Comparison of image quality control tools for Diffusion Tensor Images

Bilan Liu[1], Tong Zhu[2], and Jianhui Zhong[2]
[1]*Electrical and Computer Engineering, University of Rochester, Rochester, NY, United States,* [2]*Imaging Sciences, University of Rochester, Rochester, NY, United States*

Introduction: Diffusion tensor imaging (DTI) [1] is susceptible to numerous detrimental artifacts that may impair the reliability and validity of the obtained data. The quality of images acquired in DTI is critical for image interpretation, diagnostic accuracy and efficiency. Although many Quality Control (QC) tools are being widely used, there is still no agreement on image quality control routine for diffusion images. An objective comparison between existing QC tools is helpful for identifying the pros and cons for each, and for determining the situation when each should be applied and why. This study aimed to quantitatively compare the effectiveness of several popular QC tools including DTI studio (JHU) [2], DTIprep (Chapel Hill, NC) [4] and Tortoise (Bethesda, MD) [5]. Both synthetic and human in vivo data were used to quantify the adverse effect of major DTI artifacts to the tensor calculation as well as the effectiveness of different QC software in correcting and compensating these artifacts.
Method: Three QC tools were evaluated here: DTI studio, DTIprep and Tortoise using both synthetic DWI data simulated by Monte Carlo simulations and in vivo human DTI data. For synthetic data, simulated FA/MD values were used as golden standard, and for data in vivo, DTI parameters calculated from artifact-free dataset of the same subject were used as references. **Data acquisition:** *Simulation data:* synthetic data were simulated with characteristics similar to brain tissue using Monte Carlo simulations, with FA=0.7, Trace $=700\times10^{-6}mm^2/s$ ($\lambda1>>\lambda2=\lambda3$) in White matter (WM) and FA=0.2, Trace= $700\times10^{-6}mm^2/s$($\lambda1=\lambda2>>\lambda3$) in gray matter (GM). Two major DTI-specific artifacts were simulated in this study: (1) Motion artifact, which often leads to signal variation across the image. Motion-induced artifacts were simulated with DWI intensity decreased by 90% within the slice. (2) Cardiac pulsation, which also lead to significant brain motion and signal drop mainly in the vicinity of ventricles. Motion-induced artifacts were simulated with DWI intensity decreased by 50% near the area prone to cardiac artifact. Adverse effects of DTI-specific artifacts to the accuracy and precision of tensor-derived parameters are often complicated by several confounding factors. Therefore, for each artifact type, Monte Carlo simulations were performed using combination of each of following conditions: (a) 3 different diffusion gradient schemes with the number of diffusion directions=12, 30, 60.In all instances the ratio of the number of DWIs to the number of b0 images (N_{DWI}: N_{b0}) in each set was maintained at 6:1[3]. (b) 3 different level of SNRs =25, 50, 125. (c) For motion artifacts, 3 different levels of corrupted datasets were considered, with 8%, 17%, and 33% of total diffusion gradient directions contaminated with motion artifacts. For each level of corrupted datasets, simulations with corrupted gradients either evenly or unevenly distributed in hemisphere were also simulated and investigated. As a result, a total of 54 datasets were simulated for motion artifact. Another 9 datasets for cardiac artifact were also simulated. *In vivo human DTI data:* Two repeated scans of the same subject were obtained using a 1.5T MR unit (General Electric), and with the following imaging protocol: FOV=24cm, TR/TE=10s/66.3ms, data matrix size: 256*256, b-value=1000$seconds/mm^2$, 30 DWIs with 5 b0 images. One repeated scan with both motion and cardiac artifacts and the other repeat scan was artifact-free through careful visual inspection.
Data analysis: Diffusion tensors and tensor derived quantities (FA/MD) were calculated using different QC tools. Data results were processed offline using MATLAB (The Mathworks, MA, USA). To assess potential variations in the accuracy and precision of DT derived parameters, the following measures were computed over each dataset: 1) Mean FA/MD values were calculated from 10000 QCed MC Samples. 2) Standard deviation of FA/MD values were also calculated from 10000 QCed MC Samples. The differences in accuracy (ΔFA and ΔMD) and in precision ($\Delta\sigma(FA)$ and$\Delta\sigma(MD)$) from QCed diffusion weighted data were reported and analyzed for both simulation data and human in vivo data.
Results: We evaluated the performances of three QC tools applied to each of the 63 datasets. Our results show that none of the tools strictly subsumes another. The overall performances of different QC tools are shown in Table1. Figure 1 shows measured mean FA, FA SD in WM of simulation data as a function of percentage of corrupted gradients, with corrupted gradients evenly (blue dot) or unevenly (green dot) distributed over the hemisphere. The loss in accuracy increases with increasing ratio of corrupted gradients for both DTIstudio and Tortoise. The seemly stable results from DTIprep were obtained by discarding the bad gradients, which yields error to tensor calculation when the number of remaining gradients is small. Figure2 shows measured FA mean, FA SD on synthetic datasets contaminated by cardiac artifact as a function of number of diffusion directions (SNR=50). The increase of total number of diffusion gradients leads to increase in both accuracy and precision of DT derived parameters. Figure3 shows measured mean FA/Trace with error bar showing the STD of FA/MD of human in vivo dataset. Results show good performance of both tools in the presence of artifacts in real data.

Table1: Conclusion of performances of different QC tools

	Tortoise	DTIPrep	DTIstudio
Fully automated	Partially	Fully automated	No
Image information check	No	Yes	No
Outlier detection	Slice-based/ Pixel based	Gradient-based	Slicet-based/pixel-based
Input/output data type	Philips Par/Rec, Dicom, FSL Nifti, Analyze	DICOM /NRRD	Works for most of the data types.
stability	unstable when the corrupted data ratio >30%	Stable when the remaining number of gradients ≥6	Become unstable when the corrupted data ratio >30%
Efficiency	Moderate	Fastest	Fast
Ease-of-use	Moderate	Easy	Easiest

Table1*:* Conclusion of performances of different QC tools. ***Figure1****:* Measured FA mean, FA SD profile for synthetic datasets as a function of percentage of corrupted gradients. (SNR=50, Number of Directions=30.) ***Figure2:*** Measured FA mean, FA SD on synthetic datasets corrupted by cardiac artifact as a function of number of diffusion directions. (SNR=50). ***Figure3:*** Measured mean FA/Trace with error bars showing the STD of FA/MD of human in vivo datasets. Blue bar-artifact-free dataset, Red bar - dataset with artifact.

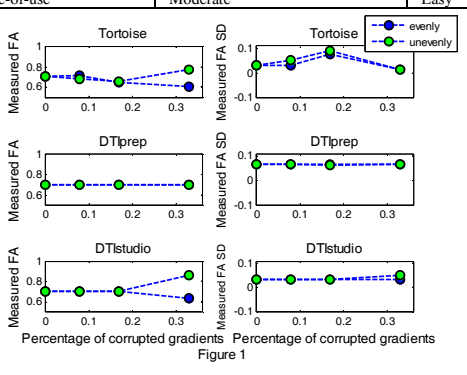
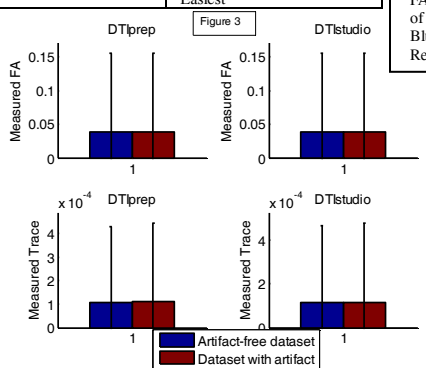


Figure 1

Figure 3



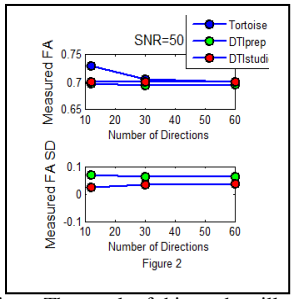


Figure 2

Conclusion: In this study, we present an experimental protocol for the evaluation of QC tools for DT acquisitions. The result of this study will help DTI users choose among different QC tools and to develop pipelines for Quality Control of images in diffusion tensor MRI studies.

References: [1] Basser PJ et al. *NMR Biomed* 2002; 15:456 – 467. [2] Hangyi Jiang et al. *COMPUT METH PROG BIO*, Vol. 81, Issue 2, 106-116 [3] Jones DK, et al. *Magn Reson Med* 1999; 42:515–525. [4] Zhexing Liu, et al, Proc. *of SPIE* Vol. 7628 76280J-1. [5] C. Pierpaoli1, et al, *ISMRM 18th annual meeting*.