Extraction of nonlinear features in MEG and functional MRI data of human brain

[1]Amir Assadi, Liya Wang
University of Wisconsin - Madison

Short Abstract

We have determined an new method for quantifying the nonlinearities of brain imaging data sets to allow comparison of measures of nonlinearity across comparable experiments. We have an efficient algorithm for finding nonlinear features of high dimensional data from variations in two-dimensional projections and slices. By computing the second order statistics of these features, we are able to construct an analogue of the Riemann Curvature Tensor, thereby describing the nonlinearity of the data set. We demonstrate the usefulness of this technique on real MEG and fMRI data.

Extended Abstract

1 Purpose

Working with massive "real world" data is inherently difficult because we often do not know important factors that are significant in any pattern recognition and feature extraction tasks. These factors range from the signal to noise ratio in the data, lack of a statistical distribution for the uncontaminated data, to situations such as not having a robust statistical model for the noise. A simplifying assumption is often helpful, but we must be certain to assess the impact of the assumptions that are made along the way on the final outcome of the analysis. Bayesian methods have the advantage that they make such assumptions explicit in the model. Nonetheless, the Bayesian models will reach their limits in the absence of better estimates on the distribution of data or at least, partial knowledge of such distributions. We propose a method for obtaining relevant probabilistic distributions by extracting information about nonlinearity in those representations of data that lend themselves to certain operations that we call for short "local-to-global integration" and "local linearization". We apply this technique to fMRI data to describe certain aspects in nonlinear dynamics of the brain function as reported by the fMRI acquisition.

2 Method

From differential geometry, we know that the Riemannian curvature tensor (RCT) gives a comprehensive description of the nonlinearity in the underlying spaces endowed with certain abstractions of the infinitesimal calculus. A possible approach to description of nonlinearity in data is to look for a concept similar to this marvelous discovery of Riemann. In particular, we confine our study to data sets that can be represented as a collection of points in the Euclidean space. In this context, we define a new geometric structure that lends itself to associate to the data at various points something that is analogous to the Riemann Curvature Tensor, that we refer to as "data Curvature Tensor" (dCT) . If the data is sampled sufficiently finely from the points on a Riemannian manifold in the standard sense, then the Riemann Curvature Tensor and the dCT are within the approximation bounds defining the sampling of data from the manifold. Thus, the Riemann curvature tensor, and hence nonlinearities of the manifold can be recovered by computing the dCT from available data points. We prove a result that under reasonable circumstances, data a collection of two-dimensional subspaces that linearly approximate local statistics of data suffices for estimation of dCT. Therefore, in the dCT approach, we must determine the nonlinearity of features in the prescribed family of two-dimensional planes in order to have the basic estimates from which we calculate estimates for other nonlinear features of the data. We consider, therefore, estimation of nonlinearities of data in the plane.

In this paper, we propose a solution for describing the nonlinear dependencies of data in the family of two-dimensional linearization. With this in mind, we must remark that extracting the feature from a general data set is an ill-posed problem in the mathematical sense, and we make certain assumptions to transform it to a well-defined estimate. Thus, it is important to recover such assumptions when we have implicit knowledge from the system that gives rise to the data, here, the MRI instrumentation and basic facts about neurophysiology of the brain underlying the acquired signals. The technical aspect

outside of mathematics is to justify the realistic nature of such estimates from fMRI data. We propose to a large-scale data analysis experiment that will either validate our hypotheses, or provide alternative estimates that could be used for analysis of nonlinearity in fMRI in the dCT framework.

3 Results

In our paper we demonstrate these techniques when applied to both synthetic and real data from fMRI of human brain. The outcome describes the nonlinearity of the temporal relationship between regions of the human brain during a motor activation task, and in turn, we show what the results reveal about nonlinear dynamics in fMRI.

4 New or breakthrough aspects of work

We have determined a new and efficient method for describing the nonlinearity of a data set. Our algorithms that recover high-dimensional nonlinearity from nonlinear features in a family of two-dimensional planes is bound tom have many applications.

5 Conclusion

We have proposed a novel approach to identification of the high-dimensional nonlinearity of a data set that embraces ideas from both differential geometry, probability, and learning theory.