https://doi.org/10.2352/EI.2023.35.11.HPCI-237

Clinical Validation of Rapid GPU-enabled DTI Tractography of the Brain

Felix Liu; UCSF, San Francisco, CA, USA,
Vanitha Sankaranarayanan; UCSF, San Francisco, CA, USA,
Javier Villanueva-Meyer; UCSF, San Francisco, CA, USA,
Shawn Hervey-Jumper; UCSF, San Francisco, CA, USA,
James Hawkins; UCSF, San Francisco, CA, USA,
Pablo Damasceno; Janssen Pharmaceutical, Boston, MA, USA,
Mauro Bisson; NVIDIA, Santa Clara, CA,
Josh Romero; NVIDIA, Santa Clara, CA,
Thorsten Kurth; NVIDIA, Santa Clara, CA,
Massimiliano Fatica; NVIDIA, Santa Clara, CA,
Eleftherios Garyfallidis; Indiana University, Bloomington, IN, USA,
Ariel Rokem; University of Washington, eScience Institute, Seattle, WA, USA,
Jason C. Crane; UCSF, San Francisco, CA, USA,
Sharmila Majumdar; UCSF, San Francisco, CA, USA

Abstract

Diffusion tensor imaging (DTI) is a non-invasive magnetic resonance imaging (MRI) modality used to map white matter fiber tracts for a variety of clinical applications; one of which is aiding preoperative assessments for tumor patients. DTI requires numerical computations on multiple diffusion weighted images to calculate diffusion tensors at each voxel and probabilistic tracking to construct fiber tracts, or tractography. Greater accuracy in tractography is possible with higher imaging resolution, more diffusion gradients and advanced reconstruction algorithms. However, tractography for larger scans and the complex reconstruction required are often computationally intensive, taking up to 40 minutes of CPU time on state-of-the-art hardware. Parallel GPU computations can improve time for the resource-intensive tractography. A collaborative team from DIPY, NVIDIA, and UCSF recently developed a tool, GPUStreamlines, for GPU-enabled tractography which has been expanded to support the constant solid angle (CSA) reconstruction algorithm. This GPU-enabled tractography was applied to MRIs of brains with and without presence of lesions, with demonstrated 24-30 fold increases in processing speed. The tracts were also similar to those derived on CPU, with bundle adjacency similarity scores of 0.82 to 0.88. This work demonstrates that results from GPU-enabled CSA tractography in patients with and without brain tumors are comparable to the existing CSA tractography process.

Introduction

Diffusion tensor imaging (DTI) is a non-invasive magnetic resonance imaging (MRI) modality used to map white matter fiber tracts for a variety of clinical applications; one of which is aiding preoperative assessments for brain tumor patients [1]. DTI requires numerical computations on multiple diffusion-weighted MR images (dMRI) to calculate diffusion tensors at each voxel and probabilistic tracking between voxels to delineate fiber tract pathways, or tractography [2].

However, DTI fails to accurately map the fiber tract orientations in voxels containing crossing fibers [3]. Within a typical dMRI voxel, multiple fiber orientations may exist whereas a diffusion tensor may only represent a single orientation. Greater accuracy in tractography is possible by increasing the spatial resolution of dMRI or by increasing the number of dMRI acquisitions in combination with advanced diffusion reconstruction algorithms capable of distinguishing multiple intravoxel fiber orientations. For example, Q-Ball Imaging (QBI) with probabilistic tracking has demonstrated superior fiber tract maps compared to conventional DTI [4]. However, larger scans and advanced reconstruction is often computationally intensive.

At the University of California San Francisco (UCSF) Medical Center, over 700 dMRI exams are acquired yearly for white matter fiber tractography and preoperative planning. Existing analysis at UCSF utilizes in-house software based on libraries from the Diffusion Imaging in Python project, or DIPY (dipy.org) to perform probabilistic Bootstrap tracking on QBI with the Orientation Probability Density Transform (OPDT) model [1]. The post-processing for tractography is computationally expensive and may require to 40 minutes or more of CPU time, even on state-of-the-art hardware. Constant Solid Angle (CSA) is a model for diffusion reconstruction with wide support among the DIPY community and may offer more accurate white matter fiber tract delineations [5] [6]. However, the CSA model results in a 3-5 fold increase to computation times for tractography while results are often needed in a timely manner for patient care.

Parallel GPU computations can improve processing time for the resource-intensive CSA-based tractography [7]. A collaborative team from DIPY, NVIDIA (NVIDIA Corporation, Santa Clara, CA), and UCSF recently developed a tool, GPUStreamlines (github.com/dipy/GPUStreamlines), for GPU-enabled tractography which has been expanded to support the CSA reconstruction algorithm. This GPU-enabled tractography was applied to MRIs of brains with and without presence of lesions, with 24-30 fold increases in processing speed. The present work demonstrates that GPU-enabled CSA tractography in normal controls and patients is comparable to tractography using the existing CPU-based CSA reconstruction.

Background

In a pulsed gradient spin echo sequence, the MR signal is attenuated due to dephasing from random Gaussian displacement of the protons in the gradient direction [4]. The relationship between the average probability of displacement and the MR signal decay can be modeled as diffusion probability density functions (PDFs), where the diffusion propagator P describes the probability of displacement at a given space and time. For tractography, the fiber orientations are of most interest and are modeled as orientation distribution functions (ODFs), defined as

$$\Psi(\mathrm{u}) = \int_0^\infty P(ru)r^2dr \tag{1}$$

where Ψ is a function on a sphere, r is the relative displacement, and u is a unit vector [6]. With single-shell HARDI data, the ODFs may be efficiently estimated using spherical harmonics and the CSA model, giving

$$\Psi(\theta, \varphi) = \sum_{j=1}^{R} 2\pi \frac{c_j}{S_0} P_{l(j)}(0) Y_j(\theta, \varphi) \tag{2}$$

where S_0 is the non-diffusion weighted imaging data, $P_{l(j)}$ is the Legendre polynomial of order l, Yj is the spherical harmonic basis, and c_j is the spherical harmonic coefficient. Finally, probabilistic residual bootstrap tracking may be applied to follow ODFs in a stepwise manner.

DIPY provides accessible open-source libraries in Python for CSA Bootstrap probabilistic tracking [6]. These operations can be additionally implemented with the CUDA platform to allow rapid computations for a large number of tractography seeds in parallel on graphics processing units (GPUs).

Methods

Patient Selection

Twenty-four retrospective patients with brain diffusion MRI exams were also selected for the study. Ten patients had tumors present in the brain exams, and 14 patients did not have tumors present.

Image Acquisition

Whole-brain images were acquired on 3T GE MR scanners (General Electric, Waukesha, WI) with high angular diffusion imaging (HARDI) sequences comprised of diffusion-encoding gradients in 55 directions and diffusion-weighting b-value of 2000 s/mm^2, 70 ms TE, 8400 ms TR, and 1.1x1.1x2.2 mm voxel size. Additional acquisitions without diffusion weighting at b-value 0 s/mm^2 were also obtained. The imaging protocol was approved by the UCSF institutional review board.

Post-processing

Image volumes were corrected for eddy current distortion using FSL (www.fmrib.ox.ac.uk/fsl) and resliced to 1.38x1.38x1.38 mm isotropic voxels. An expert reader (VS) segmented regions of interests (ROI) for tractography seeding in the cerebral peduncle (CP) and lateral geniculate nucleus (LGN) from the anatomical images using Trackvis (www.trackvis.org). Target ROIs were also segmented for the motor tract from the cerebral peduncle and the optic radiation from the lateral geniculate nucleus - two pathways from the eloquent cortex.

GPUStreamlines

Prior work by teams from DIPY, NVIDIA, and UCSF implemented residual bootstrap tractography with the OPDT model as GPUStreamlines, a multi-GPU parallelizable version. GPUStreamlines was implemented on NVIDIA's CUDA application programming interface (API) and is compatible with DIPY. In the current work, the existing GPU-implementation was adapted to include capabilities for bootstrap tractography with the CSA model. The software is available at https://github.com/dipy/GPUStreamlines.

Tractography

Fiber tracts were mapped by tracking streamlines from the seeding ROIs with the standard CSA model and probabilistic Bootstrap direction getter [2] from DIPY 1.4 (dipy.org) on 12 CPUs (AMD Opteron 3680) in parallel, and with GPUStreamlines on 1 GPU (24GB NVIDIA TITAN RTX GPU). All tracking was performed with seed density of 7, step size of 0.5 mm, and fractional anisotropy threshold stopping criteria of 0.1.

For preoperative planning images, streamlines are pruned by expert readers to isolate relevant pathways. In order to emulate processing by expert readers, the streamlines were programmatically pruned by only selecting streamlines that reach or pass through the target ROIs for the motor tract and optic radiation pathways. The streamlines were further pruned by removing streamlines with lengths less than 30 mm or more than 200 mm.

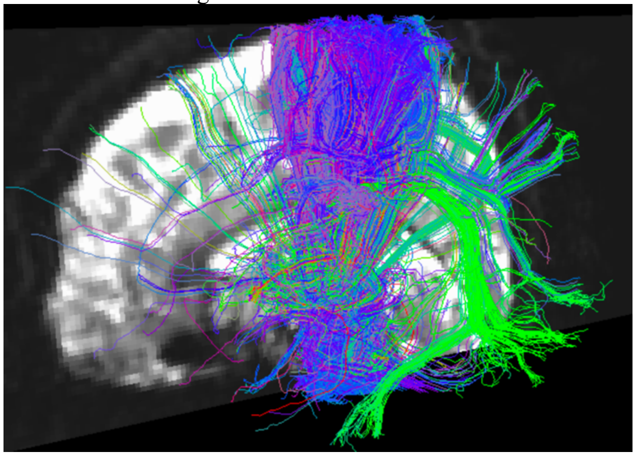

Figure 1. Motor pathway streamlines generated with GPUStreamlines and tracked starting in seeding voxels from ROI in cerebral peduncle.

Analysis

Binary masks of the tracts were generated by segmenting all voxels containing or intersected by fiber tract streamlines. The binary masks from CPU-enabled and GPU-enabled tracking were compared using voxel-wise DICE coefficient [8], recall, and precision. Streamlines were compared before and after pruning to assess the impact of removing irrelevant streamlines: binary masks were generated on streamlines from the full tractography results, and on streamlines remaining after targeting pathway ROIs and limiting streamline lengths. CPU-enabled tracking was considered the gold standard in all analyses.

The sets of streamlines from GPU-enabled tracking were compared with corresponding sets from CPU-enabled tracking using Bundle Adjacency (BUAN) shape similarity scores [9]. Streamlines were down sampled to 20 points each and the minimum average direct-flip (MDF) distance, a measure of the average distance between each pair of streamlines, was calculated. Streamlines with 0.75 mm or less MDF distance were clustered together using QuickBundles [10].

The centroids of each cluster were calculated and collected as a set of average streamlines. The MDF distances between the corresponding CPU-enabled tracking and the GPU-enabled tracking sets were calculated for assessing streamline similarity. BUAN shape similarity scores were computed as the fraction of streamlines with MDF distances less than 5 mm between the 2 sets. BUAN shape similarity scores approaching 1.0 represent two highly similar sets of streamlines, while scores near 0.0 represent two highly dissimilar sets of streamlines.

Results

Bootstrap tractography computation times with CSA were significantly faster in the GPU process compared to the CPU process. The times varied across patients, depending on the seeding ROI and on the patient's tumor status. For tracking seeded in the cerebral peduncle overall, CSA-GPU tractography was processed in 80 seconds on average while CSA-CPU processing on 12 CPUs averaged 32 minutes. Processing was faster in patients without tumors (58 seconds for CSA-GPU and 17 minutes for CSA-CPU) than in patients with tumors (111 seconds for CSA-GPU and 54 minutes for CSA-CPU). Tracking seeded in the LGN averaged 3 minutes on GPU and 95 minutes on CPU. Processing was similarly faster for patients without tumors (2.5 minutes on GPU and 76 minutes on CPU) compared to patients with tumors (3.7 minutes on GPU and 120 minutes on CPU).

The mean average Dice coefficient, recall, and precision ranged from 0.39-0.48, 0.95-0.97, and 0.24-0.32 for binary masks generated for streamlines tracked from the cerebral peduncle. The same metrics in masks after streamlines were pruned for the motor pathway resulted in average Dice, recall and precision ranged from 0.49-0.58, 0.92, 0.34-0.43.

For masks generated from LGN ROIs, average Dice, recall, and precision were 0.54-0.56, 0.98-0.99, and 0.38-0.40. Masks of the optics radiations had ranges of 0.61, 0.94-0.96, and 0.45-0.46.

The mean BUAN similarity scores ranged from 0.82 to 0.88 in the motor pathways and 0.87 in the optic radiations.

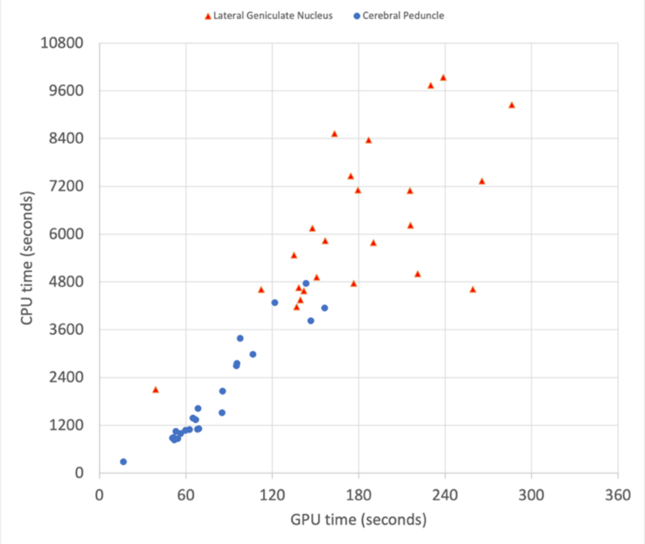

Figure 2. Computation times in seconds for CSA tractography on 12 CPUs in parallel and 1 GPU with GPUStreamlines. Times for tractography seeded from cerebral peduncle are in blue circles and lateral geniculate nuclei are in orange triangles.

Table 1. Voxel-wise comparisons of fiber tract binary masks between CPU-enabled and GPU-enabled tractography

Tracts	Tumors	Average DICE	Average Recall	Average Precision
Cerebral peduncle, all	Absent	0.386	0.951	0.244
Cerebral peduncle, all	Present	0.479	0.967	0.320
Lateral geniculate nucleus, all	Absent	0.542	0.989	0.375
Lateral geniculate nucleus, all	Present	0.564	0.981	0.401
Motor pathway only	Absent	0.493	0.925	0.337
Motor pathway only	Present	0.583	0.928	0.427
Optic radiation only	Absent	0.607	0.958	0.446
Optic radiation only	Present	0.611	0.944	0.455

Table 2. Bundle adjacency shape similarity scores between CPU-enabled and GPU-enabled tractography

Tracts	Tumors	Average BUAN
Motor pathway only	Absent	0.821
Motor pathway only	Present	0.875
Optic radiation only	Absent	0.870
Optic radiation only	Present	0.874

Discussion

GPU processing demonstrated 24-fold improvement on average in speed compared to processing on 12 CPUs in parallel. The improvement in speed increases to 30-fold in patients with tumors and for streamlines seeded in the LGN; these two scenarios exhibited the longest tractography times in this study. Patients with tumors may have higher complexity in their neural anatomy due to the displacement of white matter tissue and may have more complicated fiber tract pathways. Tractography seeded in the LGN resulted in higher numbers of fiber tract streamlines compared to those seeded in the cerebral peduncle, and thus also resulted in longer total computation times.

The overall reduction in computation time is reduced to approximately 2-4 minutes per ROI. With a DWI data preprocessing step data for motion and noise that requires approximately 10 minutes per exam, the total computation time for tractography would total 16 to 22 minutes for 3 ROIs. This is faster than the current tractography processes that requires 40 minutes. Assuming that pruning by expert readers requires approximately 10 minutes per ROI, DTI exams may be fully processed and ready for clinicians to review in less than 1 hour.

The low Dice and precision scores in comparing the binary masks indicate the resultant tracts from CPU and GPU processes are not identical. However, the extremely high recall scores show that fiber tract from the GPU-enabled GPUStreamlines process includes nearly all those derived from the standard DIPY process on CPUs. When examining the masks from streamlines pruned for the motor pathway or optic radiations only, the Dice and precision scores improved, and the recall scores decreased slightly. A proportion of the streamlines from the GPU process may therefore be irrelevant for preoperative purposes, at least for the motor and optic radiation pathways.

The BUAN shape similarity scores also suggest most of the GPU-derived tracts follow similar nerve fiber bundles from the CPU process, with over 80% of the bundles deviating an average MDF distance of 5 mm or less.

Some variation in tracts is expected considering the differences in tractography and the probabilistic nature of the reconstruction algorithms. With the probabilistic bootstrap direction getter and CSA model, repeat tractography with sample data on CPU hardware often produces streamlines with Dice coefficient scores of 0.80 to 0.85. Forcing the permutations of the probabilistic tracking to be identical result in tracts from CPU and GPU processes to be nearly equivalent. Thus, the quality of random number generation may be a factor. Additionally, the data in this study was acquired in vivo and represent real-world scenarios with noise and physiological distortions, which likely contribute to greater variability. Appropriately denoising the diffusion data before applying tractography may result in more robust or consistent streamlines. Observationally, the CSA-GPU tracts exhibit a greater number of fiber tracts and it is possible they represent the true nerve fibers more accurately.

There are a several limitations in the current work. Binary masks require manual pruning by expert readers before display for preoperative planning; although streamlines were programmatically pruned by targeting ROIs in this study, it is not equivalent to expertly pruned tracts. It is conceivable that expertly pruned streamlines from both CSA-CPU and CSA-GPU processes would match more closely to each other. This study did not directly correlate results from CSA-CPU or CSA-GPU processes to the current UCSF methodology using OPDT modeling. Lastly, the results have not yet been assessed by neuroradiologist or neurosurgeons whose expertise is critical for determining clinical validity.

Further studies are needed to assess the computational improvements, accuracy and validity of the GPU-enabled CSA tractography. Ongoing work at UCSF will aim to address these limitations in order to characterize the differences observed in the GPU-enabled CSA tracking results and to fully validate the fiber tract streamlines in clinical patients.

References

[1] J. I. Berman, M. S. Berger, P. Mukherjee and R. G. Henry, "Diffusion-tensor imaging—guided tracking of fibers of the pyramidal tract combined with intraoperative cortical stimulation mapping in patients with gliomas," *Journal of neurosurgery,* vol. 101, no. 1, pp. 66-72, 2004.

[2] J. I. Berman, S. Chung, P. Mukherjee, C. P. Hess, E. T. Han and R. G. Henry, "Probabilistic streamline q-ball tractography using the residual bootstrap," *NeuroImage,* vol. 39, pp. 215-222, 2008.

[3] C. P. Hess, P. Mukherjee, E. T. Han, D. Xu and D. B. Vigneron, "Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis," *Magnetic Resonance in Medicine,* vol. 56, no. 1, pp. 104-117, 2006.

[4] D. S. Tuch, "Q-Ball Imaging," *Magnetic Resonance in Medicine,* vol. 52, p. 1358 –1372, 2004.

[5] I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil and N. Harel, "Reconstruction of the Orientation Distribution Function in Single- and Multiple-Shell q-Ball Imaging Within Constant Solid Angle," *Magnetic Resonance in Medicine,* vol. 64, p. 554–566, 2010.

[6] E. Garyfallidis, B. Amirbekian, M. Brett, A. Rokem, S. Van Der Walt, M. Descoteaux and I. Nimmo-Smith, "Dipy, a library for the analysis of diffusion MRI data," *Frontiers in neuroinformatics,* vol. 8, 2014.

[7] A. Rokem, M. Bisson, J. Romero, T. Kurth, M. Fatica, P. Damasceno, X. Xie, A. Richie-Halford, S. Koudoro and E. Garyfallidis, "GPU-accelerated diffusion MRI tractography in DIPY," in *International Society for Magnetic Resonance in Medicine*, 2021.

[8] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology,* vol. 26, no. 3, pp. 297-302, 1945.

[9] B. Q. Chandio, S. L. Risacher, F. Pestilli, D. Bullock, F.-C. Yeh, S. Koudoro, A. Rokem, J. Harezlak and E. Garyfallidis , "Bundle analytics, a computational framework for investigating the shapes and profiles of brain pathways across populations," *Scientific Reports,* vol. 10, p. 17149, 2020.

[10] E. Garyfallidis, M. Brett, M. M. Correia, G. B. Williams and I. Nimmo-Smith, "QuickBundles, a method for tractography simplification," *Frontiers in Neuroscience,* vol. 6, 2012.

Author Biography

Felix Liu received his BS in biomedical engineering from Boston University (2008) and his MS in Bioimaging from Boston University (2010). He has since worked at the University of California, San Francisco (UCSF). His work includes research into imaging biomarkers, clinical deployment, and deep learning of various areas including osteoarthritis, musculoskeletal imaging, brain diffusion, and breast cancer.