Preprocessing Pipeline for fNIRS Data in Children

Caterina Piazza[1(✉)], Andrea Bacchetta[1], Alessandro Crippa[1], Maddalena Mauri[1], Silvia Grazioli[1], Gianluigi Reni[1], Maria Nobile[1], and Anna Maria Bianchi[2]

[1] Scientific Institute, IRCCS E. Medea, Bosisio Parini, LC, Italy
caterina.piazza@lanostrafamiglia.it
[2] Department of Electronics Information and Bioengineering, Politecnico di Milano, Milan, Italy

Abstract. Functional near-infrared spectroscopy (fNIRS) is a non-invasive neuroimaging technique, largely used in paediatric research. However, there is not a standardized and widely accepted protocol for fNIRS data processing with potential effects on the reliability and replicability of the obtained results. The present study is within this framework aiming at the identification of an adequate pre-processing pipeline to be used for the analysis of children fNIRS datasets. The performance of five different motion correction techniques, based on the principal component analysis and on the wavelet filtering, was evaluated by analyzing fNIRS data recorded in 22 typically developing children (mean age 11.4 years). The results showed that the wavelet analysis combined with a moving average filter achieved the best performance, suggesting that this technique might become a gold-standard approach for motion artifacts correction in fNIRS children's datasets.

Keywords: Functional Near-Infrared Spectroscopy · Children · Motion correction · Principal component analysis · Wavelet filtering

1 Introduction

Functional near-infrared spectroscopy (fNIRS) is an optical neuroimaging technique that measures the changes in tissue hemodynamic related to brain activities [1]. Thanks to its non-invasiveness, portability and ease of use, fNIRS is particularly suitable for studies that involve infants, children and hard-to-test clinical populations (e.g. [2–6]). Although the use of this technique has largely increased over the last decades, no standardized and widely accepted signal processing pipelines for fNIRS data are available and suboptimal data processing might affect the obtained results [7]. Only recently, some studies tried to address this issue evaluating different preprocessing methods using both simulated and real data [8–11]. Since, motion artifacts still remain one of the major problems for the interpretation of fNIRS data, these studies are mainly devoted to assess different motion correction techniques.

One of the most common ways to handle with motion artifacts is the rejection of the trials where they are detected [12]. However, this solution is not suitable when the number of motion artifacts is significantly higher than the number of trials, as it often

J. Henriques et al. (Eds.): MEDICON 2019, IFMBE Proceedings 76, pp. 235–244, 2020.
https://doi.org/10.1007/978-3-030-31635-8_28

happens in pediatric studies. Thus, other methods have been proposed. Some of them are based on complementary measures of motion artifacts that allow their direct removal (e.g. [13, 14]). Whereas, others methods consist in data processing techniques that correct the motion artifacts by using computational approaches (e.g. principal component analysis (PCA) [15], spline interpolation [1] and wavelet analysis [16]). These solutions are the most commonly used, since they can be applied to every experimental design and they do not require the use of additional instrumentations. However, fNIRS signal analysis remains largely data- and subjects-dependent and the lack of standardized methods is particularly problematic in pediatric research.

The present study aims at identifying an efficient pre-processing pipeline for the analysis of children fNIRS datasets. Particular attention is given to the evaluation of motion artifact correction methods.

Among these methods, PCA has been traditionally used in the literature as a signal denoising technique (e.g. [17, 18]) and approaches based on the wavelet filtering seem to be particularly promising [8–10]. Given this preliminary evidence, five different motion artifact correction techniques based on the PCA and on the wavelet analysis were used and compared.

2 Methods

2.1 NIRS Data Acquisition

Twenty-two typically developing children (21 males, mean age 11.4 ± 1.9) were recruited in the present study. All children were Caucasian, had normal or corrected-to-normal vision, no history of medical treatment and were not taking any medication. Written informed consent to participate to the study was obtained from all parents prior to the inclusion in the protocol; the study was approved by the Ethics Committee of the Scientific Institute IRCCS E. Medea (Bosisio Parini, LC, Italy).

The participants performed a computerized variant of the visuospatial n-back working memory task (n-back VSWM), modified from Cui and colleagues [19].

The paradigm consisted of three blocks: 0-Back (0B), 1-Back (1B) and 2-Back (2B), presented in the following order: rest – 0B – 1B – 2B – rest – 1B – 0B – 2B – rest – 2B – 1B – 0B – rest. The first rest period lasted 60 s, while the other three rests lasted 30 s each. During the rest, children passively viewed an image on a black screen. Each experimental epoch included 32 stimuli, lasting 0.5 s each, with a 1.5 s interstimulus interval. The stimulus was a clown's face displayed in different locations in a 3 × 3 matrix.

NIRS data were acquired using a continuous-wave NIRS device (NIRx DYNOT compact 9–32, NIRx Medizintechnik, Berlin, DE). 32 channels were used as result of the coupling of 8 emitters and 24 detectors. The channels were placed on the bilateral frontotemporal areas centered at F3 and F4, according to the International 10–20 system. The source-detector distance was 2.7 cm (Fig. 1). Near-infrared spectroscopy recording was performed at two wave lengths (760 and 830 nm).

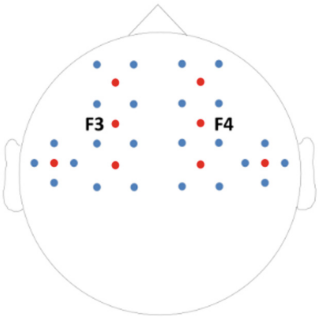

Fig. 1. Sources (red dots) and detectors (blue dots) locations on a simplified 10–20 system.

2.2 Data Pre-processing and Artifact Correction Techniques

NIRS data pre-processing was performed using the Homer2 v2.3 software [20]. Figure 2 depicts the processing pipeline.

First, the channels whose optical intensity signal (d) was too weak, too strong or had a high standard deviation (SD) were discarded (i.e., with mean(d) < 0.03 or > 3; SD(d)/mean(d) > 15). On average 2.6 channels (SD = 2.6, range = 0–9) per subject were rejected.

Then, the raw data were converted into changes in optical density data.

After the conversion, five different motion correction techniques were separately applied to the data. The algorithms used were based on the PCA and on the wavelet filtering.

Specifically, the PCA [15] was performed changing the percentage of variance removed. Two different values of threshold on the variance were used: 90% (PCA90) and 80% (PCA80), as suggested by Hu et al. [10], and Brigadoi et al. [9], respectively.

The third tested technique was the wavelet-based motion artifact correction approach proposed by Molavi and Dumont [16]. This technique is based on the hypothesis that the measured signal is a linear combination of the physiological hemodynamic signal and the artifacts, and that the detailed wavelet coefficients, obtained after the application of the discrete wavelet transform, have a Gaussian probability distribution. Because the hemodynamic response is slower and smoother than the motion artifacts, the distribution of the wavelet coefficients that represents the physiological components of the signal should be centered around zero with a low variance, whereas those corresponding to motion artifact should be the outliers of the probability distribution. The applied algorithm detected the outliers by means of a probability threshold and set them to zero before the wavelet reconstruction. In the present work, a threshold of 0.1 was used [8, 16].

The fourth technique applied consisted in the wavelet denoising described above preceded by a Moving Average filter (MA+Wavelet) performed over 5 s data windows [10].

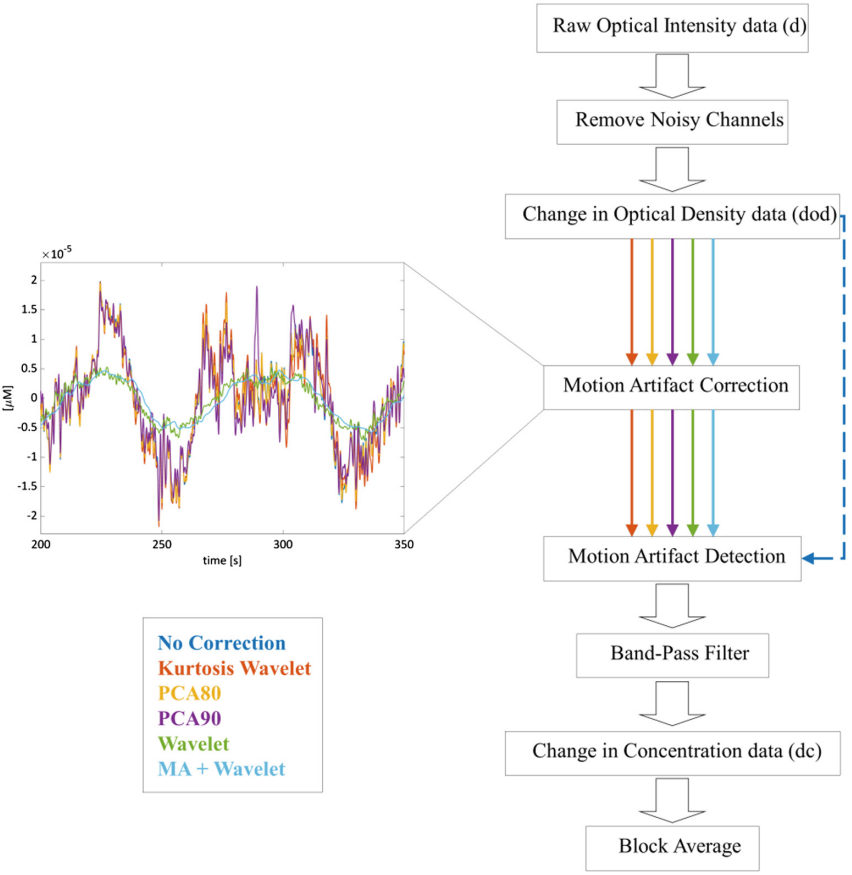

Fig. 2. Right panel: pre-processing pipeline of analysis. The dotted arrow refers to the processing stream without applying any technique for artifact correction (No Correction). Left panel: an example of the application of the different motion correction techniques used vs. uncorrected data is shown. The graph refers to dc data for one representative channel of one subject.

The last algorithm tested was the kurtosis-based wavelet filtering (Kurtosis Wavelet), which is an iterative motion correction method based on the evaluation of the fourth standardized moment (kurtosis) of the discrete wavelet transform coefficients distribution for the chosen decomposition level [21]. For this study, the kurtosis was set to 3.3 as suggested by Chiarelli and colleagues [21].

Finally, we performed an additional processing stream without applying any technique for artifact correction.

For each of the processing stream, the motion detection algorithm implemented in Homer2 was applied. The algorithm identifies the data-points that exceed a threshold in change of amplitude (AMPtresh = 0.6) or in change of standard deviation (STDEV-thresh = 50) within a predefined time interval (1 s).

Data were then filtered with a third order Butterworth 1 Hz low-pass filter, followed by a fifth order Butterworth 0.1 Hz high-pass filter, to further enhance the signal-to-noise-ratio.

Finally, the optical density data were converted into concentration changes (|HbO| and |HbR|) through the Modified Beer-Lambert Law. For the data conversion, the differential path length factor (DPF) was appropriately calculated accordingly to the age of the tested subjects [22]. Specifically, the DPF was set to 5.7 and 5.1 for the 760 nm and the 830 nm wavelength, respectively.

To obtain the task-related hemodynamic responses, data were divided in blocks accordingly to the task conditions (0 Back, 1 Back, and 2 Back) and averaged.

2.3 Metrics of Comparison

Five metrics of comparison were selected in order to quantitatively evaluate the performance of the different motion artifacts correction methods.

The first metric consisted is the percentage of data-points identified through the motion artifact detection algorithm implemented in Homer2 that was ran after the different motion correction techniques, in respect to the artifactual data-points identified in the uncorrected signals.

The second metric was the area under the curve (AUC) computed on the mean hemodynamic response function (HRF) for the first 1.5 s after stimulus onset ($AUC_{0–1.5}$). In particular, it can be assumed that the lower the AUC is, the better is the performance of the artifact correction method applied [9]. The 1.5 value was selected taking into account the time span in which the hemodynamic curve should become significant in children data [10].

The third metric was the ratio (AUC ratio) between the AUC of the mean HRF between 1.5 and 3 s after the stimuli presentation ($AUC_{1.5–3}$) and the $AUC_{0–1.5}$ reported above. This assumes that children hemodynamic response reaches its maximum between 1.5 and 3 s after the stimulus onset [10].

The fourth metric was the mean of the standard deviation (SD) of the hemodynamic data for each subject (within-subject SD). The assumption made is that the variability on the signal is mainly caused by motion artifacts [9].

The last metric used was the mean of the hemodynamic data across all subjects for a given channel (between-subjects SD). This index considers the variability between subjects [9].

For the last two metrics, it can be assumed that lower the standard deviation is, the better the motion artifacts correction technique is performed.

The above-mentioned metrics were computed for both HbO and HbR data, except for the percentage of corrected data-points (first metric), that was calculated on the optical density data.

2.4 Statistical Analysis

The statistical analysis was performed with the IBM SPSS, Statistics 21, software. For each metric, data were averaged across channels and conditions, thus obtaining a unique value for each subject. For between-subjects SD data were averaged across conditions and subjects, thus obtaining a unique value for each channel. The statistical analysis was done separately for HbO and HbR.

To check the distribution of each metric of comparison the Kolmogorov-Smirnov test was used. Since most of the variables showed a non-normal distribution ($p < .050$), the Friedman test was applied for each metric. Significant interaction effects were investigated using Wilcoxon pairwise comparisons.

3 Results

The median of the percentage of the motion artifact removed per method in respect to the number of artifacts identified in the uncorrected data is showed in Table 1. MA+Wavelet, Wavelet and PCA90 showed the best performances, while Kurtosis Wavelet removed less than 1% of the motion artifacts. The Friedman test identified a main effect of technique ($p < .001$) and post hoc comparisons showed statistical differences between Kurtosis Wavelet vs. MA+Wavelet, Wavelet and PCA90 ($p < .001$), PCA80 vs. MA+Wavelet ($p = .001$) and PCA80 vs. Wavelet ($p = .003$).

Table 1. Median, 25th and 75th percentiles of the percentage of motion artifacts removed by each technique in respect to the number of artifacts identified in the uncorrected data

	Median [%]	25th Perc. [%]	75th Perc. [%]
MA+Wavelet	100	100	100
Kurtosis Wavelet	0.75	0.34	1.23
Wavelet	100	100	100
PCA80	82.79	62.90	99.34
PCA90	100	96.27	100

A summary of the other metric results for HbO and HbR, is reported in Figs. 3 and 4. A main effect of technique was found for all the metrics (all Friedman test $p < .002$). The pattern of the results for HbO and HbR are similar.

Considering $AUC_{0–1.5}$, MA+Wavelet and Wavelet exhibited the lower values with low variability as well, compared to the other techniques and to uncorrected data (Fig. 3), as indicated by the Wilcoxon test. Specifically, MA+Wavelet showed $AUC_{0–1.5}$ lower values than the ones obtained with all the other techniques but Wavelet; Wavelet showed $AUC_{0–1.5}$ lower values than the ones obtained with Kurtosis Wavelet and PCA80.

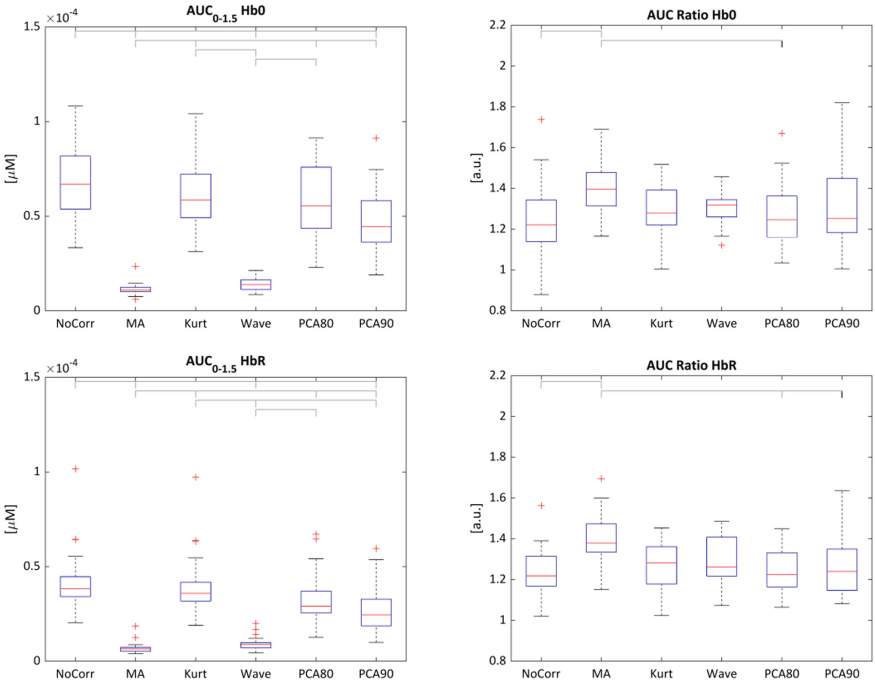

Fig. 3. Box plots of AUC ratio and $AUC_{0–1.5}$ computed for all techniques (from left to right: No correction, MA+Wavelet, Kurtosis Wavelet, Wavelet, PCA80 and PCA90) and for both HbO and HbR. The line in the box plot indicates the median, while the two extremities of the box plot represent the first and third quartile. Crosses indicate outliers. The lines linking the different techniques represent the significant statistical differences calculated with the Wilcoxon test (grey $p < .010$, black $p < .030$).

For AUC ratio, MA+Wavelet produced the highest values corresponding to a better performance (Fig. 3). The statistical analysis shows that only the MA+Wavelet method differed from the other techniques (No Correction, PCA80 and PCA90, the latter only for HbR).

The within-subject SD resulted to be lower in the data corrected with MA+Wavelet and Wavelet (Fig. 4). All techniques but the Kurtosis Wavelet significantly reduced the within-subject SD with respect to uncorrected data. Moreover, data corrected with MA +Wavelet and Wavelet had a significantly lower within-subject SD than data corrected with all other methods.

For the between-subject SD, the same pattern of statistically significant results was observed (Fig. 4).

Overall, the results showed that Wavelet and MA+Wavelet techniques outperformed the other motion artifact correction techniques with respect to the present dataset. However, the MA+Wavelet technique had better results for the AUC ratio metric compared to the Wavelet one.

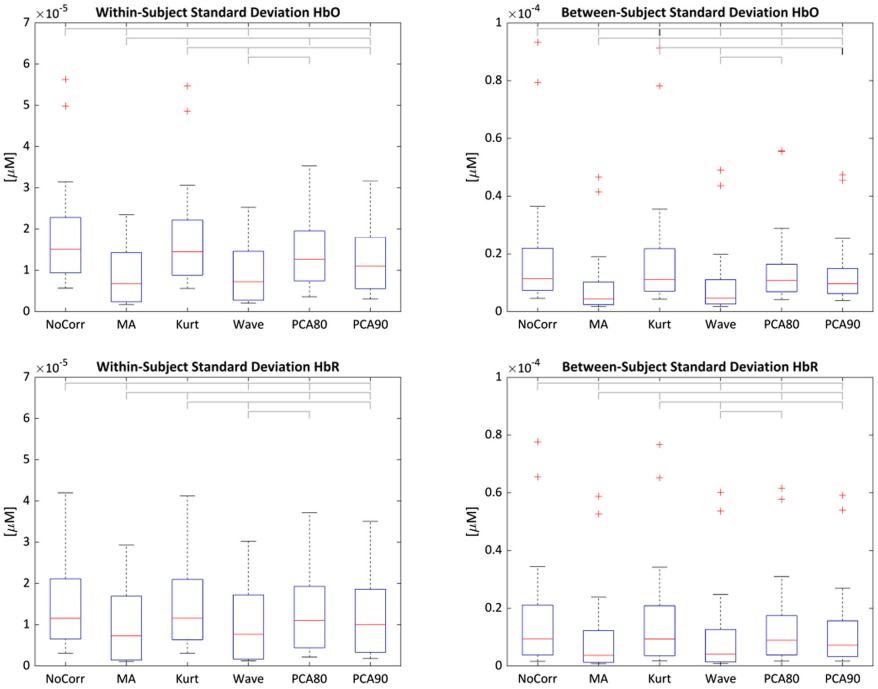

Fig. 4. Box plots of within-subject and between-subject SD computed for all techniques (from left to right: No correction, MA+Wavelet, Kurtosis Wavelet, Wavelet, PCA80 and PCA90) and for both HbO and HbR. The line in the box plot indicates the median, while the two extremities of the box plot represent the first and third quartile. Crosses indicate outliers. The lines linking the different techniques represent the significant statistical differences calculated with the Wilcoxon test (grey $p < .010$, black $p < .030$).

4 Discussion

The goal of this study was to identify the most efficient and reliable pipeline of analysis for the preprocessing of children fNIRS data. The performance of five different motion correction methods was evaluated by analyzing fNIRS data recorded in a group of children. Specifically, two PCA-based techniques and three wavelet-based methods were used. The methods' comparison was performed using five different metrics in order to qualitatively and quantitatively compare the performances of the motion correction techniques.

Previous studies reported that PCA effectively reduce motion artifact [8], although its performance heavily depends on the specific dataset [9]. In this study, we tested the PCA with two different variance thresholds. The PCA80 is expected to be more conservative and should identify only the most pronounced artifacts, whereas the PCA90 should remove more artifacts, with the risk of discharging also part of the brain hemodynamic response. As expected, our results showed a better performance of the PCA90 compared to the PCA80. In particular, the PCA80 obtained poor results in most

of the metrics, over performing only the uncorrected data in the within-subject SD of the HbR response and in the between-subject SD of both HbO and HbR responses. Conversely, the PCA90 showed good results in all the metrics, except for the AUC ratio. However, both the PCAs failed to efficiently remove the motion artifacts in a subset of participants. A possible reason for this failure might be that the PCA is a multi-channel approach. This assumes that the artifacts have to be present in multiple channels to be discharged, and this was not always the case for our dataset.

The Kurtosis Wavelet technique yielded minimal improvement in comparison to the non-corrected data and resulted to be the worst correction method among those tested. This is in contrast with the results obtained by Hocke and colleagues [11]. They showed that the Kurtosis Wavelet performed efficiently, although overcompensating the motion artifacts. The discrepancy between the two studies could be due either to the different metrics used for the method comparison or to the differences in the datasets analyzed (children vs. adults; n-back VSWM task vs. finger-tapping task).

As previously reported in the literature [8, 9], our results acknowledge the wavelet technique as the most promising method for artifact correction. Indeed, it significantly reduced the motion artifacts and obtained very good results in all the metrics of comparison, especially in the $AUC_{0–1.5}$ and in the two SD-based metrics.

The MA+Wavelet technique achieved a better performance than the other techniques considered. It efficiently removed all the identified motion artifacts for every subject and achieved good results in all the metrics of comparison. The present results support and expand the previous literature [10], indicating the reliability of this method. The main disadvantage of this technique is that the motion correction process must be applied twice and thus may cause an over-reduction of the magnitude of the signal.

Future investigations of this work should extend the application of the present approach to independent data-sets and to the study of different type of motion artifacts. Nevertheless, if replicated in larger datasets, the present findings could suggest a promising, standardized approach as the gold-standard method for motion artifacts correction in fNIRS children's datasets. This will allow to obtain a more reliable reconstruction of the hemodynamic response, thus contributing to a better understanding/interpretation of children's cognitive states. Moreover, replicability and result comparisons between different studies will be made easier and more consistent.

References

1. Scholkmann, F., et al.: A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. Neuroimage **85**, 6–27 (2014)
2. Gallagher, A., et al.: Near-infrared spectroscopy as an alternative to the Wada test for language mapping in children, adults and special populations. Epileptic Disord. **9**(3), 241–255 (2007)
3. Crippa, A., et al.: The utility of a computerized algorithm based on a multi-domain profile of measures for the diagnosis of attention deficit/hyperactivity disorder. Front. Psychiatry **8**, 189 (2017)

4. Mauri, M., et al.: Light up ADHD: I. cortical hemodynamic responses measured by functional near infrared spectroscopy (fNIRS): special section on translational and neur. J. Affect. Disord. **234**, 358–364 (2018)
5. Grazioli, S., et al.: Light up ADHD: II. neuropharmacological effects measured by near infrared spectroscopy: is there a biomarker? J. Affect. Disord. **244**, 100 (2019)
6. McDonald, N.M., et al.: Infant brain responses to social sounds: a longitudinal functional near-infrared spectroscopy study. Dev. Cogn. Neurosci. **36**, 100638 (2019)
7. Pfeifer, M.D., Scholkmann, F., Labruyère, R.: Signal processing in functional near-infrared spectroscopy (fNIRS): methodological differences lead to different statistical results. Front. Hum. Neurosci. **11**, 641 (2018)
8. Cooper, R., et al.: A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy. Front. Neurosci. **6**, 147 (2012)
9. Brigadoi, S., et al.: Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data. Neuroimage **85**, 181–191 (2014)
10. Hu, X., et al.: Comparison of motion correction techniques applied to functional near-infrared spectroscopy data from children. J. Biomed. Opt. **20**(12), 126003 (2015)
11. Hocke, L., et al.: Automated processing of fNIRS data—a visual guide to the pitfalls and consequences. Algorithms **11**(5), 67 (2018)
12. Selb, J.J., et al.: Effect of motion artifacts and their correction on near-infrared spectroscopy oscillation data: a study in healthy subjects and stroke patients. J. Biomed. Opt. **20**(5), 056011 (2015)
13. Robertson, F.C., Douglas, T.S., Meintjes, E.M.: Motion artifact removal for functional near infrared spectroscopy: a comparison of methods. IEEE Trans. Biomed. Eng. **57**(6), 1377–1387 (2010)
14. Virtanen, J., Noponen, T.E., Meriläinen, P.: Comparison of principal and independent component analysis in removing extracerebral interference from near-infrared spectroscopy signals. J. Biomed. Opt. **14**(5), 054032 (2009)
15. Zhang, Y., et al.: Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. J. Biomed. Opt. **10**(1), 011014 (2005)
16. Molavi, B., Dumont, G.A.: Wavelet-based motion artifact removal for functional near-infrared spectroscopy. Physiol. Meas. **33**(2), 259 (2012)
17. Wilcox, T., et al.: Using near-infrared spectroscopy to assess neural activation during object processing in infants. J. Biomed. Opt. **10**(1), 011010 (2005)
18. Zhang, X., et al.: Signal processing of functional NIRS data acquired during overt speaking. Neurophotonics **4**(4), 041409 (2017)
19. Cui, X., et al.: A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. Neuroimage **54**(4), 2808–2821 (2011)
20. Homer2. https://homer-fnirs.org/. Accessed Mar 2019
21. Chiarelli, A.M., et al.: A kurtosis-based wavelet algorithm for motion artifact correction of fNIRS data. Neuroimage **112**, 128–137 (2015)
22. Duncan, A., et al.: Measurement of cranial optical path length as a function of age using phase resolved near infrared spectroscopy. Pediatr. Res. **39**(5), 889–894 (1996)