the user selects “labels” from the menu in the upper left, which loads a volumetric collection of over 200 subregions listed in the menu on the right. Selecting a label from this menu, or clicking on a specific region of either of the lower two images, loads the label and identifies the region of interest. Other protocols provide different 3D views of the data and allow such things as different comparisons in the other species, between other contrasts, etc.

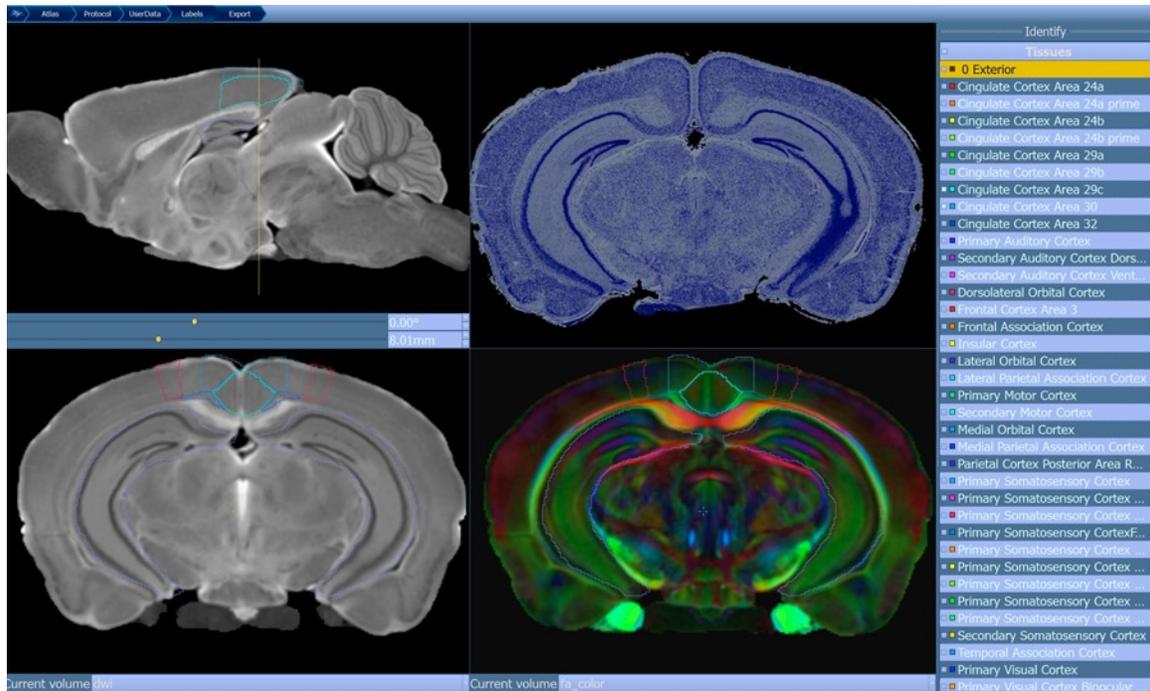


Figure 8. Screen capture of viewer that allows users interactive exploration of the multidimensional mammalian brain atlas. This example shows the mouse brain sagittal diffusion-weighted image (DWI) (upper left) with the navigation tools that allow selection of a longitudinal slice and its angle. The bottom two panels show the chosen slice, selected from two of seven different image contrasts, each providing a unique display of the underlying anatomy. The user interactively navigates the isotropic arrays to match the plane for their test set, in this case a conventional Nissl section (upper right).

4. DISCUSSION

Our first demonstration of magnetic resonance microscopy¹² used 2D encoding that yielded single 256x256 arrays, ~ 130 kB. The addition of isotropic 3D encoding³ moved the data size to tens of MBs. Diffusion tensor imaging added more data and dimensions⁵. Applications to population studies increased the dimensionality yet again, so that a typical study can yield hundreds of gigabytes of data—data volumes more than 6 orders of magnitude beyond our initial work. The specific niche in which we have worked, magnetic resonance histology, has matured from a laboratory curiosity to a robust methodology with applications in neuroscience, genetics, and toxicology^{6, 13-17}. Image-processing pipelines have been critical for all of these developments. These pipelines are a sometimes subtle innovation and sometimes obvious and profound. By conceptualizing the need for linked, dedicated pipelines tuned with hardware and software for the task and data structures, we have enabled continuing growth and innovation.

Imaging pipelines will continue to define new opportunities for MR histology. Recent work yielded what we believe to be the highest resolution, most comprehensive MR-based connectome of the mouse¹⁸. By acquiring diffusion-weighted images with 120 different gradient directions with isotropic resolution at 43 microns, the effects of volume averaging and ambiguities on crossing and merging fibers could be minimized. A pipeline has been constructed to yield a two-dimensional matrix describing the connectivity between the many subregions of the brain. Probabilistic tractography of the whole brain required nearly a week of dedicated time on the cluster. But, the resulting tractography set provides what we believe to be the most accurate measure of connectivity in the mouse brain. Application of this methodology in mouse models of disease promises new insight into conditions such as dementia and autism.

5. CONCLUSION

Over the last 30 years, magnetic resonance microscopy and magnetic resonance histology have moved from Professor Lauterbur's concept to a reality. Technical innovation in radiofrequency coils, pulse sequence design, and novel reconstruction methods have brought us to spatial resolution that is 500,000-times that of clinical MRI. This maturing technology has opened a number of fascinating new applications in high-throughput phenotyping, toxicology, and neuroscience. All of these have depended heavily on the use of sophisticated imaging pipelines. While room still exists for technological advances in data acquisition, there is no doubt that much of the future work will depend on efficient image-processing pipelines to accelerate the pathways to discovery.

6. ACKNOWLEDGMENTS

The Duke Center for In Vivo Microscopy is an NIH/NIBIB national Biomedical Technology Resource Center, supported by P41 EB015897. Work presented in this paper was also supported by NIH/Office of the Director 1S10 OD010683-01 and NIH/NIA K01 AG041211.

REFERENCES

- [1] Lauterbur, P. C., "Image formation by induced local interactions - examples employing nuclear magnetic resonance," *Nature*, 242, 190-191 (1973).
- [2] Johnson, G. A., Benveniste, H., Black, R. D., Hedlund, L. W., Maronpot, R. R., and Smith, B. R., "Histology by magnetic resonance microscopy," *Magn. Reson. Q.*, 9(1), 1-30 (1993).
- [3] Suddarth, S. A., and Johnson, G. A., "Three-dimensional MR microscopy with large arrays," *Magn. Reson. Med.*, 18(1), 132-141 (1991).
- [4] Sharief, A. A., and Johnson, G. A., "Enhanced T2 contrast for MR histology of the mouse brain," *Magn. Reson. Med.*, 56(4), 717-725 (2006).
- [5] Jiang, Y., and Johnson, G. A., "Microscopic diffusion tensor imaging of the mouse brain," *Neuroimage*, 50(2), 465-471 (2010).
- [6] Johnson, G. A., Badea, A., Brandenburg, J., Cofer, G., Fubara, B., Liu, S., and Nissanov, J., "Waxholm space: An image-based reference for coordinating mouse brain research," *Neuroimage*, 53(2), 365-372 (2010).
- [7] Calabrese, E., Badea, A., Coe, C. L., Lubach, G. R., Shi, Y., Styner, M. A., and Johnson, G. A., "A diffusion tensor MRI atlas of the postmortem rhesus macaque brain," *Neuroimage*, 117, 408-416. (2015).
- [8] Johnson, G. A., Calabrese, E., Badea, A., Paxinos, G., and Watson, C., "A multidimensional magnetic resonance histology atlas of the Wistar rat brain," *Neuroimage*, 62(3), 1848-1856 (2012).
- [9] Calabrese, E., Badea, A., Watson, C., and Johnson, G. A., "A quantitative magnetic resonance histology atlas of postnatal rat brain development with regional estimates of growth and variability," *Neuroimage*, 71, 196-206 (2013).
- [10] Paxinos, G., Watson, C., Calabrese, E., Badea, A., and Johnson, G. A., [An MRI/DTI atlas of the rat brain] Academic Press, Elsevier, (2015).
- [11] Calabrese, E., Du, F., Garman, R. H., Johnson, G. A., Riccio, C., Tong, L. C., and Long, J. B., "Diffusion tensor imaging reveals white matter injury in a rat model of repetitive blast-induced traumatic brain injury," *J. Neurotrauma*, 31(10), 938-950 (2014).
- [12] Johnson, G. A., Thompson, M. B., Gewalt, S. L., and Hayes, C. E., "Nuclear magnetic resonance imaging at microscopic resolution," *J. Magn. Reson.*, 68, 129-137 (1986).
- [13] Badea, A., Nicholls, P. J., Johnson, G. A., and Wetsel, W. C., "Neuroanatomical phenotypes in the reeler mouse," *Neuroimage*, 34(4), 1363-1374 (2007).
- [14] Badea, A., Johnson, G. A., and Williams, R. W., "Genetic dissection of the mouse brain using high-field magnetic resonance microscopy," *Neuroimage*, 45(4), 1067-1079 (2009).
- [15] Badea, A., Johnson, G. A., and Jankowsky, J. L., "Remote sites of structural atrophy predict later amyloid formation in a mouse model of Alzheimer's disease," *Neuroimage*, 50 (2), 416-427 (2010).
- [16] Badea, A., Gewalt, S., Avants, B. B., Cook, J. J., and Johnson, G. A., "Quantitative mouse brain phenotyping based on single and multispectral MR protocols," *Neuroimage*, 63(3), 1633-1645 (2012).

- [17] Johnson, G. A., Calabrese, E., Little, P. B., Hedlund, L., Qi, Y., and Badea, A., "Quantitative mapping of trimethyltin injury in the rat brain using magnetic resonance histology," *Neurotoxicology*, 42, 12-23 (2014).
- [18] Calabrese, E., Badea, A., Cofer, G., Qi, Y., and Johnson, G. A., "A diffusion MRI tractography connectome of the mouse brain and comparison with neuronal tracer data," *Cereb. Cortex*, 25(11), 4628-4637 (2015).