NeuroImage 85 (2014) 181–191

Contents lists available at ScienceDirect

NeuroImage

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

Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data

Sabrina Brigadoi [a,*], Lisa Ceccherini [a], Simone Cutini [b], Fabio Scarpa [a], Pietro Scatturin [a], Juliette Selb [c], Louis Gagnon [c], David A. Boas [c], Robert J. Cooper [d]

[a] Department of Developmental Psychology, University of Padova, Italy
[b] Department of General Psychology, University of Padova, Italy
[c] Department of Radiology, Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA, USA
[d] Biomedical Optics Research Laboratory, Department of Medical Physics and Bioengineering, University College London, UK

ARTICLE INFO

Article history:
Received 29 January 2013
Revised 17 April 2013
Accepted 18 April 2013
Available online 29 April 2013

Keywords:
Functional near-infrared spectroscopy
fNIRS
Motion artifact
Hemodynamic response
Motion correction

ABSTRACT

Motion artifacts are a significant source of noise in many functional near-infrared spectroscopy (fNIRS) experiments. Despite this, there is no well-established method for their removal. Instead, functional trials of fNIRS data containing a motion artifact are often rejected completely. However, in most experimental circumstances the number of trials is limited, and multiple motion artifacts are common, particularly in challenging populations. Many methods have been proposed recently to correct for motion artifacts, including principle component analysis, spline interpolation, Kalman filtering, wavelet filtering and correlation-based signal improvement. The performance of different techniques has been often compared in simulations, but only rarely has it been assessed on real functional data. Here, we compare the performance of these motion correction techniques on real functional data acquired during a cognitive task, which required the participant to speak aloud, leading to a low-frequency, low-amplitude motion artifact that is correlated with the hemodynamic response. To compare the efficacy of these methods, objective metrics related to the physiology of the hemodynamic response have been derived. Our results show that it is always better to correct for motion artifacts than reject trials, and that wavelet filtering is the most effective approach to correcting this type of artifact, reducing the area under the curve where the artifact is present in 93% of the cases. Our results therefore support previous studies that have shown wavelet filtering to be the most promising and powerful technique for the correction of motion artifacts in fNIRS data. The analyses performed here can serve as a guide for others to objectively test the impact of different motion correction algorithms and therefore select the most appropriate for the analysis of their own fNIRS experiment.

Introduction

Functional near-infrared spectroscopy (fNIRS) is a non-invasive neuroimaging technique, which uses light in the near-infrared range to infer cerebral activity. From the changes in intensity of light directed from a source fiber into the tissues of the head and back-scattered to a detector fiber positioned several centimeters from the source, concentration changes of oxy-hemoglobin (HbO) and deoxy-hemoglobin (HbR) can be computed (Boas et al., 2002; Jöbsis, 1977). fNIRS is becoming more and more common in the study of infants (Lloyd-Fox et al., 2010; Taga et al., 2011; Wilcox et al., 2010), cognition (Cutini et al., 2012; Köchel et al., 2011; Tupak et al., 2012), motor tasks (Brigadoi et al., 2012; Perrey, 2008) and in studies with difficult and hard-to-test populations, e.g. stroke patients (Lin et al., in press; Muehlschlegel et al., 2009; Obrig and Steinbrink, 2011). Although the improvement in fNIRS technology has been significant in recent years, effectively coupling the sources and the detectors to the head can be problematic and motion artifacts are often a significant component of the measured fNIRS signal. Indeed, every movement of the head causes a decoupling between the source/detector fiber and the scalp, which is reflected in the measured signal, usually as a high-frequency spike and a shift from the baseline intensity. In order to properly estimate the hemodynamic response function (HRF), motion artifacts should be detected and removed.

A common and simple way to solve the issue of motion artifacts is to reject all trials where a motion artifact has been detected. However, this approach is only suitable if the number of motion artifacts detected is low and the number of trials is high, otherwise the risk is that too few trials will be accepted, resulting in a very noisy

* Corresponding author at: Department of Developmental Psychology (DPSS), University of Padova, Via Venezia 8, 35131, Padova, Italy. Fax: +39 049 8276547.
E-mail address: sabrina.brigadoi@studenti.unipd.it (S. Brigadoi).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.04.082

hemodynamic response. fNIRS is particularly suited for examining challenging populations (e.g. infants, clinical patients, children) who might not be easily investigated with fMRI. However, in these populations the number of functional trials is almost always strictly limited, and therefore trial rejection might not be feasible.

Several methods have been proposed to solve this issue. Some methods require a complementary measure of the motion artifact to aid in its removal, e.g. with a short-separation fNIRS channel (Robertson et al., 2010), or with an accelerometer (Virtanen et al., 2011). Others rely on the inherent changes in the amplitude and frequency of the data due to the artifact and act as post-processing techniques. The latter group does not require a complementary measure and thus can be used with every experimental paradigm, making it the most general solution. Among these approaches are principal component analysis (PCA) (Zhang et al., 2005), Kalman filtering (Izzetoglu et al., 2010), correlation-based signal improvement (CBSI) (Cui et al., 2010), wavelet filtering (Molavi and Dumont, 2012) and spline interpolation (Scholkmann et al., 2010).

Motion artifacts can have different shapes, frequency content and timing. They can be high amplitude, high frequency spikes, easily detectable in the data-series or they can have lower frequency content and be harder to distinguish from normal hemodynamic fNIRS signals. Motion artifacts can be generally classified into three categories, spikes, baseline shifts and low-frequency variations. They can be isolated events or they can be temporally correlated with the HRF. Therefore, it is likely that the efficacy of each motion artifact correction technique will vary with the type of motion artifact and that the best technique to apply is data-dependent. One way to estimate the performance of a motion correction technique or to compare different techniques is to simulate motion artifacts (Scholkmann et al., 2010) or to ask participants to move their head purposely to create a motion artifact (Izzetoglu et al., 2010; Robertson et al., 2010). However, real motion artifacts are complex and variable, and thus difficult to simulate. Furthermore, motion artifacts are not only due to the movement of the head, but also due to the movement of the eyebrows or the jaw, for example. The most suitable approach to quantifying the performance of different motion artifact correction techniques is to use real, resting-state fNIRS data, which are contaminated with real motion artifacts, and add a simulated HRF to these data (Cooper et al., 2012). Knowing the true hemodynamic response, it is possible to compute the MSE (mean-squared error) and the Pearson's correlation coefficient (R^2) between the simulated and the recovered HRF, and hence to have a quantitative measure to compare the different performances.

The next step towards establishing a standard approach for the correction of motion artifacts in fNIRS data is to compare the performance of multiple motion correction approaches on real cognitive data. To that end, the aim of this paper is to compare the performance of five motion correction techniques: PCA, spline interpolation, wavelet filtering, Kalman filtering and CBSI, on real data acquired during a cognitive linguistic paradigm. This data-series has been specifically chosen because it contains a particular type of motion artifact, a task-related, low frequency artifact with amplitude comparable with that of the HRF. These characteristics make artifact detection and correction especially challenging. In most cases to date, motion correction techniques have been tested, with great success, on high frequency spike artifacts occurring randomly throughout the data-series, but their ability to isolate and correct artifacts which more closely resemble normal physiological fNIRS signals has not been assessed. As the true HRF in these data is unknown, we use parameters related to a physiologically plausible HRF to compare the performance of each motion correction technique. We also compare the performance of each correction technique with the results obtained by rejecting all trials where a motion artifact was detected and the results obtained by simply including all trials and ignoring the motion artifact altogether.

Materials and methods

fNIRS data

Twenty-two students of the University of Padova (10 males, mean age 25.54 ± 3.14) took part in the experiment, after providing written informed consent. The data of one participant was discarded because she was unable to correctly perform the task, while the data of three others was discarded because of poor SNR in every channel (likely due to a large mass of hair). Therefore, the total number of participants considered in the following analysis is 18. Each participant was comfortably seated in front of an LCD computer monitor at a viewing distance of approximately 60 cm in a dimly lit room. The paradigm consisted of a color-naming of a non-color word task; the participant was asked to say aloud the color of the text of a word appearing on the screen. The study consisted of 4 different stimulus conditions, with 40 trials per condition presented to the participants, leading to a total of 160 trials, divided into two sets of 80. Each word was presented on the screen until the subject started to pronounce the color of the word (~850 ms). The inter-stimulus interval varied among 10, 11 or 12 s. The experiment was approved by the ethical committee of the University of Padova.

The fNIRS data was acquired with a multi-channel, frequency-domain NIR spectrometer (ISS Imagent™, Champaign, Illinois) equipped with 32 laser diodes (16 emitting light at 690 nm and 16 at 830 nm) and 4 photo-multiplier tubes. Source and detector fibers were positioned on the participants' head using a probe-placement method based on a physical model of the head surface (Cutini et al., 2011) so that frontal and premotor areas were sampled (Fig. 1a) (for more details on the positions of sources and detectors see Cutini et al. (2008)). Each source fiber carried light at both of the two different wavelengths; five source fibers were placed around each detector fiber, at a distance of 3 cm. Therefore, a total of 20 channels per wavelength (10 per hemisphere) were measured for each participant. The sampling frequency was set to approximately 7.8 Hz.

The data acquired during this experiment contained a particular type of motion artifact, which was caused by the participants' jaw movement induced by the vocal response. The opening and closing of the mouth caused an abrupt displacement of the sources and detectors positioned on the participant's head, thus producing a motion artifact in the data series that was correlated with the evoked cerebral response (present in the first 1–2 s after stimulus onset). The shape and duration of this artifact (Fig. 1b) differ from the more common spike-like artifacts because it is slower and correlated with the hemodynamic response. Given that its amplitude is comparable to the hemodynamic response elicited by cortical activity, the artifact is more difficult to detect.

It is also important to note that not all participants and all channels presented this type of artifact; participants with less hair tended to have the fiber holder placed more tightly to the head and hence this type of artifact was less common. The artifact was also channel-specific, appearing more commonly in the most anterior channels (see Fig. 1a). The fact that the motion artifact is not observed on all channels simultaneously is likely to affect the performance of the motion artifact correction, since some methods inherently require unwanted signal components to be apparent in multiple channels. While this hypothesis may be reasonable in many cases, as motion artifacts are often due to movement of the whole head, this is not the case for this data series. Therefore, it is likely that the correction methods which work on a channel-by-channel basis will perform better than those that work on all channels all together. Below we describe the motion correction techniques compared in the present work.