Comparison of motion correction techniques applied to functional near-infrared spectroscopy data from children

Xiao-Su Hu
Maria M. Arredondo
Megan Gomba
Nicole Confer
Alexandre F. DaSilva
Timothy D. Johnson
Mark Shalinsky
Ioulia Kovelman
Comparison of motion correction techniques applied to functional near-infrared spectroscopy data from children

Xiao-Su Hu, Maria M. Arredondo, Megan Gomba, Nicole Confer, Alexandre F. DaSilva, Timothy D. Johnson, Mark Shalinsky, and Ioulia Kovelman

1 Introduction

Functional near-infrared spectroscopy (fNIRS) is an emerging optical brain imaging technique that is becoming an increasingly popular method for pediatric brain imaging. fNIRS measures the hemodynamic changes that effectively reflect brain activities occurring while people perform a wide range of mental tasks. It can provide both topographic and tomographic brain images. Specifically, fNIRS monitors concentration variations of oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) in the blood of the brain, through the skull, with absorption changes of near-infrared light at wavelengths between 700 and 1000 nm.

Similar to electroencephalography (EEG), fNIRS allows for noninvasive observations of awake infants and children, but unlike EEG, fNIRS provides complimentary hemodynamic response data. Compared with functional magnetic resonance imaging, fNIRS devices are usually portable, silent, and use near-infrared light to detect brain activity. These aspects of fNIRS technology make it a useful tool for researchers interested in child and infant development. Therefore, the use of this method with pediatric populations has grown threefold in the past decade.

Motion artifacts caused by participants’ head, jaw, eyebrow, or body movement can distort the data analysis. A major impediment of interpreting fNIRS brain imaging data of child participants is that children’s data typically have more motion artifacts than the fNIRS data of adults. The field has put forth a variety of methods for addressing the fNIRS motion artifact problem in general. The most direct method is trial rejection (for event-related design) or block rejection (for block design), which excludes all the data sections that contain motion artifacts from further analysis. Sib et al. compared several motion correction techniques on synthetic data and found trial rejection was the best method for preserving the original shape of the data. However, in the context of pediatric brain imaging, the major problem with this method is that the entire data samples tend to be relatively small, as children and infants have a short attention span and experimental designs can be as short as a few minutes. Thus, researchers have attempted methods that allow them to retain more of the collected data. These methods can be divided into two categories: the first category includes hardware-based methods (which need additional sensor or special optode setup), including short separation channel-based independent component analysis, using specially designed collodion-fixed fibers, accelerometer-based regression, or rescaling. The difficulty with these methods is that additional hardware may further complicate and increase participant setup time, which is always a problem with young participants. The second category includes software-based methods, such as principal...
component analysis (PCA),\textsuperscript{35} spline interpolation,\textsuperscript{36} wavelet,\textsuperscript{37} correlation-based signal improvement (CBSI),\textsuperscript{38} and moving average (MA), plus Kalman filter and recursive least-square based methods for real-time preprocessing.\textsuperscript{34,39}

There have been a few prior studies that have attempted to explore the relative efficacy of different artifact correction approaches using simulated adult fNIRS data. These studies concluded that the wavelet method was best suited for the simulated adult fNIRS data.\textsuperscript{8,29} Yet, it remains largely unknown whether the same is true of real pediatric fNIRS data from children subjects. Thus, in the present study, we aimed to improve the field’s methodology for child fNIRS brain imaging by directly comparing the efficacy of six software-based offline motion correction methods on real fNIRS data acquired from children (ages 6 to 12) completing a language task.

2 Methods

2.1 Participants

Twelve children (eight females, age range: 6.8 to 12.6 years, mean age = 9.9 years, SD = 1.75) took part in the study. Participants were excluded from the imaging analysis if they did not complete all the necessary tasks or did not fit this study’s inclusion criteria (see fNIRS analysis section below for more details). Since this paper discusses signal qualities, we also revealed the hair and skin color of all participants (Table 1, for more info on participants see Arredondo et al.\textsuperscript{46}) The University of Michigan Institutional Medical Review Boards reviewed and approved the study. Parents provided consent and children provided assent to participate in the study.

2.2 Materials and Procedure

The children completed an auditory grammatical judgment language task during the fNIRS data collection. The task was based on the Test of Early Grammatical Impairment,\textsuperscript{41} and it included three conditions: (1) control sentences with developmentally atypical errors (e.g., “Yesterday he bake a cake”), (2) experimental sentences with developmentally typical (optional infinitive) errors (e.g., “He is tallest in class”), and (3) correct sentences with no grammatical errors (e.g., “She is the prettiest cat”). A total of 60 trials were presented; each condition had 20 trials, and no trials were repeated. Children were instructed to respond by pressing buttons in a button-box. If a sentence did not contain any mistakes, the children were instructed to use their right hand to press the right button, and if the sentence contained mistakes, children were instructed to use their left hand to press the left button.

The task was designed to be rapid event-related, and it was randomized using OptSeq2.\textsuperscript{42} Jittered rest time in between trials varied, and it totaled 90 s. Each sentence played for \(\sim 4\) s, which was followed by a 2-s display of a question mark in the center of the screen that allowed time for the participant to respond. The children could respond any time between the start of the sentence, which was then followed by an additional response time period that presented a question mark in the middle of the screen. A fixation cross was displayed in the center of the screen between jittered rest periods. The approximate total time of the task was 456 s (7.6 min). The task was presented using E-Prime 2 (Psychology Software Tools) on a 23-in. Philips 230E wide LCD screen connected to a Dell Optiplex 780 desktop computer. Sound was played via a Creative Inspire T12 2.0 multimedia speaker system.

2.3 Functional Near-Infrared Spectroscopy Recordings

We used a TechEN-CW6 system with 690 and 830 nm wavelengths. The setup included one emitter and three detectors spaced 2.7 cm apart, yielding three data channels sampled at 10 Hz (Fig. 1). We examined brain activation in the left inferior frontal gyri (a language-related area). The probe localization was established and applied consistently for each participant using the international 10–10 transcranial system positioning;\textsuperscript{43} Fz, Cz, and preauricular were measured on the participants’ head using costume made caps (foam). Two caps were built to fit the different head sizes of the age range. Each cap was equipped with optode holders into

Table 1 Participants’ information (hair and skin colors).

<table>
<thead>
<tr>
<th>Participant</th>
<th>Hair color</th>
<th>Skin tone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>Medium</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>Light</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Light</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Light</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>Medium</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>Light</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>Light/medium</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>Light</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>Light</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>Light</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>Medium</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>Light</td>
</tr>
</tbody>
</table>

Hair color: 0, light blond; 1, dirty blond; 2, light brown; 3, medium brown; 4, dark brown; 5, light black; and 6, dark black.

![Fig. 1 Measurement location and channel distribution.](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics/126003-2/2015/December/20(12)/Hu-et-al.-Comparison-of-motion-correction-techniques-applied-to-functional-data/journalofbiomedicaloptics-126003-2.png)
which the optodes were plugged in. An additional wrapping band was used to secure the optodes in place.

2.4 Data Processing

Data processing was completed using Homer2 fNIRS processing package and an fNIRS data analysis tool (fNIRSDAT, a homemade software for general linear model (GLM) regression-based individual and group-level analysis) based on MATLAB (MathWorks, Natick, Massachusetts). In order to understand the performance of different methods, we categorized the motion artifacts in our data into four types: type A—defined as a spike with a standard deviation of 50 that was over the mean within a second; type B—a peak with a standard deviation of 100 from the mean during a time portion ranging from 1 to 5 s; type C—a gentle slope between 5 and 30 s with a standard deviation of 300 from the mean; type D—a slow baseline shifting longer than 30 s with a standard deviation of 500 from the mean.

The raw data were first converted into optical density change (delta OD) using the optical density change calculation function (hmrIntensity2D in Homer2 software package) in Homer2. We then slightly modified the motion artifact identification function (hmrMotionArtifactbyChannel in Homer2 software package) and applied it to the delta OD data to detect the number of the type A motion artifact. The parameters were selected as SDThresh = 50, tMotion = 1 (SDThresh, standard deviation threshold; tMotion, motion artifact time period); these parameters guaranteed the identification of the type A motion artifact (50 standard deviations to the mean within 1 s). However, the motion artifact identification function in Homer2 is only designed for identification of type A motion artifact (sudden move generated motion artifact). In this study, we also wanted to study the effect of type B, C, and D motion artifacts. Therefore, we used a home-made trend detection and detrending function using an MA algorithm to track the trend in fNIRS signal at different frequencies to identify type B, C, and D motion artifacts. This process also located the motion artifacts in the time series, thus providing the prior information for PCA and spline techniques.

We then applied five different motion correction techniques (PCA, spline, MA, wavelet, as well as MA and wavelet) to the delta OD series after the first round of motion artifact detection process (except CBSI method; this method needs HbO and HbR data for motion correction; therefore, a different processing stream was applied for the CBSI method). Then, the motion artifact detection process was applied again in order to count the number of motion artifacts of different types. The motion corrected delta OD data were further filtered with a third-order Butterworth low-pass filter of ∼0.5 Hz to prevent the physiological noise interference (e.g., cardiac noises) and high-frequency measurement noises; the filter coefficients are automatically designed by the MakeFilter function in MATLAB (coefficients: A: [1 −2.37 1.93 −0.53] B: [0.0029 0.0087 0.0087 0.0029]). Then the filtered OD data were converted to HbO and HbR concentration data by the modified Beer-Lambert law (MBLL) (different path length factor is selected to be 6 for both 690 and 830 nm wavelengths). Finally, the HbO and HbR data derived from previous steps went through a block averaging process and were regressed with a GLM. A flow-chart of the processing steps can be found in Fig. 2.

For CBSI method, we first applied a low-pass filter of (∼0.5 Hz) followed by the MBLL conversion to the delta OD data after the first round of motion detection process. The derived HbO and HbR data were then passed through the CBSI motion correction, and the processed HbO and HbR data were sent to the block averaging process and GLM regression, as well as converted back to delta OD data for the motion artifacts counting process (see Fig. 2).

2.5 Motion Artifacts Correction Methods

2.5.1 Spline

The spline interpolation method used in this paper is a channel-by-channel approach, first proposed by Scholkmann and coworkers. This method will only act on the motion artifacts detected from the previous step. This method was integrated in Homer2. This method uses a cubic spline function to model the detected motion artifact, and the resulting spline interpolation is then subtracted from the original signal to correct the motion artifact. The model of the spline function can be described using the following equation:

\[ Y_i(p) = a_i + b_ip + c_ip^2 + d_ip^3, \]

where \( a, b, c, \) and \( d \) are parameters, \( i \) is the time index, and \( p \) is a parameter \( p \in [0,1] \), which determines the degree of the spline function. The values of \( a, b, c, \) and \( d \) at different times are estimated using the cspars function in MATLAB. In Homer2, the spline interpolation depends on a parameter \( p \); if \( p = 0 \), the interpolation will be a straight line, while if \( p = 1 \), it will be a cubic interpolation. In this study, the parameter \( p \) was set to 0.99, the same value used by both Scholkmann et al. and Cooper et al.
2.5.2 Principal component analysis

PCA is a statistical procedure that uses an orthogonal transformation to convert a set of variables into a set of values of linearly uncorrelated variables called principal components. In fNIRS context, suppose we have correlated $N_{ch}$ channels and $N_t$ measured data time points. PCA is able to extract $N_{ch}$ linearly uncorrelated components from the original data set. The order of these components is related to the variance of the original data that the component accounts for. Thus, the first component will account for the largest proportion of the variance of the data. Since motion artifacts are often much larger in amplitude than normal physiological fNIRS signals, they should constitute a large proportion of the variance of the data. This process can be expressed as

$$t = w \cdot x,$$

where $x$ is the data matrix with dimension $N_{ch} \times N_t$, $w$ is the weight matrix with dimension $N_{ch} \times N_{ch}$, and $t$ is the extracted uncorrelated matrix with the same dimension as the data matrix. Hence, removing the several components with largest variance from the converted signal should correct for the motion artifacts in the original data. In Homer2, the PCA function requires one parameter that controls the variance to be removed. In this study, this parameter is selected to be 90%.

2.5.3 Wavelet

The wavelet-based motion artifact removal proposed by Molavi and Dumont is a channel-by-channel approach designed to correct for motion artifacts. Briefly, the one-dimensional discrete wavelet transform is applied to every channel data series for a number of levels of decomposition, given by the duration of the time series. Thus, the details of the signal on each level are estimated as approximation coefficients. This process can be expressed as

$$y(t) = \sum_{j=j_0}^{j} \sum_{k} w_{jk} \psi_{jk}(t),$$

where $y(t)$ is the signal, $\psi_{jk}(t) = 2^{j/2} \psi(2^j t - k)$ is the wavelet function (Daubechies 5 wavelet is used in Homer2, consistent with Ref. 37), $j$ and $k$ are the dilation and translation parameters, and $w_{jk}$ are the approximation and detail coefficients.

With the assumptions that the measured signal is a linear combination of the physiological signal of interest and the artifacts, that the detail wavelet coefficients have a Gaussian probability distribution, and that the hemodynamic response is smoother and slower than motion artifacts, the expectation is that the coefficients accounting for the evoked response will be centered around zero and with low variance, while the outliers in the Gaussian distribution are the coefficients accounting for the motion artifacts.

Therefore, setting these outlying coefficients to zero and then reconstructing the signal with the inverse discrete wavelet transform should remove the corresponding motion artifacts in the temporal time series. Outliers are detected using a probability threshold $\alpha$; if the probability of a given wavelet detail coefficient is less than $\alpha$, then this coefficient is assumed not to belong to the Gaussian distribution and is, hence, considered an outlier and set to zero. In Homer2, the parameter $\alpha$ is the tuning parameter of wavelet filtering. In this study, it was set to 0.1, the same value used by Cooper et al. and Molavi and Dumont.

2.5.4 Correlation-based signal improvement

CBSI is a channel-by-channel approach developed by Cui et al. to reduce motion artifacts caused by the movement of the head. It is based on the hypothesis that HbO and HbR should be negatively correlated during functional activation, but they become more positively correlated when a motion artifact occurs. The measured HbO and HbR signals, $x$ and $y$, respectively, can be described as

$$\begin{cases} x = x_0 + \alpha \cdot F + \text{Noise} \\ y = y_0 + F + \text{Noise} \end{cases},$$

where $x_0$ and $y_0$ are the true HbO and HbR signals to be estimated, $F$ is the motion artifact, with identical effects on both chromophores (but for a constant weighting $\alpha$), and Noise is the remaining high-frequency white noise, easily removed with a low-pass filter. To compute $x_0$ and $y_0$, two assumptions are required: the correlation between $x_0$ and $y_0$ should be close to -1 and the correlation between the artifact $F$ and the true signal $x_0$ should be close to 0. This leads to the following equations for the computation of the true HbO and HbR signals:

$$\begin{cases} x_0 = (x - \alpha \cdot y)/2 \\ y_0 = -(1/\alpha) \cdot x_0 \end{cases},$$

where $\alpha = \text{std}(x)/\text{std}(y)$, std is the standard deviation. The approach taken by Cui et al. also assumes that the ratio between HbO and HbR when no artifact is present is the same as when an artifact occurs. In this study, the CBSI method was implemented in Homer2 and no parameter was required for this method.

2.5.5 Moving average

MA is a calculation to analyze data points by creating a series of averages of different subsets on time course of the full data set. It is a type of finite impulse filter. In this study, the simple moving average (SMA) algorithm was applied to the data channel-by-channel. It was simply the calculated mean of the values before and after $n$ data with a given time point in the data time series from a channel. The process can be described as

$$\text{SMA} = \frac{P_{t-(n-1)} + \cdots + P_{t-1} + P_t + P_{t+1} + \cdots + P_{t+n}}{2n},$$

where $t$ is the time and $n$ is the number of points in the MA at every time point. In this study, $n$ is selected to be 25 (5 s data); thus, the resulting cut-off frequency was 0.064 Hz. In this study, the MA algorithm is designed for the low-frequency drift in the data. First, it is applied to each channel data to get the drift, and then the derived drift is subtracted from the original signal.

2.6 Metrics of Comparison

We defined five metrics in order to quantitatively evaluate the performance of the different motion correction methods.

The first metric is defined as the amount variation of different motion artifacts (the number of motion artifacts identified before
subtracted by the number of motion artifacts identified after). The numbers of four kinds of motion artifacts, identified before and after motion correction methods application, were calculated and compared for data on both 690- and 830-nm wavelengths.

The second and third metrics are the $R$ values and $t$ values generated during the GLM-based regression. The $R$ value is the residual of the GLM-based regression. It indicates the similarity between the predesigned model and the recorded data. Therefore, a smaller $R$ value indicates a better performance of a motion correction technique. A $t$ value is the ratio between the regression coefficient beta and the residual. An increasing $t$ value can be caused by either an increase of the beta value or a decrease of the residual, thus also indicating a promising performance of a motion correction method. In the current study, we expected the experimental conditions to have the highest responses in the language-related cortical area. Therefore, we only focused on the experimental condition relevant $t$ values.

The fourth and fifth metrics are the area under the averaged hemodynamic response curve between 0 and 1.5 s, and the ratio of the area under the curve between 1.5 and 3 s to the area under the curve between 0 and 1.5 s. These two metrics are important parameters of the block-averaged response. These are concepts similar to those adopted in Ref. 29. However, the participants in this study are children, not adults. Therefore, we assumed that the time period it takes for a hemodynamic curve to reach its peak is shorter than that of an adult; thus, we used 1.5 s instead of 2 s.

### 3 Results

The quantity variations of different motion artifacts are displayed in Fig. 3.

For type A motion artifacts, the motion correction methods we applied, which produced the best results for both 690 and 830 nm wavelengths, were wavelet, CBSI, and the wavelet and MA combination, while the MA and PCA methods produced poor results for both wavelengths. For type B motion artifacts, the wavelet method alone, as well as the wavelet and MA combination, attained the best results for both wavelengths. For type C motion artifacts, the spline and wavelet and MA combination each performed better than other methods for both wavelengths. For type D motion artifacts, spline, MA, and the wavelet and MA combination outperformed the other methods for the 690 nm data, while all methods achieved the same performance compared to each other in 830 nm data, yet wavelet and CBSI performed better.

The GLM-based regression results are presented in Fig. 4. The $t$ values (averaged across three channels) of the HbO signal

![Figure 3](https://example.com/f3.png)

**Fig. 3** The amount of variation of four types of motion artifacts after applying motion correction techniques.

* MC: motion correction
from all channels increased after we applied CBSI, wavelet, MA, and the wavelet and MA combination, but decreased after we applied PCA and spline. The \( t \) values of the HbR signal changed in accordance with the HbO signal, but with a weaker magnitude. The \( R \) values of both HbO and HbR decreased after we applied wavelet, MA, CBSI, and the wavelet and MA combination.

Figure 5 shows the AUC\(_{0.15} \) results of the HbO and HbR data. All of the methods we applied achieved similar performances, except the spline and wavelet methods. Wavelet performed better than the other methods with a 65% decrease in HbO data and 64% increase in HbR data. Spline presented a less promising performance, showing only a 46% decrease in HbO data and a 46% increase in HbR data. The AUC\(_{\text{ratio}} \) results are shown in Fig. 6. The MA method achieved the best performance for both HbO and HbR data, while the spline and CBSI methods performed worse than other methods for both HbO and HbR data.

4 Discussion

The goal of the study was to examine the efficiency of standard motion correction methods toward developmental fNIRS data. Specifically, we compared the performances of six motion correction techniques on fNIRS data derived from children 6 to 12 years old. The motion correction techniques we selected for comparison did not require extra equipment or a real-time data acquisition environment. Importantly, we did not consider the trial rejection method. The comparison metrics we selected aimed to qualitatively and quantitatively compare the performances of the motion correction abilities of these different techniques.

Previous studies reported that spline interpolation can achieve promising \(^{28,36,47}\) and less promising results.\(^{29}\) In theory, the spline technique can reduce all four types of motion artifacts equally. However, in our study, using this technique yielded minimal improvement in the data, in comparison to the data before motion correction techniques were applied in all metrics. The reason for this is that the spline technique uses the preidentified motion artifacts to subtract a generated cubic function. However, the motion artifact shapes vary in fNIRS data, thus causing unstable results. Vinette et al.\(^{47}\) tested the spline interpolation method on real long-term fNIRS data measured from patients with epilepsy. Their results showed spline interpolation method is good at reducing baseline shift motion artifacts. However, they also pointed out the limitations of the spline interpolation method, requiring different parameters for different participants and lower performance when brain response signal and motion artifact overlapped. Thus, in our study, the use of same parameters for all participants and possible overlapping between task evoked responses and motion artifacts may lead to the less promising motion correction results.

Cooper et al. showed that using PCA produced better results than using no motion correction technique,\(^{28}\) and Brigadoi et al. claimed the performance of PCA is variable depending on the data set.\(^{29}\) In the current study, we examined a data set acquired from children. The performance of the PCA technique in our study was not promising. Our data show that PCA successfully reduced type B, C, and D motion artifacts, but did not reduce type A motion artifacts. Although the PCA technique performed fairly well in the area under the curve (AUC)-related metrics, the GLM-based regression-related metrics decreased after applying PCA for both HbO and HbR. The reason for this may be that the PCA technique is a multichannel approach; thus, it requires that an artifact be presented in multiple channels (to be identified as a principal component), which is not always the case. The basic assumption of the PCA technique, that the components need to be independent, was violated. This may have additionally contributed to the poor performance we observed.

The CBSI technique performed well with our data. It decreased the number of all motion artifacts, except type B. The AUC\(_{0.15} \) and AUC\(_{\text{ratio}} \) metrics jointly showed an improvement in both averaged HbO and HbR data. The GLM-related metrics increased in \( t \) value, and also slightly decreased in residual for both HbO and HbR data. However, the results of CBSI can be unstable, especially because an assumption of the CBSI method is that the HbO and HbR data are always negatively correlated, which may not always be the case. According to our data, in most cases, when the HbO signal increases, the HbR signal decreases, but sometimes HbR increases or remains unchanged. Furthermore, this is a reason why fNIRS researchers have reported that the HbO signal is more reliable than the HbR signal in fNIRS studies.\(^{48}\)
In our study, the wavelet technique achieved a promising result, as it greatly reduced type A and type B motion artifacts, and barely reduced type C and type D motion artifacts. It also achieved very good results in two AUC-related metrics, and slightly increased the $t$ value and $R$ value for both HbO and HbR data. An advantage of the wavelet technique is that it does not require any prior motion artifact detection algorithm. Wavelet decomposes the signal into components on different time resolutions and then detects and fixes motion artifacts, thus eliminating the need for detection algorithms. Finally, the corrected component signals are restored to form the motion corrected signal. Brigadoi et al.\textsuperscript{29} concluded that the wavelet technique achieved the best results in their study. However, according to our data, wavelet showed a weaker ability in reducing type C and type D motion artifacts. Therefore, we introduced the MA technique to work jointly with the wavelet technique to enhance the motion correction results.

Our data showed that using MA accomplished decent performance on reducing type C and type D motion artifacts. It also achieved good results for all other four metrics. The problem with the MA method is that it has a weak ability to reduce type A and type B motion artifacts. Therefore, the MA method may not work independently when type A and type B motion artifacts are significantly dominant in the data.

The wavelet and MA methods independently yielded some of the best results in reducing type A/B and type C/D motion artifacts. Figure 5 shows the AUC\textsubscript{0.1.5} variation before and after applying different motion correction techniques. Figure 6 compares the AUC\textsubscript{0.0.3} before and after applying different motion correction techniques.
artifacts. Due to this, we applied wavelet and MA combined methods to our data to observe the cumulative results. As expected, the wavelet and MA combination method attained the best performance of reducing motion artifacts in our data. The method greatly reduced all four kinds of motion artifacts and acquired very good AUC-related metrics. The GLM regression $t$ value for HbR signal we obtained was not as high as CBSI method. The $R$ value we obtained was the smallest among all methods, indicating that this motion corrected data had the smallest deviation from the modeled data. As stated above, the reason for this is that the wavelet method and the MA method can, respectively, handle high-frequency and low-frequency noises. However, the disadvantage of using these combined methods is that the motion correction process must be applied two times, which may cause an overload of the magnitude of the signal and can lead to low regression coefficients and low $t$ values.

In general, the spline technique was fairly successful at reducing all types of motion artifacts, but leads to unsatisfactory results. This was possibly due to improper correction of motion artifacts. The PCA method reduced different motion artifacts less successfully than other methods in this study, thus leading to promising results in all metrics except AUC$_{0.1}$. The reason for this could be the absence of motion artifacts in the channels in our study. If so, PCA could not catch the motion artifacts as a principal component. CBSI demonstrated good results for reducing type A and type C motion artifacts, wavelet reduced type A and type B motion artifacts better than other methods, MA corrected type C and type D motion artifacts better than other methods, and the combination of wavelet and MA effectively reduced all types of motion artifacts. These latter four techniques achieved similar results in AUC$_{0.1}$, indicating they reduced the curve from 0 to 1.5 s. MA and wavelet combination and MA performed better than other methods for the AUC$_{ratio}$ metric. Interestingly, both MA and wavelet combination and MA achieved promising results for $t$ value and $R$ value metrics. This indicates that type C and type D motion artifacts affect GLM metrics more than type A and type B motion artifacts. Table 2 provides a performance summarization for all the methods applied in the current study.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Motion artifact reduction</th>
<th>General linear model (GLM) $t$ value</th>
<th>GLM $R$ value</th>
<th>AUC$_{0.1}$</th>
<th>AUC$_{1.5}$</th>
<th>AUC$<em>{1.5} - 3 / AUC</em>{0.1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principal component analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spline</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moving average (MA)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation-based signal improvement</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA and wavelet</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note: ✓ indicates the good performance in certain category.

5 Conclusion

In this study, we applied six different techniques on fNIRS data recorded from children completing a language task and compared the performance of these techniques. We compared the differences in amount of motion artifacts, AUC$_{0.1}$, AUC$_{ratio}$, GLM regression-based $t$ values and $R$ values before and after motion correction application. Our data revealed that the MA and wavelet combination produced the largest improvement in the computed metrics. We also found that type C and type D motion artifacts affect GLM-related metrics more than type A and type B motion artifacts.

Acknowledgments

The authors thank the University of Michigan Center for Human Growth and Development and the University of Michigan Office of Research. We also thank participating families and members of the Language & Literacy Laboratory for their help in data collection, analyses, and manuscript preparation.

References


Xiao-Su Hu is a research investigator at the INIRS laboratory at the Center for Human Growth and Development, University of Michigan. His research interests are methodology development for neuroimaging techniques (fNIRS & EEG) and their multidisciplinary applications.

Maria M. Arredondo is a PhD student in developmental psychology at the University of Michigan, and is a recipient of an NSF Graduate Research Fellowship. Her research interests include understanding bilingual children’s cognitive mechanisms as applied during language acquisition and academic achievement; her methods include using experimental and neuroimaging techniques (fMRI, fNIRS).

Alexandre F. DaSilva is the director of H.O.P.E. (Headache & Orofacial Pain Effort), which is a multidisciplinary collaborative team to investigate the brain as a research and therapeutic target for chronic trigeminal pain disorders. He is an assistant professor at the University of Michigan School of Dentistry. He has collaborated with his colleagues on novel neuroimaging techniques (PET and MRI-based, non-invasive brain stimulation (HD)-tDCS, mobile technology, and advanced 3D-neuronavigation projects, all related to trigeminal pain.

Timothy D. Johnson is professor of biostatistics at the University of Michigan, School of Public Health. His research interests are in Bayesian modeling, computational statistics, the interface of Bayesian methods and big data, and statistical image analysis. He received his PhD in biostatistics at UCLA in 1997. He is a fellow of the American Statistical Association.

Ioulia Kovelman is an assistant professor at the University of Michigan. She is the director of the Language & Literacy Laboratory in the Center for Human Growth and Development. She studies bilingual language and reading acquisition in typically developing children and those with learning impairments. She also works to improve fNIRS neuroimaging method for the study of child development.