Future Generation Computer Systems 43–44 (2015) 110–119

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Towards adaptive synchronization measurement of large-scale non-stationary non-linear data

Chang Cai [a], Ke Zeng [b], Lin Tang [a], Dan Chen [d,f,*], Weizhou Peng [a], Jiaqing Yan [e], Xiaoli Li [b,c]

[a] School of Computer Science, China University of Geosciences, Wuhan, 430074, China
[b] State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China
[c] Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing 100875, China
[d] Collaborative and Innovative Center for Educational Technology, China
[e] Institute of Electrical Engineering, Yanshan University, Qinhuangdao, 066004, China
[f] School of Computer, Wuhan University, Wuhan, 430072, China

HIGHLIGHTS

- We designed an approach to measure the synchronization strength of non-stationary nonlinear data against phase differences.
- We demonstrated that the synchronization analysis was an effective indicator of an epileptic focus location.
- We developed a parallelized approach with general-purpose computing on the graphics processing unit (GPGPU), and it largely improved the scalability of data processing.

ARTICLE INFO

Article history:
Received 4 February 2014
Received in revised form
1 May 2014
Accepted 1 June 2014
Available online 30 June 2014

Keywords:
Non-stationary nonlinear data
Synchronization
Adaptive analysis
EEG
Epilepsy

ABSTRACT

Synchronization measurement of non-stationary nonlinear data is an ongoing problem in the study of complex systems, e.g., neuroscience. Existing methods are largely based on Fourier transform and wavelet transform, and there is a lack of methods capable of (1) measuring the synchronization strength of multivariate data by adapting to non-stationary, non-linear dynamics, and (2) meeting the needs of sophisticated scientific or engineering applications. This study proposes an approach that measures the synchronization strength of bivariate non-stationary nonlinear data against phase differences. The approach (briefed as AD-PDSA) relies on adaptive algorithms for data decomposition. A parallelized approach was also developed with general-purpose computing on the graphics processing unit (GPGPU), which largely improved the scalability of data processing, namely, GAD-PDSA. We developed a model on the basis of GAD-PDSA to verify its effectiveness in analyzing multi-channel, event-related potential (ERP) recordings against Daubechies (DB) wavelet with reference to the Morlet wavelet transform (MWT). GAD-PDSA was applied to an EEG dataset obtained from epilepsy patients, and the synchronization analysis manifested an effective indicator of epileptic focus localization.

1. Introduction

Synchronization measurement plays an important role in the study of the interacting dynamics of complex systems [1]. The recordings of the activities of such systems are generally non-stationary and non-linear. Typical examples include electroencephalogram (EEG) and magnetoencephalogram (MEG), which normally consist of simultaneous recordings of tens to hundreds of data channels. Synchronization measurements of such non-stationary, non-linear data remain a challenging issue.

A number of methods have been developed to measure the synchronization of a bivariate signal, e.g., a cross-correlation function, advanced in the 1950s; the use of linear synchronization [2] to analyze time delay and synchronization; the coherence [3] quantities of the synchronization of a bivariate signal in the frequency do-

* Corresponding author at: School of Computer, Wuhan University, Wuhan, 430072, China.
E-mail addresses: dan.chen@ieee.org (D. Chen), xiaoli@bnu.edu.cn (X. Li).

http://dx.doi.org/10.1016/j.future.2014.06.003

main; phase synchronization, including the phase lock, the Hilbert transform, and the wavelet transform, and the phase synchronization references [4,5].

Since the synchronization occurring at different frequencies between two or more signals indicates the interacting dynamics of a complex system [6], signals need to be decomposed into several components consisting of basic functions to obtain the frequencies in order to estimate the synchronization. Fourier analysis has been widely used in the decomposition of interesting signals [7–9] and provides accurate estimates of their spectra with the assumption that these signals are stationary. However, Fourier analysis is not well-suited for non-stationary data since the spectra of non-stationary data change with time. Fortunately, wavelet-based methods have been proposed as powerful alternatives to Fourier-based methods for the synchronization estimates of non-stationary signals. Time-course synchronization, based on wavelet, enables the extraction of the temporal variation in the synchronization among non-stationary signals and is thus a better candidate for the measurement of non-stationary dynamic signal interaction.

However, wavelet-based synchronization methods generally suffer from a lack of energy–time–frequency correlation due to the limited length of the basic wavelet function. This inherent problem inevitably misleads the calculation of time-course synchronization. Furthermore, the selection of the center frequency of wavelet transform relies on a priori knowledge of the frequency characteristics of the target components. A blind full spectrum scan (if at all possible) will only result in very difficult interpretations [10]. Moreover, some wavelet transforms aided by Fourier transform, taking the most commonly used Morlet wavelet as an example, cannot make a physically meaningful interpretation of the non-linear signal because of the spurious harmonic components that cause energy spreading induced by the linear Fourier transform. Apparently, the wavelet analysis is not adaptive in nature [11]. There is a pressing need for an approach that adapts to the dynamics of non-stationary non-linear data.

We consider that the key to solving this problem lies with the decomposition of the data. The premise to achieve adaptability, in contrast to using these linear methods, is the viability of automatically exploring the structure of non-stationary, non-linear data and generating the necessary adaptive bases from the data. Ensemble empirical mode decomposition (EEMD) and local mean decomposition (LMD) are two prevalent methods in this direction. EEMD can break down complicated data without a basic function into a series of embedded oscillatory intrinsic mode functions (IMFs), which is particularly suited to noisy data. Similarly, LMD decomposes amplitude and frequency modulated signals into a small set of product functions (PFs), each of which is the product of an envelope signal and a frequency modulated signal from which a time-varying instantaneous phase and instantaneous frequency can be derived. In this study, we propose an adaptive decomposition based phase difference synchronization analysis (AD-PDSA) method.

The decomposed components with EEMD or LMD are further processed via the Hilbert transform [12] to obtain the instantaneous phase of each component. The synchronization strength of bivariate data series, i.e., two IMFs/PFs of the same original data segment or two different ones, is measured by the statistical analysis of the ratio between the phase differences of each series at the same time point based on Cauchy–Schwarz inequality [13]. The value of the strength ranges from 0 to 1, where 0 denotes no synchrony and 1 denotes perfect synchrony. The synchronization measurement can eventually be extended to quantify the global synchronization of multiple components from the same data segment or multivariate data.

Another challenge is to bridge the gap between the proposed method and sophisticated scientific or engineering applications. Both EEMD and LMD algorithms must repetitively process iterative computations. The precision of outputs heavily depends on the number of iterations, which should be large enough to eliminate the iterative error. The number of simultaneously recorded neurons has approximately doubled every seven years over the last five decades [14]. Analysis of neural data in tens or hundreds of channels has become common within the neuroscience community. Due to the rapidly growing scales and sizes of neural data and the high complexity of decomposition algorithms, how to ensure the scalability of the proposed method is an important issue. Modern cyber infrastructures have played an important role in solving compute-intensive [15–18] and data-intensive scientific problems [19–22]. In this study, this goal was achieved by gearing contemporary computing technologies rather than altering the fundamental theory of data decomposition.

A parallelized method of successfully adapting the technology of general-purpose computation on graphics processing units (GPGPU) was developed to significantly enhance the scalability of the original sequential version. The GPGPU-enabled AD-PDSA (GAD-PDSA) explores the very fine-grained parallelism of LMD with original parallel algorithms. It can operate on neural data streams with a large (up to one thousand) number of channels. The scalability of the approach adapts well to the advances in recording technologies of neural activities. Compared to the AD-PDSA, the execution time of the new approach is a polynomial function of the data size while that for conventional CPU-based platforms conforms to an exponential function. This will result insignificant differences when handling very large data. We developed a model on the basis of GAD-PDSA to verify its effectiveness in analyzing multi-channel, event-related potential (ERP) recordings against Daubechies (DB) wavelet with the Morlet wavelet transform (MWT) reference. AD-PDSA was then applied to experimental datasets to analyze the potential synchronization of EEG data obtained from epilepsy patients. Epilepsy is defined as spontaneous clinical seizures caused by paroxysmal, abnormally synchronous neuronal activity. The electrical symptoms of this abnormal activity are believed to uniquely define and reveal the mechanisms of an underlying abnormal neural function and structure. The localization of the initial seizure discharge is an attempt to find the region that generates the abnormal neural activity. Therefore, the analysis of ictal EEG (scalp or intracranial) is an effective standard for the identification of the epileptic focus localization. The results indicated that (1) the new method is adaptively suitable for the synchronization analysis of non-stationary and non-linear series, (2) it can detect a reliable global phase correlation amongst series components. GAD-PDSA has also improved scalability compared to the original serial AD-PDSA method.

The remainder of this paper is organized as follows. In Section 2, we propose a new method, AD-PDSA, which is based on the PDSA and EEMD or LMD. Section 3 details the design of the GPGPU-enabled AD-PDSA. Experiments and results are presented in Section 4. We conclude the paper with a summary in Section 5.

2. Methods

This section first introduces the algorithm of the phase difference synchronization analysis (PDSA) and the adaptive decomposition-based PDSA (AD-PDSA). Following this, an extended AD-PDSA for global analysis is introduced. Then, the GPGPU is briefly described, and the levels of the AD-PDSA as well as the parallelization of the AD-PDSA are discussed in the next part.

2.1. Adaptive decomposition-based phase difference synchronization analysis (AD-PDSA)

Let $X = (x_1, x_2, \ldots, x_T), Y = (y_1, y_2, \ldots, y_T)$ denote the bivariate data (e.g., activities of two brain regions), T is the number of data channels. The flow of the PDSA is presented in Fig. 1.