Systematic Evaluation of Linear and Nonlinear DTI Estimation Methods: An Open Framework

B. A. Landman[1], S. Mori[1,2], and J. L. Prince[1,3]

[1]Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States, [2]The Russell H. Morgan Department of Radiology and Radiological Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States, [3]Electrical and Computer Engineering, Johns Hopkins University, Baltimore, Maryland, United States

Introduction

Diffusion tensor imaging (DTI) provides contrasts sensitive to tissue orientation and micro-architecture. A basic premise of DTI is that the tensor formalism (e.g., assumption of Gaussian diffusion) meaningfully represents diffusion processes, and thus the derived contrasts are relevant. Yet, not all tensors represent physically possible processes (e.g., those with negative eigenvalues); typical log-linear mean squared error methods can result in these non-physical solutions. Various nonlinear tensor estimation frameworks have been developed to prevent these problems [1-3], while regularization and robust tensor estimation methods employ spatial correlations to lessen the effects of noise [4-5]. Despite the emergence of new methods, little evidence has been presented to provide equivalent experimental comparisons or enable well-informed selection of the appropriate tensor estimation method for a particular task.

This study investigates the impact of tensor estimation method on derived contrasts as a function of SNR, presents a framework to evaluate methods, and establishes an open web portal to enable collaborative evaluation of methods. Tensor estimation accuracy as a function of SNR is mapped through multiple simulations, and trends in bias and variability are exposed. These data provide guidance as to the level of uncertainty introduced between studies conducted with differing tensor estimation schemes and/or SNR levels.

Methods

A high resolution, high SNR dataset with 15 repetitions of 30 diffusion weighted (DW) and one scanner average of 5 minimally weighted (b0) acquisition at 1.5T was used as ground truth to assess the effects tensor estimation method. Simulations were performed with modeled noise as previously described [6,7]. An axial slice at the level of the lateral ventricles was selected to be representative of human brain anatomy (see inlays **Fig. 1**). SNR is reported with respect to the b0.

Four methods of tensor estimation were compared to demonstrate the practical consequences of applying a few of the many widely available choices. First, we directly applied the Stejskal-Tanner tensor model [8] to the estimation of tensor coefficients (LL-MMSE). Next, two *ad hoc* methods for reducing the impact of negative eigenvalues were evaluated: (1) we replaced any DW values that were greater than the b0 value with the b0 value prior to LL-MMSE tensor estimation, denoted LL-MMSE (clip DWI), and (2) we reduced the impact of non-physical solutions by replacing negative eigenvalues with zero-values, denoted as LL-MMSE (clip eigenvalues). Finally, we evaluated the non-linear fitting method provided with the AFNI toolkit (National Institutes of Health, Bethesda, MD) [2]. Analyses were performed with a combination of custom Matlab (MathWorks, Natick, MA) scripts and AFNI tools.

For each SNR and method combination, the bias, variability, and root mean square (RMS) error for fractional anisotropy (FA) and mean diffusivity (MD) were computed. For comparison, the average RMS errors for tensors estimated from single *in vivo* acquisitions are shown. To provide a single quantitative measure of the differences between tensor estimation methods, we computed the average RMS error over 25 dB to 40 dB. This interval was chosen to correspond to typical achievable range of SNR in an *in vivo* clinical DTI setting. To enable direct comparisons at arbitrary SNR, a sigmoid function was fit to the FA and MD RMS errors over the same SNR levels (functional fits are reported online).

Results and Discussion

The low SNR limiting mean values of FA and MD depended on tensor estimation method (**Fig. 1**). At very low SNR (<0 dB), there was little contribution from the underlying ground truth data on the estimated contrasts. At low SNR (0-20 dB), the bias in underlying contrasts rapidly increased with decreasing SNR, while the rate of change was much reduced at moderate SNR (20-40 dB). At high SNR (>40 dB), there was little change with SNR or dependence on estimation method. The positive definite method showed reduced bias in FA (**Fig. 1A**), but increased bias in MD (**Fig. 1B**). In terms of RMS error, the positive definite method out performed the three LL-MMSE methods for estimating FA, but showed reduced performance in MD estimation (**Table 1**). The modified LL-MMSE algorithms demonstrated improvements in all FA measures over the LL-MMSE method and nominal MD improvements.

Many factors influence tensor estimation accuracy and bias, including diffusion weighting schemes, number of b0 acquisitions, spatially varying SNR, and artifact. This simple framework provides a clinically relevant and easy to interpret metric to guide selection of tensor estimation procedure based on the tradeoffs between bias, variability, and CPU time.

Data, source code, and detailed results (including bias, variability, and RMS error) are presented on the project webpage, which is accessible through http://iacl.ece.jhu.edu/~bennett/. Submissions from the community of evaluations of other algorithms are encouraged and will be incorporated into the community resource.

Table 1. Mean Tensor Estimation Error (25 to 40 dB) (FA: [FA]x1000. MD: mm^2/s x 1000)

	RMS Error		Bias		Variability		CPU Time
Estimation Method	*FA*	*MD*	*FA*	*MD*	*FA*	*MD*	*Relative*
LL-MMSE	64.6	45.8	38.6	-8.48	50.8	44.8	1
LL-MMSE (clip DWI)	64.5	45.7	38.5	-8.45	50.8	44.7	1
LL-MMSE (clip eigenvalues)	64.5	45.8	38.6	-8.45	50.8	44.8	1
Positive Definite MMSE	61.2	46.8	35.7	-15.22	48.8	44.0	8.4

Figure 1A. Fractional Anisotropy

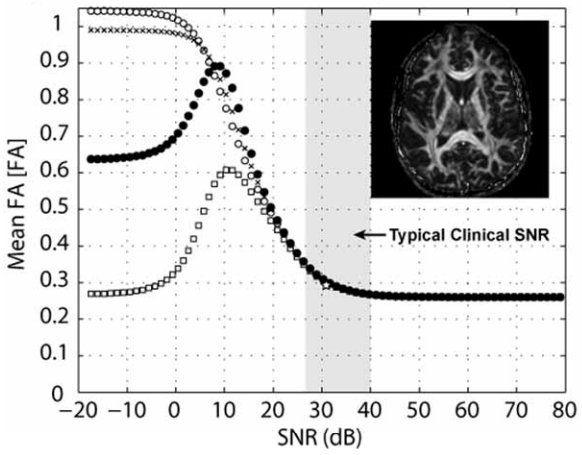
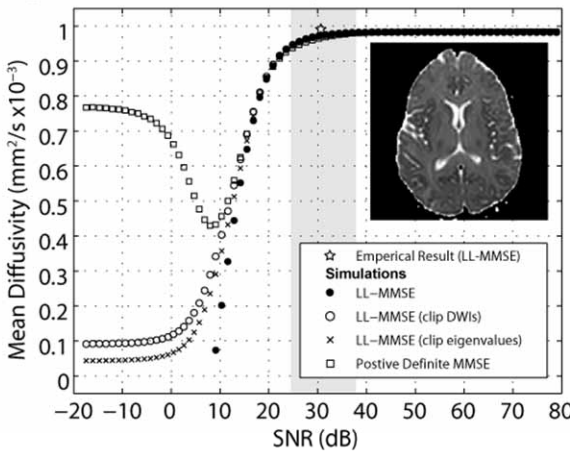


Figure 1B. Mean Diffusivity



References: [1] Tschumperlé, D. et al., 2003 9th International Conference on Computer Aided Systems Theory [2] Cox, R. et al., 2006. ISMRM [3] Niethammer, M., et al., 2006 28th IEEE EMBS [4] Mangin, J.F., et al., 2002 Med Image Anal 6(3):191 [5] Chang, L. C. et al., 2005 MRM 53(5):1088 [6] Farrell, J.A.D. et al., 2006 Under review [7] Landman, B.A. et al., 2006 Under review [8] Stejskal, E. O. et al., 1965 J Phys Chem 42:288