Journal of Neuroscience Methods 206 (2012) 94–101

Contents lists available at SciVerse ScienceDirect

Journal of Neuroscience Methods

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

Computational Neuroscience

The extraction of LRP via functional data analysis techniques

Yan-Bin Zhao[a,b], Jian Tao[b,*], Ning-Zhong Shi[b], Ming Zhang[c], Jie Sui[d]

[a] School of Mathematical Sciences, Harbin Normal University, Harbin, Heilongjiang 150025, China
[b] Key Laboratory of Applied Statistics of MOE, School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin 130024, China
[c] Department of Psychology, Northeast Normal University, Changchun, Jilin 130024, China
[d] Department of Psychology, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Received 6 October 2009
Received in revised form 6 February 2012
Accepted 7 February 2012

Keywords:
Lateralized readiness potential
Onset latency measurement
Functional data analysis

ABSTRACT

A new strategy based on functional data analysis (FDA) techniques is proposed to extract the lateralized readiness potential (LRP), which treats electroencephalographic data as functional data. This FDA-based method combines longitudinal information from each trial (time series data) with cross-sectional information from all trials at a fixed time point (cross-sectional data). The comparison results show that the FDA-based LRP is closer to the assumed true LRP and is more robust against a reduction in the number of trials than the traditional average-based LRP. Furthermore, the results indicate that the onset of an FDA-based LRP is more accurate than that of an average-based LRP under several measuring criteria.

1. Introduction

The technique of event-related potential (ERP) has been used in many ways for the development of neuroelectrophysiology since the 1960s. In particular, the lateralized readiness potential (LRP) has become an especially useful temporal marker within the area of chronopsychophysiology (Van der Molen et al., 1991). By definition, the LRP reflects hand-specific lateralization processes. Thus, by examining the effects of experimental manipulations on the onset latencies of LRPs, researchers can assess whether the manipulations influence processes prior to the onset of hand-specific lateralization, posterior to, or both (e.g., Miller et al., 1998; Osman and Moore, 1993; Smulders et al., 1995).

A variety of methods to estimate the onset of the LRP have been proposed. The popular methods can be grouped into three general types as follows:

- Baseline-deviation (BD) method (Osman and Moore, 1993). The onset of the LRP is defined as the first point in time when the LRP consistently exceeds some value, say, when the LRP value exceeds the mean plus a multiple of the standard deviation (SD) of the LRP during some baseline period.
- Criterion-based (CB) method (Smulders et al., 1996). This method identifies the onset of the LRP as the first point in time when the LRP exceeds some arbitrary value. The criterion is defined in one of two ways: (1) as a certain proportion of the maximum value of the LRP observed in the condition (i.e. "relative-criterion" method, e.g., 30% of the height of the peak), or (2) as a certain fixed value used for all conditions (i.e., a "fixed-criterion" method, e.g., 0.5 mV).
- Segment regression (SR) method (Schwarzenau et al., 1998). This method is the most recent development and operates in a different manner to both methods mentioned above. This method defines the onset of the LRP as the "break-point" between two intersecting straight lines that are fit to the LRP waveform. In general, one line is fit to the putative preonset segment of the LRP, whereas the other is fit to the segment that rises to the peak, and the two lines are found using least-squares techniques—minimizing the root mean squared difference between the fitted lines and the LRP—and the time of the intersection is taken as the estimated onset of the LRP (Mordkoff and Gianaros, 2000).

One major limitation of the chronopsychophysiological methods is that, it is difficult to determine the LRP onset latency, because of the low signal-to-noise ratio in the LRPs obtained by averaging over a limited number of trials, which is often the practical limit in response-time studies with unpracticed subjects and a number of conditions per subject. Because of the inaccuracies in the LRP onset latency measurements computed from noisy single-subject LRPs, Miller et al. (1998) presented a jackknifing method that is based on grand-average LRPs. This method focuses on estimation of the between-subject variability in LRP onset latency differences under different conditions rather than on the absolute LRP onset under a single condition.

* Corresponding author. Tel.: +86 431 13504475649.
E-mail address: taoj@nenu.edu.cn (J. Tao).

0165-0270/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2012.02.009

The extraction of a relatively pure LRP with a limited number of trials is important in every LRP study. Generally, the LRP was computed from the ERP difference waves between *C*3 and *C*4. Traditional methods usually compute the ERP based on averaging trials together in any given condition; more specifically, at each recorded time, observations from different trials are cross-sectionally averaged together. It is important to note that the cross-sectional averaging captures the mean trend, while the change between trials is reflected by its variation measure, such as the cross-sectional variance. However, the cross-sectional averaging method (averaged-based method) completely ignores the dependence structure within trials and its impacts on the mean and covariance/correlation estimation. In this paper, we will introduce a new method that is based on a functional data analysis method (FDA-based method) to extract the ERP. In contrast to the average-based method, where the observations are considered as finite sets of discrete points, the longitudinal recordings in the FDA-based method are treated as smooth curves (Ramsay and Silverman, 2001; see also Miller, 1990, for a discussion of the discrete versus continuous nature of cognitive processes). More importantly, simulation comparison shows that the FDA-based method effectively estimates both the mean and the onset of ERP by employing the longitudinal recordings from all trials and thus makes full use of the correlation (or dependence) structure among the recordings obtained from the same trial.

2. Extracting LRP based on functional smoothing

The LRP is usually extracted by averaging potentials across epochs, and the raw data are processed through five steps: (a) filtering; (b) removing artifacts from the electroencephalogram (EEG) data; (c) cutting epochs from the continuous recording; (d) correcting baseline (i.e., subtraction of average voltage during the base line from the whole epoch); and (e) averaging. Instead of using the last averaging step in the above traditional procedure, our proposed method uses FDA techniques.

Functional data analysis assumes that the curve being estimated is smooth, so we assume the existence of a function x that gives rise to the observed data and consider the observed data as *functional data*. Here, the term 'functional' refers to the intrinsic structure of the data rather than to its explicit form (Ramsay and Silverman, 2005). In practice, the data are usually observed and recorded discretely as n airs (t_i, y_i), where y_i is the observation at time t_i, possibly blurred by the noise. We express this notation as follows:

$$y_i = x(t_i) + \varepsilon_i,$$

where the noise, disturbance, error, perturbation or otherwise exogenous term ε_i contributes a roughness to the raw data. Furthermore, we assume that the random curves or functions are independently sampled from a mixture of stochastic process X with a continuous and smooth marginal mean $\mu(t)$, and a truncated Karhunen–Loève expansion of the random function exists such that

$$X(t) = \mu(t) + \sum_{j=1}^{k} \xi_j(X)\rho_j(t),$$

where ρ_j is the eigen-function and $\xi_j(X)$ is the random coefficient; for details about the ρ_j, $\xi_j(X)$, and k, refer to Chiou and Li (2007) and Hu et al. (2009). In this paper, we focus on estimating the mean function $\mu(t)$, which represents the intrinsic structure of the entire data set, and the ERP that we need.

Suppose that n trials were used to obtain the ERP, and let $\{(t_{il}, y_{il}) : i = 1, \ldots, n, l = 1, \ldots, m\}$ be the observations, where $y_{il} = y_i(t_{il})$. The estimated value of the function at point t must be influenced mostly by observations near t. This feature has not yet been taken into consideration by any estimator discussed above. In this paper we consider estimators where the local dependence is made more explicit by means of local weighting functions. The estimation method based on local polynomial regression techniques is summarized as follows.

We apply local linear regression to the pooled data of n trials to estimate the mean function $\mu(t)$ as follows:

$$\min_{(\beta_0,\beta_1)} \sum_{i,l} [y_{il} - \{\beta_0 + \beta_1(t_{il} - t)\}]^2 K_h(t_{il} - t).$$

The resulted estimate of $\mu(t)$ is then

$$\hat{\mu}(t) = \hat{\beta}_0 = \frac{\sum_{i,l} y_{il} K_h(t_{il} - t)[S_2 - (t_{il} - t)S_1]}{\sum_{i,l} K_h(t_{il} - t)[S_2 - (t_{il} - t)S_1]},$$

where

$$S_k = \sum_{i,l} K_h(t_{il} - t)(t_{il} - t)^k, \quad k = 1, 2,$$

and $K_h(\cdot) = (1/h)K(\cdot/h)$ is a known Gaussian kernel function with bandwidth h. The optimal bandwidth h can be determined by the cross-validation method (see Appendix A for details). For more general discussions on this topic, please refer to Fan and Gibels (1996), Yao et al. (2003, 2005), and Chiou and Li (2007).

3. Experimental design and data processing

Eight subjects from the Northeast Normal University participated in this study. We recorded event-related potentials from subjects when they performed a letter discrimination task. In the task, each trial started with the visual presentation of a letter (X or T) for 150 ms. Subsequently, a blank screen was presented for a variable time ranging from 950 to 2050 ms. Subjects had to judge whether the letter was 'X' or 'T' with the right or left index finger. They were instructed to respond as quickly and accurately as possible by pressing one of the two keys on the keyboard within a time limit of 500 ms. The two letters were randomly presented in each block. The hand for each judgment was counterbalanced across subjects.

For ERP data analyses, epochs where noise forced the A/D converter to saturation were removed. The sampling was performed at 500 Hz. For stimulus-locked analyses, the EEG was epoched off-line, with epochs ranging from 100 ms before stimulus onset until 900 ms after the onset of the imperative stimulus. The band-pass settings were 0–70 Hz and the data were low-pass filtered (off-line) at a half-power cut-off frequency of 0.1–40 Hz. Only trials with correct responses were included in further analyses. Traditionally, the LRP was computed from the ERP difference waves between *C*3 and *C*4, using the double subtraction method (De Jong et al., 1988; Smid et al., 1987):

$$LRP = \frac{1}{2}(A(C3 - C4)\text{righthand} - A(C3 - C4)\text{lefthand}),$$

where $A(C3–C4)$ is the average of $C3 - C4$. In this paper, instead of using $A(C3–C4)$, we utilized the mean function of $C3 - C4$ (denoted by $F(C3–C4)$), which was estimated by the FDA-based method.

First, the corresponding raw data underwent the first four steps in the processing pipeline, including filtering noise, removing artifacts from the EEG data, cutting epochs from the continuous recording, and subtraction of average voltage during the base line from the whole epoch. To examine the efficiency of our proposed approach, we made a comparison of the FDA-based results with the traditional average-based results in Fig. 1; the LRP obtained by the FDA-based method is much smoother (solid line) than the LRP obtained by the average-based analysis (dotted line). As a