Journal of Neuroscience Methods 273 (2016) 107–119

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

DiffusionKit: A light one-stop solution for diffusion MRI data analysis

Sangma Xie[a,b], Liangfu Chen[a,b], Nianming Zuo[a,b,*], Tianzi Jiang[a,b,c,d,e]

[a] Brainnetome Center, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[c] CAS Center for Excellence in Brain Science, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[d] Queensland Brain Institute, University of Queensland, Brisbane, QLD 4072, Australia
[e] Key Laboratory for NeuroInformation of the Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China

HIGHLIGHTS

- DiffusionKit has a full pipeline for (pre-)processing and visualization of diffusion MRI data.
- DiffusionKit has cross-platform support and a small installation size without 3rd party dependency.
- DiffusionKit has both a GUI interface and command-line functions that enable easy operation and batch processing.

ARTICLE INFO

Article history:
Received 10 March 2016
Received in revised form 18 June 2016
Accepted 23 August 2016
Available online 24 August 2016

Keywords:
Diffusion MRI
DTI
HARDI
Anatomical connectivity
DiffusionKit

ABSTRACT

Background: Diffusion magnetic resonance imaging (dMRI) techniques are receiving increasing attention due to their ability to characterize the arrangement map of white matter *in vivo*. However, the existing toolkits for dMRI analysis that have accompanied this surge possess noticeable limitations, such as large installation size, an incomplete pipeline, and a lack of cross-platform support.
New method: In this work, we developed a light, one-stop, cross-platform solution for dMRI data analysis, called **DiffusionKit**. It delivers a complete pipeline, including data format conversion, dMRI preprocessing, local reconstruction, white matter fiber tracking, fiber statistical analyses and various visualization schemes. Furthermore, DiffusionKit is a self-contained executable toolkit, without the need to install any other software.
Results: The DiffusionKit package is implemented in C/C++ and Qt/VTK, is freely available at http://diffusion.brainnetome.org and https://www.nitrc.org/projects/diffusionkit. The website of DiffusionKit includes test data, a complete tutorial and a series of tutorial examples. A mailing list has also been established for update notification and questions and answers.
Comparison with existing methods: DiffusionKit provides a full-function pipeline for dMRI data analysis, including data processing, modeling and visualization. Additionally, it provides both a graphical user interface (GUI) and command-line functions, which are helpful for batch processing. The standalone installation package has a small size and cross-platform support.
Conclusions: DiffusionKit provides a complete pipeline with cutting-edge methods for dMRI data analysis, including both a GUI interface and command-line functions. The rich functions for both data analysis and visualization will facilitate and benefit dMRI research.

1. Introduction

Brain network studies have become a significant approach for understanding how different functional parcellations of the brain interact with each other. Currently, there are two critical methods to inspect brain networks, functional networks and anatomical networks (Jiang, 2013), in which the anatomical networks consist of the neural bases and physical connections for the functional networks (Sporns et al., 2005). To date, diffusion magnetic resonance imaging (diffusion MRI, also known as dMRI) has been recognized as an irreplaceable technique to investigate the anatomical connectivity of the brain *in vivo*, and based on the data collected by dMRI, mathematical models have been applied to characterize the dis-

* Corresponding author.
E-mail address: nmzuo@nlpr.ia.ac.cn (N. Zuo).

http://dx.doi.org/10.1016/j.jneumeth.2016.08.011
0165-0270/

tribution of the water molecules constrained by the white matter microstructure (Zuo et al., 2012).

The brain connectivity and network analyses based on dMRI data consist of a series of image processing and modeling steps (Sotiropoulos et al., 2013). First, each separate volume of the diffusion-weighted imaging (DWI) series is corrected for the distortions induced by eddy-current and susceptibility, both of which are potential sources of off-resonance field, and for head motion, which blurs the images and affects the image alignment (Le Bihan et al., 2006). Additionally, to avoid processing of the area outside of the brain tissue, a skull stripping step is strongly recommended. Second, for each of the voxels, the diffusion model is estimated from the DWI series. According to the sampling scheme in the q-space, there are a variety of estimation strategies to dissect the white matter propagation within a voxel. The most traditional and popular strategy is diffusion tensor imaging (DTI) (Basser et al., 1994; Le Bihan et al., 2001). Recently, for the deciphering of crossing fibers, the decomposition-based spherical polar Fourier imaging (SPFI) method and the deconvolution-based constrained spherical deconvolution (CSD) method have emerged as two representative families of high angular resolution diffusion imaging (HARDI) techniques (Xie et al., 2015). Thus, voxel-based analysis and tract-based spatial statistics (TBSS) (Smith et al., 2006) can be performed, such as for fractional anisotropy (FA), mean diffusivity (MD), relative anisotropy (RA) (Assaf and Pasternak, 2008; Le Bihan et al., 2001). Third, once maps of white matter distribution within voxels are obtained, the white matter fiber connections can be tracked by a number of tractography approaches (Mori and van Zijl, 2002). Finally, to construct the connectivity between two brain areas, specific attributes can be derived from the connecting fiber bundles, such as the mean fractional anisotropy (FA), mean diffusivity (MD), number of fibers, and volumes occupied by the fiber bundles (Sporns, 2011; Sporns et al., 2005). The researchers who want to implement the entire pipeline should be acquainted with corresponding algorithms and proficient in programming.

Meeting the requirements of researchers from multidisciplinary backgrounds, a large number of toolkits have been developed, including data preprocessing (FSL; SPM) and diffusion tensor or orientation distribution function modeling (FSL; MRtrix; TrackVis), fiber tracking and visualization (DTITool; TrackVis). However, these tools have mainly focused on specific steps, and they do not provide a full set of components, consisting of data conversion, data (pre-)processing, fibertracking, network construction, image/fiber view and visualization. Recently, to address these drawbacks, a few integrated tools have been developed, such as PANDA (Cui et al., 2013), ExploreDTI (Leemans et al., 2009), MedInria (Toussaint et al., 2007), Dipy (Garyfallidis et al., 2014) and MIPAV (MIPAV). For a more detailed survey, please see the Table. 1. However, these new tools have their own limitations. First, some of them are MATLAB (The Mathworks, Inc.) based. MATLAB, an interpreted language, will execute more slowly than complied languages (Chapman, 2015) and is not powerful enough in rendering surface of the brain and numerous fibers. Second, a large number of third-party dependencies for installation can result in a large software size. Third, there is no complete pipeline or cross-platform support.

In this work, we developed a light but comprehensive solution, called DiffusionKit, for dMRI data (pre-)processing and visualization. The remainder of the manuscript is organized as follows. In the Material and methods section, we introduce the overall design of DiffusionKit and the functions of the main modules with explanations of the principles and implementations. Then, results and some examples are presented in the Results and discussion section to validate the given functions. In the Conclusions section, the main modules and features are summarized, and comparisons with existing tools are also tabulated.

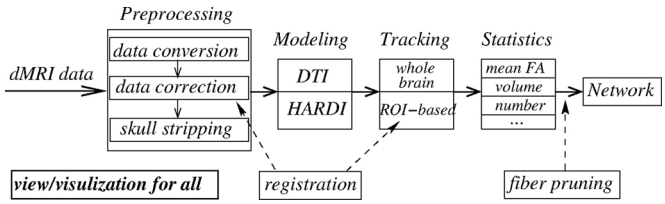

Fig. 1. The overall design framework and the main modules of DiffusionKit.

2. Material and methods

2.1. Overall design

The overall design framework is illustrated in Fig. 1. DiffusionKit was developed as a cross-platform framework, using C/C++ for computation, VTK (VTK) for visualization, and Qt for graphical user interface (GUI) design. Both GPU and CPU computing have been implemented to achieve a high frame rate for rendering complex scenes, particularly whole-brain tractography. The project was managed using the compiler-independent CMake (CMake), which is compatible with a variety of compilers, such as GCC/G++ and MS Visual Studio. Fig. 2 shows the main entry page of DiffusionKit, which consists of two core modules: Processing and Visualization. Each function is standalone, and the main features are simply organized by the GUI. Such a design fully facilitates batch processing for large datasets by means of scripting, and it also supports future enrichment of functions of the pipeline.

2.2. Data preprocessing

Before modeling the distribution function within voxels, the DWI images are corrected and aligned for precise modeling. First, for the convenience of processing in DiffusionKit, the raw data are converted from DICOM images to a single 4D NIfTI image (to save storage space, DiffusionKit uses files in zipped format, e.g., nii.gz, by default). This step utilizes *dcm2nii*, developed by Dr. Chris Rorden (MRIcron), which is fast and stable and has been well tested within the community (Jenkinson et al., 2012). Occasionally this program fails to extract the gradient table from the DICOM image series (some MRI scanners or the associated PACS systems use specific techniques to arrange the keyword dictionary of the DICOM), so we have provided a temporary amendment using Matlab code on the FAQ section of the website: http://diffusion.brainnetome.org. Hence, three files are generated: one is the 4D DWI volume series; and the other two are the b-value table and gradient direction table. These three files are used in several subsequent steps.

The DWI volume series (as a 4D zipped NIfTI file) are then corrected for the distortions induced by off-resonance field and the misalignment caused by subject motion. The off-resonance effects are usually caused by the eddy currents of switching the diffusion encoding gradients and the susceptibility distribution of the imaged subjects, resulting in the deterioration of images due to blurring, spatial distortion, local signal artifacts, etc. The motion effects also cause image blurring and geometric misalignment (Andersson and Sotiropoulos, 2016; Bernstein et al., 2004). Advanced correction mechanism for susceptibility-induced distortion when data acquisition with different phase-encode parameters is becoming increasingly popular. To include the correction method using different phase-encode information, we have exported the functions of *topup*, *applytopup*, *eddy* and *eddy_combine* from FSL (Andersson and Sotiropoulos, 2016; Andersson et al., 2003; Smith et al., 2004), compiled them on both Linux and Windows platforms, and packed the executable files into DiffusionKit. Unfortunately, most clinical acquisitions do not currently meet the requirement (two or more acquisitions where the parameters are different so